

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
(Mestrado)

Método de Pontos Interiores com Corretor Quasi-Newton

Pedro Henrique da Silva Pinto

Orientador: Prof. Dr. Francisco Nogueira Calmon Sobral

Maringá - PR

2024

¹O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Método de Pontos Interiores com Corretor Quasi-Newton

Pedro Henrique da Silva Pinto

Dissertação apresentada ao Programa de Pós-Graduação em Matemática do Departamento de Matemática, Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito parcial para obtenção do título de Mestre em Matemática.

Área de concentração: Matemática Aplicada

Orientador: Prof. Dr. Francisco Nogueira
Calmon Sobral

Maringá - PR

2024

Dados Internacionais de Catalogação na Publicação (CIP)
(Biblioteca Setorial BSE-DMA-UEM, Maringá, PR, Brasil)

P659m Pinto, Pedro Henrique da Silva
Método de pontos interiores com corretor Quase-Newton / Pedro Henrique da Silva Pinto. -- Maringá, 2024.
vi 121 f. : il.

Orientador: Prof. Dr. Francisco Nogueira Calmon Sobral.
Dissertação (mestrado) - Universidade Estadual de Maringá, Centro de Ciências Exatas, Programa de Pós-Graduação em Matemática - Área de Concentração: Matemática Aplicada, 2024.

1. Métodos de pontos interiores. 2. Método de Broyden. 3. Método Quasi-Newton. 4. Preditor-corretor. I. Sobral, Francisco Nogueira Calmon, orient. II. Universidade Estadual de Maringá. Centro de Ciências Exatas. Programa de Pós-Graduação em Matemática - Área de Concentração: Matemática Aplicada. III. Título.

CDD 22.ed. 519.72

Edilson Damasio CRB9-1.123

PEDRO HENRIQUE DA SILVA PINTO

MÉTODO DE PONTOS INTERIORES COM CORRETOR QUASI-NEWTON

Dissertação apresentada ao Programa de Pós-Graduação em Matemática do Departamento de Matemática, Centro de Ciências Exatas da Universidade Estadual de Maringá, como parte dos requisitos necessários para a obtenção do título de Mestre em Matemática tendo a Comissão Julgadora composta pelos membros:

COMISSÃO JULGADORA:

Prof. Dr. Francisco Nogueira Calmon Sobral - UEM (Presidente)

Prof. Dr. Luiz Rafael dos Santos - UFSC

Prof. Dr. Leonardo Delarmelina Secchin - UFES

Aprovada em: 29 de fevereiro de 2024.

Local de defesa: Bloco F67 – Auditório do Departamento de Matemática.

AGRADECIMENTOS

Agradeço primeiramente a Deus por tudo o que Ele me proporcionou.

Agradeço aos meus pais por todo o amor e carinho, por sempre me apoiarem nos meus estudos e por contribuírem com a minha formação ao longo da vida.

Aos meus familiares pelo amor e carinho, por sempre estarem do meu lado e por torcerem por mim.

Agradeço ao meu orientador, Prof. Dr. Francisco Nogueira Calmon Sobral, por tudo o que me ensinou e por me guiar durante toda a minha jornada, desde os projetos de iniciação científica do ensino médio até o mestrado.

Aos meus professores, por seus ensinamentos e por sempre estarem dispostos a me ajudar com as minhas dúvidas.

Aos meus amigos do PMA pelo carinho e por me ajudarem em momentos difíceis realizando grupos de estudo, discussões, entre outras coisas.

Agradeço a CAPES pelo auxílio financeiro, possibilitando a minha dedicação exclusiva aos estudos.

RESUMO

Neste trabalho, foram estudados métodos de pontos interiores em problemas de programação linear com o objetivo de propor um algoritmo preditor-corretor que utilizasse uma abordagem Quasi-Newton, na expectativa de que isso tornasse o método mais barato computacionalmente. Para a elaboração deste algoritmo, primeiramente foi estudada a teoria dos métodos de pontos interiores. Em seguida, estudou-se a teoria de convergência do método de Broyden, um famoso método Quasi-Newton, para auxiliar na elaboração de uma teoria de convergência para o algoritmo proposto. Com base em dois métodos preditores-corretores encontrados na literatura, um algoritmo preditor-corretor é proposto no qual o passo preditor, de modo usual, realiza um passo de Newton, enquanto que o passo corretor realiza alguns passos Quasi-Newton. Inspirando-se na demonstração da convergência local do método de Broyden, provou-se a convergência local do algoritmo proposto. Uma implementação computacional deste algoritmo foi feita na linguagem Julia para que aspectos práticos do algoritmo pudessem ser observados. Foram realizados testes tanto em problemas pequenos quanto problemas grandes. Nos problemas pequenos, os resultados foram excelentes, tendo ocorrido a convergência para a solução em todos os problemas testados. Já em problemas grandes presentes na biblioteca NETLIB, o algoritmo também foi capaz de resolver a maioria deles, embora tenha tido um desempenho inferior a uma implementação computacional de um método preditor-corretor fornecida pelo pacote da linguagem Julia chamado Tulip.

Palavras-chave: Método de Pontos Interiores, Método de Broyden, preditor-Corretor, Quasi-Newton

ABSTRACT

In this work, it was studied interior point methods in linear programming problems with the objective of proposing a predictor-corrector algorithm that uses a Quasi-Newton approach. In order to elaborate this algorithm, firstly the theory of interior point methods was studied. Then, the local convergency theory of the Broyden's method, a famous Quasi-Newton method, was studied to assist in the elaboration of a local convergency theory for the proposed algorithm. Based in two predictor-corrector algorithms found in the literature, a predictor-corrector algorithm was proposed, in which the predictor step, as usual, performs one Newton step, while the corrector step performs a few Quasi-Newton steps. Taking inspiration from the proof of the local convergence of the Broyden's method, it was proved the local convergence of the proposed algorithm. A computational implementation of this algorithm was made in Julia language so that practical aspects of the algorithm could be observed. Tests were made both in small problems and big problems. In small problems, the results were excellent, where convergence to the solution has occurred in all problems tested. In problems from the NETLIB test set, the algorithm was also able to solve most of them, although it performed worse than a computational implementation of a predictor-corrector algorithm provided by a Julia language package named Tulip.

Key words: Interior Point Methods, Broyden's Method, Predictor-Corrector, Quasi-Newton

CONTEÚDO

Introdução	1
1 Métodos de Pontos Interiores Primais-Duais em Problemas de Programação Linear	8
1.1 O Caminho Central	11
1.2 Boa Definição do Caminho Central	13
1.3 Vizinhanças do Caminho Central	20
1.4 Algoritmo de Passos Longos	21
1.5 Algoritmo Preditor-Corretor	30
1.6 Algoritmo Preditor-Corretor de Mehrotra	36
2 Método de Broyden	42
2.1 Convergência do Método de Broyden	43
3 Método Preditor-Corretor Newton-Quasi-Newton	62
3.1 Boa Definição do Algoritmo 6	64
3.2 Distância entre um Ponto Viável e um Ponto do Caminho Central	71
3.3 Convergência Local do Algoritmo Proposto	75
4 Implementação do Algoritmo 7	86
4.1 Implementação em Problemas Pequenos	91
4.2 Implementação Eficiente do Método de Broyden	99
4.3 Implementação em Problemas da NETLIB	102
5 Conclusões	116

Referências Bibliográficas

119

INTRODUÇÃO

Considere o problema a seguir:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.a:} \quad & Ax = b \\ & x \geq 0, \end{aligned} \tag{1}$$

em que $c, x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ e $A \in \mathbb{R}^{m \times n}$ é uma matriz com posto linha completo. Problemas como este são chamados problemas de programação linear. De acordo com [1], o matemático Dantzig propôs, em meados da década de 1940, um algoritmo chamado Simplex para resolver problemas de programação linear. Este método geralmente é muito eficiente para a maioria dos problemas, executando cerca de $2m$ à $3m$ iterações para alcançar a solução (ver [2], pág. 388). Contudo, em [3], Klee e Minty mostraram um exemplo no qual o método Simplex precisou executar 2^n iterações para alcançar a solução, verificando que sua complexidade era exponencial.

O mal desempenho do Simplex em problemas patológicos como o apresentado por Klee e Minty motivou outros pesquisadores a desenvolverem métodos com uma performance garantidamente melhor. Em [4], o matemático Karmarkar apresentou um método novo, o qual foi verificado ter complexidade polinomial. O artigo de Karmarkar foi o estopim para uma onda de pesquisas sobre uma nova classe de métodos conhecida como métodos de pontos interiores. Entre os algoritmos de pontos interiores, temos os métodos primais, duais, e primais-duais. Uma década após a publicação do artigo de Karmarkar, os algoritmos primais-duais se mostraram os mais importantes e úteis na classe de métodos de pontos interiores (ver o prefácio de [1]).

Considere as seguintes condições:

$$\begin{aligned}
 A^T \lambda + s &= c, \\
 Ax &= b, \\
 x_i s_i &= 0, \quad i = 1, 2, \dots, n, \\
 (x, s) &\geq 0.
 \end{aligned} \tag{2}$$

As condições acima são chamadas condições de Karush–Kuhn–Tucker, ou simplesmente, condições KKT. As condições KKT são condições necessárias e suficientes para que um dado ponto seja uma solução de um problema de programação linear [2]. Os métodos de pontos interiores primais-duais procuram explorar essas condições para encontrar uma solução para o problema (1). O primeiro fato relacionado a essas condições que é explorado por esses métodos é a existência de uma curva denominada caminho central, denotada por \mathcal{C} , que é formada pelos pontos $(x_\tau, \lambda_\tau, s_\tau)$ tais que

$$\begin{aligned}
 A^T \lambda + s_\tau &= c, \\
 Ax_\tau &= b, \\
 (x_\tau)_i (s_\tau)_i &= \tau, \quad i = 1, 2, \dots, n, \\
 (x_\tau, s_\tau) &\geq 0,
 \end{aligned} \tag{3}$$

em que $\tau > 0$. As sequências de pontos geradas pelos métodos de pontos interiores primais-duais procuram manter-se próximas ao caminho central. O caminho central, por sua vez, guia os iterandos do método de pontos interiores em direção à solução conforme τ se aproxima de zero. Em cada passo do método, é fixado um parâmetro $\sigma \in (0, 1)$, o qual tem como função regular a proximidade dos iterandos com relação ao caminho central. O parâmetro σ é chamado parâmetro centralizador. Além do parâmetro σ , fixa-se $\mu > 0$ e depois é considerada uma função $F_{\sigma\mu} : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ tal que $F_{\sigma\mu}(x, \lambda, s) = 0$ se e somente se as seguintes condições são satisfeitas:

$$\begin{aligned}
 A^T \lambda + s &= c, \\
 Ax &= b, \\
 x_i s_i &= \sigma\mu, \quad i = 1, 2, \dots, n.
 \end{aligned} \tag{4}$$

Dentre os métodos de otimização que existem para a resolução de problemas do tipo $f(x) = 0$, onde $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma função continuamente diferenciável, temos o método de Newton. Geralmente, os métodos de pontos interiores utilizam o método de Newton para a obtenção das direções. No entanto, o método de Newton não é executado até a

convergência para uma solução de $F_{\sigma\mu}(x, \lambda, s) = 0$. Ao invés disso, uma única iteração de Newton é executada. Neste caso, encontra-se a direção d dada pelo sistema

$$J(x, \lambda, s) d = -F_{\sigma\mu}(x, \lambda, s), \quad (5)$$

onde J representa a Jacobiana de $F_{\sigma\mu}$, e depois executa-se um passo nessa direção. A direção d obtida pelo sistema de Newton também é chamada de *direção afim-escala*. Desse modo, o próximo ponto obtido irá satisfazer (4) apenas de forma aproximada. Na iteração seguinte, os valores de μ e σ são alterados, visando sempre a redução de μ . Alguns algoritmos de pontos interiores tomam pontos iniciais viáveis, tanto para a formulação primal quanto dual, embora existam algoritmos que não exijam a viabilidade desse ponto. Nosso interesse é estudar os métodos que utilizam pontos viáveis. Sendo assim, para o caso viável, é exigido que o ponto inicial satisfaça

$$\begin{aligned} Ax &= b \\ A^T \lambda + s &= c \\ (x, s) &> 0. \end{aligned} \quad (6)$$

Depois, procura-se manter essas mesmas condições válidas para os próximos iterandos. Em [2] (veja o Teorema 14.3), prova-se que locomover-se na direção de Newton associada a (5) não altera as condições de viabilidade. Sendo assim, os algoritmos de pontos interiores procuram manter a positividade estrita para os próximos pontos tomando passos de Newton menores, caso necessário. Dentro de determinadas hipóteses, é possível mostrar que este procedimento gera uma sequência de pontos viáveis (no sentido primal-dual) $w_k = (x_k, \lambda_k, s_k)$ satisfazendo $(x_k, s_k) > 0$ tal que a medida de complementaridade $\mu_k = \frac{x_k^T s_k}{n} \rightarrow 0$, e conseqüentemente, esses pontos se aproximam cada vez mais de satisfazer as condições KKT do problema (1).

Para resolver um problema do tipo $f(x) = 0$ pelo método de Newton, é necessário calcular a Jacobiana de f em cada iteração. Uma outra forma de resolver tais problemas é através da utilização de métodos Quasi-Newton. Os métodos Quasi-Newton evitam o cálculo de Jacobianas fazendo a construção de aproximações para essas matrizes (ver [2], pág. 279). Essas aproximações são atualizadas a cada iteração de modo que elas imitem o comportamento da Jacobiana verdadeira.

Em [5], considerou-se problemas de programação quadráticos e foi proposta uma mudança na abordagem usual dos métodos de pontos interiores. Esta mudança consistiu em

substituir o método de Newton por métodos Quasi-Newton. Foi observado que o custo computacional gerado por essa mudança se compara ao custo de calcular um corretor em um método de pontos interiores usual. Através de experimentos numéricos, observou-se que essa estratégia foi capaz de reduzir o número de fatorações de matrizes.

Em [6], foi estudada a complexidade dos métodos de pontos interiores Quasi-Newton em problemas de programação linear. O algoritmo utilizado alternava entre passos de Newton e passos Quasi-Newton. A complexidade foi estudada em ambos os casos em que se admite ou não pontos inviáveis. Em ambos os casos, foi provado que a complexidade é polinomial, onde o grau dos polinômios nos resultados de complexidade são maiores em comparação aos resultados de complexidade relativos ao método de Newton.

Em [7], o método de Newton foi trocado por um método Quasi-Newton chamado método de Broyden, e a convergência local do método de pontos interiores com essa modificação foi provada. Também foram consideradas versões não suaves das condições KKT e uma variação não suave do método de Broyden, assim como resultados de convergência associados a esses métodos no contexto de pontos interiores. Ainda, foi proposta uma modificação envolvendo o método de Broyden em um algoritmo que utiliza métodos locais em sua operação. Os experimentos realizados mostraram que houve uma melhora no desempenho de convergência global para o problema testado.

Neste trabalho, é estudada a teoria por trás dos métodos de pontos interiores primais-duais, em especial métodos preditores-corretores. Esses métodos executam dois tipos de passos, o passo preditor e o passo corretor. No passo preditor, escolhe-se $\sigma = 0$ em $F_{\sigma\mu}$, depois calcula-se a direção de Newton d dada pelo sistema (5), e em seguida, move-se nessa direção. A principal vantagem do passo preditor é que ele promove uma redução considerável de μ . Contudo, o novo ponto tende a se afastar do caminho central, o que é indesejável, pois quanto mais próximo o ponto está do caminho central, maiores poderão ser os passos preditores sem que a restrição $(x, s) \geq 0$ seja violada (ver [2], pág. 408).

Aproximar o ponto do caminho central é exatamente o que o passo corretor procura realizar. No passo corretor, resolve-se o sistema (5) para $\sigma = 1$, obtendo uma direção d , e em seguida o passo nessa direção é realizado. Essa escolha de σ implica que o valor de μ não será alterado (ver [1], pág. 92), no entanto a direção d possibilita encontrar um novo ponto mais próximo do caminho central. Ao executar os passos preditor e corretor de forma alternada, os métodos preditores-corretores aproveitam das vantagens de ambos

os passos.

O mais famoso algoritmo preditor-corretor é o algoritmo preditor-corretor de Mehrotra [8]. Uma das vantagens desse método sobre os métodos primais-duais usuais, além do fato de existir um corretor que aproxima os pontos do caminho central, é que ele possui uma escolha inteligente de σ que se adapta ao longo das iterações (ver [1], pág. 14). De forma grosseira, esta escolha permite que o passo corretor promova reduções consideráveis de μ quando o ponto está próximo do caminho central, enquanto que também permite aproximar o ponto do caminho central, quando este está longe. Este algoritmo será abordado em mais detalhes na Seção 1.6 do Capítulo 1.

Outro método que aplica correções de centralidade é o método de múltiplos corretores de centralidade [9]. Este método toma como base algoritmos primais-duais inviáveis, modificando-os para aplicar múltiplas correções com a finalidade de melhorar a centralidade do ponto atual (isto é, aproximá-lo do caminho central). De acordo com [9], o uso de correções múltiplas acarreta em uma redução de 25% à 40% do número de iterações em comparação com métodos preditores-corretores de segunda ordem largamente utilizados.

O objetivo do presente trabalho foi propor um algoritmo preditor-corretor no qual o passo preditor executasse, de modo usual, um passo na direção afim-escala, enquanto que o passo corretor executasse uma ou mais iterações de um método Quasi-Newton. Como mencionado anteriormente, em [5] foi feita uma substituição do método de Newton por um método Quasi-Newton. O método Quasi-Newton utilizado foi o método de Broyden em sua versão “ruim”. O Lema 2 de [5] mostra um fato interessante sobre uma direção d_k encontrada após k iterações deste método. Considerando especificamente os problemas de programação linear e a função $F_{\sigma\mu}$ apresentada anteriormente, o Lema 2 de [5] diz que dado $v \in \mathbb{R}^{2n+m}$, calcular uma direção $d_k = H_k v$, onde H_k^{-1} foi obtida pelo método de Broyden “ruim” após l atualizações e $H_{k-l} = J(x_{k-l}, \lambda_{k-l}, s_{k-l})^{-1}$, é equivalente a resolver um sistema da forma

$$J(x_{k-l}, \lambda_{k-l}, s_{k-l})d_k = v + (0, 0, v_3),$$

onde $v_3 \in \mathbb{R}^n$ (ainda, o Lema 2 de [5] fornece uma fórmula para v_3). Ou seja, apenas a terceira componente do vetor v precisa ser modificada para calcular a direção do método de Broyden “ruim” em uma determinada iteração k . Como é apontado em [5], esta estrutura se assemelha muito aos métodos preditores-corretores ou ainda aos múltiplos corretores de centralidade, o que, de acordo com [5], reforça o argumento de que o custo de computar

um passo Quasi-Newton é menor do que o passo de Newton.

Este resultado motivou a escolha do método de Broyden em nosso trabalho como o método Quasi-Newton a ser utilizado no passo corretor do algoritmo que desejamos propor. Com base nos algoritmos preditores-corretores de [2] (Algoritmo 14.3, pág. 411) e [1] (Algoritmo PC, pág. 93) e também na teoria de convergência do método de Broyden encontrada em [10] (Capítulo 8, Seção 8.2), um algoritmo preditor-corretor com corretor Quasi-Newton foi proposto, e sua convergência local foi demonstrada. Uma implementação do algoritmo proposto foi feita na linguagem de programação Julia, e experimentos numéricos foram realizados em problemas pequenos e problemas maiores. Foi possível resolver todos os problemas pequenos testados. Já os problemas maiores, ocorreram alguns problemas envolvendo a convergência do método de Broyden e a presença de Jacobianas não inversíveis por conta de seu mal condicionamento. Nos casos em que o método de Broyden não convergiu, foi aplicado um método alternativo, enquanto que as Jacobianas não inversíveis foram corrigidas para que fosse possível as inverter.

Foi escolhido um pacote da linguagem Julia chamado Tulip [11] para que seus resultados pudessem ser comparados com o algoritmo proposto. Este pacote implementa o algoritmo preditor-corretor de Mehrotra [8] em conjunto aos múltiplos corretores de centralidade [12]. O algoritmo proposto foi capaz de resolver a maioria dos problemas com essas modificações, embora tenha tido um desempenho inferior em comparação ao Tulip.

O trabalho está organizado da seguinte forma: no Capítulo 1, o método de pontos interiores primal-dual será formalmente apresentado, bem como alguns conceitos básicos relacionados a ele, como por exemplo a medida de dualidade μ , o parâmetro centralizador σ , o caminho central \mathcal{C} , incluindo suas vizinhanças mais conhecidas $\mathcal{N}_2(\theta)$ e $\mathcal{N}_{-\infty}(\gamma)$. Ainda nesse capítulo, serão apresentados os Algoritmos de Passos Longos [2] e preditores-corretores de ambos [2] e [1], e seus resultados de convergência. No Capítulo 2, o método de Broyden será apresentado. Além disso, uma sequência de resultados preliminares será provada para auxiliar na demonstração da convergência local do método de Broyden, sendo este o principal resultado do capítulo. No Capítulo 3, uma primeira versão do algoritmo proposto será exibida. Também serão apresentados resultados que justificam seus passos, alguns resultados auxiliares e por fim a demonstração de sua convergência local. Depois, será mostrado que com uma pequena modificação neste algoritmo, preserva-se a convergência local, enquanto que permite-se a execução de um número finito de passos

de Broyden. No Capítulo 4, serão realizados experimentos em problemas pequenos e problemas maiores utilizando uma implementação do algoritmo proposto feita na linguagem Julia. Também será apresentado um resultado de utilidade prática envolvendo a resolução de sistemas no método de Broyden, com o intuito tornar as implementações mais eficientes de um ponto de vista computacional. O algoritmo proposto será comparado com o algoritmo Tulip, um algoritmo de pontos interiores que implementa os múltiplos corretores de centralidade na linguagem Julia. Por fim, as conclusões serão apresentadas no Capítulo 5.

MÉTODOS DE PONTOS INTERIORES PRIMAIS-DUAIS EM PROBLEMAS DE PROGRAMAÇÃO LINEAR

Considere o seguinte problema de programação linear na forma canônica, isto é,

$$\begin{aligned} \min \quad & c^T x \\ \text{s.a:} \quad & Ax = b \\ & x \geq 0, \end{aligned} \tag{1.1}$$

em que $c, x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ e $A \in \mathbb{R}^{m \times n}$ é uma matriz com posto linha completo. O problema (1.1) está relacionado ao seguinte problema ([1], pág 3):

$$\begin{aligned} \max \quad & b^T \lambda \\ \text{s.a:} \quad & A^T \lambda + s = c \\ & s \geq 0, \end{aligned} \tag{1.2}$$

em que $\lambda \in \mathbb{R}^m$ e $s \in \mathbb{R}^n$. O problema (1.2) é chamado problema dual associado a (1.1), enquanto que o problema (1.1) é nomeado primal para distingui-lo do dual. Existe uma teoria de dualidade que descreve a relação entre os problemas primais e duais (esta teoria pode ser encontrada no Capítulo 6 de [13]). Os métodos de pontos interiores primais-duais procuram resolver simultaneamente os problemas (1.1) e (1.2), isto é, encontrar um ponto $(x^*, \lambda^*, s^*) \in \mathbb{R}^{2n+m}$ que soluciona ambos os problemas primal e dual.

As soluções dos problemas (1.1) e (1.2) são caracterizadas pelas condições abaixo:

$$\begin{aligned} A^T \lambda + s &= c, \\ Ax &= b, \\ x_i s_i &= 0, \quad i = 1, 2, \dots, n, \\ (x, s) &\geq 0. \end{aligned} \tag{1.3}$$

As condições (1.3) são chamadas condições de Karush-Kuhn-Tucker, ou simplesmente, KKT. As condições KKT são uma consequência direta do Teorema A.1 de [1] aplicado sobre o problema primal. Os métodos primais-duais procuram por soluções para os problemas (1.1) e (1.2) aplicando variações do método de Newton sobre as igualdades presentes nas condições KKT, tomando sempre o cuidado de escolher direções e tamanho de passo adequados para que $(x, s) > 0$ em cada iteração.

Considere a seguinte função:

$$F : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$$

$$F(x, \lambda, s) = \begin{bmatrix} A^T \lambda + s - c \\ Ax - b \\ XSe \end{bmatrix}, \quad (1.4)$$

$$\text{onde } X = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{bmatrix}, S = \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_n \end{bmatrix} \text{ e } e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Note que se $F(x, \lambda, s) = 0$, temos as igualdades presentes nas condições (1.3).

Para auxiliar a medir o quão desejável é um ponto $(x, \lambda, s) \in \mathbb{R}^{2n+m}$, temos a **medida de complementaridade**, μ , que é definida como a média dos produtos $x_i \cdot s_i$, $i = 1, 2, \dots, n$, isto é,

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i s_i = \frac{x^T s}{n}. \quad (1.5)$$

Quando um ponto (x, λ, s) é viável tanto para (1.1) quanto para (1.2), um critério de otimalidade conhecido da teoria de programação linear é que $c^T x = b^T \lambda$ (ver [13], Corolário 6.1), ou de forma equivalente, que $c^T x - b^T \lambda = 0$. A diferença $c^T x - b^T \lambda$ é chamada gap de dualidade. Para pontos viáveis para os problemas primal e dual, tal medida pode ser calculada da seguinte forma:

$$c^T x - b^T \lambda = x^T c - b^T \lambda = x^T (A^T \lambda + s) - (Ax)^T \lambda = x^T A^T \lambda + x^T s - x^T A^T \lambda = x^T s.$$

Portanto, ao definir μ como em (1.5), temos que μ nada mais é do que uma fração do gap de dualidade, significando que μ pode ser usado para saber se o método está convergindo.

Suponha que $(x, \lambda, s) \in \mathbb{R}^{2n+m}$. A direção de busca $(\Delta x, \Delta \lambda, \Delta s) \in \mathbb{R}^{2n+m}$ no método de pontos interiores é a direção de Newton, também chamada de direção afim-escala,

obtida resolvendo o seguinte sistema:

$$J(x, \lambda, s) \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = -F(x, \lambda, s),$$

em que J é a Jacobiana de F . Substituindo a Jacobiana de F , temos

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} r_c \\ r_b \\ r_\mu \end{bmatrix}, \quad (1.6)$$

em que $r_c = c - A^T \lambda - s$, $r_b = b - Ax$ e $r_\mu = -XSe$.

Segundo [2], não é incomum que um passo completo nessa direção viole $(x, s) \geq 0$. Portanto, é feito o teste da razão nesta direção, a procura de um $\alpha \in (0, 1]$ tal que o ponto $(x, \lambda, s) + \alpha(\Delta x, \Delta \lambda, \Delta s)$ satisfaça a condição $(x, s) > 0$. Outro problema recorrente é que nem sempre é possível tomar α muito próximo de 1 sem que a condição $(x, s) > 0$ seja violada. Isso acaba fazendo com que passos extremamente pequenos sejam dados. Uma estratégia adotada pela maioria dos métodos primais-duais para contornar este problema é a seguinte: ao invés de mirar diretamente na solução, mirar em um ponto cuja média dos produtos $x_i s_i$, $i = 1, \dots, n$, é menor que a média no ponto atual. Basicamente, um passo de Newton é dado na direção de um ponto no qual $x_i s_i = \sigma \mu$, para algum parâmetro $\sigma \in [0, 1]$, onde μ é a medida de complementaridade do ponto atual. Isso é equivalente a definir F da seguinte forma:

$$F_{\sigma\mu} : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$$

$$F_{\sigma\mu}(x, \lambda, s) = \begin{bmatrix} A^T \lambda + s - c \\ Ax - b \\ XSe - \sigma\mu e \end{bmatrix}. \quad (1.7)$$

Depois resolver o seguinte sistema:

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe + \sigma\mu e \end{bmatrix}. \quad (1.8)$$

Por fim, encontrar $\alpha \in (0, 1]$ de forma análoga ao que era feito anteriormente, isto é, realizando uma busca linear com o objetivo de calcular um ponto novo da forma $(x, \lambda, s) +$

$\alpha(\Delta x, \Delta \lambda, \Delta s)$, cujo coeficiente de dualidade é menor. Os algoritmos que adotam esta estratégia são chamados Algoritmos Primais-Duais Seguidores de Caminho. O parâmetro σ é chamado parâmetro centralizador, e pode ser interpretado como a taxa de redução que deseja-se obter na medida de complementaridade ao aplicar este passo. A seguir, o Algoritmo 1 apresenta a estrutura geral dos métodos primais-duais seguidores de caminho.

Algoritmo 1 Primal-Dual Seguidor de Caminho

1: Dado (x_0, λ_0, s_0) , com $(x_0, s_0) > 0$;

2: **para** $k \leftarrow 1$ **até** n **faça**

3: Escolha $\sigma_k \in [0, 1]$ e resolva

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S_k & 0 & X_k \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \\ \Delta s_k \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -X_k S_k e + \sigma_k \mu_k e \end{bmatrix}, \quad (1.9)$$

onde $\mu_k = \frac{(x_k)^T s_k}{n}$;

4: Defina

$$(x_{k+1}, \lambda_{k+1}, s_{k+1}) = (x_k, \lambda_k, s_k) + \alpha_k (\Delta x_k, \Delta \lambda_k, \Delta s_k), \quad (1.10)$$

escolhendo α_k de modo que $(x_{k+1}, s_{k+1}) > 0$.

5: **fim para**

1.1 O Caminho Central

Considere os seguintes conjuntos:

$$\begin{aligned} \mathcal{F} &= \{(x, \lambda, s) \mid Ax = b, A^T \lambda + s = c, (x, s) \geq 0\} \\ \mathcal{F}_0 &= \{(x, \lambda, s) \mid Ax = b, A^T \lambda + s = c, (x, s) > 0\} \end{aligned} \quad (1.11)$$

Dizemos que \mathcal{F} é o conjunto viável e \mathcal{F}_0 o conjunto estritamente viável. O caminho central \mathcal{C} é um conjunto de pontos estritamente viáveis que satisfazem as seguintes condições, para algum $\tau > 0$:

$$\begin{aligned} A^T \lambda + s &= c, \\ Ax &= b, \\ x_i s_i &= \tau, \quad i = 1, 2, \dots, n, \\ (x, s) &> 0. \end{aligned} \quad (1.12)$$

Na verdade, é possível mostrar que existe um único ponto do caminho central para cada $\tau > 0$ fixado se, e somente se, $\mathcal{F}_0 \neq \emptyset$ (veja a Seção 1.2 deste capítulo). Como consequência, podemos parametrizar o caminho central em função de τ , podendo assim defini-lo como

$$\mathcal{C} = \{(x_\tau, \lambda_\tau, s_\tau) \in \mathcal{F}_0 \mid x_i s_i = \tau, \tau > 0\}. \quad (1.13)$$

Ao reescrever o problema (1.1) com uma formulação de barreiras logarítmicas, as condições (1.12) passam a ser suas condições de otimalidade.

Mais especificamente, considere o seguinte problema:

$$\begin{aligned} \min \quad & c^T x - \tau \sum_{i=1}^n \ln x_i \\ \text{s.a:} \quad & Ax = b. \end{aligned} \quad (1.14)$$

As condições KKT do problema (1.14) podem ser escritas da seguinte forma:

$$\begin{aligned} c_i - \frac{\tau}{x_i} &= A_i^T \lambda, \quad i = 1, 2, \dots, n, \\ Ax &= b, \end{aligned} \quad (1.15)$$

onde $\lambda \in \mathbb{R}^m$ é o multiplicador de Lagrange da restrição de igualdade. Mostraremos mais adiante, no Lema 1.1, que (1.15) implica (1.12). Como a função objetivo de (1.14) é estritamente convexa, essas condições são necessárias e suficientes para a otimalidade (ver [2], pág. 398).

Note que quando τ se aproxima de 0, as equações (1.12) se aproximam de (1.3). Desse modo, o caminho central pode ser seguido para manter x e s positivos e reduzir os produtos $x_i s_i$, $i = 1, 2, \dots, n$.

No caso em que $(x, \lambda, s) \in \mathcal{F}$, temos que $r_b = 0$ e $r_c = 0$, e portanto o sistema (1.8), o qual resolvemos para determinar a direção de Newton, terá a seguinte forma:

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -XSe + \sigma \mu e \end{bmatrix}. \quad (1.16)$$

O parâmetro σ tem um papel muito importante sobre a direção de Newton. Por exemplo, se $\sigma = 1$, o passo de Newton estará tentando se aproximar do ponto do caminho central \mathcal{C} cuja medida de complementaridade é μ . Isso normalmente não reduz μ significativamente, mas de acordo com [2], permite dar um passo maior na próxima iteração, afim de obter um decréscimo razoável. Direções onde $\sigma = 1$ são chamadas direções centralizadoras. Por

outro lado, se $\sigma = 0$, o passo de Newton tentará alcançar diretamente a solução. Muitos algoritmos de pontos interiores tentam se aproveitar dessa propriedade centralizadora de σ , escolhendo valores intermediários no intervalo $(0, 1)$ na tentativa de equilibrar a centralidade e a redução da medida μ .

1.2 Boa Definição do Caminho Central

Nesta seção, será mostrado que o caminho central está bem definido e é dado por (1.13). Desse modo, existe uma parametrização $f_c : \mathbb{R}_+ \rightarrow \mathbb{R}^{2n+m}$ para o caminho central. Além disso, mostraremos que f_c é contínua. Os resultados apresentados nesta seção e suas demonstrações basearam-se em [7]. A continuidade de f_c será explorada mais adiante na Seção 3.1.

Para todo $\tau > 0$ fixado, considere o problema (1.14). Analisando suas condições de otimalidade, temos o seguinte lema:

Lema 1.1. *Um vetor $x^* \in \mathbb{R}^n$ é solução de (1.14) se, e somente se, existem $\lambda^* \in \mathbb{R}^m$ e $s^* \in \mathbb{R}^n$ tais que x^* , λ^* , e s^* satisfazem*

$$\begin{aligned} A^T \lambda^* + s^* &= c \\ Ax^* &= b \\ X^* S^* e &= \tau e \\ (x^*, s^*) &> 0. \end{aligned} \tag{1.17}$$

Demonstração. Para problemas de minimização com restrições de igualdade [14], temos que se x^* é solução de (1.14), então ao considerar $f(x) = c^T x - \tau \sum_{i=1}^n \ln x_i$, podemos expressar o gradiente de f da seguinte forma:

$$\nabla f(x^*) = A^T \lambda^*,$$

para algum $\lambda^* \in \mathbb{R}^m$. Como $\nabla f(x^*) = c - \tau(X^*)^{-1}e$, segue que

$$c - \tau(X^*)^{-1}e = A^T \lambda^* \Leftrightarrow A^T \lambda^* + \tau(X^*)^{-1}e = c \Leftrightarrow A^T \lambda^* + s^* = c,$$

onde $s^* = \tau(X^*)^{-1}e$. Então, para todo $i = 1, \dots, n$, temos que

$$s_i^* = \tau(x_i^*)^{-1} \Rightarrow x_i^* s_i^* = \tau \Rightarrow X^* S^* e = \tau e.$$

Além disso, $x^* > 0$, pela definição de f . Logo, $s_i^* > 0$, para todo $i = 1, \dots, n$, pois $\tau > 0$.

Como x^* é viável, $Ax^* = b$. Portanto, as condições do enunciado se verificam.

Por outro lado, se x^*, λ^* e s^* satisfazem as condições (1.17), temos que

$$X^* S^* e = \tau e \Rightarrow x_i^* s_i^* = \tau \Rightarrow s_i^* = \tau (x_i^*)^{-1}, \forall i = 1, \dots, n,$$

pois $x > 0$. Sendo assim, $\nabla f(x^*) = c - \tau (X^*)^{-1} e = c - s^* = A^T \lambda^*$, e portanto x^* satisfaz a condição de primeira ordem para o problema de minimização com restrições lineares de igualdade. Além disso, a hessiana de f ,

$$\nabla^2 f(x^*) = \begin{bmatrix} \tau x_1^{-2} & 0 & \cdots & 0 \\ 0 & \tau x_2^{-2} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \tau x_n^{-2} \end{bmatrix}$$

é definida positiva, pois se $y = (y_1, \dots, y_n)^T \neq 0$, temos que

$$y^T \nabla^2 f(x^*) y = \begin{bmatrix} \tau x_1^{-2} y_1 & \cdots & \tau x_n^{-2} y_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^n \tau x_i^{-2} y_i^2 > 0.$$

Em particular, $\nabla^2 f(x^*)$ é definida positiva em vetores no núcleo de A , e portanto x^* é minimizador de (1.14). \square

Dado $\tau > 0$, o Lema 1.1 nos dá uma correspondência entre a solução do problema (1.14) e um ponto do caminho central. Basicamente, podemos utilizar a solução do problema (1.14) para encontrar um ponto no caminho central que satisfaz a condição $XSe = \tau e$. O contrário também é verdade, isto é, se tivermos um ponto do caminho central tal que $XSe = \tau e$, então sabemos determinar a solução de (1.14).

No estudo de métodos de pontos interiores primais-duais, é natural assumir a existência de pontos interiores no conjunto viável do problema primal-dual considerado. No entanto, a existência de tais pontos é garantida ao assumir que vale a condição de Slater (ver [1], pág. 165) para ambos os problemas primal e dual. Seguirá da condição de Slater para o problema primal que existem pontos satisfazendo $Ax = b$ e $x > 0$. Analogamente, assumindo esta condição para o problema dual, garantimos que existem pontos satisfazendo $A^T \lambda + s = c$ e $s > 0$. Uma consequência de se assumir as condições de Slater para ambos os problemas primal e dual será que $\mathcal{F}_0 \neq \emptyset$, o que motiva o próximo lema.

Lema 1.2. Dado $\tau > 0$, existem únicos $x^* \in \mathbb{R}^n$, $\lambda^* \in \mathbb{R}^m$ e $s^* \in \mathbb{R}^n$ satisfazendo

$$\begin{aligned} A^T \lambda^* + s^* &= c \\ Ax^* &= b \\ X^* S^* e &= \tau e \\ (x^*, s^*) &> 0 \end{aligned} \tag{1.18}$$

se, e somente se, $\mathcal{F}_0 \neq \emptyset$.

Demonstração. Se existem x^*, λ^*, s^* satisfazendo as condições (1.18), então $\mathcal{F}_0 \neq \emptyset$, pois $A^T \lambda^* + s^* = c$, $Ax^* = b$ e $(x^*, s^*) > 0$. Por outro lado, se $\mathcal{F}_0 \neq \emptyset$, então existe $(x_0, \lambda_0, s_0) \in \mathcal{F}_0$. Para $\tau > 0$ dado, considere o problema (1.14), onde $f_\tau : W \rightarrow \mathbb{R}$, dada por $f_\tau(x) = c^T x - \tau \sum_{i=1}^n \ln x_i$, é a função objetivo de (1.14), e $W = \{x \in \mathbb{R}^n \mid x_i > 0, \forall i = 1, \dots, n\}$.

Dado $d \in \mathbb{R}$, queremos mostrar que $\{x \in W \mid f_\tau(x) \leq d\}$ é limitado. Primeiramente, note que como $c = A^T \lambda^* + s^*$, podemos escrever

$$\begin{aligned} f_\tau(x) &= (A^T \lambda^* + s^*)^T x - \tau \sum_{i=1}^n \ln x_i = (\lambda^*)^T Ax + (s^*)^T x - \tau \sum_{i=1}^n \ln x_i \\ &= (\lambda^*)^T b + (s^*)^T x - \tau \sum_{i=1}^n \ln x_i = (\lambda^*)^T b + \sum_{i=1}^n (s_i^* x_i - \tau \ln x_i) \\ &= (\lambda^*)^T b + \tau \sum_{i=1}^n h_i(x_i), \end{aligned}$$

onde $h_i(x_i) = s_i^* x_i - \tau \ln x_i$, para todo $i = 1, \dots, n$. Note que

$$\frac{dh_i}{dx_i}(x_i) = s_i^* - \frac{\tau}{x_i} = 0 \Leftrightarrow x_i = \frac{\tau}{s_i^*},$$

para todo $i = 1, \dots, n$. Além disso,

$$\frac{d^2 h_i}{dx_i^2}(x_i) = s_i^* + \frac{\tau}{x_i^2} > 0,$$

e portanto $x_i = \frac{\tau}{s_i^*}$ é um minimizador de $h_i, \forall i = 1, \dots, n$.

Se $x_i > \frac{\tau}{s_i^*}$, temos que $\frac{dh_i}{dx_i}(x_i) > 0$, e conseqüentemente, h_i é crescente em $\left(\frac{\tau}{s_i^*}, +\infty\right)$. Agora, se $x_i < \frac{\tau}{s_i^*}$, então $s_i^* < \frac{\tau}{x_i}$, e portanto h_i será decrescente em $\left(0, \frac{\tau}{s_i^*}\right)$. Em outras palavras, a função h_i é limitada inferiormente, para todo $i = 1, \dots, n$, e conforme x_i se afasta do minimizador de h_i , h_i cresce. Note ainda que cada h_i é ilimitada superiormente.

Agora suponha, por absurdo, que dado $d \in \mathbb{R}$, o conjunto $\{x \in W \mid f_\tau(x) \leq d\}$ seja ilimitado, ou equivalentemente, que não exista uma bola centrada na origem que o contenha. Seja $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$, onde \bar{x}_i é o minimizador de h_i , para todo $i = 1, \dots, n$.

Seja $B(0, \frac{R}{2})$ uma bola centrada na origem contendo \bar{x} . Como $\{x \in W \mid f_\tau(x) \leq d\}$ é ilimitado, tome uma sequência de pontos x_k nesse conjunto tal que $x_k \notin B(0, kR)$, onde $k = 1, 2, \dots$. Note que $\|x_k\| \rightarrow \infty$, e cada $[x_k]_i$ se afasta de \bar{x}_i , o minimizador de h_i . Neste caso, temos que

$$\sum_{i=1}^n h_i([x_k]_i) \rightarrow \infty$$

implicando que $f_\tau(x_k) \rightarrow \infty$, e portanto, para algum $k_0 \in \mathbb{N}$, $f_\tau(x_{k_0}) > d$, contradizendo o fato de que $x_{k_0} \in \{x \in W \mid f_\tau(x) \leq d\}$. Portanto, o conjunto $\{x \in W \mid f_\tau(x) \leq d\}$ precisa estar contido em uma bola centrada na origem, isto é, ele é limitado. Em particular, $\{x \in W \mid f_\tau(x) \leq f_\tau(x_0)\}$ é limitado.

Note que f_τ é contínua, pois é soma de funções contínuas. Como $(-\infty, f_\tau(x_0)]$ é fechado em \mathbb{R} , segue que

$$\{x \in W \mid f_\tau(x) \leq f_\tau(x_0)\} = f_\tau^{-1}((-\infty, f_\tau(x_0)])$$

é fechado em W . Mostremos que tal conjunto é fechado em \mathbb{R}^n . De fato, seja $x' = (x'_1, x'_2, \dots, x'_n)$ um ponto no fecho de $f_\tau^{-1}((-\infty, f_\tau(x_0)])$ em \mathbb{R}^n tal que $x' \notin W$. Então existe $i \in \{1, \dots, n\}$ tal que $x'_i = 0$. Como x' pertence ao fecho de $f_\tau^{-1}((-\infty, f_\tau(x_0)])$, então existe uma sequência de pontos $y_k = (y_1^k, \dots, y_n^k) \in f_\tau^{-1}((-\infty, f_\tau(x_0)])$ tais que $y_k \rightarrow x'$. Ou seja, $y_i^k \rightarrow 0$. Neste caso, $h_i(y_i^k) \rightarrow +\infty$, e como h_j é limitada inferiormente, para todo $j = 1, 2, \dots, n$, segue que

$$\sum_{i=1}^n h_i(y_i^k) \rightarrow +\infty,$$

implicando que $f_\tau(y_k) = (\lambda^*)^T b + \tau \sum_{i=1}^n h_i(y_i^k) \rightarrow +\infty$. Neste caso, existiria $k_0 \in \mathbb{N}$ tal que $f_\tau(y_{k_0}) > f_\tau(x_0)$, contradizendo a hipótese.

Portanto, $x' \in W$ e consequentemente, o fecho de $f_\tau^{-1}((-\infty, f_\tau(x_0)])$ no \mathbb{R}^n coincide com o seu fecho em W . Uma vez que $f_\tau^{-1}((-\infty, f_\tau(x_0)])$ é fechado em W , este conjunto deve coincidir com seu fecho em W . Neste caso, $f_\tau^{-1}((-\infty, f_\tau(x_0)])$ é fechado no \mathbb{R}^n . Por também ser limitado, segue que $f_\tau^{-1}((-\infty, f_\tau(x_0)])$ é compacto.

Considere o conjunto $K = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0, f_\tau(x) \leq f_\tau(x_0)\}$. Note que $K = \{x \in \mathbb{R}^n \mid Ax = b\} \cap \{x \in \mathbb{R}^n \mid x \geq 0\} \cap \{x \in W \mid f_\tau(x) \leq f_\tau(x_0)\}$. Considere $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que $g(x) = Ax$, que é linear e portanto contínua. Temos que $\{x \in \mathbb{R}^n \mid Ax = b\} = g^{-1}(\{b\})$, e portanto $\{x \in \mathbb{R}^n \mid Ax = b\}$ é fechado. Além disso, $\{x \in W \mid x \geq 0\}$ é fechado e $\{x \in \mathbb{R}^n \mid f_\tau(x) \leq f_\tau(x_0)\}$ também é fechado. Portanto

K é fechado. Como $K \subset \{x \in W \mid f_\tau(x) \leq f_\tau(x_0)\}$, o qual é compacto, segue que K é compacto.

Considere $f_\tau|_K : K \rightarrow \mathbb{R}$, a qual é contínua, pois é restrição de uma função contínua. Como K é compacto, segue que existe $x^* \in K$ tal que

$$f_\tau(x) \geq f_\tau(x^*), \quad \forall x \in K.$$

Seja $x \in W$ tal que $x \notin K$. Como $x_0 \in K$, uma vez que $f_\tau(x_0) \leq f_\tau(x_0)$ e $(x_0, \lambda_0, s_0) \in \mathcal{F}_0$, temos que

$$f_\tau(x) > f_\tau(x_0) \geq f_\tau(x^*),$$

isto é, x^* é um minimizador de $f_\tau : W \rightarrow \mathbb{R}$. Além disso, $x^* \in K$ implica que $Ax^* = b$, isto é, x^* é solução de (1.14), considerando $\tau > 0$ dado. Pelo Lema 1.1, existem $\lambda^* \in \mathbb{R}^m$ e $s^* \in \mathbb{R}^n$ satisfazendo (1.17).

Provemos que x^* , λ^* e s^* são únicos. Como (1.14) é um problema estritamente convexo, segue que (1.14) tem no máximo uma solução, e portanto x^* é único. Note que $X^*S^*e = \tau e \Leftrightarrow x_i^*s_i^* = \tau, \forall i = 1, \dots, n$, e portanto $s_i = \frac{\tau}{x_i^*}, \forall i = 1, \dots, n$. Como x^* é único, temos que s^* é unicamente determinado por x^* , e portanto s^* é único. Uma vez que A tem posto linha completo, segue que AA^T é inversível. Sendo assim, temos da igualdade $A^T\lambda^* + s^* = c$ que

$$AA^T\lambda^* = A(c - s^*) \Rightarrow \lambda^* = (AA^T)^{-1}A(c - s^*).$$

Neste caso, λ^* é unicamente determinado por s^* , sendo portanto único. Logo, x^* , λ^* e s^* são únicos. \square

Quando $\mathcal{F}_0 \neq \emptyset$, o Lema 1.2 nos permite concluir que existe uma correspondência biunívoca entre $\tau \in \mathbb{R}_+$ e um ponto $(x_\tau, \lambda_\tau, s_\tau) \in \mathcal{C}$. O próximo lema nos dá condições para que a Jacobiana de F_τ seja inversível.

Lema 1.3. *Considere a matriz*

$$\begin{bmatrix} 0 & A^T & I_n \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix},$$

onde $A \in \mathbb{R}^{m \times n}$ possui posto linha completo, $I_n \in \mathbb{R}^{n \times n}$ é a matriz identidade, $S, X \in \mathbb{R}^{n \times n}$ são matrizes diagonais, com elementos da diagonal principal estritamente positivos. Então essa matriz é não singular.

Demonstração. Considere

$$S = \begin{bmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_n \end{bmatrix}, \quad X = \begin{bmatrix} x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n \end{bmatrix} \quad \text{e} \quad M = \begin{bmatrix} 0 & A^T & I_n \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix}.$$

Provemos que $\ker M = \{0\}$, isto é, $Mx = 0 \Rightarrow x = 0$. Considere então o sistema

$$\begin{bmatrix} 0 & A^T & I_n \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Do último bloco do sistema, temos que

$$\begin{aligned} Su + Xw &= 0 \\ \Leftrightarrow Su &= -Xw \\ \Leftrightarrow u &= -S^{-1}Xw. \end{aligned} \tag{1.19}$$

Do primeiro bloco de matrizes, temos que

$$\begin{aligned} A^T v + w &= 0 \\ \Leftrightarrow A^T v &= -w. \end{aligned} \tag{1.20}$$

Do segundo bloco de matrizes, temos que $Au = 0 \Leftrightarrow u^T A^T = 0^T$. Novamente pelo primeiro bloco,

$$\begin{aligned} A^T v + w &= 0 \\ \Rightarrow u^T A^T v + u^T w &= 0 \\ \Rightarrow 0^T v + u^T w &= 0 \\ \Rightarrow u^T w &= 0 \\ \sum_{i=1}^n u_i \cdot w_i &= 0. \end{aligned} \tag{1.21}$$

De (1.19), temos que $u_i = -s_i^{-1}x_i w_i$, $\forall i = 1, \dots, n$, e portanto, (1.21) se reescreve da seguinte forma:

$$-\sum_{i=1}^n \frac{x_i}{s_i} \cdot w_i^2 = 0.$$

Como cada parcela $\frac{x_i}{s_i} \cdot w_i^2 \geq 0$, temos que

$$\frac{x_i}{s_i} \cdot w_i^2 = 0, \quad \forall i = 1, \dots, n,$$

e conseqüentemente, $w_i = 0$, para todo $i = 1, \dots, n$. Como $u_i = -s_i^{-1}x_iw_i$, $\forall i = 1, \dots, n$, segue que $u = 0$. Além disso, como A tem posto linha completo, segue que as colunas de A^T são linearmente independentes. Neste caso, temos de (1.20) que $-w$ é uma combinação linear das colunas de A^T , onde $-w = 0$. Isso só pode acontecer se $v = 0$. Sendo assim, u , v e w são nulos, implicando que $\ker M = \{0\}$. Deste modo, M é inversível, e portanto não singular. \square

Um importante teorema da matemática é o Teorema da Função Implícita. Este teorema nos permitirá definir uma parametrização contínua do caminho central. A seguir, será enunciada uma versão desse teorema disponível em [15] (Teorema 2.87), a qual será suficiente obter tal parametrização contínua. Uma versão mais geral do Teorema da Função Implícita pode ser encontrada no livro [16] (Teorema 9.2).

Teorema 1.4 (Teorema da Função Implícita). *Seja $F : \mathbb{R}^k \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ uma função de classe C^1 . Suponha que $F(x_0, y_0) = 0$ e*

$$\det \left[\frac{\partial F}{\partial y}(x_0, y_0) \right] \neq 0.$$

Então, existe um aberto $W \subset \mathbb{R}^k$ e $\phi : W \rightarrow \mathbb{R}^m$ função de classe C^1 tais que

- (a) $x_0 \in W$ e $\phi(x_0) = y_0$;
- (b) $F(x, \phi(x)) = 0, \forall x \in W$.

Suponha que $\mathcal{F}_0 \neq \emptyset$. Pelo Lema 1.2, para cada $\tau > 0$, existe um único vetor $(x_\tau, \lambda_\tau, s_\tau)$, com $x_\tau \in \mathbb{R}^n$, $\lambda_\tau \in \mathbb{R}^m$ e $s_\tau \in \mathbb{R}^n$, tal que

$$\begin{aligned} A^T \lambda_\tau + s_\tau &= c \\ Ax_\tau &= b \\ X_\tau S_\tau e &= \tau e \\ (x_\tau, s_\tau) &> 0. \end{aligned}$$

Portanto, existe uma função $f_c : \mathbb{R}_+ \rightarrow \mathbb{R}^{2n+m}$ tal que $f_c(\tau) = (x_\tau, \lambda_\tau, s_\tau)$. Mostremos que f_c é contínua. Considere $G : \mathbb{R}_+ \times \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ tal que

$$G(\tau, (x, \lambda, s)) = \begin{bmatrix} A^T \lambda + s - c \\ Ax - b \\ XSe - \tau e \end{bmatrix}.$$

Para $\tau_0 > 0$ fixado, como $\mathcal{F}_0 \neq \emptyset$, existe $(x_{\tau_0}, \lambda_{\tau_0}, s_{\tau_0})$ satisfazendo (1.17), e portanto,

$$G(\tau_0, (x_{\tau_0}, \lambda_{\tau_0}, s_{\tau_0})) = 0. \text{ Além disso, } \frac{\partial G}{\partial(x, \lambda, s)}(\tau_0, (x_{\tau_0}, \lambda_{\tau_0}, s_{\tau_0})) = \begin{bmatrix} 0 & A & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix}, \text{ que é}$$

não singular pelo Lema 1.3, e portanto $\det \frac{\partial G(\tau_0, (x_{\tau_0}, \lambda_{\tau_0}, s_{\tau_0}))}{\partial(x, \lambda, s)} \neq 0$. Consequentemente, pelo Teorema da Função Implícita, temos que existe um aberto $W_{\tau_0} \subset \mathbb{R}_+$ e uma função de classe C^1 $\phi_{\tau_0} : W_{\tau_0} \rightarrow \mathbb{R}^{2n+m}$ tal que

- (a) $\tau_0 \in W_{\tau_0}$ e $\phi_{\tau_0}(\tau_0) = (x_{\tau_0}, \lambda_{\tau_0}, s_{\tau_0})$;
- (b) $G(\tau, \phi_{\tau_0}(\tau)) = 0, \forall \tau \in W_{\tau_0}$.

O item (b) implica que $\phi_{\tau_0}(\tau)$ satisfaz (1.17), para todo $\tau \in W_{\tau_0}$. Pelo Lema 2, temos que

$$\phi_{\tau_0}(\tau) = (x_\tau, \lambda_\tau, s_\tau) = f_c(\tau), \forall \tau \in W_{\tau_0}.$$

Portanto, $f_c|_{W_{\tau_0}} : W_{\tau_0} \rightarrow \mathbb{R}^{2n+m}$ é contínua. Note que $W_\tau \subset \mathbb{R}_+$ e W_τ é aberto, para todo $\tau > 0$. Além disso,

$$\mathbb{R}_+ = \cup_{\tau>0} \{\tau\} \subset \cup_{\tau>0} W_\tau \subset \mathbb{R}_+,$$

e portanto, repetindo o argumento para $\tau_0 = \tau$, temos que f_c é contínua em todo ponto de $\mathbb{R}_+ = \cup_{\tau>0} W_\tau$.

Por fim, observe que a função f_c associa cada $\tau \in \mathbb{R}_+$ à um ponto $(x_\tau, \lambda_\tau, s_\tau) \in \mathbb{R}^{2n+m}$ satisfazendo (1.12). Sendo assim, seguirá da definição de \mathcal{C} e do Lema 1.2 que f_c é bijetora, donde podemos concluir que $f_c(\mathbb{R}_+) = \mathcal{C}$.

1.3 Vizinhanças do Caminho Central

Os algoritmos de pontos interiores normalmente restringem seus iterandos a uma vizinhança do caminho central \mathcal{C} ([2], pág. 399). Essas vizinhanças são planejadas com o

intuito de prevenir que os iterandos aproximem-se da fronteira das restrições $(x, s) \geq 0$. Por sua vez, isso garante que haja um progresso mínimo ao mover-se em direção ao conjunto de soluções ([1], pág. 9). As duas vizinhanças comuns do caminho central \mathcal{C} , no caso de pontos viáveis, são

$$\mathcal{N}_2(\theta) = \{(x, \lambda, s) \in \mathcal{F}_0 \mid \|XSe - \mu e\|_2 \leq \theta\mu\},$$

para algum $\theta \in [0, 1)$, e

$$\mathcal{N}_{-\infty}(\gamma) = \{(x, \lambda, s) \in \mathcal{F}_0 \mid x_i s_i \geq \gamma\mu, \forall i = 1, \dots, n\},$$

para algum $\gamma \in (0, 1]$. Segundo [2], é comum que θ e γ sejam escolhidos como 0.5 e 10^{-3} , respectivamente.

Quando um ponto pertence a $\mathcal{N}_{-\infty}(\gamma)$, cada produto $x_i s_i$ precisa valer ao menos um pequeno múltiplo γ de μ , que é a média dos produtos $x_i s_i$, com $i = 1, 2, \dots, n$. Uma característica da vizinhança $\mathcal{N}_{-\infty}(\gamma)$ é que para valores pequenos de γ (próximos de 0), esta vizinhança contém quase toda a região viável \mathcal{F} . Já a vizinhança $\mathcal{N}_2(\theta)$ é mais restritiva, pois haverão pontos de \mathcal{F}_0 fora de $\mathcal{N}_2(\theta)$ mesmo com θ próximo de 1.

Segundo [1] (pág. 9), exigir que todos os iterandos permaneçam em uma dessas vizinhanças faz com que os métodos seguidores de caminhos reduzam cada produto $x_i s_i$ para 0 em uma taxa mais ou menos parecida.

Nos métodos primais-duais, essas vizinhanças costumam ter um formato cônico em torno do caminho central. Além disso, para valores grandes de μ ela tende a ser muito larga, e conforme $\mu \rightarrow 0$ ela vai se tornando muito fina, por conta da exigência de que $(x, s) > 0$ ([1], pág. 10).

1.4 Algoritmo de Passos Longos

Um caso particular do Algoritmo 1 é o algoritmo de passos longos, enunciado a seguir. Este algoritmo procura dar o maior passo α_k possível de modo que o iterando se mantenha na vizinhança $\mathcal{N}_{-\infty}(\gamma)$. De acordo com [2], este algoritmo é capaz de progredir rapidamente quando γ é tomado muito pequeno.

Algoritmo 2 Passos Longos

Dado γ , σ_{\min} , σ_{\max} , com $\gamma \in (0, 1)$, $0 < \sigma_{\min} \leq \sigma_{\max} < 1$, e $(x_0, \lambda_0, s_0) \in \mathcal{N}_{-\infty}(\gamma)$;

para $k = 0, 1, 2, \dots$ **faça**:

Escolha $\sigma_k \in [\sigma_{\min}, \sigma_{\max}]$ e resolva

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S_k & 0 & X_k \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \\ \Delta s_k \end{bmatrix} = \begin{bmatrix} -r_c^k \\ -r_b^k \\ -X_k S_k e + \sigma_k \mu_k e \end{bmatrix}, \quad (1.22)$$

onde $\mu_k = \frac{(x_k)^T s_k}{n}$;

Encontre o maior valor $\alpha_k \in [0, 1]$ satisfazendo

$$(x_k, \lambda_k, s_k) + \alpha_k (\Delta x_k, \Delta \lambda_k, \Delta s_k) \in \mathcal{N}_{-\infty}(\gamma);$$

Defina

$$(x_{k+1}, \lambda_{k+1}, s_{k+1}) = (x_k, \lambda_k, s_k) + \alpha_k (\Delta x_k, \Delta \lambda_k, \Delta s_k). \quad (1.23)$$

fim

A seguir, serão provados alguns resultados técnicos que nos permitirão concluir a ordem do número de iterações necessárias para que o Algoritmo 2 obtenha um ponto tal que $\mu_k < \varepsilon$, para alguma tolerância $\varepsilon \in (0, 1)$ dada. Esses resultados são os Lemas 1.5 e 1.6, que correspondem aos Lemas 14.1 e 14.2, respectivamente, do livro [2]. As provas para estes lemas foram inteiramente baseadas neste livro.

Lema 1.5. *Sejam $u, v \in \mathbb{R}^n$ tais que $u^T v \geq 0$. Então,*

$$\|Uv\|_2 \leq 2^{-\frac{3}{2}} \|u + v\|_2^2,$$

$$\text{onde } U = \begin{bmatrix} u_1 & 0 & \cdots & 0 \\ 0 & u_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_n \end{bmatrix} \text{ e } V = \begin{bmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_n \end{bmatrix}.$$

Demonstração. Primeiramente, note que dados $\alpha, \beta \in \mathbb{R}$ tais que $\alpha\beta \geq 0$, temos da desigualdade da médias geométrica e aritmética que

$$\sqrt{|\alpha\beta|} \leq \frac{1}{2} |\alpha + \beta| \quad (1.24)$$

Além disso,

$$\begin{aligned} 0 \leq u^T v &= \sum_{u_i v_i \geq 0} u_i v_i + \sum_{u_i v_i < 0} u_i v_i = \sum_{i \in \mathcal{P}} |u_i v_i| - \sum_{i \in \mathcal{N}} |u_i v_i| \\ &\Rightarrow \sum_{i \in \mathcal{N}} |u_i v_i| \leq \sum_{i \in \mathcal{P}} |u_i v_i|, \end{aligned} \quad (1.25)$$

onde \mathcal{P} é o conjunto dos índices i para os quais $u_i v_i \geq 0$, e \mathcal{N} o conjunto dos índices i onde $u_i v_i < 0$. Agora, como $\|\cdot\|_2 \leq \|\cdot\|_1$, temos que

$$\begin{aligned} \|UVe\| &= \left(\|[u_i v_i]_{i \in \mathcal{P}}\|_2^2 + \|[u_i v_i]_{i \in \mathcal{N}}\|_2^2 \right)^{\frac{1}{2}} \leq \left(\|[u_i v_i]_{i \in \mathcal{P}}\|_1^2 + \|[u_i v_i]_{i \in \mathcal{N}}\|_1^2 \right)^{\frac{1}{2}} \\ &= \left(\left(\sum_{i \in \mathcal{P}} |u_i v_i| \right)^2 + \left(\sum_{i \in \mathcal{N}} |u_i v_i| \right)^2 \right)^{\frac{1}{2}}. \end{aligned} \quad (1.26)$$

Pela inequação (1.25), temos de (1.26) que

$$\begin{aligned} \|UVe\| &\leq \left(2 \left(\sum_{i \in \mathcal{P}} |u_i v_i| \right)^2 \right)^{\frac{1}{2}} = \sqrt{2} \sum_{i \in \mathcal{P}} |u_i v_i| \stackrel{(1.24)}{\leq} \sqrt{2} \sum_{i \in \mathcal{P}} \left(\frac{1}{2} |u_i + v_i| \right)^2 \\ &= 2^{-\frac{3}{2}} \sum_{i \in \mathcal{P}} |u_i + v_i|^2 \leq 2^{-\frac{3}{2}} \sum_{i=1}^n |u_i + v_i|^2 = 2^{-\frac{3}{2}} \|u + v\|_2^2. \end{aligned}$$

□

O resultado a seguir é verdadeiro para qualquer iteração k do Algoritmo 2, portanto o índice k será omitido.

Lema 1.6. *Se $(x, \lambda, s) \in \mathcal{N}_{-\infty}(\gamma)$, então*

$$\|\Delta X \Delta S e\| \leq 2^{-\frac{3}{2}} \left(1 + \frac{1}{\gamma} \right) n \mu.$$

Demonstração. Primeiramente, note que das duas primeiras linhas de (1.8), segue que

$$0 \Delta x + A^T \Delta \lambda + I \Delta s = -r_c = 0 \quad \text{e} \quad A \Delta x + 0 \Delta \lambda + 0 \Delta s = -r_b = 0$$

$$\Leftrightarrow A^T \Delta \lambda + I \Delta s = 0 \quad \text{e} \quad A \Delta x = 0$$

$$\Rightarrow \Delta x^T (A^T \Delta \lambda + I \Delta s) = 0 \Rightarrow (A \Delta x)^T \Delta \lambda + \Delta x^T \Delta s = 0 \Rightarrow \Delta x^T \Delta s = 0.$$

Agora, multiplicando a terceira linha do sistema (1.8) por $(XS)^{-\frac{1}{2}}$, e denotando $D = X^{\frac{1}{2}} S^{-\frac{1}{2}}$, temos que

$$(XS)^{-\frac{1}{2}} \cdot (S \Delta x + X \Delta s) = (XS)^{-\frac{1}{2}} \cdot (-X S e + \sigma \mu e)$$

$$\begin{aligned} &\Leftrightarrow X^{-\frac{1}{2}}S^{\frac{1}{2}}\Delta x + X^{\frac{1}{2}}S^{-\frac{1}{2}}\Delta s = (XS)^{-\frac{1}{2}} \cdot (-XSe + \sigma\mu e) \\ &\Leftrightarrow D^{-1}\Delta x + D\Delta s = (XS)^{-\frac{1}{2}} \cdot (-XSe + \sigma\mu e). \end{aligned} \quad (1.27)$$

Como $(D^{-1}\Delta x)^T(D\Delta s) = \Delta x^T(D^{-1})^T D\Delta s = \Delta x^T D^{-1}D\Delta s = \Delta x^T \Delta s = 0$, temos pelo Lema 1.5, quando $u = D^{-1}\Delta x$ e $v = D\Delta s$, que

$$\begin{aligned} \|\Delta X \Delta S e\| &= \|I \Delta X \Delta S e\| = \|(D^{-1}D)\Delta X \Delta S e\| = \|(D^{-1}\Delta X)(D\Delta S)e\| \\ &\leq 2^{-\frac{3}{2}}\|D^{-1}\Delta x + D\Delta s\|^2 = 2^{-\frac{3}{2}}\|(XS)^{-\frac{1}{2}} \cdot (-XSe + \sigma\mu e)\|^2, \end{aligned} \quad (1.28)$$

onde a última igualdade segue de (1.27). Note que

$$\begin{aligned} \|(XS)^{-\frac{1}{2}} \cdot (-XSe + \sigma\mu e)\|^2 &= \|- (XS)^{\frac{1}{2}}e + (XS)^{-\frac{1}{2}}\sigma\mu e\|^2 \\ &= \left[-(XS)^{\frac{1}{2}}e + (XS)^{-\frac{1}{2}}\sigma\mu e \right]^T \left[-(XS)^{\frac{1}{2}}e + (XS)^{-\frac{1}{2}}\sigma\mu e \right] \\ &= e^T XSe - 2e^T \sigma\mu e + e^T (XS)^{-1} \sigma^2 \mu^2 e. \end{aligned}$$

Como $Ue = u$, para todo vetor $u \in \mathbb{R}^n$, temos que

$$\begin{aligned} \|(XS)^{-\frac{1}{2}} \cdot (-XSe + \sigma\mu e)\|^2 &= e^T XSe - 2e^T \sigma\mu e + e^T (XS)^{-1} \sigma^2 \mu^2 e \\ &= (Xe)^T Se - 2\sigma\mu e^T e + \sigma^2 \mu^2 e^T X^{-1} S^{-1} e = (Xe)^T Se - 2\sigma\mu e^T e + \sigma^2 \mu^2 (X^{-1}e)^T S^{-1} e \\ &= x^T s - 2\sigma\mu n + \sigma^2 \mu^2 \sum_{i=1}^n \frac{1}{x_i} \cdot \frac{1}{s_i} = x^T s - 2\sigma\mu n + \sigma^2 \mu^2 \sum_{i=1}^n \frac{1}{x_i s_i}. \end{aligned}$$

Como $(x, \lambda, s) \in \mathcal{N}_{-\infty}(\gamma)$, segue que $x_i s_i \geq \gamma\mu$, para todo $i \in \{1, 2, \dots, n\}$. Consequentemente,

$$\begin{aligned} \|(XS)^{-\frac{1}{2}} \cdot (-XSe + \sigma\mu e)\|^2 &= x^T s - 2\sigma\mu n + \sigma^2 \mu^2 \sum_{i=1}^n \frac{1}{x_i s_i} \\ &\leq x^T s - 2\sigma\mu n + \sigma^2 \mu^2 \sum_{i=1}^n \frac{1}{\gamma\mu} = \mu n - 2\sigma\mu n + \sigma^2 \mu^2 \frac{n}{\gamma\mu} \\ &= \mu n - 2\sigma\mu n + \frac{\sigma^2}{\gamma} \mu n = \left(1 - 2\sigma + \frac{\sigma^2}{\gamma}\right) \mu n \end{aligned} \quad (1.29)$$

Segue de (1.28) e (1.29) que

$$\|\Delta X \Delta S e\| \leq 2^{-\frac{3}{2}} \left(1 - 2\sigma + \frac{\sigma^2}{\gamma}\right) \mu n \leq 2^{-\frac{3}{2}} \left(1 + \frac{1}{\gamma}\right) \mu n,$$

onde a última desigualdade se deve pelo fato de que $0 \leq \sigma \leq 1$. \square

Definição 1.7. *Seja $x^* \in \mathbb{R}^n$ e $(x_k)_{k \in \mathbb{N}}$ uma sequência em \mathbb{R}^n tal que $x_k \rightarrow x^*$. Se existir $c \in [0, 1)$ tal que*

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = c,$$

dizemos que $(x_k)_{k \in \mathbb{N}}$ converge q -linearmente para x^ . Ainda, se $c = 0$, dizemos que $(x_k)_{k \in \mathbb{N}}$ converge q -superlinearmente para x^* .*

O Teorema 1.8 descreve como se dá o decréscimo de μ em uma determinada iteração k do Algoritmo 2. Este teorema, bem como sua demonstração, baseiam-se no Teorema 14.3 de [2].

Teorema 1.8. *Dados os parâmetros γ , σ_{\min} , e σ_{\max} no Algoritmo 2, existe uma constante δ que não depende de n tal que*

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n}\right) \mu_k, \quad (1.30)$$

para todo $k \geq 0$.

Demonstração. Primeiramente, vamos provar que $(x_k(\alpha), \lambda_k(\alpha), s_k(\alpha)) = (x_k, \lambda_k, s_k) + \alpha(\Delta x_k, \Delta \lambda_k, \Delta s_k) \in \mathcal{N}_{-\infty}(\gamma)$, $\forall \alpha \in \left[0, 2^{\frac{3}{2}} \gamma \frac{1 - \gamma \sigma_k}{1 + \gamma n}\right]$. Como uma consequência disso, o valor α_k no algoritmo 2 deverá satisfazer

$$\alpha_k \geq 2^{\frac{3}{2}} \gamma \frac{1 - \gamma \sigma_k}{1 + \gamma n}, \quad (1.31)$$

pela definição de α_k .

Para cada $i = 1, \dots, n$, temos do Lema 1.6 que

$$|[\Delta x_k]_i [\Delta s_k]_i| \leq \|\Delta X_k \Delta S_k e\|_2 \leq 2^{-\frac{3}{2}} \left(1 + \frac{1}{\gamma}\right) n \mu_k. \quad (1.32)$$

Da terceira linha do sistema (1.22), temos o seguinte sistema:

$$\begin{bmatrix} S_k & 0 & X_k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -X S e + \sigma_k \mu_k e \end{bmatrix}. \quad (1.33)$$

Fixada a i -ésima linha deste sistema, temos que

$$[s_k]_i [\Delta x_k]_i + [x_k]_i [\Delta s_k]_i = -[x_k]_i [s_k]_i + \sigma_k \mu_k. \quad (1.34)$$

Deste modo,

$$[x_k]_i(\alpha) [s_k]_i(\alpha) = ([x_k]_i + \alpha [\Delta x_k]_i) ([s_k]_i + \alpha [\Delta s_k]_i)$$

$$\begin{aligned}
&= [x_k]_i [s_k]_i + \alpha([s_k]_i [\Delta x_k]_i + [x_k]_i [\Delta s_k]_i) + \alpha^2 [\Delta x_k]_i [\Delta s_k]_i \\
&= [x_k]_i [s_k]_i + \alpha(-[x_k]_i [s_k]_i + \sigma_k \mu_k) + \alpha^2 [\Delta x_k]_i [\Delta s_k]_i = (1-\alpha)[x_k]_i [s_k]_i + \alpha \sigma_k \mu_k + \alpha^2 [\Delta x_k]_i [\Delta s_k]_i
\end{aligned} \tag{1.35}$$

Como $[\Delta x_k]_i [\Delta s_k]_i \geq -|[\Delta x_k]_i [\Delta s_k]_i|$, segue que

$$[x_k]_i(\alpha)[s_k]_i(\alpha) \geq (1-\alpha)[x_k]_i [s_k]_i + \alpha \sigma_k \mu_k - \alpha^2 |[\Delta x_k]_i [\Delta s_k]_i|.$$

Por (1.32), temos que

$$\begin{aligned}
[x_k]_i(\alpha)[s_k]_i(\alpha) &\geq (1-\alpha)[x_k]_i [s_k]_i + \alpha \sigma_k \mu_k - \alpha^2 2^{-\frac{3}{2}} \left(1 + \frac{1}{\gamma}\right) n \mu_k \\
&\geq (1-\alpha)\gamma \mu_k + \alpha \sigma_k \mu_k - \alpha^2 2^{-\frac{3}{2}} \left(1 + \frac{1}{\gamma}\right) n \mu_k,
\end{aligned} \tag{1.36}$$

onde a última inequação segue do fato de que $(x_k, \lambda_k, s_k) \in \mathcal{N}_{-\infty}(\gamma)$.

Veja que

$$\begin{aligned}
\mu_k(\alpha) &= \frac{x_k^T(\alpha) s_k(\alpha)}{n} = \frac{\sum_{i=1}^n [x_k]_i(\alpha) [s_k]_i(\alpha)}{n} \\
&\Leftrightarrow n \mu_k(\alpha) = \sum_{i=1}^n [x_k]_i(\alpha) [s_k]_i(\alpha).
\end{aligned}$$

Da equação (1.35), temos que

$$\begin{aligned}
n \mu_k(\alpha) &= \sum_{i=1}^n [(1-\alpha)[x_k]_i [s_k]_i + \alpha \sigma_k \mu_k + \alpha^2 [\Delta x_k]_i [\Delta s_k]_i] \\
&= (1-\alpha) \sum_{i=1}^n [x_k]_i [s_k]_i + \sum_{i=1}^n \alpha \sigma_k \mu_k + \alpha^2 \sum_{i=1}^n [\Delta x_k]_i [\Delta s_k]_i \\
&= (1-\alpha) n \mu_k + n \alpha \sigma_k \mu_k + \alpha^2 \Delta x^T \Delta s.
\end{aligned}$$

Na demonstração do Lema 1.6 foi mostrado que $\Delta x^T \Delta s = 0$, logo

$$\begin{aligned}
n \mu_k(\alpha) &= (1-\alpha) n \mu_k + n \alpha \sigma_k \mu_k \\
\Rightarrow \mu_k(\alpha) &= (1-\alpha) \mu_k + \alpha \sigma_k \mu_k = \mu_k - \alpha \mu_k + \alpha \sigma_k \mu_k = \mu_k - \alpha(1-\sigma_k) \mu_k \\
&\Rightarrow \mu_k(\alpha) = (1-\alpha(1-\sigma_k)) \mu_k.
\end{aligned} \tag{1.37}$$

De (1.36) e (1.37), temos que a desigualdade $[x_k]_i(\alpha)[s_k]_i(\alpha) \geq \gamma \mu_k(\alpha)$ será verdadeira se ocorrer

$$\begin{aligned}
(1-\alpha)\gamma \mu_k + \alpha \sigma_k \mu_k - \alpha^2 2^{-\frac{3}{2}} \left(1 + \frac{1}{\gamma}\right) n \mu_k &\geq \gamma(1-\alpha + \alpha \sigma_k) \mu_k \\
\Leftrightarrow \alpha \sigma_k \mu_k - \alpha^2 2^{-\frac{3}{2}} \left(1 + \frac{1}{\gamma}\right) n \mu_k &\geq \gamma \alpha \sigma_k \mu_k \Leftrightarrow (1-\gamma) \alpha \sigma_k \mu_k \geq \alpha^2 2^{-\frac{3}{2}} \left(1 + \frac{1}{\gamma}\right) n \mu_k
\end{aligned}$$

$$\begin{aligned} \Leftrightarrow (1 - \gamma)\sigma_k\mu_k \geq \alpha 2^{-\frac{3}{2}} \left(1 + \frac{1}{\gamma}\right) n\mu_k &\Leftrightarrow (1 - \gamma)\sigma_k\mu_k \geq \alpha 2^{-\frac{3}{2}} \left(\frac{1 + \gamma}{\gamma}\right) n\mu_k \\ &\Leftrightarrow \alpha \leq 2^{\frac{3}{2}}\gamma \frac{1 - \gamma \sigma_k}{1 + \gamma n}. \end{aligned}$$

Resta mostrar que o ponto $(x_k(\alpha), \lambda_k(\alpha), s_k(\alpha))$ pertence a \mathcal{F}_0 para que possamos concluir está em $\mathcal{N}_{-\infty}(\gamma)$. Dos dois primeiros blocos do sistema (1.22), temos

$$A^T \Delta \lambda_k + \Delta s_k = 0 \quad \text{e} \quad A \Delta x_k = 0.$$

Então,

$$Ax_k(\alpha) = A(x_k + \alpha \Delta x_k) = Ax_k + \alpha A \Delta x_k = b + 0 = b$$

e

$$\begin{aligned} A^T x_k(\alpha) + s_k(\alpha) &= A^T(\lambda_k + \alpha \Delta \lambda_k) + s_k + \alpha \Delta s_k = A^T \lambda_k + s_k + \alpha(A^T \Delta \lambda_k + \Delta s_k) \\ &= c + 0 = c. \end{aligned}$$

Além disso, se $\alpha \leq 2^{\frac{3}{2}}\gamma \frac{1 - \gamma \sigma_k}{1 + \gamma n}$, temos que $[x_k]_i(\alpha)[s_k]_i(\alpha) \geq \gamma\mu_k(\alpha)$, para todo $i = 1, \dots, n$. Como $\gamma > 0$, teremos que $[x_k]_i(\alpha)[s_k]_i(\alpha) > 0$, para todo $i \in \{1, \dots, n\}$, se $\mu_k(\alpha) > 0$. Mas $\mu_k(\alpha) = (1 - \alpha(1 - \sigma_k))\mu_k > 0$ se $(1 - \alpha(1 - \sigma_k)) > 0$, pois $\mu_k > 0$. Sendo assim,

$$1 - \alpha(1 - \sigma_k) > 0 \Leftrightarrow 1 > \alpha(1 - \sigma_k) \Leftrightarrow \frac{1}{1 - \sigma_k} > \alpha,$$

pois $\sigma_k < 1$, e conseqüentemente, $1 - \sigma_k > 0$. Como $\alpha \in [0, 1]$, segue que

$$\alpha \leq 1 < \frac{1}{1 - \sigma_k},$$

e conseqüentemente, $\mu_k(\alpha) > 0$, provando que $[x_k]_i(\alpha)[s_k]_i(\alpha) > 0$, para todo $i \in \{1, \dots, n\}$. Em particular, $[x_k]_i(\alpha) \neq 0$ e $[s_k]_i(\alpha) \neq 0$, para todo $i = 1, \dots, n$.

Para provar a positividade estrita de $[x_k]_i(\alpha)$ e $[s_k]_i(\alpha)$ suponha, por absurdo, que $[x_k]_i(\alpha) < 0$, para algum $i \in \{1, \dots, n\}$. Segue de $[x_k]_i(\alpha)[s_k]_i(\alpha) > 0$ que $[s_k]_i(\alpha) < 0$. Note que

$$[x_k]_i(\alpha) = [x_k]_i + \alpha[\Delta x_k]_i < 0 \Rightarrow \alpha[\Delta x_k]_i < -[x_k]_i < 0 \Rightarrow [\Delta x_k]_i < 0,$$

pois $\alpha \geq 0$. Analogamente, temos que $[\Delta s_k]_i < 0$. Agora, reescreva a equação (1.34) da seguinte forma:

$$[s_k]_i([x_k]_i + [\Delta x_k]_i) = \sigma_k\mu_k - [x_k]_i[\Delta x_k]_i. \quad (1.38)$$

Como $[\Delta x_k]_i < 0$ e $\alpha \in [0, 1]$, segue que

$$[x_k]_i(\alpha) = [x_k]_i + \alpha[\Delta x_k]_i \geq [x_k]_i + [\Delta x_k]_i.$$

Sendo assim,

$$[s_k]_i[x_k]_i(\alpha) \geq [s_k]_i([x_k]_i + [\Delta x_k]_i) = \sigma_k \mu_k - [x_k]_i[\Delta x_k]_i > 0,$$

pois $[x_k]_i > 0$ e $[\Delta x_k]_i < 0$. Mas $[s_k]_i[x_k]_i(\alpha) > 0$ implica que $[s_k]_i$ e $[x_k]_i(\alpha)$ têm mesmo sinal. Neste caso, teríamos que $[x_k]_i(\alpha) > 0$, contradizendo a hipótese de que $[x_k]_i(\alpha) < 0$. Portanto, $[x_k]_i(\alpha) > 0$, para todo $i = 1, \dots, n$. Decorre de $[x_k]_i(\alpha)[s_k]_i(\alpha) > 0$ que $[s_k]_i(\alpha) > 0$, para todo $i = 1, \dots, n$. Isso conclui a prova de que $(x_k(\alpha), \lambda_k(\alpha), s_k(\alpha)) \in \mathcal{F}_0$. Portanto, podemos afirmar que $(x_k(\alpha), \lambda_k(\alpha), s_k(\alpha)) \in \mathcal{N}_{-\infty}(\gamma)$, $\forall \alpha \in \left[0, 2^{\frac{3}{2}}\gamma \frac{1 - \gamma \sigma_k}{1 + \gamma n}\right]$, e consequentemente, vale a inequação (1.31).

Por fim,

$$\begin{aligned} \mu_{k+1} &= \frac{x_k(\alpha)^T s_k(\alpha)}{n} = \frac{x_k^T s_k + \alpha_k (x_k^T \Delta s_k + s_k^T \Delta x_k) + \alpha_k^2 \Delta x_k^T \Delta s_k}{n} \\ &= \frac{x_k^T s_k + \alpha_k (x_k^T \Delta s_k + s_k^T \Delta x_k)}{n}, \end{aligned}$$

pois $\Delta x_k^T \Delta s_k = 0$. Somando as equações (1.34) para cada $i \in \{1, 2, \dots, n\}$, temos que

$$s_k^T \Delta x_k + x_k^T \Delta s_k = -x_k^T s_k + n\sigma_k \mu_k,$$

e portanto,

$$\begin{aligned} \mu_{k+1} &= \frac{x_k^T s_k + \alpha_k (-x_k^T s_k + \sigma_k \mu_k)}{n} = \frac{x_k^T s_k}{n} - \alpha_k \frac{x_k^T s_k}{n} + \alpha_k \sigma_k \mu_k \\ &= \mu_k - \alpha_k \mu_k + \alpha_k \sigma_k \mu_k = \mu_k - \alpha_k (1 - \sigma_k) \mu_k = (1 - \alpha_k (1 - \sigma_k)) \mu_k. \end{aligned} \quad (1.39)$$

Se $\alpha_k \geq 2^{\frac{3}{2}}\gamma \frac{1 - \gamma \sigma_k}{1 + \gamma n}$, então

$$\mu_{k+1} \leq \left(1 - \frac{2^{\frac{3}{2}}}{n} \gamma \frac{1 - \gamma}{1 + \gamma} \sigma_k (1 - \sigma_k)\right) \mu_k.$$

Agora, observe que a função $\sigma \mapsto \sigma(1 - \sigma)$, $\sigma \in [\sigma_{\min}, \sigma_{\max}]$, é uma função quadrática côncava para para baixo, pois $\sigma(1 - \sigma) = -\sigma^2 + \sigma$. Neste caso, esta função admitirá seu mínimo em um dos seus pontos extremos σ_{\min} ou σ_{\max} . Sendo assim,

$$\sigma_k(1 - \sigma_k) \geq \min\{\sigma_{\max}(1 - \sigma_{\max}), \sigma_{\min}(1 - \sigma_{\min})\}, \forall \sigma_k \in [\sigma_{\min}, \sigma_{\max}].$$

Deste modo, se definirmos

$$\delta = 2^{\frac{3}{2}}\gamma \frac{1-\gamma}{1+\gamma} \min\{\sigma_{\max}(1-\sigma_{\max}), \sigma_{\min}(1-\sigma_{\min})\},$$

temos que

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n}\right) \mu_k.$$

□

No próximo teorema, o qual corresponde ao Teorema 14.4 de [2], provaremos que $\mu_k \rightarrow 0$, e daremos uma estimativa do número de iterações necessário para atingir uma determinada tolerância $\varepsilon > 0$. Observe que, da forma como foi definido δ na demonstração do Teorema 1.8, temos que $\delta < 2^{\frac{3}{2}}$. Como $2^{\frac{3}{2}} \approx 2.83$, para $n > 3$, temos que $\frac{\delta}{n} < 1$. Sendo assim, para $n > 3$, o Teorema 1.8 nos diz que μ_k converge q -linearmente à 0.

Teorema 1.9. *Dado $\varepsilon > 0$ e $\gamma \in (0, 1)$, suponha que o ponto inicial no Algoritmo 2 satisfaz $(x_0, \lambda_0, s_0) \in \mathcal{N}_{-\infty}(\gamma)$. Então, existe um índice K com $K = O(n \log \frac{1}{\varepsilon})$ tal que*

$$\mu_k \leq \varepsilon \mu_0, \quad \forall k \geq K.$$

Demonstração. Do Teorema 1.8, temos que

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n}\right) \mu_k,$$

para todo $k \geq 0$. Tomando o logaritmo de ambos os lados,

$$\log \mu_{k+1} \leq \log \left(\left(1 - \frac{\delta}{n}\right) \mu_k \right) = \log \left(1 - \frac{\delta}{n}\right) + \log \mu_k.$$

Aplicando a desigualdade acima k vezes, temos

$$\log \mu_k \leq \log \left(1 - \frac{\delta}{n}\right) + \log \mu_{k-1} \leq \dots \leq k \log \left(1 - \frac{\delta}{n}\right) + \log \mu_0.$$

Temos a seguinte estimativa para a função log:

$$\log(1 + \beta) \leq \beta, \quad \forall \beta > -1. \tag{1.40}$$

Por essa estimativa, temos que

$$\log \mu_k \leq k \frac{\delta}{n} + \log \mu_0 \Leftrightarrow \log \mu_k - \log \mu_0 \leq k \left(-\frac{\delta}{n}\right) \Leftrightarrow \log \frac{\mu_k}{\mu_0} \leq k \left(-\frac{\delta}{n}\right).$$

Observe que se $k \left(-\frac{\delta}{n}\right) \leq \log \varepsilon$, temos que $\log \frac{\mu_k}{\mu_0} \leq \log \varepsilon$, implicando $\frac{\mu_k}{\mu_0} \leq \varepsilon$. Mas,

$$k \left(-\frac{\delta}{n}\right) \leq \log \varepsilon \Leftrightarrow k \geq \left(-\frac{n}{\delta}\right) \log \varepsilon \Leftrightarrow k \geq \frac{n}{\delta} \log \frac{1}{\varepsilon}.$$

Quando $\varepsilon < 1$, temos que $\log \frac{1}{\varepsilon} = -\log \varepsilon = |\log \varepsilon|$, e conseqüentemente,

$$k \geq \frac{n}{\delta} |\log \varepsilon|.$$

Definindo $K = \frac{n}{\delta} \log \frac{1}{\varepsilon}$, teremos

$$\mu_k \leq \varepsilon \mu_0,$$

sempre que $k \geq K$.

□

1.5 Algoritmo Preditor-Corretor

Nessa seção, os métodos preditores-corretores serão abordados. Os métodos preditores-corretores basicamente acrescentam um passo a mais nos algoritmos primais-duais convencionais para garantir que os iterandos se mantenham próximos do caminho central. Isso distancia os iterandos da fronteira das restrições $(x, s) \geq 0$, permitindo que passos maiores sejam dados sem violar estas restrições.

Segundo [1] (pág. 85), muitos aspectos desse tipo de algoritmo haviam sido descritos por diversos autores antes mesmo de sua criação. Um algoritmo preditor-corretor foi formalmente apresentado e analisado pela primeira vez em [17]. Este algoritmo corresponde ao Algoritmo 3 (Algoritmo PC de [1]), que será apresentado mais adiante. O Algoritmo 3 também é conhecido como “preditor-corretor Mizuno-Todd-Ye”.

Neste algoritmo, existem dois tipos de passos: os passos preditor e corretor. Os passos preditor e corretor são dados em iterações distintas e de forma alternada. Mais especificamente, em uma iteração se realiza o passo preditor, no qual é tomado $\sigma_k = 0$ em (1.22) para obter a direção afim-escala, e em seguida é feito um deslocamento nessa direção. Na iteração seguinte, aplica-se o passo corretor, em que toma-se $\sigma_k = 1$ em (1.22) para obter a direção afim-escala e em seguida move-se nesta direção. Quando fazemos $\sigma_k = 0$, estamos mirando diretamente na solução do sistema KKT (equação (1.3)). Isto acarreta em uma redução de μ por um fator $(1 - \alpha_k)$. No entanto, como esta escolha de σ_k

é muito agressiva, o próximo iterando tende a se desviar do caminho central, e portanto, perde-se a centralidade.

Já a escolha de $\sigma_k = 1$ faz com que miremos no ponto do caminho central cuja medida de complementaridade é μ_k . Isso não proporciona redução alguma de μ , porém essa escolha em geral permite obter o iterando seguinte mais próximo do caminho central, o que promove uma melhora na centralidade. Um iterando bem centralizado permite dar passos maiores sem que as restrições de não negatividade sejam violadas, o que potencializa a redução de μ no próximo passo preditor. A seguir, o Algoritmo Preditor-Corretor será apresentado:

Algoritmo 3 Algoritmo Preditor-Corretor

Dado $(x_0, \lambda_0, s_0) \in \mathcal{N}_2(0.25)$:

para $k = 0, 1, 2, \dots$ **faça**:

se k for par **faça**:

Resolva (1.22) com $\sigma_k = 0$ para obter $(\Delta x_k, \Delta \lambda_k, \Delta s_k)$;

Escolha α_k como o maior valor de $\alpha \in [0, 1]$ tal que

$$(x_k(\alpha), \lambda_k(\alpha), s_k(\alpha)) = (x_k, \lambda_k, s_k) + \alpha(\Delta x_k, \Delta \lambda_k, \Delta s_k) \in \mathcal{N}_2(0.5);$$

Defina $(x_{k+1}, \lambda_{k+1}, s_{k+1}) = (x_k, \lambda_k, s_k) + \alpha_k(\Delta x_k, \Delta \lambda_k, \Delta s_k)$;

senão:

Resolva (1.22) com $\sigma_k = 1$ para obter $(\Delta x_k, \Delta \lambda_k, \Delta s_k)$;

Defina $(x_{k+1}, \lambda_{k+1}, s_{k+1}) = (x_k, \lambda_k, s_k) + (\Delta x_k, \Delta \lambda_k, \Delta s_k)$;

fim

fim

Os Lemas 1.10, 1.12, 1.43 e 1.13, a seguir, foram encontrados no livro [1], e correspondem aos Lemas 5.4, 5.5, 5.7 e 5.8, respectivamente. Suas demonstrações seguem o mesmo raciocínio apresentado em [1]. Esses lemas dizem respeito a uma iteração qualquer do Algoritmo 3, portanto os índices k serão omitidos. O próximo lema é um resultado análogo ao Lema 1.6 para a vizinhança $\mathcal{N}_2(\theta)$.

Lema 1.10. *Se $(x, \lambda, s) \in \mathcal{N}_2(\theta)$, então*

$$\|\Delta X \Delta S e\| \leq \left(\frac{\theta^2 + n(1 - \sigma)^2}{2^{\frac{3}{2}}(1 - \theta)} \right) \mu.$$

Os lemas abaixo formalizam os passos do algoritmo.

Lema 1.11. *Se $(x, \lambda, s) \in \mathcal{N}_2(\theta)$, temos*

$$\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \leq |1 - \alpha| \|XSe - \mu\| + \alpha^2 \|\Delta X \Delta Se\| \quad (1.41)$$

$$|1 - \alpha| \theta \mu + \alpha^2 \left(\frac{\theta^2 + n(1 - \sigma)^2}{2^{\frac{3}{2}}(1 - \theta)} \right) \mu, \quad (1.42)$$

onde $(x(\alpha), \lambda(\alpha), s(\alpha)) = (x, \lambda, s) + \alpha(\Delta x, \Delta \lambda, \Delta s)$ e $\mu(\alpha) = \frac{x(\alpha)^T s(\alpha)}{n}$.

Demonstração. Na demonstração do Teorema 1.8 (ver a equação (1.37)) foi provado que $\mu(\alpha)$ é dado por $\mu(\alpha) = (1 - \alpha(1 - \sigma))\mu$. Além disso, do terceiro bloco de (1.8), temos que $s_i \Delta x_i + x_i \Delta s_i = -x_i s_i + \sigma \mu$, para todo $i = 1, 2, \dots, n$. Desse modo,

$$\begin{aligned} x_i(\alpha)s_i(\alpha) - \mu(\alpha) &= x_i s_i + \alpha(s_i \Delta x_i + x_i \Delta s_i) + \alpha^2 \Delta x_i \Delta s_i - (1 - \alpha(1 - \sigma))\mu \\ &= x_i s_i(1 - \alpha) + \alpha \sigma \mu + \alpha^2 \Delta x_i \Delta s_i - (1 - \alpha + \alpha \sigma)\mu \\ &= x_i s_i(1 - \alpha) + \alpha^2 \Delta x_i \Delta s_i - (1 - \alpha)\mu, \forall i = 1, \dots, n. \end{aligned}$$

Tomando a norma do vetor $X(\alpha)S(\alpha)e - \mu(\alpha)e$, temos

$$\begin{aligned} \|X(\alpha)S(\alpha)e - \mu(\alpha)e\| &= \|[x_i(\alpha)s_i(\alpha) - \mu(\alpha)]_{i=1}^n\| = \|[x_i s_i(1 - \alpha) + \alpha^2 \Delta x_i \Delta s_i - (1 - \alpha)\mu]_{i=1}^n\| \\ &\leq |1 - \alpha| \|XSe - \mu e\| + \alpha^2 \|\Delta X \Delta Se\| \\ &\leq |1 - \alpha| \theta \mu + \alpha^2 \left(\frac{\theta^2 + n(1 - \sigma)^2}{2^{\frac{3}{2}}(1 - \theta)} \right) \mu, \end{aligned}$$

onde a última desigualdade decorre do Lema 1.10. \square

Lema 1.12. *Suponha que $(x, \lambda, s) \in \mathcal{N}_2(0.25)$ e seja $(\Delta x, \Delta \lambda, \Delta s)$ a direção obtida em (1.8) quando $\sigma = 0$. Então, $(x(\alpha), \lambda(\alpha), s(\alpha)) \in \mathcal{N}_2(0.5)$, para todo $\alpha \in [0, \bar{\alpha}]$, onde*

$$(x(\alpha), \lambda(\alpha), s(\alpha)) = (x, \lambda, s) + \alpha(\Delta x, \Delta \lambda, \Delta s)$$

e

$$\bar{\alpha} = \min \left(\frac{1}{2}, \left(\frac{\mu}{8\|\Delta X \Delta Se\|} \right)^{\frac{1}{2}} \right).$$

Consequentemente, o comprimento do passo preditor é de pelo menos $\bar{\alpha}$, e o novo valor de μ é ao menos $(1 - \bar{\alpha})\mu$.

Demonstração. Da inequação (1.41) e da hipótese sobre α , temos que

$$\begin{aligned} \|X(\alpha)S(\alpha)e - \mu(\alpha)e\| &\leq (1 - \alpha)\|XSe - \mu e\| + \frac{\mu}{8\|\Delta X \Delta Se\|}\|\Delta X \Delta Se\| \\ &= (1 - \alpha)\|XSe - \mu e\| + \frac{\mu}{8} = (1 - \alpha)\|XSe - \mu e\| + \frac{1}{8(1 - \alpha)}(1 - \alpha)\mu. \end{aligned}$$

Como $(x, \lambda, s) \in \mathcal{N}_2(0.25)$ e , temos que

$$\begin{aligned} \|X(\alpha)S(\alpha)e - \mu(\alpha)e\| &\leq (1 - \alpha)0.25\mu + \frac{1}{8(1 - \alpha)}(1 - \alpha)\mu \\ &= \frac{1}{4}(1 - \alpha)\mu + \frac{1}{8(1 - \alpha)}(1 - \alpha)\mu. \end{aligned}$$

Como $\alpha < \frac{1}{2}$, temos que $4 \leq 8(1 - \alpha)$, e conseqüentemente,

$$\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \leq \frac{1}{4}(1 - \alpha)\mu + \frac{1}{4}(1 - \alpha)\mu = \frac{1}{2}(1 - \alpha)\mu.$$

Quando $\sigma = 0$, temos que $\mu(\alpha) = (1 - \alpha)\mu$, logo

$$\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \leq \frac{1}{2}\mu(\alpha).$$

A verificação de que as condições de viabilidade do ponto $(x(\alpha), \lambda(\alpha), s(\alpha))$ são satisfeitas é realizada de forma idêntica a demonstração do Teorema 1.8. Portanto, resta mostrar apenas que $x_i(\alpha)$ e $s_i(\alpha)$ são estritamente positivos, para todo $i = 1, 2, \dots, n$. Observe que

$$\mu(\alpha) - x_i(\alpha)s_i(\alpha) \leq |x_i(\alpha)s_i(\alpha) - \mu(\alpha)| \leq \|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \leq \frac{1}{2}\mu(\alpha), \forall i = 1, 2, \dots, n,$$

logo,

$$\frac{1}{2}\mu(\alpha) \leq x_i(\alpha)s_i(\alpha), \forall i = 1, 2, \dots, n.$$

Como $\frac{1}{2}\mu(\alpha) = \frac{1}{2}(1 - \alpha(1 - \sigma))\mu = \frac{1}{2}(1 - \alpha)\mu \leq \frac{1}{2} \cdot \frac{1}{2}\mu = \frac{1}{4}\mu > 0$, pois $\mu > 0$, temos que

$$x_i(\alpha)s_i(\alpha) > 0, \forall i = 1, 2, \dots, n.$$

Conseqüentemente, pelos mesmos argumentos do Teorema 1.8, temos que $x_i(\alpha)$ e $s_i(\alpha)$ devem ser estritamente positivos, para $i = 1, 2, \dots, n$.

□

Observe que se $\theta = 0.25$ e $\sigma = 0$, temos pelo Lema 1.10 que

$$\begin{aligned} \|\Delta X \Delta S e\| &\leq \left(\frac{(0.25)^2 + n(1-0)^2}{2^{\frac{3}{2}}(1-0.25)} \right) \mu \Leftrightarrow \frac{1}{\|\Delta X \Delta S e\|} \geq \left(\frac{0.75 \cdot 2^{\frac{3}{2}}}{(0.25)^2 + n} \right) \frac{1}{\mu} \\ &\Leftrightarrow \frac{1}{\|\Delta X \Delta S e\|} \geq \left(\frac{3}{4} \frac{2\sqrt{2}}{\left(\frac{1}{4}\right)^2 + n} \right) \frac{1}{\mu} \Leftrightarrow \frac{1}{\|\Delta X \Delta S e\|} \geq \left(\frac{3}{4} \frac{2\sqrt{2}}{\frac{1}{16} + n} \right) \frac{1}{\mu} \\ &\Leftrightarrow \frac{1}{8\|\Delta X \Delta S e\|} \geq \left(\frac{1}{8} \frac{3}{4} \frac{2\sqrt{2}}{\frac{1}{16} + n} \right) \frac{1}{\mu} \Leftrightarrow \frac{1}{8\|\Delta X \Delta S e\|} \geq \left(\frac{1}{16} \frac{3\sqrt{2}}{\frac{1}{16} + n} \right) \frac{1}{\mu} \\ &\Leftrightarrow \frac{\mu}{8\|\Delta X \Delta S e\|} \geq \frac{3\sqrt{2}}{1 + 16n}. \end{aligned}$$

Como $1 \leq n \leq 16n + 1$, segue que

$$\frac{\mu}{8\|\Delta X \Delta S e\|} \geq \frac{3\sqrt{2}}{n} \geq \frac{0.16}{n}.$$

Assim sendo, segue pelo Lema 1.12 que

$$\bar{\alpha} \geq \min \left\{ \frac{1}{2}, \left(\frac{0.16}{n} \right)^{\frac{1}{2}} \right\} = \min \left\{ \frac{1}{2}, \frac{0.4}{\sqrt{n}} \right\} = \frac{0.4}{\sqrt{n}},$$

onde a última igualdade se deve ao fato de que $n \geq 1$. A desigualdade acima implica que $\bar{\alpha}$ possui uma cota inferior. Uma vez que o passo corretor ocorrerá sempre em iterações de índices pares, esta limitação inferior de $\bar{\alpha}$ implicará que

$$\mu_{k+1} = (1 - \alpha)\mu_k \leq \left(1 - \frac{0.4}{\sqrt{n}} \right) \mu_k, \quad \forall k = 0, 2, 4, \dots \quad (1.43)$$

Portanto, o passo preditor toma um ponto na vizinhança $\mathcal{N}_2(0.25)$ e encontra um ponto em $\mathcal{N}_2(0.5)$, que é uma vizinhança mais ampla do caminho central. Além disso, a medida de complementaridade no ponto encontrado é inferior em comparação ao ponto antigo. O lema a seguir mostra que o passo corretor toma um ponto na vizinhança $\mathcal{N}_2(0.5)$ e encontra um ponto na vizinhança $\mathcal{N}_2(0.25)$, que é uma vizinhança menos ampla do caminho central, mantendo a medida de complementaridade inalterada.

Lema 1.13. *Suponha que $(x, \lambda, s) \in \mathcal{N}_2(0.5)$ e seja $(\Delta x, \Delta \lambda, \Delta s)$ a direção obtida em (1.8) quando $\sigma = 1$. Então, temos que*

$$(x(1), \lambda(1), s(1)) \in \mathcal{N}_2(0.25) \text{ e } \mu(1) = \mu,$$

onde $(x(1), \lambda(1), s(1)) = (x, \lambda, s) + (\Delta x, \Delta \lambda, \Delta s)$, e $\mu(1)$ é a medida de complementaridade de $(x(1), \lambda(1), s(1))$.

Demonstração. Como $\sigma = 1$, temos que $\mu(1) = (1 - \alpha(1 - \sigma))\mu = \mu$. Uma vez que $\alpha = 1$ e $\theta = 0.5$, temos da desigualdade (1.42) que

$$\begin{aligned} \|X(1)S(1)e - \mu(1)e\| &\leq |1 - 1| \cdot 0.5\mu + 1^2 \left(\frac{(0.5)^2 + n(1 - 1)^2}{2^{\frac{3}{2}}(1 - 0.5)} \right) \mu \\ &= \frac{1}{4} \cdot \frac{1}{2^{\frac{1}{2}}} \mu \leq \frac{1}{4} \mu = \frac{1}{4} \mu(1). \end{aligned}$$

A verificação de que as condições de viabilidade do ponto $(x(1), \lambda(1), s(1))$ são satisfeitas é realizada de forma idêntica a demonstração do Teorema 1.8. Para mostrar que $x_i(\alpha)$ e $s_i(\alpha)$ são estritamente positivos, para todo $i = 1, 2, \dots, n$, veja que $\|X(1)S(1)e - \mu(1)e\| \leq \frac{1}{4}\mu(1)$ implica que

$$\begin{aligned} \mu(1) - x_i(1)s_i(1) &\leq \frac{1}{4}\mu(1), \forall i = 1, \dots, n \\ \Leftrightarrow 0 &< \frac{3}{4}\mu \leq x_i(1)s_i(1), \forall i = 1, \dots, n, \end{aligned}$$

provando o resultado. □

O teorema a seguir demonstra a convergência do Algoritmo 3.

Teorema 1.14. *Dado $\varepsilon > 0$, suponha que o ponto inicial no Algoritmo 3 seja $(x_0, \lambda_0, s_0) \in \mathcal{N}_2(0.25)$ e satisfaça*

$$\mu_0 \leq \frac{1}{\varepsilon^\kappa},$$

para algum $\kappa > 0$. Então, existe um índice K , com $K = O(\sqrt{n} \log \frac{1}{\varepsilon})$ tal que

$$\mu_k \leq \varepsilon, \forall k \geq K.$$

Demonstração. Primeiramente, temos de (1.43) e do Lema 1.13 que

$$\mu_{k+2} = \mu_{k+1} \leq \left(1 - \frac{0.4}{\sqrt{n}}\right) \mu_k, \forall k = 0, 2, 4, \dots \quad (1.44)$$

Se tomarmos o logaritmo de ambos os lados em (1.44), temos que

$$\log(\mu_{k+2}) \leq \log\left(\left(1 - \frac{0.4}{\sqrt{n}}\right) \mu_k\right) = \log\left(1 - \frac{0.4}{\sqrt{n}}\right) + \log(\mu_k), \forall k = 0, 2, 4, \dots$$

Aplicando a desigualdade acima recursivamente, temos que

$$\log(\mu_k) \leq \log\left(1 - \frac{0.4}{\sqrt{n}}\right) + \log(\mu_k) \leq 2 \log\left(1 - \frac{0.4}{\sqrt{n}}\right) + \log(\mu_{k-2})$$

$$\leq \dots \leq \frac{k}{2} \log \left(1 - \frac{0.4}{\sqrt{n}} \right) + \log(\mu_0), \forall k = 2, 4, 6, \dots$$

Por hipótese, temos que $\mu_0 \leq \frac{1}{\varepsilon^\kappa}$, então pela desigualdade acima, segue que

$$\begin{aligned} \log(\mu_k) &\leq \frac{k}{2} \log \left(1 - \frac{0.4}{\sqrt{n}} \right) + \log(\mu_0) = \frac{k}{2} \log \left(1 - \frac{0.4}{\sqrt{n}} \right) + \log \left(\frac{1}{\varepsilon^\kappa} \right) \\ &= \frac{k}{2} \log \left(1 - \frac{0.4}{\sqrt{n}} \right) + \log \left(\left(\frac{1}{\varepsilon} \right)^\kappa \right) = \frac{k}{2} \log \left(1 - \frac{0.4}{\sqrt{n}} \right) + \kappa \log \left(\frac{1}{\varepsilon} \right). \end{aligned}$$

Podemos utilizar a estimativa (1.40) dada na demonstração do Teorema 1.9 para concluir que

$$\log(\mu_k) \leq \frac{k}{2} \left(-\frac{0.4}{\sqrt{n}} \right) + \kappa \log \left(\frac{1}{\varepsilon} \right).$$

Veja que se $\frac{k}{2} \left(-\frac{0.4}{\sqrt{n}} \right) + \kappa \log \left(\frac{1}{\varepsilon} \right) \leq \log(\varepsilon)$, temos que $\mu_k \leq \varepsilon$. Mas,

$$\begin{aligned} \frac{k}{2} \left(-\frac{0.4}{\sqrt{n}} \right) + \kappa \log \left(\frac{1}{\varepsilon} \right) &\leq \log(\varepsilon) \Leftrightarrow -\frac{0.2k}{\sqrt{n}} \leq \log(\varepsilon) - \kappa \log \left(\frac{1}{\varepsilon} \right) \\ \Leftrightarrow -\frac{0.2k}{\sqrt{n}} &\leq -\log(1) + \log(\varepsilon) - \kappa \log \left(\frac{1}{\varepsilon} \right) \Leftrightarrow -\frac{0.2k}{\sqrt{n}} \leq -\log \left(\frac{1}{\varepsilon} \right) - \kappa \log \left(\frac{1}{\varepsilon} \right) \\ &\Leftrightarrow -\frac{0.2k}{\sqrt{n}} \leq -(1 + \kappa) \log \left(\frac{1}{\varepsilon} \right) \Leftrightarrow \frac{0.2k}{\sqrt{n}} \geq (1 + \kappa) \log \left(\frac{1}{\varepsilon} \right) \\ &\Rightarrow \frac{k}{\sqrt{n}} \geq (1 + \kappa) \log \left(\frac{1}{\varepsilon} \right) \Rightarrow k \geq (1 + \kappa) \sqrt{n} \log \left(\frac{1}{\varepsilon} \right). \end{aligned}$$

Definindo $K = (1 + \kappa) \sqrt{n} \log \left(\frac{1}{\varepsilon} \right)$, temos o desejado.

□

Se compararmos o Teorema 1.14 com o Teorema 1.9, podemos perceber que a ordem do número de iterações não é a mesma. Na verdade, como $n \geq 1$, temos que $\sqrt{n} \leq n$, ou seja, a ordem proveniente do Teorema 1.14 é menor. Neste caso, a teoria de complexidade favorece o Algoritmo 3 quando o comparamos com o Algoritmo 2.

1.6 Algoritmo Preditor-Corretor de Mehrotra

Nesta seção, apresentaremos o algoritmo preditor-corretor de Mehrotra [8]. Este algoritmo é a base da maioria dos softwares que lidam com problemas de programação linear via métodos de pontos interiores, segundo [1] (pág. 14) e [2] (pág. 393). O método preditor-corretor de Mehrotra busca compensar o erro obtido pelo método de Newton ao

fazer uma aproximação linear dos produtos $x_i s_i$, $i = 1, \dots, n$. Considere a direção afim $(\Delta x_a, \Delta \lambda_a, \Delta s_a)$, a qual é dada pelo seguinte sistema

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x_a \\ \Delta \lambda_a \\ \Delta s_a \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -X S e \end{bmatrix}. \quad (1.45)$$

Se um passo completo for dado nessa direção, temos que

$$(x_i + \Delta x_a^i)(s_i + \Delta s_a^i) = x_i s_i + x_i [\Delta s_a]_i + s_i [\Delta x_a]_i + [\Delta x_a]_i [\Delta s_a]_i.$$

Pelo terceiro bloco de (1.45), temos que $x_i [\Delta s_a]_i + s_i [\Delta x_a]_i = -x_i s_i$, logo

$$(x_i + \Delta x_a^i)(s_i + \Delta s_a^i) = [\Delta x_a]_i [\Delta s_a]_i.$$

Observe que o valor de $(x_i + \Delta x_a^i)(s_i + \Delta s_a^i)$ não é 0, como o esperado. Ao invés disso, temos que $x_i s_i = [\Delta x_a]_i [\Delta s_a]_i$. Na tentativa de corrigir este erro e obter $x_i s_i$ mais próximo de 0, podemos encontrar outro passo $(\Delta x_c, \Delta \lambda_c, \Delta s_c)$ visando minimizar $[\Delta x_a]_i [\Delta s_a]_i$ ao resolver o seguinte sistema:

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x_c \\ \Delta \lambda_c \\ \Delta s_c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\Delta X_a \Delta S_a e \end{bmatrix}. \quad (1.46)$$

Em muitos casos, segundo [2], o passo combinado $(\Delta x_a, \Delta \lambda_a, \Delta s_a) + (\Delta x_c, \Delta \lambda_c, \Delta s_c)$ reduz a medida de complementaridade de forma mais eficiente do que simplesmente tomar a direção afim. Dessa forma, pode ser mais interessante dar um passo que seja uma combinação linear de $(\Delta x_a, \Delta \lambda_a, \Delta s_a)$ e $(\Delta x_c, \Delta \lambda_c, \Delta s_c)$.

Em [18], explorou-se o uso de passos compostos pela combinação linear de duas direções em métodos de pontos interiores para a solução de problemas de programação linear. Métodos de pontos interiores que fazem uma combinação linear de diversas direções, procurando encontrar uma combinação linear adequada, isto é, encontrar tamanhos de passos apropriados para cada direção, são chamados métodos de busca multidimensional.

Para [18] (pág. 215), o algoritmo de Mehrotra [8] está de certa forma relacionado com as buscas multidimensionais (mais especificamente bidimensionais), no sentido de que em um único passo, duas direções são calculadas (uma que melhora o custo e outra recentralizadora), mas ambas são somadas para formarem uma única direção.

O que diferencia o algoritmo de Mehrotra dos métodos de buscas multidimensionais, destaca [18] (pág. 216), é o fato de que ele adota uma busca linear na direção combinada ao invés de uma busca multidimensional para determinar uma combinação linear adequada.

Uma característica marcante do método preditor-corretor de Mehrotra é o uso inteligente do parâmetro centralizador σ . De forma sucinta, se o passo afim proporciona uma redução significativa da medida de complementaridade, é razoável escolher σ próximo de 0. Por outro lado, se não for possível reduzir a medida de complementaridade consideravelmente nesta direção sem violar a não negatividade de x_i e s_i , a escolha de σ próximo de 1 irá melhorar a centralidade do próximo iterando, permitindo que na próxima iteração seja possível dar um passo maior. O algoritmo preditor-corretor de Mehrotra propõe uma escolha de σ que se autoajusta de acordo com a situação.

O Algoritmo de Mehrotra, que ainda será apresentado (Algoritmo 4), realiza o cálculo dos maiores tamanho de passo α_a^p e α_a^d que podem ser dados na direção afim para ambos os problemas primal e dual, respectivamente, de forma separada. Isso é feito por meio de uma busca linear, determinando α_a^p e α_a^d da seguinte forma:

$$\alpha_a^p = \min \left\{ 1, \min \left\{ -\frac{x_i}{[\Delta x_a]_i} \mid [\Delta x_a]_i < 0 \right\} \right\}, \quad \text{e} \quad (1.47)$$

$$\alpha_a^d = \min \left\{ 1, \min \left\{ -\frac{s_i}{[\Delta s_a]_i} \mid [\Delta s_a]_i < 0 \right\} \right\}. \quad (1.48)$$

Depois, calcula-se a medida de complementaridade que seria obtida caso fosse dado o passo nesta direção para os valores α_a^p e α_a^d calculados, isto é,

$$\mu_a = (x + \alpha_a^p \Delta x_a)^T (s + \alpha_a^d \Delta s_a) / n. \quad (1.49)$$

O valor σ é escolhido de acordo com a seguinte heurística, que tem origem em [8], e de acordo com [2] (pág. 408), não possui uma justificativa analítica sólida, mas que mostrou bons resultados na prática:

$$\sigma = \left(\frac{\mu_a}{\mu} \right)^3.$$

O passo no qual calcula-se a direção afim é chamado passo preditor. Neste passo, resolve-se o sistema (1.45), calcula-se os tamanhos de passo α_a^p e α_a^d , e μ_a e σ . Isto fornece os dados necessários para resolver o sistema (1.46). O passo em que resolvemos este sistema é chamado passo corretor. O Algoritmo 4 combina o lado direito de ambos os sistemas

(1.45) e (1.46) e encontra uma direção $(\Delta x, \Delta \lambda, \Delta s)$ satisfazendo o seguinte sistema:

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe - \Delta X_a \Delta S_a e + \sigma \mu e \end{bmatrix}. \quad (1.50)$$

O sistema (1.50) encontra uma direção que leva em consideração o erro na aproximação linear $\Delta X_a \Delta S_a$, e também a centralidade através do termo $\sigma \mu e$, a qual é controlada pela constante σ .

Encontrada a direção $(\Delta x, \Delta \lambda, \Delta s)$, é importante determinar o maior passo possível tal que x e s permaneçam não negativos. Isto é feito considerando os maiores valores α^p e α^d tais que $x_k + \alpha^p \Delta x_k \geq 0$ e $s_k + \alpha^d \Delta s_k \geq 0$. Considerando essas inequações apenas no caso em que $\Delta x_i < 0$ e $\Delta s_i < 0$ e depois impondo a igualdade, podemos isolar α e definir

$$\alpha_{\max}^p = \min \left\{ -\frac{[x_k]_i}{[\Delta x_k]_i} \mid [\Delta x_k]_i < 0 \right\} \text{ e } \alpha_{\max}^d = \min \left\{ -\frac{[s_k]_i}{[\Delta s_k]_i} \mid [\Delta s_k]_i < 0 \right\}.$$

Os valores α_{\max}^p e α_{\max}^d assim definidos são os maiores tamanhos de passo para os problemas primal e dual, respectivamente, tais que o próximo iterando respeita as restrições de não negatividade. Tomando $\alpha_k^p \in (0, \alpha_{\max}^p)$ e $\alpha_k^d \in (0, \alpha_{\max}^d)$, podemos definir o próximo iterando como sendo

$$x_{k+1} = x_k + \alpha_k^p \Delta x_k, \text{ e } (\lambda_{k+1}, s_{k+1}) = (\lambda_k, s_k) + \alpha_k^d (\Delta \lambda_k, \Delta s_k).$$

O Algoritmo 4 escolhe α_k^p e α_k^d da seguinte forma:

$$\alpha_k^p = \min \{1, \eta_k \alpha_{\max}^p\} \text{ e } \alpha_k^d = \min \{1, \eta_k \alpha_{\max}^d\}, \quad (1.51)$$

onde $\eta_k \in [0.9, 1.0)$ é escolhido de modo que $\eta_k \rightarrow 1$ conforme os iterandos convergem para a solução, para acelerar a convergência.

O ponto inicial do Algoritmo 4 é tomado da seguinte forma [8]: resolve-se os seguintes problemas de programação linear:

$$\begin{array}{ll} \min & \frac{1}{2} x^T x \\ \text{s.a. :} & Ax = b \end{array} \quad \text{e} \quad \begin{array}{ll} \min & \frac{1}{2} s^T s \\ \text{s.a. :} & A^T \lambda + s = c. \end{array}$$

Resolvendo os problemas acima, encontra-se pontos \tilde{x} e $(\tilde{\lambda}, \tilde{s})$ tais que o ponto $(\tilde{x}, \tilde{\lambda}, \tilde{s})$ é viável e a norma de \tilde{x} e \tilde{s} são as menores possíveis para pontos viáveis.

As soluções para ambos os problemas podem ser explicitadas da seguinte forma [2]:

$$\tilde{x} = A^T(AA^T)^{-1}b, \quad \tilde{\lambda} = (AA^T)Ac \quad \text{e} \quad \tilde{s} = c - A^T\tilde{\lambda}.$$

Como \tilde{x} e \tilde{s} não necessariamente possuem todas as entradas positivas, é feita uma correção da seguinte forma:

$$\hat{x} = \tilde{x} + \delta_x e \quad \text{e} \quad \hat{s} = \tilde{s} + \delta_s e,$$

onde

$$\delta_x = \max\{-\frac{3}{2} \min_i \tilde{x}_i, 0\} \quad \text{e} \quad \delta_s = \max\{-\frac{3}{2} \min_i \tilde{s}_i, 0\}.$$

A correção acima garante que todas as entradas de \hat{x} e \hat{s} são não negativas. Uma segunda correção é realizada a fim de evitar que estes vetores possuam entradas muito próximas de 0:

$$x_0 = \hat{x} + \hat{\delta}_x e \quad \text{e} \quad s_0 = \hat{s} + \hat{\delta}_s e, \quad (1.52)$$

onde

$$\hat{\delta}_x = \frac{1}{2} \frac{\hat{x}^T \hat{s}}{e^T \hat{s}} \quad \text{e} \quad \hat{\delta}_s = \frac{1}{2} \frac{\hat{x}^T \hat{s}}{e^T \hat{x}}.$$

Por meio deste procedimento, encontra-se um ponto (x_0, λ_0, s_0) que satisfaz de forma aproximada as restrições de viabilidade $Ax = b$ e $A^T\lambda + s = c$, e também as condições de positividade $x > 0$ e $s > 0$, evitando que x e s possuam componentes excessivamente grandes. O Algoritmo 4, logo abaixo, foi encontrado em [2] (Algoritmo 14.3).

Algoritmo 4 Algoritmo Preditor-Corretor de Mehrotra [8]

Calcule (x_0, λ_0, s_0) como em (1.52);

para $k = 0, 1, 2, \dots$ **faça**:

Faça $(x, \lambda, s) = (x_k, \lambda_k, s_k)$ e resolva (1.45) para $(\Delta x_a, \Delta \lambda_a, \Delta s_a)$;

Calcule α_a^p, α_a^d e μ_a como em (1.47), (1.48) e (1.49);

Defina $\sigma = (\mu_a/\mu)^3$;

Resolva (1.50) para $(\Delta x, \Delta \lambda, \Delta s)$;

Calcule α_k^p e α_k^d de (1.51);

Defina

$$x_{k+1} = x_k + \alpha_k^p \Delta x;$$

$$(\lambda_{k+1}, s_{k+1}) = (\lambda_k, s_k) + \alpha_k^d (\Delta \lambda, \Delta s);$$

fim

Em geral, métodos preditores-corretores executam uma iteração do método de Newton em ambos os passos preditor e corretor. Uma vez que em nenhum desses passos o sistema $F_{\sigma_k \mu_k}(w) = 0$ é resolvido de forma exata, uma pergunta que pode surgir é se há a necessidade de se utilizar uma direção tão boa quanto a direção de Newton, especialmente no passo corretor, que visa apenas melhorar a centralidade do próximo ponto. Neste sentido, estamos interessados em estudar métodos preditores-corretores no qual o passo corretor executa um método Quasi-Newton.

Em [5], é apresentada uma maneira inteligente de implementar o método de Broyden em sua versão “ruim” visando reduzir seu custo computacional. Além disso, ao utilizar o método de Broyden “ruim” no lugar do método de Newton para realizar os passos no método de pontos interiores, foi observado por [5] que o método se comportou naturalmente de forma parecida com os métodos preditores-corretores. Com base nesses fatos, optamos por estudar o método de Broyden em sua versão “boa”, que muitas vezes é referido apenas como o método de Broyden. No próximo capítulo, abordaremos este método em mais detalhes.

MÉTODOS DE BROYDEN

Seja $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma função para a qual se deseja encontrar x^* tal que $F(x^*) = 0$. O método de Newton resolve este tipo de problema considerando direções d_k da forma $d_k = -[J(x_k)]^{-1}F(x_k)$. No contexto de pontos interiores, quando $F : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ é dada por

$$F(x, \lambda, s) = \begin{bmatrix} r_c \\ r_b \\ XSe \end{bmatrix},$$

onde $r_c = A^T\lambda + s - c$ e $r_b = Ax - b$, isto origina o sistema (1.6). O método de Broyden, por sua vez, realiza uma aproximação de $J(x_k)$ por uma matriz B_k , tomando direções da forma $d_k = -B_k F(x_k)$. Mais precisamente, dada uma aproximação $B_k \in \mathbb{R}^{n \times n}$ para $J(x_k)$, considera-se o seguinte modelo para F :

$$M_{k+1}(x) = F(x_{k+1}) + B_{k+1}(x - x_{k+1}),$$

onde $B_{k+1} \in \mathbb{R}^{n \times n}$ é uma aproximação para $J(x_{k+1})$ que deseja-se encontrar. Depois, é exigido que a seguinte equação seja satisfeita:

$$F(x_k) = F(x_{k+1}) + B_{k+1}(x_k - x_{k+1}),$$

ou equivalentemente,

$$B_{k+1}(x_{k+1} - x_k) = F(x_{k+1}) - F(x_k). \tag{2.1}$$

A equação (2.1) é chamada equação secante. Costuma-se denotar $s_k = x_{k+1} - x_k$ e $y_k = F(x_{k+1}) - F(x_k)$ e escrever a equação secante como segue:

$$B_{k+1}s_k = y_k. \tag{2.2}$$

Quando $n > 1$, existem muitas possibilidades de escolha de B_{k+1} satisfazendo a equação secante. O método de Broyden busca fazer a escolha que minimiza $\|M_{k+1} - M_k\|_F$, onde

$\|\cdot\|_F$ denota a norma de Frobenius. Em outras palavras, procura-se minimizar a diferença entre o modelo novo e o antigo de F . Para isto, impõe-se que $B_{k+1} - B_k$ seja uma matriz de posto 1 dada por

$$B_{k+1} - B_k = \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k}.$$

Neste caso, define-se B_{k+1} como sendo

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k}. \quad (2.3)$$

A equação (2.3) foi proposta por C. Broyden em 1965, e chama-se atualização de Broyden, ou atualização secante. É importante destacar que o método de Broyden não aproxima $J(x_k)$ do zero a cada iteração. Na verdade, ele parte de uma aproximação inicial para $J(x_0)$, e faz sucessivas atualizações sobre esta aproximação inicial. Veja abaixo um algoritmo para o método de Broyden [10] para resolver $F(x) = 0$.

Algoritmo 5 Método de Broyden

Dada $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$ e $B_0 \in \mathbb{R}^{n \times n}$

para $k = 0, 1, 2, \dots$ **faça:**

 Resolva $B_k s_k = -F(x_k)$ para s_k

 Defina $x_{k+1} = x_k + s_k$

 Defina $y_k = F(x_{k+1}) - F(x_k)$

 Defina $B_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k}$

fim

2.1 Convergência do Método de Broyden

Nesta seção, será demonstrada a convergência q -linear do método de Broyden, e posteriormente, dentro de determinadas hipóteses, sua convergência q -superlinear. Os resultados apresentados e demonstrados ao longo desta seção foram baseados em [10].

Para realizar a demonstração do Teorema da Convergência Local do Método de Broyden, serão necessários alguns resultados preliminares, os quais serão apresentados a seguir. Começaremos demonstrando o Teorema 2.1 e o Lema 2.3, que podem ser encontrados em [10] como Teorema 3.1.4 e Lema 4.1.2, respectivamente.

Teorema 2.1. *Seja $\|\cdot\|$ uma norma qualquer em $\mathbb{R}^{n \times n}$ que satisfaz $\|AB\| \leq \|A\| \cdot \|B\|, \forall A, B \in \mathbb{R}^{n \times n}$, e $\|I\| = 1$ e seja $E \in \mathbb{R}^{n \times n}$. Se $\|E\| < 1$, então $(I - E)^{-1}$ existe e*

$$\|(I - E)^{-1}\| \leq \frac{1}{1 - \|E\|}. \quad (2.4)$$

Se A é não singular e $\|A^{-1}(B - A)\| < 1$, então B é não singular e

$$\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}(B - A)\|}. \quad (2.5)$$

Demonstração. Primeiramente, mostremos que $S_k = (I + E + E^2 + \dots + E^k)$ é uma sequência de Cauchy em $\mathbb{R}^{n \times n}$, e portanto converge. Sejam $p, q \in \mathbb{N}$, tais que $p > q$. Então, temos da desigualdade triangular e da propriedade que a norma satisfaz por hipótese que

$$\begin{aligned} \|S_p - S_q\| &= \|E^{p+1} + \dots + E^q\| \leq \|E^{p+1}\| + \dots + \|E^q\| \leq \|E\|^{p+1} + \dots + \|E\|^q \\ &= \|E\|^{p+1} (1 + \dots + \|E\|^{q-p-1}) \leq \|E\|^{p+1} \sum_{k=0}^{\infty} \|E\|^k. \end{aligned}$$

Como $\|E\| < 1$, segue que $\sum_{k=0}^{\infty} \|E\|^k$ converge para $\frac{1}{1 - \|E\|}$. Além disso, o termo geral desta série, $\|E\|^k$, converge para 0, e desse modo, dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que se $k > n_0$, então $\|E\|^k < (1 - \|E\|)\varepsilon$. Sendo assim, se tomarmos p e q tais que $p, q > n_0$, segue que

$$\|S_p - S_q\| \leq \|E\|^{p+1} \sum_{k=0}^{\infty} \|E\|^k < (1 - \|E\|)\varepsilon \frac{1}{1 - \|E\|} = \varepsilon,$$

e portanto S_k é uma sequência de Cauchy em $\mathbb{R}^{n \times n}$, e conseqüentemente convergente.

Agora, mostremos que $\lim_{k \rightarrow \infty} S_k$ é a inversa de $(I - E)$. De fato,

$$\begin{aligned} (I - E) \lim_{k \rightarrow \infty} S_k &= \lim_{k \rightarrow \infty} (I - E)S_k = \lim_{k \rightarrow \infty} (I - E)(I + \|E\| + \dots + \|E\|^k) \\ &= \lim_{k \rightarrow \infty} [(I + \|E\| + \dots + \|E\|^k) - (\|E\| + \|E\|^2 + \dots + \|E\|^{k+1})] \\ &= \lim_{k \rightarrow \infty} [I - \|E\|^{k+1}] = \lim_{k \rightarrow \infty} I - \lim_{k \rightarrow \infty} \|E\|^k = I - \lim_{k \rightarrow \infty} \|E\|^k, \end{aligned}$$

e como $\|E\|^k \rightarrow 0$, segue que

$$(I - E) \lim_{k \rightarrow \infty} S_k = I - 0 = I.$$

Para provar que $(\lim_{k \rightarrow \infty} S_k)(I - E) = I$ o processo é análogo. Portanto $(I - E)^{-1}$ existe e $(I - E)^{-1} = \lim_{k \rightarrow \infty} S_k$. Para obter (2.4), basta notar que

$$\|(I - E)^{-1}\| = \|\lim_{k \rightarrow \infty} S_k\| = \lim_{k \rightarrow \infty} \|S_k\| = \lim_{k \rightarrow \infty} \|I + E + \dots + E^k\| \leq \lim_{k \rightarrow \infty} \sum_{j=0}^k \|E^j\|$$

$$\leq \lim_{k \rightarrow \infty} \sum_{j=0}^k \|E\|^j = \sum_{k=0}^{\infty} \|E\|^k = \frac{1}{1 - \|E\|}.$$

Agora, Se $\|A^{-1}(B - A)\| \leq 1$, então pela primeira parte deste Lema, temos que $I - A^{-1}(B - A)$ é inversível. Seja $C = (I - A^{-1}(B - A))^{-1}$. Temos que

$$C^{-1} = I - A^{-1}(B - A) = A^{-1}A - A^{-1}(B - A) = A^{-1}(A + B - A) = A^{-1}B,$$

e portanto $B = AC^{-1}$. Sendo assim, $B^{-1} = CA^{-1}$, isto é, B é inversível. Também temos da primeira parte deste Lema que

$$\|C\| = \|(I - A^{-1}(B - A))^{-1}\| \leq \frac{1}{1 - \|A^{-1}(B - A)\|},$$

e assim,

$$\|B^{-1}\| = \|CA^{-1}\| \leq \|C\| \|A^{-1}\| \leq \frac{1}{1 - \|A^{-1}(B - A)\|} \|A^{-1}\| = \frac{\|A^{-1}\|}{1 - \|A^{-1}(B - A)\|},$$

provando (2.5). □

Definição 2.2. Dados $x, y \in \mathbb{R}^n$, com $x \neq y$, considere o seguinte conjunto:

$$(x, y) = \{x + t(y - x) \mid t \in (0, 1)\}.$$

O conjunto (x, y) é chamado segmento aberto que liga x e y .

Lema 2.3. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função continuamente diferenciável em um aberto convexo $D \subset \mathbb{R}^n$. Para qualquer $x, x + p \in D$,

$$f(x + p) = f(x) + \int_0^1 \nabla f(x + tp)^T p \, dt, \quad (2.6)$$

e existe $z \in (x, x + p)$

$$f(x + p) = f(x) + \nabla f(z)^T p. \quad (2.7)$$

Demonstração. Considere $g : \mathbb{R} \rightarrow \mathbb{R}$ tal que $g(t) = f(x + tp)$. Temos que para todo $0 \leq s \leq 1$,

$$\frac{dg}{dt}(s) = \nabla f(x + sp)^T p.$$

Por outro lado, pelo Teorema Fundamental do Cálculo, temos que

$$g(1) - g(0) = \int_0^1 g'(t) \, dt$$

$$\Leftrightarrow f(x+p) - f(x) = \int_0^1 \nabla f(x+tp)^T p \, dt.$$

Por fim, pelo Teorema do Valor Médio, segue que existe $\xi \in (0, 1)$ tal que

$$g(1) - g(0) = g'(\xi) \Leftrightarrow f(x+p) - f(x) = \nabla f(x + \xi p)^T p \Leftrightarrow f(x+p) - f(x) = \nabla f(z)^T p,$$

onde $z = x + \xi p$.

□

O Lema 2.3 nos dá a equação (2.7), que se assemelha à fórmula de Taylor. A fórmula de Taylor pode ser utilizada para aproximar uma função diferenciável f em uma vizinhança de um ponto x . Contudo, o polinômio de Taylor sozinho não é suficiente para aproximar f de forma exata, sendo necessário introduzir um resto satisfazendo uma propriedade especial para que a fórmula de Taylor descreva f de forma exata em uma vizinhança de x .

Por sua vez, a fórmula (2.7) não necessita a introdução desse resto para aproximar f em um ponto $x+p$ de forma exata, embora z dependa de $x+p$ e a demonstração do Lema 2.3 não providencie uma forma de encontrar z (pelo Teorema do Valor Médio podemos apenas afirmar que z existe). Mesmo assim, para fins teóricos, a fórmula dada pelo Lema 2.3 é de grande utilidade.

Definição 2.4. *Sejam $m, n > 0$, $G : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, seja $\|\cdot\|$ uma norma em \mathbb{R}^n e $\|\|\cdot\|\|$ uma norma em $\mathbb{R}^{m \times n}$. G é dita ser Lipschitz contínua em x se existir um aberto $D \subset \mathbb{R}^n$, $x \in D$, e uma constante L tal que para todo $v \in D$,*

$$\|\|G(v) - G(x)\|\| \leq L\|v - x\|. \quad (2.8)$$

A constante L é chamada constante de Lipschitz para G em x . Além disso, para qualquer D satisfazendo (2.8), G é chamada Lipschitz contínua em x na vizinhança D . Se (2.8) for válida para todo $x \in D$, então $G \in Lip_L(D)$, isto é, G é Lipschitz em D com constante de Lipschitz L .

Definição 2.5. *Fixada uma norma de vetores $\|\cdot\|$ em \mathbb{R}^n , podemos definir uma norma matricial $\|\|\cdot\|\|$ em $\mathbb{R}^{n \times n}$ da seguinte forma:*

$$\|\|M\|\| = \max_{v \in \mathbb{R}^n, v \neq 0} \left\{ \frac{\|Mv\|}{\|v\|} \right\}, \quad \forall M \in \mathbb{R}^{n \times n}.$$

A norma $\|\|\cdot\|\|$ é chamada norma matricial induzida por $\|\cdot\|$.

Observação 2.6. A norma matricial induzida por uma norma de vetores recebeu a notação $||| \cdot |||$ apenas para explicitar a diferença entre ela e a norma $\| \cdot \|$ na definição. Ao longo do texto, no entanto, adotaremos simplesmente $\| \cdot \|$ para nos referirmos à esta norma matricial, a menos que seja dito o contrário.

Lema 2.7. Seja $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma função continuamente diferenciável no aberto e convexo $D \subset \mathbb{R}^n$, $x \in D$, e seja J Lipschitz contínua em x na vizinhança D , com constante de Lipschitz L , onde foi utilizada uma norma de vetores, e a norma de matrizes induzida por essa norma de vetores. Então, para cada $u, v \in D$,

$$\|F(v) - F(u) - J(x)(v - u)\| \leq L \frac{\|v - x\| + \|u - x\|}{2} \|v - u\|.$$

Demonstração. Como $v = u + (v - u)$, temos que

$$\begin{aligned} F(v) - F(u) &= F(u + (v - u)) - F(u) = \int_0^1 J(u + t(v - u))(v - u) dt \\ \Leftrightarrow F(v) - F(u) - J(x)(v - u) &= \int_0^1 J(u + t(v - u))(v - u) dt - J(x)(v - u) \\ \Leftrightarrow F(v) - F(u) - J(x)(v - u) &= \int_0^1 [J(u + t(v - u)) - J(x)](v - u) dt. \end{aligned}$$

Segue de uma propriedade de integração que

$$\begin{aligned} \|F(v) - F(u) - J(x)(v - u)\| &= \left\| \int_0^1 [J(u + t(v - u)) - J(x)](v - u) dt \right\| \\ &\leq \int_0^1 \| [J(u + t(v - u)) - J(x)](v - u) \| dt \leq \int_0^1 \|J(u + t(v - u)) - J(x)\| \|v - u\| dt. \end{aligned}$$

Como J é Lipschitz contínua em x , segue que

$$\|F(v) - F(u) - J(x)(v - u)\| \leq \int_0^1 L \|u + t(v - u) - x\| \|v - u\| dt. \quad (2.9)$$

Note que $x = tx + (1 - t)x$, e portanto

$$\begin{aligned} \|u + t(v - u) - x\| &= \|u + t(v - u) - tx - (1 - t)x\| = \|u + tv - tu - tx - (1 - t)x\| \\ &= \|(1 - t)u + t(v - x) - (1 - t)x\| = \|(1 - t)(u - x) + t(v - x)\|. \end{aligned}$$

Substituindo este termo em (2.9), temos que

$$\|F(v) - F(u) - J(x)(v - u)\| \leq \int_0^1 L \|(1 - t)(u - x) + t(v - x)\| \|v - u\| dt$$

$$\begin{aligned}
&\leq L\|v - u\| \int_0^1 ((1-t)\|u - x\| + t\|v - x\|) dt \\
&\leq L\|v - u\| \left(\left(t - \frac{t^2}{2} \right) \|u - x\| + \frac{t^2}{2} \|v - x\| \right) \Big|_{t=0}^{t=1} \\
&= L \frac{\|u - x\| + \|v - x\|}{2} \|v - u\|.
\end{aligned}$$

□

Lema 2.8. *Sejam F e J nas condições do Lema 2.7. Então, para cada $x + p \in D$,*

$$\|F(x + p) - F(x) - J(x)p\| \leq \frac{L}{2} \|p\|^2. \quad (2.10)$$

Demonstração. Seja $u = x$ e $v = x + p$. Temos que $p = v - u$, e pelo Lema 2.7, segue que

$$\begin{aligned}
\|F(x + p) - F(x) - J(x)p\| &= \|F(v) - F(u) - J(x)(v - u)\| \\
&\leq L \frac{\|v - u\| + \|u - u\|}{2} \|v - u\| = \frac{L}{2} \|p\|^2.
\end{aligned}$$

□

Lema 2.9. *Considere a matriz $I - \frac{s_k s_k^T}{s_k^T s_k}$ e a norma matricial induzida pela norma euclidiana. Temos que*

$$\left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 = 1.$$

Demonstração. Primeiramente, note que $P = I - (s_k s_k^T / s_k^T s_k)$ é a projeção no espaço ortogonal a s_k . Deste modo, temos que $\|Pv\| \leq \|v\|$, para todo $v \in \mathbb{R}^n$. Neste caso,

$$\frac{\|Pv\|}{\|v\|} \leq \frac{\|v\|}{\|v\|} = 1, \quad \forall v \in \mathbb{R}^n, v \neq 0,$$

e conseqüentemente,

$$\left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 = \|P\|_2 = \max_{v \in \mathbb{R}^n, v \neq 0} \left\{ \frac{\|Pv\|}{\|v\|} \right\} \leq 1.$$

Por outro lado, escolha $\bar{v} \in \mathbb{R}^n$ tal que $\bar{v} \perp s_k$. Então,

$$\frac{\|P\bar{v}\|}{\|\bar{v}\|} = \frac{\left\| I\bar{v} - \frac{s_k s_k^T}{s_k^T s_k} \bar{v} \right\|}{\|\bar{v}\|} = \frac{\left\| I\bar{v} - s_k \frac{s_k^T \bar{v}}{s_k^T s_k} \right\|}{\|\bar{v}\|} = \frac{\|\bar{v} - 0\|}{\|\bar{v}\|} = 1,$$

e portanto

$$\left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 = \max_{v \in \mathbb{R}^n, v \neq 0} \left\{ \frac{\|Pv\|}{\|v\|} \right\} \geq \frac{\|P\bar{v}\|}{\|\bar{v}\|} = 1.$$

Como $1 \leq \|P\|_2 \leq 1$, segue que $\|P\|_2 = 1$.

□

Lema 2.10. *Sejam $u, v \in \mathbb{R}^n$ tais que $u = [u_1 \ u_2 \ \cdots \ u_n]^T$ e $v = [v_1 \ v_2 \ \cdots \ v_n]^T$. Tanto na norma de Frobenius quanto na norma induzida pela norma euclidiana, temos que*

$$\|uv^T\| \leq \|u\|_2 \|v\|_2. \quad (2.11)$$

Demonstração. Na norma de Frobenius, temos que

$$\begin{aligned} \|uv^T\|_F &= \left\| [u_i v_j]_{1 \leq i, j \leq n} \right\|_F \left(\sum_{i=1}^n \left(\sum_{j=1}^n (u_i v_j)^2 \right) \right)^{\frac{1}{2}} = \left(\sum_{i=1}^n u_i^2 \left(\sum_{j=1}^n v_j^2 \right) \right)^{\frac{1}{2}} \\ &= \left(\left(\sum_{i=1}^n u_i^2 \right) \left(\sum_{j=1}^n v_j^2 \right) \right)^{\frac{1}{2}} = \left(\sum_{i=1}^n u_i^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^n v_j^2 \right)^{\frac{1}{2}} = \|u\|_2 \|v\|_2. \end{aligned}$$

Na norma induzida pela norma euclidiana, temos que dado $x \in \mathbb{R}^n$ não nulo,

$$\frac{\|uv^T x\|_2}{\|x\|_2} = \frac{\|u\|_2 |v^T x|}{\|x\|_2} \leq \frac{\|u\|_2 \|v\|_2 \|x\|_2}{\|x\|_2} = \|u\|_2 \|v\|_2.$$

Portanto,

$$\|uv^T\|_2 = \max_{x \in \mathbb{R}^n, x \neq 0} \left\{ \frac{\|uv^T x\|_2}{\|x\|_2} \right\} \leq \|u\|_2 \|v\|_2.$$

□

O Lema 2.10 nos fornece uma desigualdade para as normas de Frobenius e induzida pela norma euclidiana que pode parecer familiar. A desigualdade provada para essas normas de matrizes é muito semelhante a desigualdade de Cauchy-Schwarz disponível para vetores. A seguir, vamos demonstrar o Lema 2.11 (correspondente ao Lema 8.2.1 de [10]), que nos dá uma estimativa para o erro cometido ao aproximarmos $J(x_{k+1})$ por B_{k+1} .

Lema 2.11. *Seja $D \subset \mathbb{R}^n$ um conjunto aberto e convexo contendo x_k e x_{k+1} , com $x_k \neq x_*$. Seja $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $J \in Lip_L(D)$, $B_k \in \mathbb{R}^{n \times n}$, e B_{k+1} definida como no método de Broyden. Então, seja na norma de Frobenius ou na norma matricial induzida pela norma euclidiana,*

$$\|B_{k+1} - J(x_{k+1})\| \leq \|B_k - J(x_k)\| + \frac{3L}{2} \|x_{k+1} - x_k\|_2. \quad (2.12)$$

Além disso, se $x_ \in D$ e assumirmos apenas que a Jacobiana J seja Lipschitz contínua em x_* (ao invés de $J \in Lip_L(D)$), então*

$$\|B_{k+1} - J(x_*)\| \leq \|B_k - J(x_*)\| + \frac{L}{2} (\|x_{k+1} - x_*\|_2 + \|x_k - x_*\|_2). \quad (2.13)$$

Demonstração. Provaremos primeiramente que vale (2.13). Seja $J_* = J(x_*)$. Temos que

$$\begin{aligned} B_{k+1} - J_* &= B_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k} - J_* = B_k - J_* + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k} \\ &= B_k - J_* + \frac{(y_k - J_* s_k + J_* s_k - B_k s_k) s_k^T}{s_k^T s_k} = B_k - J_* + \frac{[y_k - J_* s_k - (B_k - J_*) s_k] s_k^T}{s_k^T s_k} \\ &= (B_k - J_*) \left(I - \frac{s_k s_k^T}{s_k^T s_k} \right) + (y_k - J_* s_k) \frac{s_k^T}{s_k^T s_k}. \end{aligned}$$

Assim, tanto na norma de Frobenius quanto na norma induzida pela norma euclidiana, temos que

$$\begin{aligned} \|B_{k+1} - J_*\| &\leq \left\| (B_k - J_*) \left(I - \frac{s_k s_k^T}{s_k^T s_k} \right) \right\| + \left\| (y_k - J_* s_k) \frac{s_k^T}{s_k^T s_k} \right\| \\ &\leq \|B_k - J_*\| \left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 + \|y_k - J_* s_k\|_2 \left\| \frac{s_k^T}{s_k^T s_k} \right\|_2 \\ &= \|B_k - J_*\| \left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 + \|y_k - J_* s_k\|_2 \frac{1}{s_k^T s_k} \|s_k^T\|_2 \\ &= \|B_k - J_*\| \left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 + \|y_k - J_* s_k\|_2 \frac{1}{\|s_k\|_2^2} \|s_k\|_2 \\ &= \|B_k - J_*\| \left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 + \frac{\|y_k - J_* s_k\|_2}{\|s_k\|_2}. \end{aligned}$$

Pelo Lema 2.9, temos que $\|I - (s_k s_k^T)/(s_k^T s_k)\| = 1$. Além disso, $\|y_k - J_* s_k\| = \|F(x_{k+1}) - F(x_k) - J_*(x_{k+1} - x_k)\|$, e pelo Lema 2.7, segue que

$$\begin{aligned} \|B_{k+1} - J_*\| &\leq \|B_k - J_*\| + L \frac{\|x_{k+1} - x_*\|_2 + \|x_k - x_*\|_2}{2} \|x_{k+1} - x_k\|_2 \frac{1}{\|s_k\|_2} \\ &= \|B_k - J_*\| + \frac{L}{2} (\|x_{k+1} - x_*\|_2 + \|x_k - x_*\|_2), \end{aligned}$$

e a desigualdade (2.13) está provada.

Para provar (2.12) iremos proceder de forma análoga. Por simplicidade, denote por J_k a Jacobiana de F calculada em x_k , para todo $k \in \{0, 1, \dots\}$. Temos que

$$\begin{aligned} B_{k+1} - J_{k+1} &= B_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k} - J_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k} - J_{k+1} + J_k - J_k \\ &= B_k - J_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k} - J_{k+1} + J_k. \end{aligned}$$

Como anteriormente, temos dos Lemas 2.9 e 2.7 que

$$\begin{aligned} \|B_{k+1} - J_{k+1}\| &= \left\| (B_k - J_k) \left(I - \frac{s_k s_k^T}{s_k^T s_k} \right) + (y_k - J_k s_k) \frac{s_k^T}{s_k^T s_k} + J_k - J_{k+1} \right\| \\ &\leq \|B_k - J_k\| \left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 + \|y_k - J_k s_k\|_2 \left\| \frac{s_k^T}{s_k^T s_k} \right\|_2 + \|J_{k+1} - J_k\| \end{aligned}$$

$$\begin{aligned}
&\leq \|B_k - J_k\| + \|y_k - J_k s_k\|_2 \left\| \frac{s_k^T}{s_k^T s_k} \right\|_2 + \|J_{k+1} - J_k\| \\
&\leq \|B_k - J_k\| + L \frac{\|x_{k+1} - x_k\|_2 + \|x_k - x_k\|_2}{2} \|x_{k+1} - x_k\|_2 \left\| \frac{s_k^T}{s_k^T s_k} \right\|_2 + \|J_{k+1} - J_k\| \\
&= \|B_k - J_k\| + L \frac{\|x_{k+1} - x_k\|_2^2}{2 \|s_k\|_2} + \|J_{k+1} - J_k\| = \|B_k - J_k\| + \frac{L}{2} \|x_{k+1} - x_k\|_2 + \|J_{k+1} - J_k\|
\end{aligned}$$

Como $J \in \text{Lip}_L(D)$, temos que $\|J_{k+1} - J_k\| \leq L \|x_{k+1} - x_k\|_2 = \|x_{k+1} - x_k\|_2$, e portanto

$$\begin{aligned}
\|B_{k+1} - J_{k+1}\| &\leq \|B_k - J_k\| + \frac{L}{2} \|x_{k+1} - x_k\|_2 + \|x_{k+1} - x_k\|_2 \\
&\leq \|B_k - J_k\| + \frac{3L}{2} \|x_{k+1} - x_k\|_2
\end{aligned}$$

□

Os resultados apresentados até o momento são suficientes para a demonstração de que o método de Broyden está bem definido e que sua convergência é ao menos q -linear. Veremos mais adiante que, na verdade, a convergência do método de Broyden é q -superlinear. A versão do teorema de convergência do método de Broyden que apresentaremos foi encontrada em [10] (Teorema 8.2.2). No entanto, dividiremos este teorema em dois resultados, correspondentes aos Teorema 2.12 e o Corolário 2.17, os quais mostrarão as convergências q -linear e q -superlinear do método, respectivamente.

Teorema 2.12 (Convergência Local do Método de Broyden). *Seja $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuamente diferenciável em um aberto e convexo $D \subset \mathbb{R}^n$. Assuma que exista $x_* \in \mathbb{R}^n$ e $\beta > 0$, tais que $F(x_*) = 0$, $J(x_*)^{-1}$ existe e que $\|J(x_*)^{-1}\| \leq \beta$. Além disso, suponha que $\{B_k\}$ satisfaz (2.13), para alguma constante $L \geq 0$. Então, existem constantes positivas ε e δ tais que se $\|x_0 - x_*\| \leq \varepsilon$ e $\|B_0 - J(x_*)\| \leq \delta$, então o Algoritmo 5 está bem definido e converge q -linearmente para x_* .*

Demonstração. Denote por $\|\cdot\|$ a norma de vetores ou a norma de matrizes induzida pela norma euclidiana, $e_k = x_k - x_*$ e $J_* = J(x_*)$. Escolha ε e δ positivos tais que

$$6\beta\delta \leq 1, \quad (2.14)$$

$$3L\varepsilon \leq 2\delta. \quad (2.15)$$

A demonstração da convergência q -linear consiste em provar, por indução, que

$$\|B_k - J_*\| \leq (2 - 2^{-k})\delta \quad (2.16)$$

e

$$\|e_{k+1}\| \leq \frac{\|e_k\|}{2}, \quad (2.17)$$

para todo $k = 0, 1, \dots$.

Para $k = 0$, temos que

$$\|B_0 - J_*\| \leq \delta = (2 - 2^{-0})\delta.$$

Para provar (2.17) no caso em que $k = 0$, observe primeiramente que como $\|J_*^{-1}\| \leq \beta$, segue que

$$\|J_*^{-1}(B_0 - J_*)\| \leq \|J_*^{-1}\| \|B_0 - J_*\| \leq \beta\delta \leq 2\beta\delta \leq \frac{1}{3} < 1,$$

e portanto, pelo Teorema 2.1, segue que B_0 é inversível (e portanto $x_1 = x_0 - B_0^{-1}F(x_0)$ está bem definido), e

$$\|B_0^{-1}\| \leq \frac{\|J_*^{-1}\|}{1 - \|J_*^{-1}(B_0 - J_*)\|} \leq \frac{\beta}{1 - \frac{1}{3}} = \frac{3\beta}{2}. \quad (2.18)$$

Como B_0 é inversível e $F(x_*) = 0$, temos que

$$\begin{aligned} x_1 - x_* &= x_0 - x_* - B_0^{-1}F(x_0) = x_0 - x_* - B_0^{-1}(F(x_0) - F(x_*)) \\ &\Leftrightarrow B_0(x_1 - x_*) = B_0(x_0 - x_*) - F(x_0) + F(x_*), \end{aligned}$$

e somando e subtraindo J_*e_0 no lado direito da equação acima, temos

$$B_0e_1 = (B_0 - J_*)e_0 - F(x_0) + F(x_*) + J_*e_0. \quad (2.19)$$

Neste caso,

$$\begin{aligned} \|e_1\| &= \|B_0^{-1}[(B_0 - J_*)e_0 - F(x_0) + F(x_*) + J_*e_0]\| \\ &\leq \|B_0^{-1}\| \| (B_0 - J_*)e_0 - F(x_0) + F(x_*) + J_*e_0 \| \\ &\leq \|B_0^{-1}\| (\| (B_0 - J_*)e_0 \| + \| -F(x_0) + F(x_*) + J_*e_0 \|) \\ &\leq \|B_0^{-1}\| (\|B_0 - J_*\| \|e_0\| + \| -F(x_0) + F(x_*) + J_*e_0 \|). \end{aligned} \quad (2.20)$$

Pelo Lema 2.8, temos que

$$\| -F(x_0) + F(x_*) + J_*e_0 \| = \|F(x_0) - F(x_*) - J_*e_0\| \leq \frac{L\|e_0\|^2}{2}. \quad (2.21)$$

Segue de (2.20), (2.18), (2.21) que

$$\|e_1\| \leq \frac{3\beta}{2} \left(\delta \|e_0\| + \frac{L\|e_0\|^2}{2} \right) = \frac{3\beta}{2} \left(\delta + \frac{L\|e_0\|}{2} \right) \|e_0\| \leq \left(\frac{3\beta\delta}{2} + \frac{3\beta L\varepsilon}{4} \right) \|e_0\|.$$

Das inequações (2.14) e (2.15), temos que

$$\|e_1\| \leq \left(\frac{1}{2 \cdot 2} + \frac{2\beta\delta}{4} \right) \|e_0\| \leq \left(\frac{1}{4} + \frac{1}{4 \cdot 3} \right) \|e_0\| = \frac{\|e_0\|}{3} \leq \frac{\|e_0\|}{2}.$$

Agora, suponha que as equações (2.16) e (2.17) sejam válidas para $k = 0, 1, \dots, i-1$, para algum $i \in \mathbb{N}$ positivo. Para $k = i$, temos por hipótese que vale a inequação (2.13), e portanto

$$\begin{aligned} \|B_k - J_*\| &\leq \|B_{k-1} - J_*\| + \frac{L}{2} (\|e_k\| + \|e_{k-1}\|) \leq (2 - 2^{-(k-1)})\delta + \frac{L}{2} \left(\frac{\|e_{k-1}\|}{2} + \|e_{k-1}\| \right) \\ &\leq (2 - 2^{1-k})\delta + \frac{3L\|e_{k-1}\|}{4}. \end{aligned}$$

Como

$$\|e_{k-1}\| \leq \frac{\|e_{k-2}\|}{2} \leq \dots \leq \frac{\|e_0\|}{2^{k-1}} \leq \frac{\varepsilon}{2^{k-1}},$$

segue que

$$\begin{aligned} \|B_k - J_*\| &\leq (2 - 2^{1-k})\delta + \frac{3L\varepsilon}{2^{k+1}} \leq (2 - 2^{1-k})\delta + \frac{2\delta}{2^{k+1}} \leq (2 - 2^{1-k})\delta + \frac{\delta}{2^k} \\ &\leq 2\delta - \frac{2\delta}{2^k} + \frac{\delta}{2^k} \leq 2\delta - 2^{-k}\delta = (2 - 2^{-k})\delta, \end{aligned}$$

provando (2.16).

A demonstração de que vale (2.17) para $k = i$ segue o mesmo princípio do caso $k = 0$. De fato, temos que

$$\|J_*^{-1}(B_k - J_*)\| \leq \|J_*^{-1}\| \|B_k - J_*\| \leq \beta(2 - 2^{-k})\delta \leq 2\beta\delta \leq \frac{1}{3}.$$

Segue de modo análogo, pelo Teorema 2.1, que B_k é inversível (e portanto $x_{k+1} = x_k - B_k^{-1}F(x_k)$ está bem definido) e $\|B_k^{-1}\| \leq \frac{3\beta}{2}$. Procedendo exatamente como no caso $k = 0$, obtemos

$$\|e_{k+1}\| \leq \|B_k^{-1}\| (\|B_k - J_*\| \|e_k\| + \|-F(x_k) + F(x_*) + J_*e_k\|),$$

e pelo Lema 2.8,

$$\|e_{k+1}\| \leq \frac{3\beta}{2} \left((2 - 2^{-k})\delta \|e_k\| + \frac{L\|e_k\|^2}{2} \right) = \frac{3\beta}{2} \left((2 - 2^{-k})\delta + \frac{L\|e_k\|}{2} \right) \|e_k\|$$

$$\begin{aligned}
&\leq \frac{3\beta}{2} \left((2 - 2^{-k})\delta + \frac{L}{2} \frac{\varepsilon}{2^k} \right) \|e_k\| \leq \frac{3\beta}{2} \left((2 - 2^{-k})\delta + \frac{L\varepsilon}{2^{k+1}} \right) \|e_k\| \\
&\leq \frac{3\beta}{2} \left((2 - 2^{-k})\delta + \frac{2\delta}{2^{k+1}} \right) \|e_k\| = \frac{3\beta}{2} ((2 - 2^{-k})\delta + 2^{-k}\delta) \|e_k\| \leq 3\beta\delta \|e_k\| \leq \frac{\|e_k\|}{2}.
\end{aligned}$$

□

Isto prova a convergência q -linear do algoritmo de Broyden. A demonstração da convergência q -superlinear requer mais alguns resultados preliminares e será demonstrada mais adiante. A seguir, serão demonstrados os Lemas 2.13 e 2.14, correspondentes aos Lemas 4.1.16 e 8.2.3 de [10], com o objetivo de demonstrar o Teorema 2.15 (Teorema 8.2.4 de [10]). Este Teorema é conhecido como Teorema de Dennis-Moré [19].

Lema 2.13. *Sejam F e J como no enunciado do Lema 2.7, e assumamos que $J(x)^{-1}$ exista. Então, existem $\varepsilon, \alpha, \beta > 0$, onde $\alpha < \beta$, tais que*

$$\alpha \|v - u\| \leq \|F(v) - F(u)\| \leq \beta \|v - u\|,$$

para todos $u, v \in D$ satisfazendo $\max\{\|v - x\|, \|u - x\|\} \leq \varepsilon$.

Demonstração. Provemos primeiramente que existe $\beta > 0$ tal que $\|F(v) - F(u)\| \leq \beta$. De fato,

$$\begin{aligned}
\|F(v) - F(u)\| &= \|F(v) - F(u) - J(x)(v - u) + J(x)(v - u)\| \\
&\leq \|F(v) - F(u) - J(x)(v - u)\| + \|J(x)(v - u)\| \\
&\leq \|F(v) - F(u) - J(x)(v - u)\| + \|J(x)\| \|v - u\|.
\end{aligned}$$

Pelo Lema 2.7,

$$\begin{aligned}
\|F(v) - F(u)\| &\leq \frac{L}{2} (\|u - x\| + \|v - x\|) \|v - u\| + \|J(x)\| \|v - u\| \\
&= \left(\frac{L}{2} (\|u - x\| + \|v - x\|) + \|J(x)\| \right) \|v - u\|.
\end{aligned}$$

Como $\max\{\|v - x\|, \|u - x\|\} \leq \varepsilon$, temos que

$$\|F(v) - F(u)\| \leq (L\varepsilon + \|J(x)\|) \|v - u\| = \beta \|v - u\|,$$

onde $\beta = L\varepsilon + \|J(x)\|$. Para provar que existe $\alpha > 0$, com $\beta > \alpha$, satisfazendo $\alpha \|v - u\| \leq \|F(v) - F(u)\|$, note primeiramente que

$$\|v - u\| = \|J(x)^{-1} J(x)(v - u)\| \leq \|J(x)^{-1}\| \|J(x)(v - u)\|$$

$$\Rightarrow \frac{1}{\|J(x)^{-1}\|} \|v - u\| \leq \|J(x)(v - u)\|.$$

Agora, procedendo de forma parecida com o caso anterior,

$$\begin{aligned} \|F(v) - F(u)\| &= \|F(v) - F(u) - J(x)(v - u) + J(x)(v - u)\| \\ &\geq \|J(x)(v - u)\| - \|F(v) - F(u) - J(x)(v - u)\| \\ &\geq \frac{1}{\|J(x)^{-1}\|} \|v - u\| - \frac{L}{2} (\|v - x\| + \|u - x\|) \|v - u\| \\ &\geq \left(\frac{1}{\|J(x)^{-1}\|} - \frac{L}{2} (\|v - x\| + \|u - x\|) \right) \|v - u\| \\ &\geq \left(\frac{1}{\|J(x)^{-1}\|} - L\varepsilon \right) \|v - u\| = \alpha \|v - u\|, \end{aligned}$$

onde $\alpha = \frac{1}{\|J(x)^{-1}\|} - L\varepsilon$. Como $\|I\| = \|II\| \leq \|I\|\|I\| \Rightarrow 1 \leq \|I\|$, temos que

$$1 \leq \|I\| = \|J(x)^{-1}J(x)\| \leq \|J(x)^{-1}\|\|J(x)\| \Rightarrow \frac{1}{\|J(x)^{-1}\|} \leq \|J(x)\|,$$

e conseqüentemente,

$$\alpha = \frac{1}{\|J(x)^{-1}\|} - L\varepsilon \leq \|J(x)\| - L\varepsilon \leq \|J(x)\| + L\varepsilon = \beta.$$

□

Lema 2.14. *Seja $x_k \in \mathbb{R}^n$, com $x_k \neq x_*$, $k = 0, 1, \dots$. Se $(x_k)_{k \in \mathbb{N}}$ converge q -superlinearmente para x_* , então em qualquer norma $\|\cdot\|$,*

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} = 1.$$

Demonstração. Defina $s_k = x_{k+1} - x_k$ e $e_k = x_k - x_*$. Se $\|e_k\| \neq 0$, para todo $k \in \mathbb{N}$, então pela Definição 1.7, a convergência q -superlinear de $(x_k)_{k \in \mathbb{N}}$ é caracterizada pelo limite

$$\lim_{k \rightarrow \infty} \frac{\|e_{k+1}\|}{\|e_k\|} = 0.$$

Como

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} = 1 \Leftrightarrow \lim_{k \rightarrow \infty} \left| \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} - 1 \right| = 0 \Leftrightarrow \lim_{k \rightarrow \infty} \left| \frac{\|s_k\|}{\|e_k\|} - 1 \right| = 0,$$

provemos que $\lim_{k \rightarrow \infty} \left| \frac{\|s_k\|}{\|e_k\|} - 1 \right| = 0$. De fato,

$$\begin{aligned} 0 \leq \lim_{k \rightarrow \infty} \left| \frac{\|s_k\|}{\|e_k\|} - 1 \right| &= \lim_{k \rightarrow \infty} \left| \frac{\|s_k\|}{\|e_k\|} - \frac{\|e_k\|}{\|e_k\|} \right| = \lim_{k \rightarrow \infty} \frac{\left| \|s_k\| - \|e_k\| \right|}{\|e_k\|} \leq \lim_{k \rightarrow \infty} \frac{\|s_k + e_k\|}{\|e_k\|} \\ &= \lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_k + x_k - x_*\|}{\|e_k\|} = \lim_{k \rightarrow \infty} \frac{\|e_{k+1}\|}{\|e_k\|} = 0, \end{aligned}$$

e portanto, segue que $\lim_{k \rightarrow \infty} \left| \frac{\|s_k\|}{\|e_k\|} - 1 \right| = 0$. □

Teorema 2.15 (Dennis-Moré [19]). *Seja $D \subset \mathbb{R}^n$ um conjunto aberto e convexo, $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $J \in Lip_L(D)$, $x_* \in D$ e $J(x_*)$ não singular. Seja $\{B_k\}$ uma sequência de matrizes não singulares em $\mathbb{R}^{n \times n}$, e suponha que para algum $x_0 \in D$ a sequência de pontos gerada por*

$$x_{k+1} = x_k - B_k^{-1}F(x_k) \quad (2.22)$$

permanece em D , e satisfaz $x_k \neq x_$, para todo k , e $\lim_{k \rightarrow \infty} x_k = x_*$. Então $(x_k)_{k \in \mathbb{N}}$ converge q -superlinearmente para x_* em alguma norma $\|\cdot\|$, e $F(x_*) = 0$, se, e somente se,*

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - J(x_*))s_k\|}{\|s_k\|} = 0, \quad (2.23)$$

onde $s_k = x_{k+1} - x_k$.

Demonstração. Sejam $J_* = J(x_*)$ e $e_k = x_k - x_*$. Suponha que $\lim_{k \rightarrow \infty} \frac{\|(B_k - J(x_*))s_k\|}{\|s_k\|} = 0$. Queremos provar que $(x_k)_{k \in \mathbb{N}}$ converge q -superlinearmente para x_* e que $F(x_*) = 0$. Da definição de x_{k+1} , temos que

$$\begin{aligned} x_{k+1} = x_k - B_k^{-1}F(x_k) &\Leftrightarrow x_{k+1} - x_k + B_k^{-1}F(x_k) = 0 \Leftrightarrow s_k + B_k^{-1}F(x_k) = 0 \\ \Leftrightarrow B_k s_k + F(x_k) = 0 &\Leftrightarrow (B_k - J_* + J_*)s_k + F(x_k) = 0 \Leftrightarrow (B_k - J_*)s_k + F(x_k) + J_*s_k = 0 \\ &\Leftrightarrow (B_k - J_*)s_k - F(x_{k+1}) + F(x_k) + J_*s_k = -F(x_{k+1}). \end{aligned} \quad (2.24)$$

Assim, temos do Lema 2.7 que

$$\begin{aligned} 0 \leq \frac{\|F(x_{k+1})\|}{\|s_k\|} &\leq \frac{\|(B_k - J_*)s_k\|}{\|s_k\|} + \frac{\|-F(x_{k+1}) + F(x_k) + J_*s_k\|}{\|s_k\|} \\ &\leq \frac{\|(B_k - J_*)s_k\|}{\|s_k\|} + \frac{L}{2}(\|e_k\| + \|e_{k+1}\|). \end{aligned}$$

Como, por hipótese, $\lim_{k \rightarrow \infty} \frac{\|(B_k - J(x_*))s_k\|}{\|s_k\|} = 0$, e $\|e_k\| \rightarrow 0$, temos que

$$\lim_{k \rightarrow \infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} = 0,$$

pelo Teorema do Confronto. Como $\|s_k\| \rightarrow 0$, o limite acima implica que $\lim_{k \rightarrow \infty} \|F(x_{k+1})\| = 0$. Pela continuidade de F , segue que $F(x_*) = \lim_{k \rightarrow \infty} F(x_{k+1}) = 0$. Sendo assim, pelo Lema 2.13, temos que existe $\alpha > 0$ e $k_0 \in \mathbb{N}$ tais que

$$\|F(x_{k+1})\| = \|F(x_{k+1}) - F(x_*)\| \geq \alpha \|x_{k+1} - x_*\| = \alpha \|e_{k+1}\|, \quad \forall k \geq k_0.$$

Pela desigualdade triangular, $\|s_k\| = \|x_{k+1} - x_k\| \leq \|x_{k+1} - x_*\| + \|x_k - x_*\| = \|e_{k+1}\| + \|e_k\|$, e portanto,

$$\frac{\|F(x_{k+1})\|}{\|s_k\|} \geq \frac{\alpha\|e_{k+1}\|}{\|s_k\|} \geq \frac{\alpha\|e_{k+1}\|}{\|e_{k+1}\| + \|e_k\|} = \frac{\|e_k\|}{\|e_k\|} \frac{\alpha \frac{\|e_{k+1}\|}{\|e_k\|}}{\frac{\|e_{k+1}\|}{\|e_k\|} + 1} = \alpha \frac{\frac{\|e_{k+1}\|}{\|e_k\|}}{\frac{\|e_{k+1}\|}{\|e_k\|} + 1} \geq 0.$$

Defina $r_k = \frac{\|e_{k+1}\|}{\|e_k\|}$. Como $\frac{\|F(x_{k+1})\|}{\|s_k\|} \rightarrow 0$, temos novamente pelo Teorema do Confronto que $\alpha \frac{r_k}{r_k+1} \rightarrow 0$, e conseqüentemente $\frac{r_k}{r_k+1} \rightarrow 0$. Pela definição de convergência de r_k , dado $\varepsilon > 0$, existe $k_1 \in \mathbb{N}$ tal que se $k \geq k_1$, então

$$\left| \frac{r_k}{r_k+1} - 0 \right| < \frac{\varepsilon}{\varepsilon+1} \Rightarrow \frac{r_k}{r_k+1} < \frac{\varepsilon}{\varepsilon+1} \Rightarrow r_k + \varepsilon r_k < \varepsilon + \varepsilon r_k \Rightarrow r_k < \varepsilon.$$

Como $r_k \geq 0$, para todo $k \in \mathbb{N}$, segue que $\|r_k - 0\| = r_k < \varepsilon$, e portanto $r_k \rightarrow 0$, isto é,

$$\lim_{k \rightarrow \infty} \frac{\|e_{k+1}\|}{\|e_k\|} = \lim_{k \rightarrow \infty} r_k = 0.$$

No caso em que $x_k \neq x_*$, para todo $k \in \mathbb{N}$, o limite acima caracteriza a convergência q -superlinear de $(x_k)_{k \in \mathbb{N}}$, e portanto temos o desejado. Agora, suponha que $(x_k)_{k \in \mathbb{N}}$ converge q -superlinearmente para x_* e que $F(x_*) = 0$. Pelo Lema 2.13, temos que existe $\beta > 0$ e $k_0 \in \mathbb{N}$ tal que

$$\|F(x_{k+1})\| \leq \beta \|e_{k+1}\|, \quad \forall k \geq k_0.$$

Então,

$$\beta \frac{\|e_{k+1}\|}{\|e_k\|} \geq \frac{\|F(x_{k+1})\|}{\|e_k\|} = \frac{\|F(x_{k+1})\|}{\|s_k\|} \frac{\|s_k\|}{\|e_k\|} \geq 0.$$

Pelo Teorema do Confronto, segue que $\lim_{k \rightarrow \infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} \frac{\|s_k\|}{\|e_k\|} = 0$. Uma vez que $(x_k)_{k \in \mathbb{N}}$ converge q -superlinearmente, temos pelo Lema 2.14 que $\lim_{k \rightarrow \infty} \frac{\|s_k\|}{\|e_k\|} = \lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} = 1$, e conseqüentemente,

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} &= \lim_{k \rightarrow \infty} \left(\frac{\|F(x_{k+1})\|}{\|s_k\|} \frac{\|s_k\|}{\|e_k\|} \cdot \frac{1}{\frac{\|s_k\|}{\|e_k\|}} \right) \\ &= \left(\lim_{k \rightarrow \infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} \frac{\|s_k\|}{\|e_k\|} \right) \left(\lim_{k \rightarrow \infty} \frac{1}{\frac{\|s_k\|}{\|e_k\|}} \right) = 0. \end{aligned}$$

Pela equação (2.24), temos que

$$(B_k - J_*)s_k = -F(x_{k+1}) - (-F(x_{k+1}) + F(x_k) + J_*s_k),$$

e pelo Lema 2.7, segue que

$$0 \leq \frac{\|(B_k - J_*)s_k\|}{\|s_k\|} \leq \frac{\|F(x_{k+1})\|}{\|s_k\|} + \frac{L}{2}(\|e_k\| + \|e_{k+1}\|),$$

e pelo Teorema do Confronto, concluímos que

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - J_*)s_k\|}{\|s_k\|} = 0,$$

finalizando a demonstração do teorema. □

Um último resultado que precisaremos antes de demonstrar a convergência q -superlinear do método de Broyden é o Lema 2.16, que corresponde ao Lema 8.2.5 de [10]. Este resultado é bastante técnico, e será necessário em uma passagem da demonstração da convergência q -superlinear do método.

Lema 2.16. *Seja $s \in \mathbb{R}^n$ não nulo, $E \in \mathbb{R}^{n \times n}$, e denote a norma de matrizes induzida pela norma euclidiana de vetores por $\|\cdot\|$. Então,*

$$\left\| E \left(I - \frac{ss^T}{s^T s} \right) \right\|_F = \left(\|E\|_F^2 - \left(\frac{\|Es\|^2}{\|s\|^2} \right) \right)^{1/2} \quad (2.25)$$

$$\leq \|E\|_F - \frac{1}{2\|E\|_F} \left(\frac{\|Es\|}{\|s\|} \right)^2. \quad (2.26)$$

Demonstração. Lembremos que $I - (ss^T/s^T s)$ e $ss^T/s^T s$ são matrizes de projeções euclidianas. Deste modo, segundo [10] (demonstração do Lema 8.2.5), podemos aplicar o Teorema de Pitágoras e obter

$$\|E\|_F^2 = \left\| E \frac{ss^T}{s^T s} \right\|_F^2 + \left\| E \left(I - \frac{ss^T}{s^T s} \right) \right\|_F^2. \quad (2.27)$$

Da demonstração do Lema 2.10, temos que na norma de Frobenius vale $\|uv^T\|_F = \|u\|_2 \|v\|_2 = \|u\| \|v\|$, para todos $u, v \in \mathbb{R}^n$. Assim,

$$\left\| E \frac{ss^T}{s^T s} \right\|_F = \left\| Es \frac{s^T}{s^T s} \right\|_F = \frac{1}{s^T s} \left\| (Es)s^T \right\|_F = \frac{1}{s^T s} \|Es\| \|s\| = \|Es\| \frac{1}{\|s\|^2} \|s\| = \frac{\|Es\|}{\|s\|}.$$

Substituindo a equação acima em (2.27), temos que

$$\begin{aligned} 0 &\leq \left\| E \left(I - \frac{ss^T}{s^T s} \right) \right\|_F^2 = \|E\|_F^2 - \left(\frac{\|Es\|}{\|s\|} \right)^2 \\ \Rightarrow \left\| E \left(I - \frac{ss^T}{s^T s} \right) \right\|_F &= \left(\|E\|_F^2 - \left(\frac{\|Es\|}{\|s\|} \right)^2 \right)^{1/2}. \end{aligned}$$

Note que para todo $\alpha \geq |\beta| \geq 0$, temos que

$$\left(\alpha - \frac{\beta^2}{2\alpha}\right)^2 = \alpha^2 - 2\alpha \frac{\beta^2}{2\alpha} + \frac{\beta^4}{4\alpha^2} \geq \alpha^2 - \beta^2 \geq 0,$$

e portanto,

$$(\alpha^2 - \beta^2)^{\frac{1}{2}} \leq \left|\alpha - \frac{\beta^2}{2\alpha}\right|.$$

Como $\alpha - \frac{\beta^2}{2\alpha} = \frac{2\alpha^2 - \beta^2}{2\alpha} \geq \frac{2\beta^2 - \beta^2}{2\alpha} = \frac{\beta^2}{2\alpha} \geq 0$, temos que

$$(\alpha^2 - \beta^2)^{\frac{1}{2}} \leq \alpha - \frac{\beta^2}{2\alpha}.$$

Segue da desigualdade acima e de (2.25) que

$$\left\|E \left(I - \frac{ss^T}{s^T s}\right)\right\|_F \leq \|E\|_F - \frac{1}{2\|E\|_F} \left(\frac{\|Es\|}{\|s\|}\right)^2,$$

onde tomamos $\alpha = \|E\|_F$ e $\beta = \frac{\|Es\|}{\|s\|}$.

□

Por fim, a convergência q -superlinear do método de Broyden será demonstrada no Corolário 2.17, cuja demonstração é baseada na segunda parte da demonstração do Teorema 8.2.2 de [10] (pág. 183).

Corolário 2.17 (Convergência q -superlinear do método de Broyden). *Seja $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuamente diferenciável em um aberto e convexo $D \subset \mathbb{R}^n$. Assuma que exista $x_* \in \mathbb{R}^n$ e $r, \beta > 0$, tais que $B(x_*, r) \subset D$, $F(x_*) = 0$, $J(x_*)^{-1}$ existe, $\|J(x_*)^{-1}\| \leq \beta$, e $J \in \text{Lip}_L(B(x_*, r))$. Existem constantes positivas ε e δ tais que se $\|x_0 - x_*\| \leq \varepsilon$ e $\|B_0 - J(x_*)\| \leq \delta$, então o Algoritmo 5 está bem definido e converge q -superlinearmente para x_* .*

Demonstração. Lembremos que $J \in \text{Lip}_L(B(x_*, r))$ implica que J é Lipschitz contínua em x_* . Pelo Lema 2.11, segue que a sequência $\{B_k\}$ deverá satisfazer (2.13). Neste caso, as hipóteses do Teorema 2.12 são válidas, assim como a sua conclusão. Além disso, podemos resgatar passos importantes de sua demonstração que possam ser convenientes para demonstrarmos a convergência q -superlinear do método de Broyden.

Defina $E_k = B_k - J_*$. Pelo Teorema 2.15, temos que se

$$\lim_{k \rightarrow \infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0,$$

então $(x_k)_{k \in \mathbb{N}}$ converge q -superlinearmente para x_* .

Da demonstração do Lema 2.11, temos que

$$\begin{aligned} E_{k+1} &= B_{k+1} - J_* = (B_k - J_*) \left[I - \frac{s_k s_k^T}{s_k^T s_k} \right] + \frac{(y_k - J_* s_k) s_k^T}{s_k^T s_k} \\ &= E_k \left[I - \frac{s_k s_k^T}{s_k^T s_k} \right] + \frac{(y_k - J_* s_k) s_k^T}{s_k^T s_k}. \end{aligned}$$

Além disso, na demonstração do Lema 2.11 foi mostrado que

$$\left\| \frac{(y_k - J_* s_k) s_k^T}{s_k^T s_k} \right\|_F \leq \frac{L}{2} (\|e_k\| + \|e_{k+1}\|).$$

Sendo assim,

$$\begin{aligned} \|E_{k+1}\|_F &\leq \left\| E_k \left(I - \frac{s_k s_k^T}{s_k^T s_k} \right) \right\|_F + \frac{\|(y_k - J_* s_k) s_k^T\|_F}{s_k^T s_k} \\ &\leq \left\| E_k \left(I - \frac{s_k s_k^T}{s_k^T s_k} \right) \right\|_F + \frac{L}{2} (\|e_k\| + \|e_{k+1}\|). \end{aligned}$$

Na demonstração do Teorema 2.12 foi mostrado que $\|e_{k+1}\| \leq \frac{\|e_k\|}{2}$ assumindo que (2.13) era válida. Como, por hipótese, $J \in \text{Lip}_L(B(x_*, r))$, temos que (2.13) é válida, pelo Lema 2.11. Segue então do Lema 2.16 que

$$\begin{aligned} \|E_{k+1}\|_F &\leq \left\| E_k \left(I - \frac{s_k s_k^T}{s_k^T s_k} \right) \right\|_F + \frac{L}{2} (\|e_k\| + \|e_{k+1}\|) \\ &\leq \|E_k\|_F - \frac{1}{2\|E_k\|_F} \left(\frac{\|E_k s_k\|}{\|s_k\|} \right)^2 + \frac{L}{2} (\|e_k\| + \|e_{k+1}\|) \\ &\leq \|E_k\|_F - \frac{1}{2\|E_k\|_F} \left(\frac{\|E_k s_k\|}{\|s_k\|} \right)^2 + \frac{3L}{4} \|e_k\|. \end{aligned}$$

Reorganizando os termos da inequação acima, temos

$$\begin{aligned} \frac{1}{2\|E_k\|_F} \left(\frac{\|E_k s_k\|}{\|s_k\|} \right)^2 &\leq \|E_k\|_F - \|E_{k+1}\|_F + \frac{3L}{4} \|e_k\| \\ \Leftrightarrow \frac{\|E_k s_k\|^2}{\|s_k\|^2} &\leq 2\|E_k\|_F \left(\|E_k\|_F - \|E_{k+1}\|_F + \frac{3L}{4} \|e_k\| \right) \end{aligned}$$

Novamente da demonstração do Teorema 2.12, temos que

$$\|E_k\|_F = \|B_k - J_*\| \leq (2 - 2^{-k})\delta \leq 2\delta, \quad \forall k \in \mathbb{N},$$

e também que $\|e_k\| \leq \frac{\varepsilon}{2^k} \forall k \in \mathbb{N}$, logo

$$\sum_{k=0}^{\infty} \|e_k\| \leq \sum_{k=0}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon \sum_{k=0}^{\infty} \frac{1}{2^k} = \varepsilon \frac{1}{1 - \frac{1}{2}} = 2\varepsilon.$$

Com base nesses fatos, temos que

$$\frac{\|E_k s_k\|^2}{\|s_k\|^2} \leq 4\delta \left(\|E_k\|_F - \|E_{k+1}\|_F + \frac{3L}{4} \|e_k\| \right),$$

e conseqüentemente, para todo $i \in \mathbb{N}$,

$$\begin{aligned} & \sum_{k=0}^i \frac{\|E_k s_k\|^2}{\|s_k\|^2} \leq \sum_{k=0}^i \left(4\delta \left(\|E_k\|_F - \|E_{k+1}\|_F + \frac{3L}{4} \|e_k\| \right) \right) \\ &= 4\delta \left(\sum_{k=0}^i (\|E_k\|_F - \|E_{k+1}\|_F) + \frac{3L}{4} \sum_{k=0}^i \|e_k\| \right) = 4\delta \left(\|E_0\|_F - \|E_{i+1}\|_F + \frac{3L}{4} \sum_{k=0}^i \|e_k\| \right) \\ &\leq 4\delta \left(\|E_0\|_F + \frac{3L}{4} \sum_{k=0}^i \|e_k\| \right) \leq 4\delta \left(\delta + \frac{3L}{4} 2\varepsilon \right). \end{aligned}$$

Neste caso,

$$\begin{aligned} \sum_{k=0}^{\infty} \left(\frac{\|E_k s_k\|}{\|s_k\|} \right)^2 &= \sum_{k=0}^{\infty} \frac{\|E_k s_k\|^2}{\|s_k\|^2} \\ &\leq 4\delta \left(\delta + \frac{3L}{4} 2\varepsilon \right), \end{aligned}$$

e portanto, a série $\sum_{k=0}^{\infty} \left(\frac{\|E_k s_k\|}{\|s_k\|} \right)^2$ converge, e conseqüentemente seu termo geral $\left(\frac{\|E_k s_k\|}{\|s_k\|} \right)^2$ converge para 0. Neste caso,

$$\lim_{k \rightarrow \infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0,$$

e conseqüentemente, $(x_k)_{k \in \mathbb{N}}$ converge q -superlinearmente para x_* .

□

Em suma, as hipóteses do Teorema 2.12 garantem que o método de Broyden está bem definido e converge ao menos q -linearmente para uma solução x_* . Acrescentando a hipótese mais forte de que $J \in \text{Lip}_L(B(x_*, r))$, temos pelo Corolário 2.17 que a convergência do método de Broyden passa a ser q -superlinear.

No próximo capítulo, vamos propor um método preditor-corretor no qual o passo corretor substitui as direções afim-escala provenientes do método de Newton pelas direções oriundas do método de Broyden. A análise de convergência local do método de Broyden foi de extrema importância para obtermos, na Seção 3.3 do próximo capítulo, uma teoria de convergência local do método proposto.

MÉTODO PREDITOR-CORRETOR

NEWTON-QUASI-NEWTON

Neste capítulo, será proposto um algoritmo preditor-corretor no qual o passo preditor realiza um deslocamento na direção de Newton para $\sigma_k = 0$, assim como nos algoritmos 4 e 3, porém o passo corretor busca recentralizar o ponto encontrado no passo preditor por meio do método de Broyden ao invés do método de Newton.

Um dos motivos para a troca do método de Newton por métodos Quasi-Newton no passo corretor vem do fato de que a principal função do passo corretor é centralizar o ponto, o que normalmente não proporciona um decréscimo muito grande de μ . Em cada passo corretor dos métodos preditores-corretores usuais, normalmente se resolve o sistema $F_{\sigma_k \mu_k}(w) = 0$ de forma aproximada. Isso é feito dando um único passo na direção afim-escala. Embora a direção afim-escala seja muito boa, esta é uma aproximação grosseira para a solução do sistema $F_{\sigma_k \mu_k}(w) = 0$, portanto não parece uma má ideia substituí-la por outra direção, desde que essa direção cumpra o papel do passo corretor de recentralizar o ponto.

Ao escolher um método Quasi-Newton para o passo corretor, estamos trocando um método que possui convergência quadrática por um método de convergência superlinear. Contudo, como estamos encarregando o passo preditor de proporcionar decréscimos significativos de μ , onde esse passo utiliza a direção de Newton, espera-se que não haja grandes prejuízos com essa escolha, desde que o passo corretor cumpra sua função.

Como foi comentado anteriormente, foi apresentada em [5] uma forma mais barata computacionalmente de implementar o método de Broyden “ruim”. Esta estratégia pode ser replicada para o Broyden “bom”, como faremos na Seção 4.2. Em [5], o método de Newton foi completamente substituído pelo método de Broyden “ruim”, e percebeu-se que

ele se comportou de forma semelhante a métodos preditores-corretores, o que também influenciou na escolha do método de Broyden em sua versão “boa”.

Lembremos que dados $\sigma \in [0, 1]$ e $\mu > 0$, a função $F_{\sigma\mu}$ foi definida como

$$F_{\sigma\mu} : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$$

$$F_{\sigma\mu}(x, \lambda, s) = \begin{bmatrix} A^T \lambda + s - c \\ Ax - b \\ XSe - \sigma\mu e \end{bmatrix}.$$

A seguir, será apresentado o Algoritmo 6, o qual é uma versão inicial do algoritmo que vamos propor.

Algoritmo 6 Preditor com Corretor Quasi-Newton

Dado $w_0 = (x_0, \lambda_0, s_0) \in \mathcal{F}_0$ e $0 < \sigma_{\max} < 1$.

Seja $k = 0$.

Passo 1: Faça $\sigma_k = 0$ em (1.22) e encontre $d_k = (\Delta x_k, \Delta \lambda_k, \Delta s_k)$.

Passo 2: Escolha o maior $\alpha_k \in [0, 1]$ tal que $w_k + \alpha_k d_k \in \mathcal{F}_0$ e $\mu \leq \mu_k$, onde μ é a medida de complementaridade em $w_k + \alpha_k d_k$.

Passo 3: Defina $\sigma_k = \frac{\sigma_{\max}}{2}$.

Passo 4: Resolva

$$F_{\sigma_k \mu_k}(w) = 0$$

pelo Método de Broyden tomando $w_k + \alpha_k d_k$ como ponto inicial e $B_0 = J(w_k)$. Se uma solução $\tilde{w} \in \mathcal{F}_0$ for encontrada, vá para o Passo 6. Caso contrário, vá para o Passo 5.

Passo 5: Defina $\alpha_k = \frac{\alpha_k}{2}$ e $\sigma_k = \frac{\sigma_{\max} + \sigma_k}{2}$, e depois volte ao Passo 4.

Passo 6: Defina $w_{k+1} = \tilde{w}$, faça $k = k + 1$ e volte para o Passo 1.

No Passo 1 do Algoritmo 6, a direção d_k referente ao passo preditor é determinada. Esta direção consiste na direção de Newton, quando $\sigma = 0$ e $\mu = \mu_k$ em (1.7). No Passo 2, é escolhido um tamanho de passo máximo a ser dado na direção d_k . No Passo 3, é escolhido um tamanho mínimo de σ_k com base no valor de σ_{\max} . O Passo 4 do Algoritmo 6 realiza o passo na direção d_k , obtendo um ponto $w_{\alpha_k} := w_k + \alpha_k d_k$. Em seguida, resolve o sistema $F_{\sigma_k \mu_k}(w) = 0$ pelo método de Broyden tomando w_k como ponto inicial e $J(w_k)$ como aproximação inicial da jacobiana. Dessa forma, o Algoritmo 6 procura obter um ponto mais centralizado, o que corresponde ao passo corretor.

Devido a teoria de convergência do método de Broyden ser apenas local, não se pode garantir que ele irá convergir sempre, pois é necessária uma determinada proximidade com relação a uma solução para que isso ocorra. Na Seção 3.1 deste capítulo, será discutida a estratégia adotada no Algoritmo 6 nesses casos de forma mais detalhada. De forma resumida, caso não ocorra a convergência do passo corretor, isto é, a convergência do método de Broyden, o algoritmo reduz α_k e aproxima σ_k de 1 no Passo 5, e em seguida tenta novamente executar o Passo 4. Sobre certas condições, mostramos na Seção 3.1 que em algum momento a convergência do passo corretor deverá ocorrer. Quando esta convergência ocorre, o Passo 6 é executado para definir o próximo iterando, o qual é tomado como a solução encontrada pelo método de Broyden.

O algoritmo que estamos propondo neste trabalho baseia-se no Algoritmo 6, o qual foi inspirado nos Algoritmos 4 e 3. Nos algoritmos mencionados, o passo corretor consiste de um único passo de Newton, enquanto que no Algoritmo 6, o problema de anular $F_{\sigma_k \mu_k}$ para $\sigma_k \in [0, 1]$ e fixado e $\mu_k = \frac{x_k^T s_k}{n}$ é resolvido até a convergência. Em outras palavras, estamos em busca de um ponto mais próximo do caminho central para dado σ_k . Mais adiante, na Seção 3.3 deste capítulo, será proposta também uma modificação no Algoritmo 6 que permitirá interromper o método de Broyden assim que um decréscimo suficiente da medida de complementaridade ocorrer (desde que o ponto ainda esteja em \mathcal{F}_0), mantendo a convergência local do algoritmo.

3.1 Boa Definição do Algoritmo 6

Suponha que w_k seja uma aproximação para um ponto do caminho central. Sobre este ponto aplicaremos o passo preditor, isto é, será dado um passo de Newton a partir de w_k , tomando $\sigma_k = 0$, para obter um novo ponto $w_{\alpha_k} = w_k + \alpha_k d_k$, com $\alpha_k \in [0, 1]$ e d_k é a direção de Newton. Na Seção 1.5 do Capítulo 1 vimos que isso reduz μ ao custo de possivelmente perder a centralidade do ponto w_k . Sobre o ponto w_{α_k} , iremos aplicar o passo corretor, tomando agora $\sigma_k \in (0, 1]$, para obter um ponto w_{k+1} que esteja mais centralizado, ou seja, mais próximo do caminho central, utilizando dessa vez o método de Broyden.

Vamos começar nossa análise simplificando para o caso em que $w_k = (x_k, \lambda_k, s_k)$ seja um ponto sobre o caminho central, e depois vamos adaptar o raciocínio para um ponto

w_k que esteja apenas suficientemente próximo do caminho central. Antes de começarmos, vamos enunciar uma versão do Teorema da Função Inversa [15], que será de extrema importância para a nossa abordagem. Uma versão mais geral deste teorema pode ser encontrada em [16] (Teorema 8.3). Depois, será apresentado um lema no qual veremos que a Jacobiana de $F_{\sigma\mu}$ é Lipschitz com constante igual a 1.

Teorema 3.1 (Teorema da Função Inversa). *Sejam $\Omega \subset \mathbb{R}^n$ um aberto, $f : \Omega \rightarrow \mathbb{R}^n$ uma aplicação de classe C^1 e $a \in \Omega$. Se $J(a)$ é bijetora, onde J corresponde a Jacobiana de f , então*

1. *Existem abertos $V \subset \mathbb{R}^n$ contendo a e $W \subset \mathbb{R}^n$ contendo $f(a)$ tais que $f : V \rightarrow W$ é bijetora;*
2. *$f^{-1} : W \rightarrow V$ é de classe C^1 .*

Observação 3.2. *Considere $f : W \rightarrow \mathbb{R}^n$ nas hipóteses do Teorema da Função Inversa. Temos que para todo $x \in U$,*

$$f^{-1}(f(x)) = x.$$

Sendo assim, pela Regra da Cadeia, temos que

$$J^{-1}(f(x))J(x) = I_n,$$

ou equivalentemente,

$$J^{-1}(f(x)) = [J(x)]^{-1}.$$

Lema 3.3. *Dados $\sigma \in [0, 1]$ e $\mu > 0$, então a Jacobiana de $F_{\sigma\mu}$ definida em (1.7), dada por*

$$J_{\sigma\mu}(w) = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix},$$

onde $w = (x, \lambda, s) \in \mathbb{R}^{2n+m}$, é Lipschitz com constante de Lipschitz igual a 1.

Demonstração. Dados $w_1, w_2 \in \mathbb{R}^{2n+m}$, temos que

$$\left\| J_{\sigma\mu}(w_2) - J_{\sigma\mu}(w_1) \right\|_2 = \left\| \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S_2 & 0 & X_2 \end{bmatrix} - \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S_1 & 0 & X_1 \end{bmatrix} \right\|_2$$

$$\begin{aligned}
&= \left\| \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ S_2 - S_1 & 0 & X_2 - X_1 \end{bmatrix} \right\|_2 \leq \left\| \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ S_2 - S_1 & 0 & X_2 - X_1 \end{bmatrix} \right\|_F = \left\| \begin{bmatrix} x_2 - x_1 \\ 0 \\ s_2 - s_1 \end{bmatrix} \right\|_2 \\
&\leq \left\| \begin{bmatrix} x_2 - x_1 \\ \lambda_2 - \lambda_1 \\ s_2 - s_1 \end{bmatrix} \right\|_2 = \|w_2 - w_1\|_2.
\end{aligned}$$

□

Denote por ω a parametrização $f_c : \mathbb{R}_+ \rightarrow \mathbb{R}^{2n+m}$ do caminho central obtida anteriormente no final da Seção 1.2 do Capítulo 1, a qual vimos ser contínua. Digamos que no passo corretor estamos tentando alcançar o ponto $\omega(\sigma_k \mu_k)$. Para termos a garantia teórica de que alcançaremos tal ponto, de acordo com o Teorema da Convergência Local do Método de Broyden (Teorema 2.12), devemos ter que o ponto inicial, que neste caso será w_{α_k} , deverá estar suficientemente próximo de $\omega(\sigma_k \mu_k)$. Nas notações do Teorema 2.12, existirá $\varepsilon(\sigma_k \mu_k) > 0$, associado a $\omega(\sigma_k \mu_k)$, de modo que se

$$\|w_{\alpha_k} - \omega(\sigma_k \mu_k)\| \leq \varepsilon(\sigma_k \mu_k)$$

e a nossa aproximação inicial para a Jacobiana em $\omega(\sigma_k \mu_k)$ for razoável, então haverá a convergência do método de Broyden para o ponto $\omega(\sigma_k \mu_k)$.

Observação 3.4. *Dados parâmetros $\sigma \in [0, 1]$ e $\mu > 0$, a Jacobiana da função $F_{\sigma\mu} : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ definida em (1.7) não depende das constantes σ e μ , e portanto será denotada simplesmente por J , mesmo no contexto em que houver mais de uma função com parâmetros distintos.*

Para aproximar a Jacobiana em w_{α_k} , iremos utilizar a Jacobiana J de F_0 , calculada no passo preditor, referente ao ponto w_k , e vamos denotar $J(w_k)$ por J_k . Neste caso, de acordo com o Teorema 2.12, temos que existe $\delta(\sigma_k \mu_k) > 0$, também associado a $\omega(\sigma_k \mu_k)$, de modo que se

$$\|J_k - J(\omega(\sigma_k \mu_k))\| \leq \delta(\sigma_k \mu_k),$$

então a aproximação inicial para a Jacobiana é razoável e permitirá a convergência do algoritmo. Na demonstração do Teorema 2.12, as constantes ε e δ são escolhidas de modo que

$$6\beta\delta \leq 1 \text{ e } 3L\varepsilon \leq 2\delta,$$

onde L é a constante de Lipschitz da Jacobiana J , e $\beta > 0$ é uma constante escolhida arbitrariamente de modo que $||[J(\omega(\sigma_k \mu_k))]^{-1}|| \leq \beta$. Se impusermos as igualdades nas inequações acima, podemos expressar ε e δ da seguinte forma:

$$\varepsilon = \frac{2\delta}{3L} \text{ e } \delta = \frac{1}{6\beta}.$$

Se escolhermos $\beta = ||[J(\omega(\sigma_k \mu_k))]^{-1}||$, podemos escrever as seguintes funções:

$$\begin{aligned} \beta(w) &= ||[J(w)]^{-1}||, \\ \delta(w) &= \frac{1}{6\beta(w)} \text{ e} \\ \varepsilon(w) &= \frac{2\delta(w)}{3L}, \end{aligned}$$

em que w é um ponto do caminho central. Lembremos que $F_{\mu_k} : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ dada por

$$F_{\mu_k}(x, \lambda, s) = \begin{bmatrix} A^T \lambda + s - c \\ Ax - b \\ XSe - \mu_k e \end{bmatrix},$$

é diferenciável em todo o seu domínio. Pelo Lema 1.3, temos que $J_k = J(w_k)$ é um isomorfismo quando w_k é um ponto do caminho central, pois $x, s > 0$. Logo, pelo Teorema da Função Inversa (Teorema 3.1), existem vizinhanças V de w_k e W de $F_{\mu_k}(w_k)$ tais que $F_{\mu_k} : V \rightarrow W$ é um difeomorfismo de classe C^1 , isto é, ambas $F_{\mu_k} : V \rightarrow W$ e $F_{\mu_k}^{-1} : W \rightarrow V$ são de classe C^1 . Isto significa que $J_{\mu_k}^{-1} : W \rightarrow \mathbb{R}^{(2n+m) \times (2n+m)}$ é contínua.

Considere a função auxiliar $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ tal que $g(\sigma) = \sigma \mu_k$. Claramente g é contínua, pois é polinomial. Além disso, a função ω , que foi definida como f_c no início desta seção, é contínua, pois f_c é contínua, logo a composição $\omega \circ g : \mathbb{R}_+ \rightarrow \mathbb{R}^{2n+m}$ é contínua. Como $\omega \circ g(1) = \omega(g(1)) = \omega(\mu_k) = w_k$, podemos tomar um intervalo fechado I da forma $[a, 1]$, com $a \in [0, 1)$, próximo o suficiente de 1 de modo que $\omega(\sigma \mu_k) = \omega(g(\sigma)) \in V$, para todo $\sigma \in I$. Consideremos então a restrição $g : I \rightarrow g(I)$, a qual ainda é contínua. Para todo $\sigma \in I$, temos que $F_{\mu_k}(\omega(\sigma \mu_k)) \in W$. Neste caso, a composição

$$\begin{aligned} J_{\mu_k}^{-1} \circ F_{\mu_k} \circ \omega \circ g : I &\rightarrow \mathbb{R}^{(2n+m) \times (2n+m)} \\ \sigma &\mapsto J_{\mu_k}^{-1}(F_{\mu_k}(\omega(\sigma \mu_k))) \end{aligned}$$

é contínua em I . Note que

$$J_{\mu_k}^{-1}(F_{\mu_k}(\omega(\sigma \mu_k))) = [J_{\mu_k}(\omega(\sigma \mu_k))]^{-1} = [J(\omega(\sigma \mu_k))]^{-1}.$$

Sendo assim, $\beta(\omega(\sigma\mu_k)) = \|[J(\omega(\sigma\mu_k))]^{-1}\| = \|J_{\mu_k}^{-1}(F_{\mu_k}(\omega(\sigma\mu_k)))\|$. Como a função norma é contínua, temos que $\beta \circ \omega \circ g$ é contínua em I , vista como a composição de duas funções contínuas. Ainda, como $\omega(\sigma\mu_k)$ é um ponto do caminho central, para todo $\sigma \in I$, vimos que $J(\omega(\sigma\mu_k))$ é um isomorfismo (este fato segue da demonstração de que f_c é contínua, feita no final da Seção 1.2 do Capítulo 1, em que utiliza-se o Lema 1.3 para provar que tal Jacobiana é inversível em qualquer ponto do caminho central). Consequentemente, devemos ter que $\|J(\omega(\sigma\mu_k))\| \neq 0$, e também que $\|[J(\omega(\sigma\mu_k))]^{-1}\| \neq 0$. Neste caso, $\beta(\omega(\sigma\mu_k)) > 0$, para todo $\sigma \in I$, e conseqüentemente, $\delta \circ \omega \circ g$ será contínua em I , pela forma como δ foi definida. Segue da continuidade de $\delta \circ \omega \circ g$ em I que $\varepsilon \circ \omega \circ g$ é contínua em I . Portanto, temos que as composições $\bar{\beta} = \beta \circ \omega \circ g$, $\bar{\varepsilon} = \varepsilon \circ \omega \circ g$ e $\bar{\delta} = \delta \circ \omega \circ g$ são funções contínuas em I .

Observação 3.5. *Como a Jacobiana de $F_{\sigma\mu} : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ não depende das constantes σ e μ consideradas em sua definição, segue da expressão encontrada para $\bar{\beta} : I \rightarrow \bar{\beta}(I)$ que $\bar{\beta}$ também não depende de σ e μ . Além disso, as funções $\bar{\varepsilon} : I \rightarrow \bar{\varepsilon}(I)$ e $\bar{\delta} : I \rightarrow \bar{\delta}(I)$ também não dependem de σ e μ , pois estas funções foram dadas em termos de $\bar{\beta}$. Em outras palavras, podemos determinar os raios de convergência ε e δ relativos a pontos do caminho central que estão próximos de $\omega(\mu_k)$ por meio das funções contínuas $\bar{\varepsilon}$ e $\bar{\delta}$.*

Agora, note que I é um intervalo fechado, e portanto compacto. Como consequência disso, podemos concluir que as funções $\bar{\beta}$, $\bar{\varepsilon}$ e $\bar{\delta}$ admitem máximos e mínimos em I . Em particular, existirá $\varepsilon_0 = \min\{\bar{\varepsilon}(\sigma), \sigma \in I\}$. Em outras palavras, dado $\sigma \in I$, ao considerarmos $F_{\sigma\mu_k} : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ no passo corretor, o raio de convergência do método de Broyden para o ponto $\omega(\sigma\mu_k)$ é pelo menos $\varepsilon_0 > 0$. Como $\varepsilon_0 \leq \bar{\varepsilon}(\sigma) = \varepsilon(\omega(\sigma\mu_k))$, para todo $\sigma \in I$, temos que se $\|w_{\alpha_k} - \omega(\sigma\mu_k)\| \leq \varepsilon_0$, então $\|w_{\alpha_k} - \omega(\sigma\mu_k)\| \leq \varepsilon_0 \leq \bar{\varepsilon}(\sigma)$, o que significa que ε_0 servirá como raio de convergência para todo $\sigma \in I$.

Agora, considere uma sequência $(\sigma_i)_{i \in \mathbb{N}}$, em que $\sigma_i \rightarrow 1$ e $\sigma_i \in [a, 1)$, para todo $i \in \mathbb{N}$. Por exemplo, considere $(\sigma_i)_{i \in \mathbb{N}}$ tal que $\sigma_i = 1 - \frac{1}{i}(1 - a) > 0$. Note que $\sigma_i \rightarrow 1$, e $\sigma_i \in [a, 1) \subset I$, $\forall i \in \mathbb{N}$. Dessa forma, $\sigma_i\mu_k \rightarrow \mu_k$, e pela continuidade de ω em μ_k , teremos que $\omega(\sigma_i\mu_k) \rightarrow w_k$, e conseqüentemente, existirá $i_0 \in \mathbb{N}$ suficientemente grande tal que

$$\|w_k - \omega(\sigma_{i_0}\mu_k)\| \leq \frac{\varepsilon_0}{2}.$$

Além disso, se α_k na definição de w_{α_k} for tomado de modo que $\alpha_k < \frac{\varepsilon_0}{2\|d_k\|}$, teremos que

$$\|w_k - w_{\alpha_k}\| = \|w_k - (w_k + \alpha_k d_k)\| = \alpha_k \|d_k\| < \frac{\varepsilon_0}{2}.$$

Pela desigualdade triangular, segue que

$$\|w_{\alpha_k} - \omega(\sigma_{i_0} \mu_k)\| \leq \|w_{\alpha_k} - w_k\| + \|w_k - \omega(\sigma_{i_0} \mu_k)\| < \varepsilon_0.$$

Para verificar a condição $\|J_k - J(\omega(\sigma_k \mu_k))\| \leq \delta(\sigma_k \mu_k)$, lembremos que existe $\delta_0 = \min\{\delta(\omega(\sigma \mu_k)), \sigma \in I\}$. Sendo assim, dado $\sigma \in I$, se ocorrer $\|J_k - J(\omega(\sigma \mu_k))\| \leq \delta_0$, então $\|J_k - J(\omega(\sigma \mu_k))\| \leq \delta_0 \leq \delta(\sigma_k \mu_k)$. Neste caso, δ_0 servirá como raio de convergência relativo a aproximação da Jacobiana para qualquer $\sigma \in I$. Vimos anteriormente que J é Lipschitz em todo o seu domínio, e como consequência, J é contínua. Uma vez que $\omega(\sigma_i \mu_k) \rightarrow w_k$ quando $\sigma_i \rightarrow 1$, temos que existe $i_1 \in \mathbb{N}$ tal que

$$\|J_k - J(\omega(\sigma_{i_1} \mu_k))\| = \|J(w_k) - J(\omega(\sigma_{i_1} \mu_k))\| \leq \delta_0,$$

e portanto, para $\sigma = \sigma_{i_2}$, onde $i_2 = \max\{i_0, i_1\}$, temos que

$$\|w_{\alpha_k} - \omega(\sigma \mu_k)\| \leq \varepsilon(\sigma \mu_k) \text{ e } \|J_k - J(\omega(\sigma_{i_2} \mu_k))\| \leq \delta(\sigma \mu_k),$$

ou seja, existem $\sigma \in [a, 1) \subset I \subset [0, 1]$, e $\alpha_k \in (0, 1]$ tais que o ponto inicial w_{α_k} está dentro do raio de convergência do método de Broyden, e portanto o algoritmo deverá convergir para $\omega(\sigma \mu_k)$, quando tomamos $B_0 = J_k$.

Para uma direção d_k arbitrária, pode não ser possível mostrar que $w_{\alpha_k} \in \mathcal{F}_0$. No entanto, nosso interesse é escolher d_k como a direção de Newton, isto é, a direção que soluciona o sistema (1.22) com $\sigma = 0$. Uma vez que $w_k \in \mathcal{F}_0$, temos que w_k é viável. Neste caso, $r_c = r_b = 0$ e os dois primeiros blocos do sistema (1.22) implicam

$$A^T(\lambda_k + \alpha_k \Delta \lambda_k) + s_k + \alpha_k \Delta s_k = A^T \lambda_k + s_k + \alpha_k (A^T \Delta \lambda_k + \Delta s_k) = c + 0 = c$$

e

$$A(x + \alpha_k \Delta x_k) = Ax_k + \alpha_k \Delta x_k = b + 0 = b,$$

logo w_{α_k} é minimamente viável para qualquer $\alpha_k \in [0, 1]$. Uma vez que $x_k, s_k > 0$, para α_k suficientemente pequeno temos a condição de positividade requerida para que $w_{\alpha_k} \in \mathcal{F}_0$. Com isso, garantimos que o Algoritmo 6 está bem definido, no sentido de que

os Passos 4 e 5 não serão executados indefinidamente, pois o Passo 5 encontrará α_k e σ_k que proporcionarão a convergência do método de Broyden no Passo 4.

Agora, vamos supor que w_k não é mais um ponto sobre o caminho central. Lembremos que μ_k é a medida de complementaridade de w_k . Considere o ponto $\omega(\mu_k)$, o qual possui a mesma medida de complementaridade de w_k , mas está sobre o caminho central. Vamos mostrar que existem $r > 0$, $\alpha_k \in (0, 1]$ e $\sigma \in [0, 1)$ tais que se $w_k \in B(\omega(\mu_k), r)$, com $r > 0$, então o passo corretor converge para o ponto $\omega(\sigma\mu_k)$.

Como $\omega(\mu_k)$ é um ponto do caminho central, podemos proceder de forma análoga ao que fizemos anteriormente para mostrar que existem $\varepsilon_0, \delta_0 > 0$ de modo que

$$\|w_{\alpha_k} - \omega(\sigma\mu_k)\| < \varepsilon_0 \text{ e } \|J_k - J(\omega(\sigma\mu_k))\| < \delta_0$$

implicam a convergência do método de Broyden para o ponto $\omega(\sigma\mu_k)$, para todo $\sigma \in [a, 1]$, onde $a \in [0, 1)$. Tomando a mesma sequência $(\sigma_i)_{i \in \mathbb{N}}$, podemos concluir de modo análogo ao que foi feito anteriormente que existe $\sigma \in [a, 1)$ tal que

$$\|\omega(\mu_k) - \omega(\sigma\mu_k)\| < \frac{\varepsilon_0}{3} \text{ e } \|J(\omega(\mu_k)) - J(\omega(\sigma\mu_k))\| < \frac{\delta_0}{2}.$$

Além disso, se $\alpha_k < \frac{\varepsilon_0}{3\|d_k\|}$, teremos que

$$\|w_k - w_{\alpha_k}\| < \frac{\varepsilon_0}{3}.$$

Assim sendo, se $\|w_k - \omega(\mu_k)\| < \frac{\varepsilon_0}{3}$, teremos que

$$\begin{aligned} \|w_{\alpha_k} - \omega(\sigma\mu_k)\| &\leq \|w_{\alpha_k} - w_k\| + \|w_k - \omega(\sigma\mu_k)\| \\ &\leq \|w_{\alpha_k} - w_k\| + \|w_k - \omega(\mu_k)\| + \|\omega(\mu_k) - \omega(\sigma\mu_k)\| \\ &< \frac{\varepsilon_0}{3} + \frac{\varepsilon_0}{3} + \frac{\varepsilon_0}{3} = \varepsilon_0. \end{aligned}$$

Com respeito condição sobre a aproximação inicial da Jacobiana, segue do fato de que J é Lipschitz com constante de Lipschitz igual a 1 (Lema 3.3) que se $\|w_k - \omega(\mu_k)\| < \frac{\delta_0}{2}$, então

$$\|J_k - J(\omega(\mu_k))\| = \|J(w_k) - J(\omega(\mu_k))\| \leq \|w_k - \omega(\mu_k)\| < \frac{\delta_0}{2}.$$

Se $\|w_k - \omega(\mu_k)\| < \frac{\delta_0}{2}$,

$$\|J_k - J(\omega(\sigma\mu_k))\| \leq \|J_k - J(\omega(\mu_k))\| + \|J(\omega(\mu_k)) - J(\omega(\sigma\mu_k))\| < \frac{\delta_0}{2} + \frac{\delta_0}{2} = \delta_0.$$

Tomando $r = \min \left\{ \frac{\varepsilon_0}{3}, \frac{\delta_0}{2} \right\}$, temos o desejado. Portanto, existem $\alpha_k \in (0, 1]$, $\sigma \in [0, 1)$ e $r > 0$ tais que se $w_k \in B(\omega(\mu_k), r)$, então o passo corretor converge para $\omega(\sigma\mu_k)$, quando tomamos w_k como ponto inicial e J_k como a aproximação inicial da Jacobiana em w_k .

3.2 Distância entre um Ponto Viável e um Ponto do Caminho Central

O objetivo desta seção foi encontrar uma forma de estimar a distância entre um ponto $w \in \mathbb{R}^{2n+m}$ qualquer e um ponto \hat{w} no caminho central. Inicialmente, acreditávamos que poderíamos utilizar a vizinhança $\mathcal{N}_2(\theta)$ para provar o seguinte resultado:

“Seja $\hat{w} \in \mathbb{R}^{2n+m}$ um ponto no caminho central cuja medida de complementaridade é $\hat{\mu}$. Dado $\theta \in [0, 1)$, seja $w \in \mathcal{N}_2(\theta)$ cuja medida de complementaridade também é $\hat{\mu}$. Temos que existe uma constante $C(\theta)$, que depende de θ , tal que

$$\|w - \hat{w}\| \leq C(\theta)\hat{\mu}.”$$

De acordo com ([1], pág. 10), a vizinhança $\mathcal{N}_2(\theta)$ tem um formato cônico em torno do caminho central que vai se tornando mais fino conforme μ se aproxima de 0. Levando em consideração essa característica geométrica da vizinhança $\mathcal{N}_2(\theta)$, fica intuitivamente fácil de ver que exigir que w pertença a essa vizinhança implica, conforme $\hat{\mu}$ se aproxima de 0, que w e \hat{w} ficam cada vez mais próximos.

Imaginando que a fronteira de $\mathcal{N}_2(\theta)$ controlasse a distância entre w e \hat{w} , esperávamos conseguir demonstrar o resultado acima, que nos dá uma estimativa para esta distância em termos de θ e $\hat{\mu}$. Acreditávamos também que um resultado análogo pudesse valer para a vizinhança $\mathcal{N}_{-\infty}(\gamma)$. Infelizmente, não foi possível demonstrar tal resultado nem providenciar um contra-exemplo.

Embora não tenhamos conseguido demonstrar o resultado desejado, provamos o seguinte resultado, que será muito útil na demonstração de convergência do Algoritmo 7, que será apresentado mais adiante na Seção 3.3:

Lema 3.6. *Seja \hat{w} um ponto do caminho central \mathcal{C} e $\hat{\mu}$ sua medida de complementaridade*

associada. Se $w \in \mathcal{F}_0$ satisfizer $\|w - \hat{w}\| < \frac{1}{2\|[J(\hat{w})]^{-1}\|}$, então

$$\|w - \hat{w}\| \leq \frac{\|[J(\hat{w})]^{-1}\|}{1 - 2^{-\frac{5}{2}}} \|F_{\hat{\mu}}(w)\|.$$

Demonstração. Seja $\hat{w} \in \mathbb{R}^{2n+m}$ um ponto no caminho central \mathcal{C} cuja medida de complementaridade é $\hat{\mu}$. Dado $w \in \mathcal{F}_0$, gostaríamos de estimar a distância $\|w - \hat{w}\|$. Considere $F_{\hat{\mu}} : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ tal que

$$F_{\hat{\mu}}(w) = \begin{bmatrix} Ax - b \\ A^T \lambda - s - c \\ XSe - \hat{\mu}e \end{bmatrix}.$$

Fazendo $p = \hat{w} - w$, temos pelo Lema 2.3 que

$$F_{\hat{\mu}}(w + p) - F_{\hat{\mu}}(w) = \int_0^1 J(w + tp)p \, dt.$$

Uma vez que $F_{\hat{\mu}}(w + p) = F_{\hat{\mu}}(\hat{w}) = 0$, pois $\hat{w} \in \mathcal{C}$, segue que

$$-F_{\hat{\mu}}(w) = \int_0^1 J(w + tp)p \, dt.$$

Observe que

$$\begin{aligned} -F_{\hat{\mu}}(w) &= \int_0^1 J(w + tp)p \, dt = \int_0^1 (J(w + tp) + J(\hat{w}) - J(\hat{w}))p \, dt \\ &= \int_0^1 (J(w + tp) - J(\hat{w}))p \, dt + \int_0^1 J(\hat{w})p \, dt = \int_0^1 (J(w + tp) - J(\hat{w}))p \, dt + J(\hat{w})p. \end{aligned}$$

Vimos na Seção 1.2 do Capítulo 1 que a Jacobiana de $F_{\hat{\mu}}$ é inversível em cada ponto do caminho central. Como $\hat{w} \in \mathcal{C}$, segue que

$$\begin{aligned} -F_{\hat{\mu}}(w) &= \int_0^1 (J(w+tp) - J(\hat{w}))p \, dt + J(\hat{w})p \Leftrightarrow -F_{\hat{\mu}}(w) - \int_0^1 (J(w+tp) - J(\hat{w}))p \, dt = J(\hat{w})p \\ &\Leftrightarrow -\left(F_{\hat{\mu}}(w) + \int_0^1 (J(w+tp) - J(\hat{w}))p \, dt\right) = J(\hat{w})p \quad (3.1) \\ &\Leftrightarrow -[J(\hat{w})]^{-1} \left(F_{\hat{\mu}}(w) + \int_0^1 (J(w+tp) - J(\hat{w}))p \, dt\right) = p. \end{aligned}$$

Sendo assim,

$$\begin{aligned} \|w - \hat{w}\| &= \|p\| = \left\| [J(\hat{w})]^{-1} \left(F_{\hat{\mu}}(w) + \int_0^1 (J(w+tp) - J(\hat{w}))p \, dt\right) \right\| \\ &\leq \|[J(\hat{w})]^{-1}\| \left\| F_{\hat{\mu}}(w) + \int_0^1 (J(w+tp) - J(\hat{w}))p \, dt \right\| \end{aligned}$$

$$\leq \| [J(\hat{w})]^{-1} \| \left(\| F_{\hat{\mu}}(w) \| + \left\| \int_0^1 (J(w + tp) - J(\hat{w})) p dt \right\| \right). \quad (3.2)$$

Vamos denotar $x_t = x + t(\hat{x} - x)$ e $s_t = s + (\hat{s} - s)$. Temos que

$$\begin{aligned} \int_0^1 (J(w + tp) + J(\hat{w})) p dt &= \int_0^1 \left(\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S_t & 0 & X_t \end{bmatrix} - \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ \hat{S} & 0 & \hat{X} \end{bmatrix} \right) p dt \\ &= \int_0^1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ S_t - \hat{S} & 0 & X_t - \hat{X} \end{bmatrix} p dt = \int_0^1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ S_t - \hat{S} & 0 & X_t - \hat{X} \end{bmatrix} \begin{bmatrix} \hat{x} - x \\ \hat{\lambda} - \lambda \\ \hat{s} - s \end{bmatrix} dt \\ &= \int_0^1 \begin{bmatrix} 0 \\ 0 \\ E \end{bmatrix} dt, \end{aligned}$$

onde $E = (S_t - \hat{S})(\hat{x} - x) + (X_t - \hat{X})(\hat{s} - s)$. Temos que

$$\begin{aligned} E_i &= [s_i + t(\hat{s}_i - s_i - \hat{s}_i)] \cdot (\hat{x}_i - x_i) + [x_i + t(\hat{x}_i - x_i - \hat{x}_i)] \cdot (\hat{s}_i - s_i) \\ &= (s_i - \hat{s}_i)(\hat{x}_i - x_i) + t(\hat{s}_i - s_i)(\hat{x}_i - x_i) + (x_i - \hat{x}_i)(\hat{s}_i - s_i) + t(\hat{x}_i - x_i)(\hat{s}_i - s_i). \end{aligned}$$

Integrando E_i de 0 a 1 com respeito a t , temos

$$\begin{aligned} \int_0^1 E_i dt &= \left[t(s_i - \hat{s}_i)(\hat{x}_i - x_i) + \frac{t^2}{2}(\hat{s}_i - s_i)(\hat{x}_i - x_i) \right. \\ &\quad \left. + t(x_i - \hat{x}_i)(\hat{s}_i - s_i) + \frac{t^2}{2}(\hat{x}_i - x_i)(\hat{s}_i - s_i) \right] \Big|_{t=0}^{t=1} \\ &= (s_i - \hat{s}_i)(\hat{x}_i - x_i) + \frac{1}{2}(\hat{s}_i - s_i)(\hat{x}_i - x_i) + (x_i - \hat{x}_i)(\hat{s}_i - s_i) + \frac{1}{2}(\hat{x}_i - x_i)(\hat{s}_i - s_i) \\ &= 2(s_i - \hat{s}_i)(\hat{x}_i - x_i) + (\hat{s}_i - s_i)(\hat{x}_i - x_i) = 2(s_i - \hat{s}_i)(\hat{x}_i - x_i) - (\hat{s}_i - s_i)(x_i - \hat{x}_i) \\ &= (s_i - \hat{s}_i)(\hat{x}_i - x_i). \end{aligned}$$

Neste caso,

$$\int_0^1 (J(w + tp) + J(\hat{w})) p dt = \begin{bmatrix} 0 \\ 0 \\ (S - \hat{S})(\hat{X} - X)e \end{bmatrix}, \quad (3.3)$$

e portanto, segue de (3.2) que

$$\|w - \hat{w}\| \leq \| [J(\hat{w})]^{-1} \| \left(\| F_{\hat{\mu}}(w) \| + \left\| \begin{bmatrix} 0 \\ 0 \\ (S - \hat{S})(\hat{X} - X)e \end{bmatrix} \right\| \right)$$

$$\begin{aligned}
&= \|[J(\hat{w})]^{-1}\| (\|F_{\hat{\mu}}(w)\| + \|(S - \hat{S})(\hat{X} - X)e\|) \\
&= \|[J(\hat{w})]^{-1}\| \|F_{\hat{\mu}}(w)\| + \|[J(\hat{w})]^{-1}\| \|(S - \hat{S})(\hat{X} - X)e\|. \tag{3.4}
\end{aligned}$$

Note que $(s - \hat{s})^T(\hat{x} - x) = 0$, pois como $w, \hat{w} \in \mathcal{F}_0$, temos que

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x} - x \\ \hat{\lambda} - \lambda \\ \hat{s} - s \end{bmatrix} = \begin{bmatrix} c - c \\ b - b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

A demonstração deste fato se faz de modo análogo à demonstração de que $\Delta x^T \Delta s = 0$, feita no Lema 1.6. Neste caso, pelo Lema 1.5,

$$\begin{aligned}
&\|(S - \hat{S})(\hat{X} - X)e\| \leq 2^{-\frac{3}{2}} \|s - \hat{s} + \hat{x} - x\|^2 = 2^{-\frac{3}{2}} \langle s - \hat{s} + \hat{x} - x, s - \hat{s} + \hat{x} - x \rangle \\
&= 2^{-\frac{3}{2}} (\|s - \hat{s}\|^2 + \|x - \hat{x}\|^2 + 2(s - \hat{s})^T(\hat{x} - x)) = 2^{-\frac{3}{2}} (\|s - \hat{s}\|^2 + \|x - \hat{x}\|^2). \tag{3.5}
\end{aligned}$$

Suponha que $\|w - \hat{w}\| < \frac{1}{2\|[J(\hat{w})]^{-1}\|}$. Segue de (3.4) e (3.5) que

$$\begin{aligned}
&\|w - \hat{w}\| \leq \|[J(\hat{w})]^{-1}\| \|F_{\hat{\mu}}(w)\| + \|[J(\hat{w})]^{-1}\| 2^{-\frac{3}{2}} (\|s - \hat{s}\|^2 + \|x - \hat{x}\|^2) \\
&\leq \|[J(\hat{w})]^{-1}\| \|F_{\hat{\mu}}(w)\| + \|[J(\hat{w})]^{-1}\| 2^{-\frac{3}{2}} (\|s - \hat{s}\|^2 + \|x - \hat{x}\|^2 + \|\lambda - \hat{\lambda}\|^2) \\
&\leq \|[J(\hat{w})]^{-1}\| \|F_{\hat{\mu}}(w)\| + \|[J(\hat{w})]^{-1}\| 2^{-\frac{3}{2}} \|w - \hat{w}\|^2 \\
&\leq \|[J(\hat{w})]^{-1}\| \|F_{\hat{\mu}}(w)\| + \|[J(\hat{w})]^{-1}\| 2^{-\frac{3}{2}} \|w - \hat{w}\| \|w - \hat{w}\| \\
&< \|[J(\hat{w})]^{-1}\| \|F_{\hat{\mu}}(w)\| + \|[J(\hat{w})]^{-1}\| 2^{-\frac{3}{2}} \frac{1}{2\|[J(\hat{w})]^{-1}\|} \|w - \hat{w}\| \\
&= \|[J(\hat{w})]^{-1}\| \|F_{\hat{\mu}}(w)\| + 2^{-\frac{5}{2}} \|w - \hat{w}\| \\
&\Leftrightarrow (1 - 2^{-\frac{5}{2}}) \|w - \hat{w}\| \leq \|[J(\hat{w})]^{-1}\| \|F_{\hat{\mu}}(w)\|.
\end{aligned}$$

Como $1 - 2^{-\frac{5}{2}} > 0$, temos que

$$\|w - \hat{w}\| \leq \frac{\|[J(\hat{w})]^{-1}\|}{1 - 2^{-\frac{5}{2}}} \|F_{\hat{\mu}}(w)\|.$$

□

Observe que se utilizarmos o Lema 3.6 para o caso em que $w \in \mathcal{N}_2(\theta)$ e $\mu = \hat{\mu}$, temos que se $\|w - \hat{w}\| < \frac{1}{2\|[J(\hat{w})]^{-1}\|}$, então $\|F_{\hat{\mu}}(w)\| = \|XSe - \hat{\mu}e\| = \|XSe - \mu e\| \leq \theta\mu$, e conseqüentemente,

$$\|w - \hat{w}\| \leq \frac{\|[J(\hat{w})]^{-1}\|}{1 - 2^{-\frac{5}{2}}} \theta\mu.$$

Apesar da semelhança com o resultado que gostaríamos de ter provado, a distância entre w e \hat{w} , neste caso, só pode ser controlada por essa estimativa quando uma determinada proximidade entre w e \hat{w} já é satisfeita.

3.3 Convergência Local do Algoritmo Proposto

Suponha que $\lim_{\mu \rightarrow 0} f_c(\mu)$ exista e defina $w^* = \lim_{\mu \rightarrow 0} f_c(\mu)$. Considere $\bar{w} : [0, +\infty) \rightarrow \mathbb{R}^{2n+m}$ tal que

$$\bar{w}(\mu) = \begin{cases} f_c(\mu) & , \text{ se } \mu > 0 \\ w^* & , \text{ se } \mu = 0. \end{cases}$$

Neste caso, \bar{w} é contínua. Suponha que $J(w^*)$ seja um isomorfismo. Então, podemos nos basear na demonstração do Teorema da Convergência do Método de Broyden e proceder como foi feito anteriormente na Seção 3.1 deste capítulo, definindo as funções

$$\beta(w) = \|[J(w)]^{-1}\|, \quad \delta(w) = \frac{1}{6\|[J(w)]^{-1}\|} \quad \text{e} \quad \varepsilon(w) = \frac{1}{9L\|[J(w)]^{-1}\|},$$

onde $w \in \bar{w}([0, +\infty))$. De modo análogo, podemos aplicar o Teorema da Função Inversa sobre F_0 em w^* , já que $J(w^*)$ é um isomorfismo, e obter vizinhanças V de w^* e W de $F_0(w^*)$ tais que F_0 é um difeomorfismo. Como $\bar{w} : [0, +\infty) \rightarrow \mathbb{R}^{2n+m}$ é contínua, podemos tomar $I = [0, a]$, com $a > 0$, suficientemente pequeno de modo que $\bar{w}(I) \subset V$. Nesse caso, $F_0 \circ \bar{w} : I \rightarrow W$ é contínua. Como $J_0^{-1} : W \rightarrow \mathbb{R}^{(2n+m) \times (2n+m)}$ também é contínua, a seguinte composição será contínua:

$$\begin{aligned} J_0^{-1} \circ (F_0 \circ \bar{w}) : I &\rightarrow W \rightarrow \mathbb{R}^{(2n+m) \times (2n+m)} \\ \mu &\mapsto J_0^{-1}(F_0(\bar{w}(\mu))). \end{aligned}$$

Ainda,

$$J_0^{-1}(F_0(\bar{w}(\mu))) = [J(\bar{w}(\mu))]^{-1},$$

e conseqüentemente, $\beta \circ \bar{w} : I \rightarrow \mathbb{R}$ será contínua, assim como $\delta \circ \bar{w} : I \rightarrow \mathbb{R}$ e $\varepsilon \circ \bar{w} : I \rightarrow \mathbb{R}$. Pela compacidade de I , temos que $\beta \circ \bar{w}$ admite um máximo em I , que denotaremos por β_0 . Analogamente, denotaremos por δ_0 e ε_0 os mínimos de $\delta \circ \bar{w}$ e $\varepsilon \circ \bar{w}$ em I , respectivamente. No Lema 3.7 a seguir mostramos que o Algoritmo 6 está bem definido quando iniciado próximo de uma solução do problema primal-dual (problemas (1.1) e (1.2)), quando a Jacobiana é não singular na solução.

Lema 3.7. *Suponha que $J(w^*)$ exista e seja inversível. Além disso, escolha $d \in \mathbb{R}^{2n+m}$ não nulo e $\sigma \in [0, \sigma_{\max}]$, onde $\sigma_{\max} < 1$. Seja $w \in \mathcal{F}_0$, μ sua medida de complementaridade e $\|w - w^*\| < \frac{\varepsilon_0}{2}$. Então, existe $\bar{\mu} > 0$ tal que se $\mu < \bar{\mu}$, então $\|\bar{w}(\sigma\mu) - w^*\| < \frac{\varepsilon_0}{2}$. Ainda, se $0 \leq \alpha < \frac{1}{\|d\|} \left(\frac{\varepsilon_0}{2} - \|w - w^*\| \right)$, então o método de Broyden converge para $\bar{w}(\sigma\mu)$, quando tomamos $w + \alpha d$ como ponto inicial.*

Demonstração. Como \bar{w} é uma função contínua em 0, segue que existe $\bar{\mu} > 0$ tal que

$$\tilde{\mu} = |\tilde{\mu} - 0| < \bar{\mu} \Rightarrow \|\bar{w}(\tilde{\mu}) - w^*\| = \|\bar{w}(\tilde{\mu}) - \bar{w}(0)\| < \frac{\varepsilon_0}{2}.$$

Conseqüentemente, $\|\bar{w}(\sigma\mu) - w^*\| < \frac{\varepsilon_0}{2}$, pois $\sigma\mu < \mu < \bar{\mu}$.

Uma condição suficiente para que $\|w + \alpha d - w^*\| < \frac{\varepsilon_0}{2}$ é que α satisfaça $0 \leq \alpha < \frac{1}{\|d\|} \left(\frac{\varepsilon_0}{2} - \|w - w^*\| \right)$. De fato,

$$\alpha < \frac{1}{\|d\|} \left(\frac{\varepsilon_0}{2} - \|w - w^*\| \right) \Leftrightarrow \alpha\|d\| < \frac{\varepsilon_0}{2} - \|w - w^*\| \Leftrightarrow \|w - w^*\| + \alpha\|d\| < \frac{\varepsilon_0}{2}.$$

Assim, pela desigualdade triangular,

$$\|w + \alpha d - w^*\| \leq \|w - w^*\| + \alpha\|d\| < \frac{\varepsilon_0}{2}.$$

Note que dados $p_1, p_2 \in B(w^*, \frac{\varepsilon_0}{2})$, temos que $\|p_1 - p_2\| < \varepsilon_0$. Como $\bar{w}(\sigma\mu) \in B(w^*, \frac{\varepsilon_0}{2})$ e $w + \alpha d \in B(w^*, \frac{\varepsilon_0}{2})$, segue que

$$\|w + \alpha d - \bar{w}(\sigma\mu)\| < \varepsilon_0,$$

e conseqüentemente, o método de Broyden converge (ao menos q -linearmente) para $\bar{w}(\sigma\mu)$ quando tomamos $w + \alpha d$ como ponto inicial. \square

Teorema 3.8 (Convergência Local do Algoritmo 6). *Suponha que $J(w^*)$ exista e seja inversível. Seja $(w_k)_{k \in \mathbb{N}}$ a sequência gerada pelo Algoritmo 6. Então, existe $\bar{\mu} > 0$ tal que se $\|w_k - w^*\| < \frac{\varepsilon_0}{2}$ e $\mu_k < \bar{\mu}$, onde μ_k é a medida de complementaridade de w_k , então*

$$\|w_k - w^*\| \rightarrow 0 \text{ quando } k \rightarrow +\infty.$$

Demonstração. Seja $\bar{\mu} > 0$ proveniente do Lema 3.7, o qual existe por conta da continuidade da função \bar{w} em 0 e também da existência e invertibilidade da Jacobiana de F em w^* . Suponha que para algum $k \in \mathbb{N}$ tenhamos $\mu_k < \bar{\mu}$ e $\|w_k - w^*\| < \frac{\varepsilon_0}{2}$. Então, pelo Lema 3.7, $w_{k+1} := \bar{w}(\sigma_k \mu_k)$, pois em algum momento $w_k + \alpha_k d_k$ entrará na região de convergência do método de Broyden. Uma vez que $\bar{w}(\sigma_k \mu_k)$ satisfaz as condições iniciais do Lema 3.7, podemos aplicar este lema de forma recursiva, obtendo $w_{k+j} := \bar{w}(\sigma_{k+j-1} \mu_{k+j-1})$, para todo $j \in \mathbb{N}$.

Agora, como $\sigma_k \leq \sigma_{\max} < 1$, temos que $\mu_{k+1} = \sigma_k \mu_k \leq \sigma_{\max} \mu_k$, e portanto,

$$\mu_{k+j} \leq \sigma_{\max} \mu_{k+j-1} \leq \cdots \leq \sigma_{\max}^j \mu_k, \forall j \in \mathbb{N}.$$

Como $\sigma_{\max} < 1$, isto implica que $\mu_{k+j} \rightarrow 0$ conforme $j \rightarrow +\infty$. Neste caso, $\mu_k \rightarrow 0$. Sendo assim, $\sigma_k \mu_k \rightarrow 0$. Pela continuidade da função \bar{w} em 0, temos que $\bar{w}(\sigma_k \mu_k) \rightarrow \bar{w}(0) = w^*$.

Observe que

$$\|w_{k+j} - w^*\| = \|\bar{w}(\sigma_{k+j-1} \mu_{k+j-1}) - w^*\|, \forall j \in \mathbb{N}.$$

Sendo assim, temos que $\|w_{k+j} - w^*\| \rightarrow 0$, e conseqüentemente,

$$\|w_k - w^*\| \rightarrow 0.$$

□

Considere o seguinte algoritmo, que consiste no Algoritmo 6 com uma modificação no Passo 4:

Algoritmo 7 Preditor com Corretor Quasi-Newton Modificado

Dado $w_0 = (x_0, \lambda_0, s_0) \in \mathcal{F}_0$ e $0 < \sigma_{\max} < 1$.

Seja $k = 0$.

Passo 1: Faça $\sigma_k = 0$ em (1.22) e encontre $d_k = (\Delta x_k, \Delta \lambda_k, \Delta s_k)$.

Passo 2: Escolha o maior $\alpha_k \in [0, 1]$ tal que $w_k + \alpha_k d_k \in \mathcal{F}_0$ e $\mu \leq \mu_k$, onde μ é a medida de complementaridade em $w_k + \alpha_k d_k$.

Passo 3: Defina $\sigma_k = \frac{\sigma_{\max}}{2}$.

Passo 4: Resolva

$$F_{\sigma_k \mu_k}(w) = 0,$$

tomando como ponto inicial $w_k + \alpha_k d_k$ e $B_0 = J(w_k)$. Denote por $(y_{k_i})_{i \in \mathbb{N}}$ a sequência gerada pelo Método de Broyden. Se possível, após a primeira iteração, interrompa o Método de Broyden assim que

$$\mu(y_{k_i}) < \frac{(\sigma_{\max} + \sigma_k)}{2} \mu_k \quad \text{e} \quad y_{k_i} \in \mathcal{F}_0,$$

onde $\mu(y_{k_i})$ é a medida de complementaridade de y_{k_i} , faça $\tilde{w} = y_{k_i}$ e vá para o Passo 6.

Caso contrário, vá para o Passo 5.

Passo 5: Defina $\alpha_k = \frac{\alpha_k}{2}$ e $\sigma_k = \frac{\sigma_{\max} + \sigma_k}{2}$, e depois volte ao Passo 4.

Passo 6: Defina $w_{k+1} = \tilde{w}$, faça $k = k + 1$ e volte para o Passo 1.

Observe que no Passo 4 modificado, $y_{k_0} := w_k + \alpha_k d_k \in \mathcal{F}_0$, pela escolha de α_k no Passo 2. Em particular, y_{k_0} é um ponto viável, isto é, satisfaz as restrições $Ax = b$ e

$A^T \lambda + s = c$. O lema a seguir foi adaptado do Lema 1 encontrado em [5] e mostra que os passos corretores preservam a viabilidade do ponto.

Lema 3.9. *Sejam $\sigma \in [0, 1]$, $\mu > 0$, e considere $F_{\sigma\mu}$ definida em (1.7). Seja $(w_k)_{k \in \mathbb{N}}$ a sequência gerada pelo método de Broyden ao tentar resolver $F_{\sigma\mu}(w) = 0$. Suponha que na k -ésima iteração w_k seja viável, e que a aproximação B_k tenha o seguinte formato:*

$$B_k = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ M_k^1 & M_k^2 & M_k^3 \end{bmatrix},$$

em que $M_k^1, M_k^3 \in \mathbb{R}^{n \times n}$ e $M_k^2 \in \mathbb{R}^{n \times m}$. Então, a matriz B_{k+1} obtida pela atualização de Broyden é da forma

$$B_{k+1} = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ M_{k+1}^1 & M_{k+1}^2 & M_{k+1}^3 \end{bmatrix},$$

em que $M_{k+1}^1, M_{k+1}^3 \in \mathbb{R}^{n \times n}$ e $M_{k+1}^2 \in \mathbb{R}^{n \times m}$. Ainda, o ponto w_{k+1} , obtido após um passo do método de Broyden, também é viável.

Demonstração. Vamos denotar $\bar{s}_k = w_{k+1} - w_k$. Por definição,

$$B_{k+1} = B_k + \frac{(y_k - B_k \bar{s}_k) \bar{s}_k^T}{\bar{s}_k^T \bar{s}_k},$$

logo, se

$$\frac{(y_k - B_k \bar{s}_k) \bar{s}_k^T}{\bar{s}_k^T \bar{s}_k} = B_{k+1} - B_k = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ M^1 & M^2 & M^3 \end{bmatrix},$$

para matrizes $M^1, M^3 \in \mathbb{R}^{n \times n}$ e $M^2 \in \mathbb{R}^{n \times m}$, temos que B_{k+1} tem o formato que consta no enunciado deste lema.

Como w_k é um ponto viável, segue que

$$y_k = F_{\sigma\mu}(w_{k+1}) - F_{\sigma\mu}(w_k) = \begin{bmatrix} A^T \lambda_{k+1} + s_{k+1} - c \\ Ax_{k+1} - b \\ X_{k+1} S_{k+1} e - \sigma \mu e \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ X_k S_k e - \sigma \mu e \end{bmatrix} = \begin{bmatrix} A^T \lambda_{k+1} + s_{k+1} - c \\ Ax_{k+1} - b \\ v_1 \end{bmatrix},$$

onde $v_1 = X_{k+1} S_{k+1} e - X_k S_k e$.

Por outro lado, também segue da viabilidade de w_k que

$$\begin{aligned} B_k \bar{s}_k &= \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ M_k^1 & M_k^2 & M_k^3 \end{bmatrix} \begin{bmatrix} x_{k+1} - x_k \\ \lambda_{k+1} - \lambda_k \\ s_{k+1} - s_k \end{bmatrix} = \begin{bmatrix} A^T \lambda_{k+1} + s_{k+1} - (A^T \lambda_k + s_k) \\ Ax_{k+1} - Ax_k \\ v_2 \end{bmatrix} \\ &= \begin{bmatrix} A^T \lambda_{k+1} + s_{k+1} - c \\ Ax_{k+1} - b \\ v_2 \end{bmatrix}, \end{aligned}$$

onde $v_2 = M_k^1(x_{k+1} - x_k) + M_k^2(\lambda_{k+1} - \lambda_k) + M_k^3(s_{k+1} - s_k)$. Fazendo $\Delta x_k = (x_{k+1} - x_k)$, $\Delta \lambda_k = (\lambda_{k+1} - \lambda_k)$ e $\Delta s_k = (s_{k+1} - s_k)$, podemos escrever $v_2 = M_k^1 \Delta x_k + M_k^2 \Delta \lambda_k + M_k^3 \Delta s_k$. Neste caso, $y_k - B_k \bar{s}_k = [0 \ 0 \ v_3]^T$, onde $v_3 = v_1 - v_2$. Deste modo, multiplicando $y_k - B_k \bar{s}_k$ à direita por $\frac{\bar{s}_k^T}{\bar{s}_k^T \bar{s}_k}$, devemos ter que

$$\frac{(y_k - B_k \bar{s}_k) \bar{s}_k^T}{\bar{s}_k^T \bar{s}_k} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ M^1 & M^2 & M^3 \end{bmatrix},$$

como desejado. Além disso, se denotarmos $v_3 = (v_x, v_\lambda, v_s)$, temos que

$$M^1 = \frac{v_x \bar{s}_k^T}{\bar{s}_k^T \bar{s}_k}, \quad M^2 = \frac{v_\lambda \bar{s}_k^T}{\bar{s}_k^T \bar{s}_k} \quad \text{e} \quad M^3 = \frac{v_s \bar{s}_k^T}{\bar{s}_k^T \bar{s}_k}.$$

Por fim, como \bar{s}_k satisfaz $B_k \bar{s}_k = -F_{\sigma \mu_k}(w_k)$, temos dos dois primeiros blocos desse sistema que

$$\begin{aligned} A^T(\lambda_{k+1} - \lambda_k) + (s_{k+1} - s_k) &= 0 \Leftrightarrow A^T \lambda_{k+1} + s_{k+1} - (A^T \lambda_k + s_k) = 0 \\ &\Leftrightarrow A^T \lambda_{k+1} + s_{k+1} - c = 0 \Leftrightarrow A^T \lambda_{k+1} + s_{k+1} = c \end{aligned}$$

e

$$A(x_{k+1} - x_k) = Ax_{k+1} - Ax_k = Ax_{k+1} - b = 0 \Leftrightarrow Ax_{k+1} = b.$$

Portanto, w_{k+1} é um ponto viável. □

Como y_{k_0} é um ponto viável e $B_0 = J(w_k)$, seguirá pelo Lema 3.9 que os dois primeiros blocos de B_0 são preservados pela atualização de Broyden em todas as iterações, e também que todos os elementos da sequência gerada pelo método de Broyden serão pontos viáveis.

Deste modo, basta que a condição $x, s > 0$ se verifique para que um determinado y_{k_i} pertença a \mathcal{F}_0 . No caso em que y_{k_i} converge para um ponto y^* no caminho central, por exemplo, podemos interromper o método de Broyden assim que o decréscimo suficiente de μ for satisfeito juntamente com a condição $x, s > 0$, a qual deverá ser satisfeita para iterandos suficientemente próximos de y^* .

É importante destacar que o terceiro bloco de B_0 tomado dessa forma é bastante esparsa. Contudo, o Lema 3.9 não garante que essa esparsidade será preservada ao longo das iterações. De fato, se supormos que a matriz B_k no Lema 3.9 seja a Jacobiana em w_k , teremos de $B_k \bar{s}_k = -F_{\sigma\mu}(w_k)$ que

$$v_2 = -X_k S_k e + \sigma \mu e,$$

e conseqüentemente,

$$v_3 = X_{k+1} S_{k+1} e - \sigma \mu e.$$

Para que a esparsidade seja de certa forma preservada, precisaríamos que M^1 , M^2 e M^3 contivessem o maior número de zeros possível. Pelas expressões obtidas para essas matrizes, a ocorrência de zeros nelas se dará de duas formas: entradas nulas em v_3 ou entradas nulas em \bar{s}_k . Observe que se alguma entrada i de v_3 se anular, a i -ésima linha de M^1 , M^2 e M^3 deverá ser nula. Além disso, se alguma entrada i de \bar{s}_k se anular, uma coluna inteira de alguma das matrizes M^1 , M^2 ou M^3 deverá ser nula.

Embora não possamos afirmar com certeza se o método de Broyden produzirá direções \bar{s}_k com entradas nulas, sabemos que se o método de Broyden estiver convergindo, então $v_3 \rightarrow 0$. Além disso, a convergência do método de Broyden implicaria que $\bar{s}_k \rightarrow 0$, de modo que ao menos perto da solução, podemos esperar que M^1 , M^2 e M^3 tenham muitos elementos próximos de zero.

Em [5], o método de Broyden em sua versão “ruim” foi capaz de obter um resultado um pouco melhor sobre a preservação da esparsidade. A versão “ruim” do método de Broyden faz uma atualização de B_k um pouco mais geral que o método de Broyden “bom”. Esta atualização possui um vetor que pode ser escolhido para a realização da atualização de posto 1. Ainda, o método de Broyden “bom” pode ser obtido escolhendo este vetor de uma maneira bem específica. Em [5] (Lema 1), mostrou-se que uma escolha conveniente desse vetor permite anular M^2 .

Na Seção 3.2, demonstramos o Lema 3.6 que dizia respeito a distância entre um dado

ponto e um ponto do caminho central. Supondo que $J(w^*)$ exista e seja inversível, temos que o Lema 3.6 é válido para $\hat{w} = w^*$. Neste caso, se $w \in \mathcal{F}_0$ satisfizer $\|w - w^*\| < \frac{1}{2\| [J(w^*)]^{-1} \|}$, então $\|w - w^*\| \leq \frac{\| [J(w^*)]^{-1} \| \|XSe\|}{1 - 2^{-\frac{5}{2}}}$.

A seguir, temos o Teorema da Convergência Local do Método de Newton:

Teorema 3.10 (Convergência Local do Método de Newton). *Seja $F : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ uma aplicação continuamente diferenciável em um aberto e convexo $D \subset \mathbb{R}^{2n+m}$. Assuma que exista $w^* \in \mathbb{R}^{2n+m}$ e $r, \beta > 0$, tais que $B(w^*, r) \subset D$, $F(w^*) = 0$, $J(w^*)^{-1}$ exista com $\|J(w^*)^{-1}\| \leq \beta$, e $J \in Lip_L(B(w^*, r))$. Então, existe $\varepsilon > 0$ tal que para todo $w_0 \in B(w^*, \varepsilon)$, a sequência w_1, w_2, \dots gerada por*

$$w_{k+1} = w_k - J(w_k)^{-1}F(w_k), \quad k = 0, 1, 2, \dots$$

está bem definida, converge para w^ , e obedece*

$$\|w_{k+1} - w^*\| \leq \beta L \|w_k - w^*\|^2, \quad k = 0, 1, 2, \dots$$

Demonstração. Veja o Teorema 5.2.1 de [10]. □

Observação 3.11. *Embora não tenha sido explicitado no enunciado do Teorema 3.10, em sua demonstração foi provado que $\|x_{k+1} - x^*\| \leq \frac{1}{2}\|x_k - x^*\|$, para todo $k = 0, 1, 2, \dots$. Este fato será utilizado na demonstração do Corolário 3.15.*

Observe que no Lema 3.7 está subentendido que foi encontrado no Passo 2 do Algoritmo 6 um $\alpha \in [0, 1]$ tal que $\mu_\alpha \leq \mu$, onde μ_α é a medida de complementaridade de $w + \alpha d$. Para uma direção arbitrária, pode ser que isso não seja possível. No entanto, como d é tomada como a direção de Newton obtida ao resolvermos (1.8) com $\sigma = 0$, temos que $\mu_\alpha = (1 - \alpha)\mu$, donde podemos concluir que $\mu_\alpha \leq \mu$, para todo $\alpha \in [0, 1]$. Ainda, se levarmos em consideração o resultado de convergência local do método de Newton, poderíamos melhorar o resultado do Lema 3.7.

Seja ε_1 o raio de convergência para o Método de Newton proveniente do Teorema 3.10 relativo ao ponto w^* . Vamos denotar

$$\varepsilon_2 = \min \left\{ \varepsilon_1, \frac{\varepsilon_0}{2}, \frac{1}{2\| [J(w^*)]^{-1} \|} \right\},$$

onde ε_0 foi definido como no início desta seção. Podemos adaptar o Lema 3.7 para o seguinte resultado:

Lema 3.12. *Suponha que $J(w^*)$ exista, seja inversível e d seja a solução de (1.8) quando $\sigma = 0$. Seja $\sigma \in [0, \sigma_{\max}]$, onde $\sigma_{\max} < 1$, $w \in \mathcal{F}_0$, μ sua medida de complementaridade e $\|w - w^*\| < \varepsilon_2$. Existe $\bar{\mu} > 0$ tal que se $\mu < \bar{\mu}$, então $\|\bar{w}(\sigma\mu) - w^*\| < \frac{\varepsilon_2}{4}$ e $\|w + \alpha d - w^*\| < \varepsilon_2$, para todo $\alpha \in [0, 1]$. Ainda, para todo $\alpha \in [0, 1]$, temos que o método de Broyden converge para $\bar{w}(\sigma\mu)$, quando tomamos $w + \alpha d$ como ponto inicial.*

Demonstração. A demonstração da existência de $\bar{\mu} > 0$ tal que $\|\bar{w}(\sigma\mu) - w^*\| < \frac{\varepsilon_2}{4}$ é realizada de forma análoga ao que foi feito na demonstração do Lema 3.7. Agora, observe que como $\varepsilon_2 \leq \varepsilon_1$, segue que ε_2 é um raio de convergência para o método de Newton. Ainda, $\|w - w^*\| < \varepsilon_2$ implica

$$\|w + d - w^*\| \leq \frac{1}{2}\|w - w^*\| < \frac{1}{2}\varepsilon_2 < \varepsilon_2.$$

Como $w, w + d \in B(w^*, \varepsilon_2)$ e $B(w^*, \varepsilon_2)$ é convexa, segue que $w + \alpha d \in B(w^*, \varepsilon_2)$, para todo $\alpha \in [0, 1]$. Como $\varepsilon_2 < \frac{\varepsilon_0}{2}$, a convergência do método de Broyden para $\bar{w}(\sigma\mu)$ a partir de $w + \alpha d$ está garantida, para todo $\alpha \in [0, 1]$, pelo mesmo argumento utilizado na demonstração do Lema 3.7.

□

O lema a seguir explicita o valor de μ após uma iteração do método de Broyden ao tomarmos $B_0 = J(w_k)$ e $y_{k_0} = w_k + \alpha_k d_k$.

Lema 3.13. *Seja $w_k \in \mathcal{F}_0$, μ_k sua medida de complementaridade e suponha que $y_{k_0} := w_k + \alpha_k d_k \in \mathcal{F}_0$, onde d_k é obtida resolvendo (1.22) com $\sigma_k = 0$. Seja \bar{d} a direção de Broyden obtida ao fazer $\sigma = \sigma_k$ e $\mu = \mu_k$ em (1.7). Dado $\bar{\alpha} \in [0, 1]$, temos que a medida de complementaridade do ponto $y_{k_0} + \bar{\alpha}\bar{d}$ é*

$$[(1 - \alpha_k)(1 - \bar{\alpha}) + \bar{\alpha}\sigma_k]\mu_k. \quad (3.6)$$

Demonstração. Vamos denotar $w_k = (x_k, \lambda_k, s_k)$, $y_{k_0} = (\bar{x}, \bar{\lambda}, \bar{s})$, $d_k = (d_x, d_\lambda, d_s)$ e $\bar{d} = (\bar{d}_x, \bar{d}_\lambda, \bar{d}_s)$. Como d_k é a direção de Newton tomada a partir de w_k e $w_k \in \mathcal{F}_0$, temos que

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} d_x \\ d_\lambda \\ d_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -X_k S_k e \end{bmatrix}. \quad (3.7)$$

Além disso, como $B_0 = J(w_k)$, \bar{d} é a direção de Broyden tomada a partir de y_{k_0} e $y_{k_0} \in \mathcal{F}_0$, temos que

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \bar{d}_x \\ \bar{d}_\lambda \\ \bar{d}_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \sigma_k \mu_k e - \bar{X} \bar{S} e \end{bmatrix}. \quad (3.8)$$

Dos dois primeiros blocos de (3.7), temos que $d_x^T d_s = 0$. De fato, o primeiro bloco implica que $A^T d_\lambda + d_s = 0$ e o segundo que $Ad_x = 0$, logo

$$d_x^T d_s = (Ad_x)^T d_\lambda + d_x^T d_s = d_x^T A^T d_\lambda + d_x^T d_s = d_x^T (A^T d_\lambda + d_s) = 0.$$

De modo análogo, obtemos que $\bar{d}_x^T \bar{d}_s = 0$, pelos dois primeiros blocos de (3.8). Ainda, se utilizarmos o primeiro e o segundo bloco de (3.7) e (3.8) de forma alternada, podemos concluir de forma análoga que $d_x^T \bar{d}_s = 0$ e $\bar{d}_x^T d_s = 0$. Portanto, temos que

$$d_x^T d_s = \bar{d}_x^T \bar{d}_s = d_x^T \bar{d}_s = \bar{d}_x^T d_s = 0. \quad (3.9)$$

Por (3.9), temos que

$$(\bar{x} + \bar{\alpha} \bar{d}_x)^T (\bar{s} + \bar{\alpha} \bar{d}_s) = \bar{x}^T \bar{s} + \bar{\alpha} \left(\bar{x}^T \bar{d}_s + \bar{d}_x^T \bar{s} \right) + \bar{\alpha}^2 \bar{d}_x^T \bar{d}_s = \bar{x}^T \bar{s} + \bar{\alpha} \left(\bar{x}^T \bar{d}_s + \bar{d}_x^T \bar{s} \right).$$

De modo análogo, segue de (3.9) que

$$\bar{x}^T \bar{s} = (x_k + \alpha_k d_x)^T (s_k + \alpha_k d_s) = x_k^T s_k + \alpha_k (x_k^T d_s + d_x^T s_k). \quad (3.10)$$

Pela definição de y_{k_0} e (3.10), temos que

$$\begin{aligned} (\bar{x} + \bar{\alpha} \bar{d}_x)^T (\bar{s} + \bar{\alpha} \bar{d}_s) &= (x_k + \alpha_k d_x)^T (s_k + \alpha_k d_s) + \bar{\alpha} \left((x_k + \alpha_k d_x)^T \bar{d}_s + \bar{d}_x^T (s_k + \alpha_k d_s) \right) \\ &= x_k^T s_k + \alpha_k (x_k^T d_s + d_x^T s_k) + \bar{\alpha} \left((x_k + \alpha_k d_x)^T \bar{d}_s + \bar{d}_x^T (s_k + \alpha_k d_s) \right) \\ &= x_k^T s_k + \alpha_k (x_k^T d_s + d_x^T s_k) + \bar{\alpha} \left(x_k^T \bar{d}_s + \alpha_k d_x^T \bar{d}_s + \bar{d}_x^T s_k + \alpha_k \bar{d}_x^T d_s \right). \end{aligned}$$

Segue de (3.9) que

$$(\bar{x} + \bar{\alpha} \bar{d}_x)^T (\bar{s} + \bar{\alpha} \bar{d}_s) = x_k^T s_k + \alpha_k (x_k^T d_s + d_x^T s_k) + \bar{\alpha} \left(x_k^T \bar{d}_s + \bar{d}_x^T s_k \right) \quad (3.11)$$

Se somarmos as equações do terceiro bloco de (3.7), obtemos

$$x_k^T d_s + d_x^T s_k = -x_k^T s_k. \quad (3.12)$$

De modo análogo, obtemos de (3.8) que

$$x_k^T \bar{d}_s + \bar{d}_x^T s_k = n \sigma_k \mu_k - \bar{x}^T \bar{s} = n \sigma_k \mu_k - \bar{x}^T \bar{s}.$$

Como $\bar{x}^T \bar{s} = (x_k + \alpha_k d_x)^T (s_k + \alpha_k d_s)$, segue que

$$x_k^T \bar{d}_s + \bar{d}_x^T s_k = n \sigma_k \mu_k - \bar{x}^T \bar{s} = n \sigma_k \mu_k - x_k^T s_k - \alpha_k (x_k^T d_s + d_x^T s_k). \quad (3.13)$$

Substituindo (3.13) e (3.12) em (3.11), obtemos

$$(\bar{x} + \bar{\alpha} \bar{d}_x)^T (\bar{s} + \bar{\alpha} \bar{d}_s) = x_k^T s_k + \alpha_k (-x_k^T s_k) + \bar{\alpha} (n \sigma_k \mu_k - x_k^T s_k - \alpha_k (-x_k^T s_k)).$$

Dividindo a equação acima por n , temos que

$$\begin{aligned} \frac{(\bar{x} + \bar{\alpha} \bar{d}_x)^T (\bar{s} + \bar{\alpha} \bar{d}_s)}{n} &= \mu_k + \alpha_k (-\mu_k) + \bar{\alpha} (\sigma_k \mu_k - \mu_k - \alpha_k (-\mu_k)). \\ &= [1 - \alpha_k + \bar{\alpha} (\sigma_k - 1 + \alpha_k)] \mu_k = [1 - \alpha_k + \bar{\alpha} \sigma_k - \bar{\alpha} + \bar{\alpha} \alpha_k] \mu_k \\ &= [(1 - \bar{\alpha}) - \alpha_k (1 - \bar{\alpha}) + \bar{\alpha} \sigma_k] \mu_k = [(1 - \alpha_k)(1 - \bar{\alpha}) + \bar{\alpha} \sigma_k] \mu_k, \end{aligned}$$

como desejado. □

Observação 3.14. *No caso do Algoritmo 6 e sua versão modificada, o método de Broyden sempre dá passos completos. Nas notações do Lema 3.13, isto significa tomar $\bar{\alpha} = 1$. Pelo Lema 3.13, segue que em um único passo de Broyden alcança-se uma medida de complementaridade igual à $\sigma_k \mu_k$. No entanto, é importante destacar que este resultado não garante que a condição $x, s > 0$ seja satisfeita em um único passo.*

Com base no Lema 3.12, podemos provar um resultado análogo ao Teorema 3.8:

Corolário 3.15 (Convergência Local do Algoritmo 7). *Suponha que $J(w^*)$ exista e seja inversível. Seja $(w_k)_{k \in \mathbb{N}}$ a sequência gerada pelo Algoritmo 7. Então, existe $\bar{\mu} > 0$ tal que se $\|w_k - w^*\| < \varepsilon_2$ e $\mu_k < \bar{\mu}$, onde μ_k é a medida de complementaridade de w_k , então*

$$\|w_k - w^*\| \rightarrow 0 \text{ quando } k \rightarrow +\infty.$$

Demonstração. Seja $\bar{\mu} > 0$ proveniente do Lema 3.12. Suponha que para algum $k \in \mathbb{N}$, tenhamos $\|w_k - w^*\| < \varepsilon_2$ e $\mu_k < \bar{\mu}$. Pelo Lema 3.12, temos que o algoritmo encontrará $\alpha_k \in [0, 1]$ e $\sigma_k \in [0, \sigma_{\max}]$ tais que o método de Broyden converge para $\bar{\omega}(\sigma_k \mu_k)$, quando o ponto inicial tomado for $y_{k_0} = w_k + \alpha_k d_k$. Pela demonstração do Teorema 2.12, temos que

$$\|y_{k_{i+1}} - \bar{\omega}(\sigma_k \mu_k)\| \leq \frac{1}{2^{i+1}} \|y_{k_0} - \bar{\omega}(\sigma_k \mu_k)\| \leq \frac{1}{2} \|y_{k_0} - \bar{\omega}(\sigma_k \mu_k)\|, \quad \forall i = 0, 1, 2, \dots$$

Agora, veja que

$$\begin{aligned} \|y_{k_i} - w^*\| &\leq \|y_{k_i} - \bar{w}(\sigma_k \mu_k)\| + \|\bar{w}(\sigma_k \mu_k) - w^*\| \leq \frac{1}{2} \|y_{k_0} - \bar{w}(\sigma_k \mu_k)\| + \|\bar{w}(\sigma_k \mu_k) - w^*\| \\ &\leq \frac{1}{2} \|y_{k_0} - w^*\| + \frac{1}{2} \|\bar{w}(\sigma_k \mu_k) - w^*\| + \|\bar{w}(\sigma_k \mu_k) - w^*\| \\ &\leq \frac{1}{2} \|y_{k_0} - w^*\| + \frac{3}{2} \|\bar{w}(\sigma_k \mu_k) - w^*\|, \quad \forall i = 1, 2, 3, \dots \end{aligned}$$

Pelo Lema 3.12, temos que $\|y_{k_0} - w^*\| < \varepsilon_2$ e $\|\bar{w}(\sigma_k \mu_k) - w^*\| < \frac{\varepsilon_2}{4}$, logo

$$\|y_{k_i} - w^*\| < \frac{1}{2} \varepsilon_2 + \frac{3}{8} \varepsilon_2 < \frac{1}{2} \varepsilon_2 + \frac{1}{2} \varepsilon_2 < \varepsilon_2, \quad \forall i = 1, 2, 3, \dots$$

Deste modo, o próximo iterando, w_{k+1} , deverá satisfazer $\|w_{k+1} - w^*\| < \varepsilon_2$. Aplicando o Lema 3.12 de forma recursiva, obtemos

$$\|w_{k+j} - w^*\| < \varepsilon_2, \quad \forall j = 1, 2, 3, \dots$$

Sendo assim, podemos aplicar o Lema 3.6 para $\hat{w} = w^*$, implicando que

$$\|w_{k+j} - w^*\| \leq \frac{\| [J(w^*)]^{-1} \| \| X_{k+j} S_{k+j} e \|}{1 - 2^{-\frac{5}{2}}}, \quad \forall j = 1, 2, 3, \dots \quad (3.14)$$

Além disso, por construção, temos que

$$\mu_{k+1} \leq \frac{\sigma_{\max} + \sigma_k}{2} \mu_k \leq \frac{\sigma_{\max} + \sigma_{\max}}{2} \mu_k = \sigma_{\max} \mu_k.$$

De forma recursiva, obtemos

$$\mu_{k+j} \leq \sigma_{\max}^j \mu_k, \quad \forall j \in \mathbb{N}.$$

Como $0 < \sigma_{\max} < 1$, segue que $\mu_{k+j} \rightarrow 0$ conforme $j \rightarrow +\infty$. Neste caso, $x_{k+j}^i s_{k+j}^i \rightarrow 0$, para todo $j = 1, 2, \dots, n$, implicando que $\|X_{k+j} S_{k+j} e\| \rightarrow 0$. Seguirá de (3.14) que $\|w_{k+j} - w^*\| \rightarrow 0$. Sendo assim, quando $k \rightarrow +\infty$, temos

$$\|w_k - w^*\| \rightarrow 0.$$

□

Assim sendo, temos a convergência local do Algoritmo 7. No próximo capítulo, será abordada uma implementação desse algoritmo com o fim de observar o seu comportamento na prática.

IMPLEMENTAÇÃO DO ALGORITMO 7

O Algoritmo 7 foi implementado da seguinte forma. No Passo 1, fazemos $\sigma_k = 0$ e resolvemos (1.22) para encontrar d_k . Para resolver este sistema, a matriz jacobiana de $F_{\sigma_k \mu_k}$ em w_k passa por uma fatoração LU (ver [20], pág. 129). Por meio dessa fatoração, o sistema 7 é dividido em dois sistemas mais simples de resolver, onde usa-se a eliminação de Gauss (ver [20], pág. 139). No passo 2, calculamos α_k de modo análogo ao que foi feito no Algoritmo 4, isto é, calculamos o maior α_k tal que, para todo $i = 1, 2, \dots, n$, temos

$$x_i + \alpha_k \Delta x_i \geq 0 \quad \text{e} \quad s_i + \alpha_k \Delta s_i \geq 0.$$

Note que o único caso em que o ponto novo pode se tornar negativo ocorre quando $\Delta x_i < 0$ ou $\Delta s_i < 0$, pois nos outros casos seguirá de $(x_i, s_i) > 0$ que podemos tomar $0 \leq \alpha_k$ arbitrário sem desprezar estas desigualdades. Sendo assim, isolando α_k nas inequações em que $\Delta x_i < 0$ ou $\Delta s_i < 0$, podemos concluir que o maior dos α_k 's será α_{\max} dado da seguinte forma:

$$\alpha_{\max} = \min \left\{ 1, \min_{\Delta x_i < 0} \left\{ -\frac{x_i}{\Delta x_i} \right\}, \min_{\Delta s_j < 0} \left\{ -\frac{s_j}{\Delta s_j} \right\} \right\}.$$

Para evitar que o novo ponto possua x ou s com entradas nulas, α_k foi tomado da seguinte forma:

$$\alpha_k = \min \left\{ 1, 0.9 \min_{\Delta x_i < 0} \left\{ -\frac{x_i}{\Delta x_i} \right\}, 0.9 \min_{\Delta s_j < 0} \left\{ -\frac{s_j}{\Delta s_j} \right\} \right\}.$$

Além disso, vimos anteriormente na Seção 1.5 do Capítulo 1 que a medida de complementaridade do ponto $(x(\alpha_k), \lambda(\alpha_k), s(\alpha_k))$, obtido após dar um passo na direção afim-escala com tamanho de passo α_k , é dada por

$$\mu(\alpha_k) = (1 - \alpha_k(1 - \sigma_k))\mu_k = (1 - \alpha_k)\mu_k < \mu_k,$$

logo o ponto tomado está nas condições requeridas pelo Passo 2. O Passo 4 é realizado como na modificação descrita na Seção 3.3 do Capítulo 3. Na teoria presente na Seção

3.1 do Capítulo 3, vimos que sempre é possível encontrar α_k suficientemente próximo de 0 e σ_k suficientemente próximo de 1 de modo que o método de Broyden venha a convergir. Na prática, no entanto, pode demorar muito para que estes valores sejam encontrados. Além disso, a exigência de que $\sigma_k \leq \sigma_{\max} < 1$, a qual foi necessária na Seção 3.3 do Capítulo 3 para as demonstrações de convergência local do Algoritmo 6 e sua versão modificada, também pode interferir na convergência do método de Broyden. Por esse motivo, a implementação realizada adota uma estratégia alternativa quando o Algoritmo 7 falha 3 vezes consecutivas em determinar o próximo iterando durante o Passo 4 em uma única iteração, isto é, o Passo 5 foi executado 3 vezes em uma mesma iteração e não acarretou na convergência do Método de Broyden.

No método alternativo, executamos um passo preditor no qual obtemos um ponto da forma $w_k + \alpha_k d_k$, onde d_k foi encontrada no Passo 2 (direção de Newton com $\sigma = 0$ e $\mu = \mu_k$ em (1.7)) e α_k como definido anteriormente. Depois, o passo corretor alternativo consiste em 5 iterações de uma variação do método de Broyden tomando $\tilde{w} := w_k + \alpha_k d_k$ como ponto inicial, $B_0 = J(w_k)$ e $\sigma_k = \left(\frac{\mu_a}{\mu_k}\right)^3$, onde este valor de σ é o mesmo utilizado no algoritmo preditor-corrector de Mehrotra [8] (Algoritmo 4). Nesta variação do método de Broyden, o tamanho dos passos é controlado por um múltiplo $\beta \in [0, 1]$, isto é, se \bar{d} for a direção encontrada pelo método de Broyden usual, então será dado um passo da forma $\tilde{w} + \beta \bar{d}$. Este múltiplo é escolhido de forma análoga a uma busca linear, com a exigência de que $\mu(\tilde{w} + \beta \bar{d}) \leq \mu(\tilde{w})$ e que a condição $x, s > 0$ seja satisfeita.

Outra situação que ocorre na prática é a singularidade de $J(w_k)$, por conta do mal condicionamento dessa matriz. Neste caso, o passo preditor não está bem definido, pois ele é determinado a partir do método de Newton, o qual requer a invertibilidade de $J(w_k)$. Nos casos em que isso acontece, a implementação realizada faz correções na Jacobiana para que ela se torne inversível. A correção realizada se inspira na correção do método de Levenberg-Marquardt (ver [2], pág. 258), e consiste em somar pequenos múltiplos positivos da matriz identidade na Jacobiana até que ela se torne inversível. Caso esta correção seja realizada, a direção d_k encontrada no Passo 2 deixa de ser a direção de Newton, logo adota-se uma estratégia análoga a busca linear, reduzindo α_k quando necessário para que $\mu(w_k + \alpha_k d_k) \leq \mu(w_k)$ e $x, s > 0$ se verifiquem.

É importante destacar que em problemas de grande porte, o método de Broyden da forma como foi apresentado no Capítulo 2 pode ser caro computacionalmente. Levando

isso em consideração, o algoritmo implementado utiliza uma estratégia para tornar as iterações de Broyden mais baratas. Tal estratégia será discutida com mais detalhes na Seção 4.2 do Capítulo 4. O restante dos passos do Algoritmo 7 são realizados exatamente como descrito no algoritmo.

Observação 4.1. *Foi definido um limite de 5 iterações tanto para o método de Broyden executado no Passo 4 quanto para a variação do método de Broyden usada no método alternativo.*

A variação do método de Broyden que é aplicada pelo método alternativo corresponde ao Algoritmo 8.

Algoritmo 8 Variação do Método de Broyden

Dados $\sigma \in [0, 1]$, $\mu > 0$, $\tilde{w} \in \mathbb{R}^{2n+m}$, $B_0 \in \mathbb{R}^{(2n+m) \times (2n+m)}$ e i_{\max} .

Defina $i = 0$

enquanto $i < i_{\max}$ **faça**

 Considere $F_{\sigma\mu}$ definida em (1.7) e calcule a direção de Broyden \bar{s} a partir do ponto \tilde{w} por meio do seguinte sistema:

$$B_i \bar{s} = -F_{\sigma\mu}(\tilde{w}). \quad (4.1)$$

 Escolha $\beta \in [0, 1]$ tal que $\tilde{w} + \beta\bar{s}$ satisfaça $x, s > 0$ e $\bar{\mu} \leq \mu$, onde $\bar{\mu}$ é a medida de complementaridade em $\tilde{w} + \beta\bar{d}$.

 Defina $\tilde{s} = \beta\bar{s}$.

 Defina $y_i = F_{\sigma\mu}(\tilde{w} + \tilde{s}) - F_{\sigma\mu}(\tilde{w})$.

 Faça $B_{i+1} = B_i + \frac{(y_i - B_i\tilde{s})\tilde{s}^T}{\tilde{s}^T\tilde{s}}$.

 Faça $\tilde{w} = \tilde{w} + \tilde{s}$.

 Faça $i = i + 1$.

fim enquanto

Inicialmente, nossa implementação do Algoritmo 7 não possuía o método alternativo. Isso acarretou em uma infinidade de tentativas falhas de determinar um próximo iterando pelo método de Broyden em vários dos problemas testados. Foram feitos inúmeros testes e correções no código até que foi determinada a necessidade de um método alternativo. Por esse motivo, não houve tempo de se implementar o Algoritmo 8 nos moldes do Lema

4.6, que será abordado mais adiante na Seção 4.2. Sendo assim, o sistema (4.1) não é resolvido da melhor forma em nossa implementação atual.

A forma como o sistema (4.1) é resolvido atualmente é determinada pelo operador contra-barra da linguagem Julia, o qual testa a matriz B_i para tentar determinar um método eficiente de resolver o sistema. Pretendemos implementar a resolução desse sistema com base no Lema 4.6, assim como foi feito no algoritmo principal, pois essa estratégia permite reaproveitar a fatoração de $J(w_k)$ (em nosso caso, a fatoração LU) que foi calculada no passo preditor, facilitando a resolução do sistema.

Em nossa implementação do Algoritmo 8, a constante β é determinada por uma busca linear. O Algoritmo 8 foi implementado com alguns critérios de parada extras. Basicamente, o algoritmo tentará a todo custo executar i_{\max} iterações, porém será interrompido e devolverá o último iterando nos seguintes casos:

1. Se ocorrer o surgimento de entradas do tipo *Not a Number* no próximo iterando devido a erros numéricos;
2. Se a constante β encontrada for muito pequena, isto é, menor que 10^{-8} , pois nesse caso a correção já não faria muita diferença;
3. Se ocorrer algum erro ao tentar inverter B_k , possivelmente por mal condicionamento dessa matriz.

A seguir, o Algoritmo 9 expõe os passos do método alternativo de forma mais clara e objetiva.

Algoritmo 9 Preditor com Corretor Quasi-Newton Alternativo

Dado $w_0 = (x_0, \lambda_0, s_0) \in \mathcal{F}_0$ e o número máximo de iterações de Broyden i_{\max} .

Seja $k = 0$.

Passo 1: Faça $\sigma_k = 0$ em (1.22) e encontre $d_k = (\Delta x_k, \Delta \lambda_k, \Delta s_k)$.

Passo 2: Escolha o maior $\alpha_k \in [0, 1]$ tal que $w_k + \alpha_k d_k \in \mathcal{F}_0$ e $\mu_a \leq \mu_k$, onde μ_a é a medida de complementaridade em $w_k + \alpha_k d_k$.

Passo 3: Defina $\sigma_k = \left(\frac{\mu_a}{\mu_k}\right)^3$.

Passo 4: Defina $\tilde{w} = w_k + \alpha_k d_k$, $B_0 = J(w_k)$, $\sigma = \sigma_k$ e $\mu = \mu_k$.

Passo 5: Atualize \tilde{w} por meio do Algoritmo 8.

Passo 6: Faça $w_{k+1} = \tilde{w}$.

Passo 7: Faça $k = k + 1$.

Em nossa implementação do Algoritmo 9, o sistema (1.22) não chega a ser resolvido neste passo, pois ele já foi resolvido pelo algoritmo principal. Desta forma, apenas reaproveitamos a direção que já foi calculada. Além disso, α_k também é reaproveitado do algoritmo principal.

Uma vantagem de fazer o controle do tamanho de passo como no Algoritmo 8 é poder preservar a positividade estrita de x e s em cada iteração do método de Broyden, o que é um ganho com relação ao nosso algoritmo teórico (Algoritmo 7), que exige que o método de Broyden puro encontre um ponto em \mathcal{F}_0 , o que pode não ocorrer, por exemplo, se o método de Broyden convergir para um ponto que não satisfaz $(x, s) > 0$. Ao proceder como no Algoritmo 8, seremos bem menos exigentes, já que cada iterando estará em \mathcal{F}_0 .

A escolha de σ_k no método alternativo se baseia no algoritmo Mehrotra [8]. O parâmetro de centralidade σ_k foi escolhido dessa forma pois sua característica de autoajuste proporcionou um decréscimo mais rápido de μ em comparação com alguns testes com σ_k constante igual a 1, que era a ideia inicial para esse algoritmo. Observe que o Algoritmo 9 é menos exigente que o Algoritmo 7, pois não se exige a convergência do método de Broyden, apenas move-se cegamente na direção indicada por esse método. Ao ser menos exigente, estamos buscando aumentar as chances de um ponto ser aceito ao resolvermos o sistema $F_{\sigma_k \mu_k}(w) = 0$ pelo método de Broyden, para não ficarmos presos para sempre em busca de um próximo iterando, como ocorreu frequentemente em testes preliminares no Algoritmo 7 puro em problemas da NETLIB [21].

4.1 Implementação em Problemas Pequenos

Nesta seção, foram elaborados problemas simples para que o Algoritmo 7 pudesse ter seu comportamento observado. Primeiramente, vamos considerar um exemplo no qual a Jacobiana de $F_{\sigma\mu}$ seja inversível na solução e apresentar a execução de uma iteração do algoritmo de forma detalhada. Depois, a solução encontrada pelo algoritmo será apresentada.

Exemplo 4.2. *Considere o seguinte problema:*

$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.a:} \quad & 5x_1 + 7x_2 = 35 \\ & 5x_1 - 3x_2 = 10 \\ & x \geq 0. \end{aligned}$$

A solução primal-dual esperada para este problema é

$$x^* = \begin{bmatrix} 3.5 \\ 2.5 \end{bmatrix}, \quad \lambda^* = \begin{bmatrix} 0.16 \\ 0.04 \end{bmatrix} \quad e \quad s^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

A matriz Jacobiana de $F_{\sigma\mu}$ em (x^*, λ^*, s^*) é dada por

$$J(x^*, \lambda^*, s^*) = \begin{bmatrix} 0 & 0 & 5 & 5 & 1 & 0 \\ 0 & 0 & 7 & -3 & 0 & 1 \\ 5 & 7 & 0 & 0 & 0 & 0 \\ 5 & -3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.5 \end{bmatrix},$$

para quaisquer $\sigma \in [0, 1]$ e $\mu > 0$ fixados. Ao calcular o determinante desta Jacobiana, obtemos $\det(J(x^*, \lambda^*, s^*)) = 2500 \neq 0$, e portanto, $J(x^*, \lambda^*, s^*)$ é inversível.

Para a escolha do ponto inicial, foi utilizado um algoritmo que resolve o problema de anular $F_{10}(w) = (A^T\lambda + s - c, Ax - b, XSe - 10e)$ por meio do método de Newton controlando os passos para que $x, s > 0$. Este procedimento se assemelha muito ao que os métodos de pontos interiores fazem, com a diferença de que o sistema $F_{10}(w) = 0$ é resolvido de forma exata. Deste modo, encontra-se de forma aproximada um ponto no caminho central tal que $\mu = 10$. O ponto encontrado foi

$$x_0 = \begin{bmatrix} 3.4999999999023435 \\ 2.499999999941406 \end{bmatrix}, \quad \lambda_0 = \begin{bmatrix} -0.4114285715762437 \\ 0.040000000033035714 \end{bmatrix} \quad e \quad s_0 = \begin{bmatrix} 2.857142858106665 \\ 4.000000001289063 \end{bmatrix},$$

e $\mu_0 = 10.000000003041295$.

No primeiro passo do Algoritmo 7, fazemos $\sigma = 0$ e $\mu = \mu_0$ em (1.7) e calculamos a direção de Newton $d_0 = -[J(x_0, \lambda_0, s_0)]^{-1}F_{0\mu_0}(x_0, \lambda_0, s_0)$:

$$d_0 = \begin{bmatrix} 9.765\,635\,411\,440\,597 \times 10^{-11} \\ 5.859\,384\,089\,035\,302 \times 10^{-11} \\ 0.571428571590402 \\ -3.125\,000\,702\,652\,869\,6 \times 10^{-11} \\ -2.8571428581863842 \\ -4.000000001382814 \end{bmatrix}$$

No Passo 2 do Algoritmo 7, escolhemos α_0 como foi descrito no início deste capítulo, isto é,

$$\begin{aligned} \alpha_0 &= \min \left\{ 1, -0.9 \frac{2.857142858106665}{-2.8571428581863842}, -0.9 \frac{4.000000001289063}{-4.000000001382814} \right\} \\ &= \min \{1, 0.899999999748884, 0.899999999789061\} \\ &= 0.899999999748884. \end{aligned}$$

Vamos fixar $\sigma_{\max} = 1 - 10^{-4}$. Neste caso, definimos no Passo 3 $\sigma_0 = \frac{1-10^{-4}}{2} = 0.49995$. Nos passos seguintes, os valores α_0 e σ_0 serão testados e poderão sofrer reajustes para garantir a convergência do passo corretor, seguindo a ideia do resultado apresentado na Seção 3.1 do Capítulo 3.

No Passo 4, consideramos $\bar{w} = w_0 + \alpha_0 d_0$ como sendo o ponto inicial do método de Broyden, isto é, o ponto

$$\bar{x} = \begin{bmatrix} 3.49999999990234 \\ 2.499999999941402 \end{bmatrix}, \quad \bar{\lambda} = \begin{bmatrix} 0.10285714284076863 \\ 0.04000000004910705 \end{bmatrix} \quad \text{e} \quad \bar{s} = \begin{bmatrix} 0.28571428581066627 \\ 0.40000000014497683 \end{bmatrix}.$$

Neste ponto, observamos que a medida de complementaridade é 1.00000000034732. Como aproximação inicial para a Jacobiana, será escolhida a matriz $J(x_0, \lambda_0, s_0)$. Na primeira iteração do método de Broyden, será fixado $\sigma = \sigma_0 = 0.49995$ e $\mu = \mu_0$ em (1.7) e será resolvido o sistema

$$J(x_0, \lambda_0, s_0)\bar{d} = -F_{\sigma_0\mu_0}(\bar{x}, \bar{\lambda}, \bar{s}), \quad (4.2)$$

onde \bar{d} é a direção procurada no método de Broyden. Após resolver este sistema, encontramos

$$\bar{d} = \begin{bmatrix} 9.765\,841\,468\,833\,968 \times 10^{-12} \\ 5.859\,845\,941\,813\,546\,6 \times 10^{-12} \\ -0.22854285721775672 \\ -5.464\,277\,919\,031\,702 \times 10^{-12} \\ 1.1427142860770416 \\ 1.5998000004922788 \end{bmatrix}.$$

Sendo assim, o próximo iterando do método de Broyden é o ponto

$$\bar{w} + \bar{d} = \begin{bmatrix} 3.5 \\ 2.5 \\ -0.12568571437698808 \\ 0.03999999999944643 \\ 1.4284285718877079 \\ 1.9998000006372556 \end{bmatrix}.$$

Observe que a medida de complementaridade neste ponto é 4.999500001600058, ou seja, houve um retrocesso no decréscimo obtido pelo passo predictor. No entanto, o ponto obtido satisfaz

$$4.999500001600058 < \frac{0.9999 + 0.49995}{2} \cdot 10.000000003041295 = 7.499250002280743.$$

Além disso, a condição $x, s > 0$ é satisfeita e pelo sistema (4.2), temos que $\bar{d} = (\bar{d}_x, \bar{d}_\lambda, \bar{d}_s)$ satisfaz

$$\begin{cases} A\bar{d}_\lambda + \bar{d}_s = -r_c = 0 \\ A\bar{d}_x = -r_b = 0 \end{cases},$$

em que $r_c = r_b = 0$ segue do fato de o ponto inicial ser viável. Neste caso,

$$A(x + \bar{d}_x) = Ax + A\bar{d}_x = b$$

e

$$A(\lambda + \bar{d}_\lambda) + I(s + \bar{d}_s) = A\lambda + s + A\bar{d}_\lambda + \bar{d}_s = c,$$

e conseqüentemente, o ponto $\bar{w} + \bar{d}$ pertence a \mathcal{F}_0 . Sendo assim, o método de Broyden é interrompido e este ponto é escolhido como o próximo iterando do Algoritmo 7. Se não tivéssemos sido capazes de determinar tal ponto, então alguma das seguintes situações precisaria ter acontecido:

1. O método de Broyden não convergiu;
2. O método de Broyden convergiu para um ponto que não pertence a \mathcal{F}_0 (mais especificamente, $x, s > 0$ não foi satisfeito).

Independentemente do que tenha ocorrido, o Algoritmo 6 realizará uma redução no valor α_0 pela metade e redefinirá $\sigma_0 := \frac{\sigma_0 + \sigma_{\max}}{2}$, e depois repetirá o Passo 4, na tentativa de aumentar a chance de convergência do método de Broyden.

Foi descrita de forma detalhada acima a primeira iteração do Algoritmo 7 aplicada a este problema. A execução completa deste algoritmo foi realizada por meio de uma implementação computacional, e a solução encontrada por ela foi:

$$x^* = \begin{bmatrix} 3.5 \\ 2.5 \end{bmatrix}, \quad \lambda^* = \begin{bmatrix} 0.15999999946940988 \\ 0.040000000000000002 \end{bmatrix} \quad \text{e} \quad s^* = \begin{bmatrix} 2.652\,950\,441\,895\,383 \times 10^{-9} \\ 3.714\,130\,618\,653\,536\,2 \times 10^{-9} \end{bmatrix}.$$

Como critério de parada, foi adotado a exigência de que $\mu < 10^{-8}$. A medida de complementaridade no ponto encontrado é $9.285\,326\,546\,633\,84 \times 10^{-9}$. Foram executadas 30 iterações do Algoritmo 7, e em cada iteração, apenas um passo de Broyden foi necessário para satisfazer o critério de interrupção.

A seguir, apresentaremos mais alguns exemplos pequenos. Para estes problemas, não foi levado em consideração se a Jacobiana era inversível ou não na solução.

Exemplo 4.3. *Considere o seguinte problema:*

$$\begin{aligned} \min \quad & x_1 \\ \text{s.a:} \quad & x_1 + x_2 = 1 \\ & x \geq 0. \end{aligned}$$

A solução primal-dual para o problema acima é

$$x^* = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \lambda^* = 0 \quad \text{e} \quad s^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

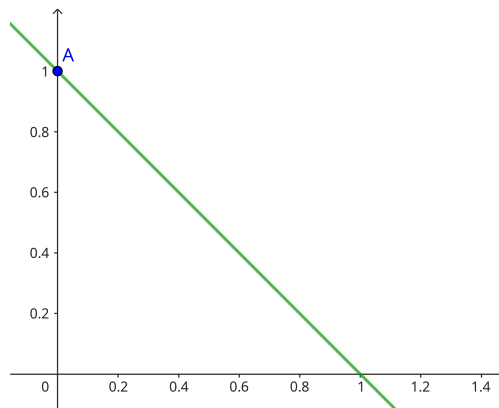


Figura 4.1: Representação gráfica da região viável do problema primal e da solução primal.

Primeiramente, escolhemos o ponto inicial em \mathcal{F}_0 como sendo o ponto (x_0, λ_0, s_0) no caminho central tal que $\mu_0 = 10$. O ponto inicial encontrado foi:

$$w_0 = \begin{bmatrix} 0.4875077655786048 \\ 0.512492235202648 \\ -19.51249192153052 \\ 20.512491922311767 \\ 19.51249192309302 \end{bmatrix}.$$

Depois, foi escolhido $\sigma_{\max} = 1 - 10^{-4}$. Ao executar o algoritmo implementado, foi obtida a seguinte solução primal-dual:

$$x^* = \begin{bmatrix} 9.285\,326\,122\,703\,666 \times 10^{-9} \\ 0.9999999907146739 \end{bmatrix}, \quad \lambda^* = -9.285\,326\,681\,434\,427 \times 10^{-9}$$

$$\text{e } s^* = \begin{bmatrix} 1.0000000092853267 \\ 9.285\,326\,681\,434\,427 \times 10^{-9} \end{bmatrix}.$$

O critério de parada adotado foi $\mu_k \leq 10^{-8}$, e foram executadas 30 iterações.

Exemplo 4.4. Considere o seguinte problema:

$$\begin{aligned} \max \quad & x_1 + x_2 \\ \text{s.a.} \quad & -x_1 + 3x_2 \leq 9 \\ & 5x_1 - 3x_2 \leq 15 \\ & x \geq 0. \end{aligned}$$

A solução primal para o problema acima é

$$x^* = \begin{bmatrix} 6 \\ 5 \end{bmatrix}.$$

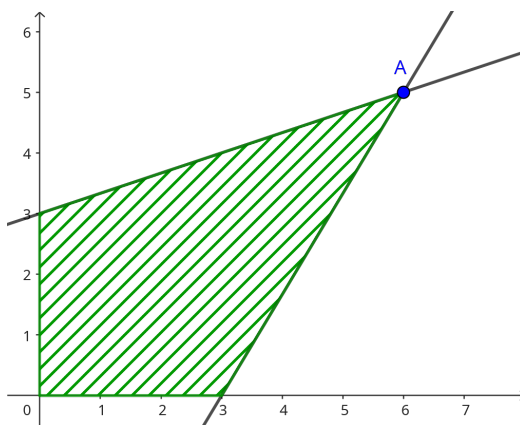


Figura 4.2: Representação gráfica da região viável do problema primal e da solução primal.

Primeiramente, é necessário reescrever o problema acima no mesmo formato de (1.1). Para isso, serão acrescentadas variáveis de folga e a função objetivo será trocada por uma equivalente para problemas de minimização. Deste modo, o problema acima ficará reescrito da seguinte forma:

$$\begin{aligned} \min \quad & -x_1 - x_2 \\ \text{s.a:} \quad & -x_1 + 3x_2 + x_3 = 9 \\ & 5x_1 - 3x_2 + x_4 = 15 \\ & x \geq 0. \end{aligned}$$

Depois disso, encontramos um ponto em \mathcal{F}_0 de modo análogo ao que foi feito no exemplo

anterior. O ponto inicial encontrado foi:

$$w_0 = \begin{bmatrix} 3.30291882661946 \\ 2.9296596213122537 \\ 3.5139399579952 \\ 7.274384721464461 \\ -2.8458086289256754 \\ -1.3746867327333967 \\ 3.0276250394288087 \\ 3.413365690139337 \\ 2.8458086304881762 \\ 1.3746867342958982 \end{bmatrix}.$$

Foi escolhido $\sigma_{\max} = 1 - 10^{-4}$. Ao executar o algoritmo implementado, foi obtida a seguinte solução primal-dual:

$$x^* = \begin{bmatrix} 5.999999989506151 \\ 4.999999991838118 \\ 1.399\ 179\ 919\ 445\ 095\ 9 \times 10^{-8} \\ 2.798\ 359\ 839\ 947\ 434\ 5 \times 10^{-8} \end{bmatrix}, \quad \lambda^* = \begin{bmatrix} -0.6666666678326499 \\ -0.33333333387745884 \end{bmatrix}$$

$$e \quad s^* = \begin{bmatrix} 1.554\ 644\ 372\ 559\ 684\ 7 \times 10^{-9} \\ 1.865\ 573\ 246\ 366\ 793\ 2 \times 10^{-9} \\ 0.6666666678326499 \\ 0.33333333387745884 \end{bmatrix}.$$

O critério de parada também foi $\mu_k \leq 10^{-8}$, e o número de iterações foi 33.

Exemplo 4.5. *Considere o seguinte problema:*

$$\begin{aligned} \min \quad & x_2 \\ \text{s.a:} \quad & x_1 + 2x_2 \leq 6 \\ & x \geq 0. \end{aligned}$$

O problema primal acima possui uma aresta ótima. Mais especificamente, qualquer ponto da forma $(t, 0)$, com $t \in [0, 6]$ soluciona o problema primal.

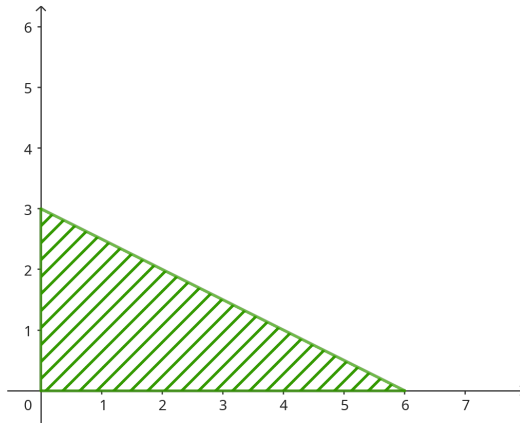


Figura 4.3: Representação gráfica da região viável do problema primal

Também será necessário escrever este problema no mesmo formato que (1.1). Para isso, uma variável de folga será acrescentada:

$$\begin{aligned} \min \quad & x_2 \\ \text{s.a:} \quad & x_1 + 2x_2 + x_3 = 6 \\ & x \geq 0. \end{aligned}$$

Após reescrever o problema, encontramos um ponto em \mathcal{F}_0 como temos feito nos exemplos anteriores. O ponto inicial encontrado foi:

$$w_0 = \begin{bmatrix} 2.0643824166964353 \\ 0.9356175827697104 \\ 2.0643824166964353 \\ -4.844063733725749 \\ 4.844063733803874 \\ 10.688127467529624 \\ 4.844063733803874 \end{bmatrix}.$$

Foi escolhido $\sigma_{\max} = 1 - 10^{-4}$. Ao executar o algoritmo implementado, foi obtida a seguinte solução primal-dual:

$$x^* = \begin{bmatrix} 2.9999999907146737 \\ 9.285\ 326\ 351\ 335\ 96 \times 10^{-9} \\ 2.9999999907146733 \end{bmatrix}, \quad \lambda^* = -3.095\ 108\ 876\ 900\ 449\ 7 \times 10^{-9} \text{ e}$$

$$s^* = \begin{bmatrix} 3.095\ 108\ 876\ 900\ 45 \times 10^{-9} \\ 1.0000000061902177 \\ 3.095\ 108\ 876\ 900\ 449\ 2 \times 10^{-9} \end{bmatrix}.$$

Este resultado foi obtido após 30 iterações, terminando com o critério de parada $\mu_k \leq 10^{-8}$. Um fato interessante sobre este problema é que a solução encontrada é o ponto médio da aresta de soluções. Dado $\tau > 0$, as condições abaixo descrevem o ponto (x, λ, s) do caminho central associado à τ :

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 6 \\ \lambda + s_1 &= 0 \\ 2\lambda + s_2 &= 1 \\ \lambda + s_3 &= 0 \\ x_i s_i &= \tau, \quad i \in \{1, 2, 3\} \\ (x, s) &> 0. \end{aligned}$$

Das condições acima, podemos deduzir que $x_i = \frac{\tau}{s_i}$, $i = 1, 2, 3$. Além disso, podemos deduzir que $s_1 = s_3 = -\lambda$. Dessa forma, temos que $x_1 = x_3 = -\frac{\tau}{\lambda}$. Como as soluções desse problema são pontos tais que $x_2 = 0$, isto significa que conforme τ se aproxima de 0, devemos ter pontos do caminho central tais que $x_1 = x_3$ e x_2 próximo de 0. Neste caso, conforme $\tau \rightarrow 0$, a equação $x_1 + 2x_2 + x_3 = 6$ se aproxima cada vez mais de $x_1 + x_3 = 6$. Como 6 é o comprimento da aresta de soluções e $x_1 = x_3$, é natural que x_1 tenda à metade desse comprimento. Por esse motivo, o caminho central dirige o método de pontos interiores ao ponto médio da aresta de soluções.

4.2 Implementação Eficiente do Método de Broyden

Nesta seção, serão apresentados alguns resultados que permitirão uma implementação com menos custo computacional do Método de Broyden. Em uma determinada iteração k do Algoritmo 7, obtemos um ponto $y_{k_0} := w_k + \alpha_k d_k$ no Passo 2. Tal ponto é tomado como ponto inicial para o Método de Broyden no Passo 4. Em cada iteração i do Método de Broyden, resolvemos o seguinte sistema:

$$B_{k_i} \bar{s}_i = -F_{\sigma_k \mu_k}(y_{k_i}), \quad (4.3)$$

onde $B_{k_0} = J_{\sigma_k \mu_k}(w_k) := J_k$ e $\bar{s}_i = y_{k_{i+1}} - y_{k_i}$. Observe que houve uma mudança de notações com relação as notações utilizadas no Capítulo 2. Isto se fez necessário, pois na teoria de pontos interiores já havia sido utilizada a letra s no problema dual. Ainda com

respeito ao método de Broyden, em cada iteração i , temos que

$$B_{k_{i+1}} = B_{k_i} + \frac{\bar{y}_i - B_{k_i} \bar{s}_i}{\bar{s}_i^T \bar{s}_i},$$

onde denotaremos $\bar{y}_i = F_{\sigma_k \mu_k}(y_{k_{i+1}}) - F_{\sigma_k \mu_k}(y_{k_i})$ para evitar confusões com a sequência $(y_{k_i})_{i \in \mathbb{N}}$ gerada pelo método de Broyden. Resolver o sistema $B_{k_i} \bar{s}_i = -F_{\sigma_k \mu_k}(y_{k_i})$ por meio da inversão da matriz B_{k_i} em geral é um processo caro computacionalmente. Sendo assim, é importante encontrar uma forma de resolver este sistema de forma mais eficiente. Vamos denotar $H_i = B_{k_i}^{-1}$, para todo $i = 0, 1, 2, \dots$. Devido a expressão de $B_{k_{i+1}}$, podemos utilizar a fórmula de Sherman – Morrison – Woodbury como em [5], obtendo a seguinte expressão para H_{i+1} :

$$H_{i+1} = H_i + \frac{(\bar{s} - H_i \bar{y}_i) \bar{s}_i^T}{\bar{s}_i^T H_i \bar{y}_i} H_i = \left(I + \frac{(\bar{s}_i - H_i \bar{y}_i) \bar{s}_i^T}{\bar{s}_i^T H_i \bar{y}_i} \right) H_i, \quad \forall i = 0, 1, 2, \dots,$$

Denotando $u_i = \bar{s}_i - H_i \bar{y}_i$ e $\rho_i = \bar{s}_i^T H_i \bar{y}_i$, podemos escrever

$$H_{i+1} = \left(I + \frac{u_i \bar{s}_i^T}{\rho_i} \right) H_i, \quad \forall i = 0, 1, 2, \dots \quad (4.4)$$

Sendo assim, segue de (4.3) que

$$\bar{s}_i = -H_i F_{\sigma_k \mu_k}(y_{k_i}).$$

Observe que H_i envolve apenas produtos de uma matriz de posto 1 e o produto com H_{i-1} . Esse processo é recursivo até que torna-se necessário o produto com H_0 . No nosso caso, $H_0 = [J(w_k)]^{-1}$, mas a inversa de $J(w_k)$ nunca é calculada. No seu lugar, a fatoração utilizada para calcular a direção de Newton é reaproveitada (em nosso caso, a fatoração LU [20]).

Para $i = 0$, y_{k_1} fica determinado da seguinte forma:

$$y_{k_1} = y_{k_0} + \bar{s}_0.$$

Para computar H_1 , que será utilizada na próxima iteração, precisamos antes calcular os termos u_0 e ρ_0 . Temos que

$$\begin{aligned} u_0 &= \bar{s}_0 - H_0 \bar{y}_0 = \bar{s}_0 - H_0 (F_{\sigma_k \mu_k}(y_{k_1}) - F_{\sigma_k \mu_k}(y_{k_0})) = \bar{s}_0 - H_0 F_{\sigma_k \mu_k}(y_{k_1}) + H_0 F_{\sigma_k \mu_k}(y_{k_0}) \\ &= \bar{s}_0 - H_0 F_{\sigma_k \mu_k}(y_{k_1}) - \bar{s}_0 = -H_0 F_{\sigma_k \mu_k}(y_{k_1}). \end{aligned}$$

Ainda,

$$\rho_0 = \bar{s}_0^T H_0 \bar{y}_0 = \bar{s}_0^T J_k^{-1} (F_{\sigma_k \mu_k}(y_{k_1}) - F_{\sigma_k \mu_k}(y_{k_0})) = \bar{s}_0^T (J_k^{-1} F_{\sigma_k \mu_k}(y_{k_1}) - J_k^{-1} F_{\sigma_k \mu_k}(y_{k_0}))$$

$$= \bar{s}_0^T(-u_0 + \bar{s}_0) = \bar{s}_0^T(\bar{s}_0 - u_0).$$

Para $i = 1$, temos que

$$\bar{s}_1 = -H_1 F_{\sigma_k \mu_k}(y_{k_1}) = -\left(I + \frac{u_0 \bar{s}_0^T}{\rho_0}\right) H_0 F_{\sigma_k \mu_k}(y_{k_1}).$$

Como y_{k_2} está determinado por \bar{s}_1 , resta encontrar u_1 e ρ_1 . Observe que o sistema $H_1 \bar{y}_1$ aparece na definição de ambos os termos que procuramos. Além disso,

$$\begin{aligned} H_1 \bar{y}_1 &= \left(I + \frac{u_0 \bar{s}_0^T}{\rho_0}\right) H_0 (F_{\sigma_k \mu_k}(y_{k_2}) - F_{\sigma_k \mu_k}(y_{k_1})) \\ &= \left(I + \frac{u_0 \bar{s}_0^T}{\rho_0}\right) H_0 F_{\sigma_k \mu_k}(y_{k_2}) - \left(I + \frac{u_0 \bar{s}_0^T}{\rho_0}\right) H_0 F_{\sigma_k \mu_k}(y_{k_1}). \\ &= \left(I + \frac{u_0 \bar{s}_0^T}{\rho_0}\right) H_0 F_{\sigma_k \mu_k}(y_{k_2}) + \bar{s}_1. \end{aligned}$$

Deste modo, temos que

$$u_1 = \bar{s}_1 - H_1 \bar{y}_1 = -\left(I + \frac{u_0 \bar{s}_0^T}{\rho_0}\right) H_0 F_{\sigma_k \mu_k}(y_{k_2}),$$

e também que

$$\rho_1 = \bar{s}_1^T H_1 \bar{y}_1 = \bar{s}_1^T \left[\left(I + \frac{u_0 \bar{s}_0^T}{\rho_0}\right) H_0 F_{\sigma_k \mu_k}(y_{k_2}) + \bar{s}_1 \right] = \bar{s}_1^T(-u_1 + \bar{s}_1) = \bar{s}_1^T(\bar{s}_1 - u_1).$$

Dessa forma, observamos o seguinte padrão:

- $u_0 = -H_0 F_{\sigma_k \mu_k}(y_{k_1})$ e $u_1 = -H_1 F_{\sigma_k \mu_k}(y_{k_2})$;
- $\rho_0 = \bar{s}_0^T(\bar{s}_0 - u_0)$ e $\rho_1 = \bar{s}_1^T(\bar{s}_1 - u_1)$.

De forma geral, para $i = 0, 1, 2, \dots$, temos que

$$u_i = -H_i F_{\sigma_k \mu_k}(y_{k_{i+1}}) \quad \text{e} \quad \rho_i = \bar{s}_i^T(\bar{s}_i - u_i).$$

Por fim, note que

$$\begin{aligned} \bar{s}_{i+1} &= -H_{i+1} F_{\sigma_k \mu_k}(y_{k_{i+1}}) = -\left(I + \frac{u_i \bar{s}_i^T}{\rho_i}\right) H_i F_{\sigma_k \mu_k}(y_{k_{i+1}}) = \left(I + \frac{u_i \bar{s}_i^T}{\rho_i}\right) (-H_i F_{\sigma_k \mu_k}(y_{k_{i+1}})) \\ &= \left(I + \frac{u_i \bar{s}_i^T}{\rho_i}\right) u_i. \end{aligned}$$

A discussão realizada acima demonstra o seguinte lema:

Lema 4.6. Na k -ésima iteração do Algoritmo 6 (ou de sua versão modificada), considere $(y_{k_i})_{i \in \mathbb{N}}$ a sequência gerada pelo Método de Broyden no Passo 4. Para todo $i = 0, 1, 2, \dots$, temos que

$$u_i = -H_i F_{\sigma_k \mu_k}(y_{k_{i+1}}), \quad \rho_i = \bar{s}_i^T (\bar{s}_i - u_i),$$

$$\bar{s}_{i+1} = \left(I + \frac{u_i \bar{s}_i^T}{\rho_i} \right) u_i \quad e$$

$$H_{i+1} = \left(I + \frac{u_i \bar{s}_i^T}{\rho_i} \right) \left(I + \frac{u_{i-1} \bar{s}_{i-1}^T}{\rho_{i-1}} \right) \dots \left(I + \frac{u_0 \bar{s}_0^T}{\rho_0} \right) H_0.$$

4.3 Implementação em Problemas da NETLIB

Para a realização de testes em problemas maiores, utilizamos o código fonte de um pacote da linguagem Julia chamado Tulip [11]. Este pacote é próprio para a resolução de problemas de programação linear por meio do método de pontos interiores, e também implementa a técnica dos múltiplos corretores de centralidade [12]. O código fonte do Tulip foi modificado para executar os passos do Algoritmo 7 exatamente como descrevemos no início deste capítulo. Assim como anteriormente, o ponto inicial é escolhido como um ponto do caminho central cujo $\mu = 10$. Esta escolha é feita exatamente como descrito na Seção 4.1 deste capítulo. Os problemas teste foram obtidos da NETLIB [21], uma popular biblioteca de problemas de programação linear. O pacote QPSReader [22] foi utilizado para fazer a leitura desses problemas. Na Tabela 4.1, constam os problemas que foram testados, e dados como o número de variáveis, o número de restrições e o número de elementos “não nulos” de A (elementos que em módulo são maiores que 10^{-8}).

Problema	Nº de Variáveis	Nº de Restrições	Nº de Elementos “não Nulos” de A
25FV47	1876	821	10705
ADLITTLE	138	56	424
AFIRO	51	27	102
AGG	615	488	2862
AGG2	758	516	4740
AGG3	758	516	4756
BANDM	472	305	2494
BEACONFD	295	173	3408
BNL1	1586	643	5532

Problema	Nº de Variáveis	Nº de Restrições	Nº de Elementos “não Nulos” de A
BNL2	4486	2324	14996
D2Q06C	5831	2171	33081
DEGEN2	757	444	4201
DEGEN3	2604	1503	25432
E226	472	223	2768
FFFFF800	1028	524	6401
ISRAEL	316	174	2443
LOTFI	366	153	1136
MAROS-R7	9408	3136	144848
QAP8	1632	912	7296
QAP12	8856	3192	38304
QAP15	22275	6330	94950
SC105	163	105	340
SC205	317	205	665
SC50A	78	50	160
SC50B	78	50	148
SCAGR25	671	471	1725
SCAGR7	185	129	465
SCFXM1	600	330	2732
SCFXM2	1200	660	5469
SCFXM3	1800	990	8206
SCORPION	466	388	1534
SCRS8	1275	490	3288
SCSD1	760	77	2388
SCSD6	1350	147	4316
SCSD8	2750	397	8584
SCTAP1	660	300	1872
SCTAP2	2500	1090	7334
SCTAP3	3340	1480	9734
SHARE1B	253	117	1179
SHARE2B	162	96	777
SHIP04L	2166	402	6380
SHIP04S	1506	402	4400

Problema	Nº de Variáveis	Nº de Restrições	Nº de Elementos “não Nulos” de A
SHIP08L	4363	778	12882
SHIP08S	2467	778	7194
SHIP12L	5533	1151	16276
SHIP12S	2869	1151	8284
STOCFOR1	165	117	501
STOCFOR2	3045	2157	9357
STOCFOR3	23541	16675	72721
TRUSS	8806	1000	27836
WOOD1P	2595	244	70216
WOODW	8418	1098	37487

Tabela 4.1: Dados dos problemas testados.

Os resultados obtidos nos testes foram agregados na Tabela 4.2. Nas colunas da Tabela 4.2, IT se refere ao número de iterações do algoritmo principal, ITB o número total de iterações de Broyden executadas, NCA o número de execuções do método alternativo, e NCJ indica quantas vezes foi necessário corrigir a Jacobiana. Na coluna motivo, serão apresentadas as razões pelas quais o programa parou. Nesta coluna, SO indica que foi encontrada a solução ótima, IT significa que o limite de iterações foi atingido e MEM significa que o problema não foi resolvido por falta de memória. Como critério de parada, os mesmos critérios do Tulip foram utilizados, os quais verificam a viabilidade primal, dual e o gap de dualidade como sugerido na literatura. Mais precisamente, o Tulip considera o problema resolvido quando

$$\rho_d = \frac{\|A^T \lambda_k + s_k - c\|}{1 + \|c\|} < \varepsilon_d,$$

$$\rho_p = \frac{\|Ax_k - b\|}{1 + \|b\|} < \varepsilon_p,$$

e

$$\rho_g = \frac{\|c^T x_k - b^T \lambda_k\|}{1 + \|b^T \lambda_k\|} < \varepsilon_g,$$

onde ε_d , ε_p e ε_g são tolerâncias dadas. Por padrão, $\varepsilon_d = \varepsilon_p = \varepsilon_g = 10^{-8}$.

Problema	IT	ITB	NCA	NCJ	ρ_d	ρ_p	ρ_g	Motivo
25FV47	100	1686	84	100	2.5536×10^{-8}	1.1278×10^{-11}	2.1807×10^{-7}	IT
ADLITTLE	12	89	3	0	1.5968×10^{-16}	3.8181×10^{-17}	1.1441×10^{-8}	SO
AFIRO	28	63	0	0	3.0278×10^{-17}	1.0888×10^{-16}	1.0677×10^{-8}	SO
AGG	4	80	4	0	1.4152×10^{-8}	4.5494×10^{-16}	2.1272×10^{-12}	SO
AGG2	4	80	4	0	2.3587×10^{-9}	4.2314×10^{-16}	1.0992×10^{-10}	SO
AGG3	3	60	3	0	4.6333×10^{-11}	2.3210×10^{-16}	1.0155×10^{-8}	SO
BANDM	100	1662	100	0	2.7743×10^{-6}	1.7142×10^{-15}	7.1219×10^{-16}	IT
BEACONFD	11	213	11	0	8.6697×10^{-9}	1.5735×10^{-11}	2.8248×10^{-9}	SO
BNL1	41	93	0	41	7.2446×10^{-13}	1.4201×10^{-16}	2.7051×10^{-11}	SO
BNL2	100	1744	100	0	0.00063973	7.3229×10^{-7}	2.3903×10^{-6}	IT
D2Q06C	100	1693	100	0	0.00030329	6.1022×10^{-9}	7.7153×10^{-14}	IT
DEGEN2	100	1635	93	19	0.075289	1.3820×10^{-9}	0.019519	IT
DEGEN3	100	1685	100	8	9.7038	4.7910×10^{-7}	1.6342	IT
E226	100	1674	100	0	2.5333×10^{-7}	8.4302×10^{-6}	8.9933×10^{-16}	IT
FFFFF800	46	905	46	0	1.3715×10^{-8}	2.5772×10^{-16}	2.0950×10^{-16}	SO
ISRAEL	81	1325	66	0	3.5219×10^{-16}	3.8085×10^{-16}	1.0881×10^{-8}	SO
LOTFI	21	235	10	0	1.1102×10^{-16}	2.9174×10^{-12}	1.0203×10^{-8}	SO
MAROS-R7	-	-	-	-	-	-	-	MEM
QAP8	33	34	0	33	1.2129×10^{-15}	6.5857×10^{-16}	9.2583×10^{-9}	SO
QAP12	-	-	-	-	-	-	-	MEM
QAP15	-	-	-	-	-	-	-	MEM
SC105	17	141	5	0	6.5737×10^{-16}	3.9592×10^{-15}	2.8449×10^{-10}	SO
SC205	12	165	7	0	7.1163×10^{-16}	9.0496×10^{-15}	9.9732×10^{-10}	SO
SC50A	23	150	5	0	9.2002×10^{-17}	6.6483×10^{-16}	5.4426×10^{-11}	SO
SC50B	26	144	3	3	1.2997×10^{-15}	5.3925×10^{-15}	1.1510×10^{-11}	SO
SCAGR25	32	433	21	0	1.1408e-15	7.9075e-16	9.9772×10^{-9}	SO
SCAGR7	16	30	0	0	1.2381×10^{-15}	2.6358×10^{-16}	1.2088×10^{-8}	SO
SCFXM1	100	1679	100	0	6.9358×10^{-7}	1.7536×10^{-15}	1.9752×10^{-16}	IT
SCFXM2	100	1699	100	0	6.9358×10^{-7}	1.8937×10^{-15}	3.0563×10^{-14}	IT
SCFXM3	100	1693	100	0	3.4679×10^{-7}	1.9285×10^{-15}	2.8864×10^{-13}	IT
SCORPION	28	36	0	28	4.7220×10^{-9}	1.8036×10^{-12}	1.1715×10^{-8}	SO
SCRS8	100	1671	100	0	6.3860×10^{-5}	4.8293×10^{-6}	2.5869×10^{-14}	IT
SCSD1	37	104	1	0	3.7007×10^{-16}	3.0090×10^{-16}	1.1569×10^{-8}	SO
SCSD6	37	93	0	0	1.1361×10^{-15}	3.3306×10^{-16}	1.1223×10^{-8}	SO
SCSD8	32	64	0	0	1.8203×10^{-15}	7.2721×10^{-16}	1.0566×10^{-8}	SO
SCTAP1	30	79	0	0	1.2452×10^{-15}	7.6481×10^{-16}	1.4635×10^{-8}	SO
SCTAP2	33	94	0	0	4.8246×10^{-16}	6.4979×10^{-16}	8.5092×10^{-9}	SO
SCTAP3	32	72	0	0	4.7150×10^{-16}	5.9682×10^{-16}	1.2239×10^{-8}	SO
SHARE1B	100	1948	100	0	0.098878	0.0064475	0.23764	IT
SHARE2B	28	39	0	0	1.3412×10^{-14}	2.8421×10^{-14}	1.4441×10^{-8}	SO
SHIP04L	20	20	0	20	1.2945×10^{-10}	1.4791×10^{-12}	1.0426×10^{-8}	SO
SHIP04S	20	20	0	20	6.4726×10^{-11}	1.1190×10^{-11}	7.4918×10^{-9}	SO
SHIP08L	21	23	0	21	2.9390×10^{-10}	1.8043×10^{-15}	8.0636×10^{-9}	SO
SHIP08S	20	20	0	20	1.1756×10^{-10}	7.6423×10^{-14}	9.4850×10^{-9}	SO
SHIP12L	21	21	0	21	1.6939×10^{-10}	1.7355×10^{-12}	1.4448×10^{-8}	SO
SHIP12S	21	21	0	21	8.4695×10^{-11}	8.1417×10^{-13}	8.0160×10^{-9}	SO
STOCFOR1	22	42	0	0	4.8373×10^{-16}	1.4437×10^{-14}	9.5428×10^{-9}	SO
STOCFOR2	29	79	0	0	4.0628×10^{-15}	2.5265×10^{-14}	8.5589×10^{-9}	SO

Problema	IT	ITB	NCA	NCJ	ρ_d	ρ_p	ρ_g	Motivo
STOCFOR3	-	-	-	-	-	-	-	MEM
TRUSS	-	-	-	-	-	-	-	MEM
WOOD1P	100	1601	100	0	4967.6649	1.8889×10^{-6}	3349.9197	IT
WOODW	-	-	-	-	-	-	-	MEM

Tabela 4.2: Resultados dos experimentos numéricos em problemas da NETLIB para o Algoritmo 7.

Para comparação, o código original do Tulip foi modificado para utilizar o mesmo ponto inicial calculado por nosso algoritmo. O Tulip oferece duas opções de algoritmos para a resolução de problemas de programação linear: o Autodual Homogêneo (HSD) e Preditor-Corretor de Mehrotra [8] (MPC) com os corretores múltiplos de centralidade [9]. O método escolhido para a comparação foi o MPC, e os resultados obtidos foram agregados na Tabela 4.3. Nesta tabela, a coluna IT representa o número de iterações e NC o número de correções aplicadas (soma do número de correções feitas pelo algoritmo Mehrotra e as correções múltiplas de centralidade). As legendas com relação ao motivo de parada seguem o mesmo padrão da Tabela 4.2.

Problema	IT	NC	ρ_d	ρ_p	ρ_g	Motivo
25FV47	21	27	9.7124×10^{-12}	1.7792×10^{-12}	1.5083×10^{-9}	SO
ADLITTLE	7	10	3.9314×10^{-11}	4.0492×10^{-11}	1.7551×10^{-9}	SO
AFIRO	10	12	3.6684×10^{-13}	1.1345×10^{-16}	3.9300×10^{-10}	SO
AGG	11	13	9.4348×10^{-9}	2.0490×10^{-15}	4.3019×10^{-9}	SO
AGG2	17	17	1.1793×10^{-9}	2.8264×10^{-9}	1.4077×10^{-12}	SO
AGG3	11	11	4.7174×10^{-9}	1.2053×10^{-12}	2.9422×10^{-11}	SO
BANDM	100	103	4.3348×10^{-8}	2.1428×10^{-15}	5.3414×10^{-16}	IT
BEACONFD	11	11	8.6697×10^{-9}	1.5735×10^{-11}	3.3181×10^{-13}	SO
BNL1	100	108	7.5991×10^{-8}	1.4201×10^{-16}	1.3789×10^{-15}	IT
BNL2	100	121	4.5770×10^{-7}	1.5891×10^{-10}	0.11438	IT
D2Q06C	100	108	7.1198×10^{-6}	7.6382×10^{-15}	2.4973×10^{-6}	IT
DEGEN2	14	18	4.9262×10^{-12}	2.8410×10^{-15}	1.8336×10^{-11}	SO
DEGEN3	19	23	3.5413×10^{-10}	4.1407×10^{-14}	2.8415×10^{-10}	SO
E226	66	70	7.9165×10^{-9}	2.0581×10^{-9}	1.7986×10^{-16}	SO

Problema	IT	NC	ρ_d	ρ_p	ρ_g	Motivo
FFFFF800	100	107	1.7073×10^{-5}	1.8040×10^{-15}	1.5049×10^{-8}	IT
ISRAEL	33	36	2.6505×10^{-13}	6.3577×10^{-16}	1.6809×10^{-10}	SO
LOTFI	23	33	7.4579×10^{-14}	2.6814×10^{-12}	7.0498×10^{-11}	SO
MAROS-R7	13	17	1.6449×10^{-11}	4.4045×10^{-16}	8.0866×10^{-12}	SO
QAP8	19	22	9.5121×10^{-16}	3.3076×10^{-15}	1.7319×10^{-11}	SO
QAP12	50	58	1.8533×10^{-13}	1.3577×10^{-12}	1.3425×10^{-10}	SO
QAP15	38	43	4.4865×10^{-15}	6.9048×10^{-12}	2.6683×10^{-11}	SO
SC105	22	23	5.2352×10^{-13}	2.9174×10^{-16}	6.3625×10^{-11}	SO
SC205	26	26	8.3943×10^{-14}	1.5966×10^{-15}	4.6905×10^{-11}	SO
SC50A	15	16	5.7223×10^{-16}	3.3241×10^{-16}	2.2104×10^{-14}	SO
SC50B	17	20	5.3306×10^{-17}	1.9384×10^{-16}	1.3009×10^{-14}	SO
SCAGR25	26	27	6.7462×10^{-10}	1.8954×10^{-14}	7.3981×10^{-10}	SO
SCAGR7	19	20	6.7749×10^{-12}	7.9976×10^{-13}	1.9318×10^{-10}	SO
SCFXM1	100	102	6.0079×10^8	8.4723×10^{-9}	4.7625×10^{-11}	IT
SCFXM2	100	105	8.6697×10^{-8}	8.4723×10^{-9}	8.8876×10^{-12}	IT
SCFXM3	100	103	8.6697×10^{-8}	1.6547×10^{-9}	5.5925×10^{-11}	IT
SCORPION	57	58	2.3610×10^{-9}	7.7721×10^{-16}	1.5732×10^{-12}	SO
SCRS8	19	29	8.9861×10^{-11}	1.0779×10^{-8}	8.5609×10^{-11}	SO
SCSD1	12	15	3.1618×10^{-16}	2.8101×10^{-15}	1.2060×10^{-11}	SO
SCSD6	13	13	1.2645×10^{-13}	2.7273×10^{-12}	4.9278×10^{-10}	SO
SCSD8	9	13	1.2262×10^{-13}	5.2939×10^{-15}	8.0362×10^{-12}	SO
SCTAP1	13	21	9.8096×10^{-16}	4.8097×10^{-14}	9.0169×10^{-12}	SO
SCTAP2	12	22	1.1154×10^{-13}	4.7773×10^{-13}	4.2245×10^{-9}	SO
SCTAP3	14	26	6.1404×10^{-16}	1.7110×10^{-15}	9.0799×10^{-12}	SO
SHARE1B	55	63	3.4802×10^{-10}	8.0550×10^{-13}	2.4640×10^{-10}	SO
SHARE2B	19	23	2.6748×10^{-12}	1.7449×10^{-13}	9.1492×10^{-11}	SO
SHIP04L	26	29	1.2945×10^{-10}	1.6776×10^{-10}	4.9923×10^{-9}	SO
SHIP04S	22	23	1.2945×10^{-10}	2.2408×10^{-10}	4.2528×10^{-10}	SO
SHIP08L	23	25	1.1756×10^{-10}	2.7781×10^{-12}	4.3259×10^{-11}	SO
SHIP08S	18	21	1.7634×10^{-10}	2.1744×10^{-13}	6.4493×10^{-12}	SO
SHIP12L	17	20	1.6939×10^{-10}	6.1025×10^{-10}	9.5517×10^{-9}	SO

Problema	IT	NC	ρ_d	ρ_p	ρ_g	Motivo
SHIP12S	15	16	2.5408×10^{-10}	2.4934×10^{-12}	3.2130×10^{-11}	SO
STOCFOR1	11	11	1.3695×10^{-13}	3.6867×10^{-12}	2.1619×10^{-10}	SO
STOCFOR2	16	22	6.7913×10^{-11}	2.6137×10^{-11}	8.6794×10^{-9}	SO
STOCFOR3	100	112	NaN	NaN	NaN	IT
TRUSS	14	20	2.2614×10^{-10}	1.8150×10^{-11}	5.6224×10^{-10}	SO
WOOD1P	100	103	0.00048828	9.2383×10^{-11}	5.1817×10^{-7}	IT
WOODW	100	105	0.00048828	1.6057×10^{-14}	1.9270×10^{-16}	IT

Tabela 4.3: Resultados dos experimentos numéricos em problemas da NETLIB para o Tulip (MPC).

Como mostra a Tabela 4.2, dos 52 problemas considerados, a implementação do Algoritmo 7 foi capaz de resolver 33, o que corresponde a aproximadamente 63% dos problemas. O número máximo de iterações foi atingido em 13 problemas, e faltou memória para resolver 6 dos problemas testados. O Tulip, por sua vez, foi capaz de resolver 41 problemas, isto é, cerca de 78% dos problemas testados, e o limite de iterações foi atingido em 11 dos problemas testados. O Tulip também não apresentou falta de memória em nenhum problema.

O número de iterações realizadas pelo algoritmo implementado normalmente fica entre 10 e 40, ainda que ocorram casos onde este número é muito abaixo de 10 ou muito acima de 40. Com relação ao Tulip, observamos o número de iterações entre aproximadamente 10 e 30, ocorrendo também alguns casos fora desta faixa.

O número de iterações de Broyden em nossa implementação em geral é bem alto. Pode-se observar isso nos problemas ADLITTLE, AGG, AGG2, AGG3, BEACONFD, FFFFF800, ISRAEL, LOTFI, SC105, SC205, SC50A e SCAGR25. Isso ocorre pois são realizadas até 3 tentativas de definir o próximo iterando no Passo 4, onde podem ser executadas até 5 iterações por tentativa. Além disso, as iterações de Broyden do método alternativo, caso seja executado, também podem variar de 1 à 5. Em alguns dos problemas, foi necessário aplicar o método alternativo em todas as iterações para ser capaz de resolvê-los. Com relação aos passos corretores do Tulip, foi observado que não é dado muito mais

do que 1 passo corretor por iteração, e por esse motivo o número de correções na maioria dos casos é muito próximo do número de iterações.

Em alguns dos problemas resolvidos por nossa implementação, foi necessário corrigir a Jacobiana diversas vezes, sendo QAP8 e SCORPION exemplos onde a correção precisou ser realizada em todas as iterações, ressaltando a importância dessa correção.

Dentre os problemas em que o limite de iterações foi atingido, nossa implementação chegou mais perto de atingir o critério de parada nos seguintes problemas: 25FV47, BANDM, E226, SCFXM1, SCFXM2 e SCFXM3. Com relação ao Tulip, os problemas BANDM, BNL1, D2Q06C, SCFXM2 e SCFXM3 foram os que mais se aproximaram de atingir o critério de parada.

Foi observado nos testes numéricos um comportamento do algoritmo principal que torna a convergência mais lenta. Este comportamento pode ser observado de forma mais clara no Exemplo 4.2. O comportamento ruim consiste basicamente de um decréscimo de μ bom alcançado pelo passo preditor seguido de um recuo (acrécimo) de μ após o passo corretor. Isso ocorre pois estamos mirando em $\sigma_k \mu_k$ no passo corretor, e se o μ alcançado pelo passo preditor for menor que este valor, não podemos aproveitar o decréscimo obtido no passo preditor. Para contornar este problema, a implementação computacional do Algoritmo 7 foi alterada para que o σ_k escolhido no Passo 2 tivesse a seguinte fórmula:

$$\sigma_k = \left(\frac{\mu_a}{\mu_k} \right)^3 \frac{\mu_a}{\mu_k}.$$

Esta escolha de σ_k implica que estamos buscando reduzir μ para

$$\sigma_k \mu_k = \left(\frac{\mu_a}{\mu_k} \right)^3 \frac{\mu_a}{\mu_k} \mu_k = \left(\frac{\mu_a}{\mu_k} \right)^3 \mu_a,$$

ou seja, estamos tentando simular o decréscimo de μ observado no algoritmo preditor-corretor de Mehrotra [8]. Além dessa mudança, reduzimos o número de iterações de Broyden para 1 tanto no algoritmo principal quanto no alternativo, com base no Lema 3.13. Com essas mudanças, foi possível resolver o Exemplo 4.2 em 3 iterações, encontrando $\mu = 9.999\,999\,517\,289\,103 \times 10^{-12}$, ou seja, houve uma redução muito grande no número de iterações com relação a implementação anterior. Os resultados dos testes em problemas da NETLIB obtidos após essas alterações constam na Tabela 4.4.

Problema	IT	ITB	NCA	NCJ	ρ_d	ρ_p	ρ_g	Motivo
25FV47	100	169	69	100	4.9504×10^{-6}	2.9230×10^{-12}	0.58522	IT
ADLITTLE	7	14	7	0	3.6004×10^{-11}	1.2682×10^{-16}	1.1719×10^{-8}	SO
AFIRO	7	13	6	0	1.6148×10^{-16}	1.2428×10^{-16}	1.1867×10^{-8}	SO
AGG	5	10	5	0	9.4348×10^{-9}	7.5823×10^{-17}	3.1150×10^{-9}	SO
AGG2	6	12	6	0	2.0274×10^{-13}	6.2365×10^{-17}	9.7459×10^{-9}	SO
AGG3	6	12	6	0	4.7174×10^{-9}	1.2473×10^{-16}	3.4706×10^{-9}	SO
BANDM	100	200	100	0	1.3871×10^{-6}	1.7678×10^{-15}	7.1219×10^{-16}	IT
BEACONFD	12	24	12	0	8.6697×10^{-9}	1.5735×10^{-11}	1.0190×10^{-9}	SO
BNL1	41	80	39	41	5.5193×10^{-14}	5.9861×10^{-16}	1.8635×10^{-12}	SO
BNL2	100	200	100	0	7.9967×10^{-6}	1.6273×10^{-7}	2.7928×10^{-7}	IT
D2Q06C	100	200	100	0	5.6959×10^{-5}	3.8139×10^{-10}	3.5554×10^{-16}	IT
DEGEN2	100	200	100	24	1349.5880	8.8406×10^{-8}	9.0911	IT
DEGEN3	100	200	100	16	17.7945	4.5670×10^{-7}	4.2674	IT
E226	100	200	100	0	6.3332×10^{-8}	1.3172×10^{-7}	7.9141×10^{-14}	IT
FFFFF800	100	200	100	0	0.0060419	9.0141×10^{-8}	5.0878×10^{-7}	IT
ISRAEL	100	200	100	0	0.046959	0.020743	0.24648	IT
LOTFI	24	48	24	0	1.1102×10^{-16}	4.1512×10^{-12}	3.9280×10^{-10}	SO
MAROS-R7	-	-	-	-	-	-	-	MEM
QAP8	14	27	13	14	9.3365×10^{-16}	2.8730×10^{-13}	9.4983×10^{-10}	SO
QAP12	-	-	-	-	-	-	-	MEM
QAP15	-	-	-	-	-	-	-	MEM
SC105	13	25	12	0	4.9913×10^{-17}	5.6560×10^{-16}	1.3489×10^{-10}	SO
SC205	17	34	17	0	3.7777×10^{-17}	2.2624×10^{-15}	9.6628×10^{-9}	SO
SC50A	14	28	14	0	2.7171×10^{-17}	2.7682×10^{-16}	3.0852×10^{-9}	SO
SC50B	11	22	11	0	5.1878×10^{-17}	1.8884×10^{-16}	4.0194×10^{-9}	SO
SCAGR25	49	98	49	0	1.3717×10^{-15}	1.0543×10^{-15}	9.6253×10^{-10}	SO
SCAGR7	8	16	8	0	9.1986×10^{-16}	2.6358×10^{-16}	4.4944×10^{-9}	SO
SCFXM1	100	200	100	0	2.6009×10^{-7}	1.5149×10^{-15}	2.9628×10^{-15}	IT
SCFXM2	100	200	100	0	6.9358×10^{-7}	2.5249×10^{-15}	1.1114×10^{-14}	IT
SCFXM3	100	200	100	0	5.2018×10^{-7}	1.9568×10^{-15}	1.3120×10^{-14}	IT
SCORPION	23	46	23	23	4.7220×10^{-9}	5.5229×10^{-12}	1.8785×10^{-12}	SO
SCRS8	100	200	100	0	1.0783×10^{-9}	1.0779×10^{-7}	1.5629×10^{-12}	IT
SCSD1	14	27	13	0	2.9605×10^{-16}	3.9230×10^{-17}	1.9343×10^{-9}	SO
SCSD6	20	40	20	0	5.6808×10^{-16}	3.3306×10^{-16}	2.1887×10^{-9}	SO
SCSD8	14	27	13	0	1.8203×10^{-15}	1.7763×10^{-15}	1.6321×10^{-9}	SO
SCTAP1	23	44	21	0	8.7721×10^{-16}	7.1376×10^{-16}	1.0118×10^{-10}	SO
SCTAP2	22	42	20	0	4.3860×10^{-16}	4.8315×10^{-16}	1.0322×10^{-9}	SO
SCTAP3	23	45	22	0	3.8378×10^{-16}	3.6687×10^{-16}	4.8991×10^{-9}	SO
SHARE1B	100	200	100	0	0.43592	0.028425	0.53769	IT
SHARE2B	15	28	13	0	1.0362×10^{-14}	1.9055×10^{-14}	1.0679×10^{-11}	SO
SHIP04L	11	21	10	11	6.4726×10^{-11}	6.2384×10^{-13}	8.3663×10^{-12}	SO
SHIP04S	10	19	9	10	6.4726×10^{-11}	5.0908×10^{-10}	3.2161×10^{-9}	SO
SHIP08L	13	25	12	13	5.8781×10^{-11}	4.0833×10^{-13}	1.5739×10^{-12}	SO
SHIP08S	10	18	8	10	5.8781×10^{-11}	2.2773×10^{-13}	1.2729×10^{-11}	SO
SHIP12L	15	30	15	15	1.6939×10^{-10}	7.9047×10^{-12}	2.0333×10^{-9}	SO
SHIP12S	13	25	12	13	1.6939×10^{-10}	4.0321×10^{-13}	4.7454×10^{-12}	SO
STOCFOR1	8	15	7	0	4.2186×10^{-16}	1.4437×10^{-14}	1.5246×10^{-10}	SO
STOCFOR2	24	47	23	0	5.2451×10^{-15}	3.2484×10^{-14}	4.4747×10^{-12}	SO

Problema	IT	ITB	NCA	NCJ	ρ_d	ρ_p	ρ_g	Motivo
STOCFOR3	-	-	-	-	-	-	-	MEM
TRUSS	-	-	-	-	-	-	-	MEM
WOOD1P	100	200	100	0	0.0034179	6.2545×10^{-12}	0.093737	IT
WOODW	-	-	-	-	-	-	-	MEM

Tabela 4.4: Resultados dos experimentos numéricos em problemas da NETLIB para o Algoritmo 7 alternativo.

Podemos perceber na Tabela 4.4, que houve uma redução no número de iterações do algoritmo em comparação à Tabela 4.2. A mudança proposta foi capaz de reduzir o número de iterações na maioria dos problemas resolvidos, havendo algumas exceções onde este número aumentou. No geral, o número de iterações passou a ficar entre 10 e 30 na maioria dos problemas resolvidos. Como foi reduzido o número máximo de iterações de Broyden para 1 em ambos os corretores principal e alternativo, naturalmente observou-se um decréscimo no número de iterações de Broyden em comparação ao observado na Tabela 4.2, porém, como ainda são executadas uma iteração de Broyden do algoritmo principal e às vezes uma do alternativo, este número ainda é alto.

Apesar de uma redução geral no número de iterações ter ocorrido, observou-se que foram resolvidos 31 problemas, cerca de 60% dos problemas testados, isto é, houve uma redução de aproximadamente 3% com relação ao observado na Tabela 4.2. Além disso, o limite de iterações foi atingido em 15 dos problemas testados, aumentando em 2 com relação aos resultados vistos na Tabela 4.2. O número de problemas não resolvidos por falta de memória foi o mesmo.

Foi observado um aumento drástico no número de problemas em que a correção alternativa foi aplicada, ocorrendo em quase todas as iterações dos problemas resolvidos. Os problemas para os quais o critério de parada chegou mais próximo de ser atingido foram: BANDM, BNL2, E226, SCFXM1, SCFXM2, SCFXM3, e SCRS8.

Por fim, como foi observado que o método alternativo obteve mais sucesso que o método principal, foram realizados testes utilizando apenas o método alternativo (Algoritmo 9), também executando apenas uma iteração de Broyden, e os resultados foram agregados na Tabela 4.5.

Problema	IT	ITB	NCJ	ρ_d	ρ_p	ρ_g	Motivo
25FV47	100	100	100	4.9504×10^{-6}	2.9231×10^{-12}	0.58438	IT
ADLITTLE	7	7	0	3.6004×10^{-11}	1.2682×10^{-16}	1.1719×10^{-8}	SO
AFIRO	7	7	0	1.6148×10^{-16}	1.2428×10^{-16}	1.1867×10^{-8}	SO
AGG	5	5	5	9.4348×10^{-9}	7.5823×10^{-17}	3.1150×10^{-9}	SO
AGG2	6	6	0	2.0274×10^{-13}	6.2365×10^{-17}	9.7459×10^{-9}	SO
AGG3	6	6	0	4.7174×10^{-9}	1.2473×10^{-16}	3.4706×10^{-9}	SO
BANDM	100	100	0	1.3871×10^{-6}	1.7678×10^{-15}	7.1219×10^{-16}	IT
BEACONFD	12	12	0	8.6697×10^{-9}	1.5735×10^{-11}	1.0190×10^{-9}	SO
BNL1	41	41	41	5.1747×10^{-14}	5.9863×10^{-16}	1.8631×10^{-12}	SO
BNL2	100	100	0	7.9967×10^{-6}	1.6273×10^{-7}	2.7928×10^{-7}	IT
D2Q06C	100	100	0	5.6959×10^{-5}	3.8139×10^{-10}	3.5554×10^{-16}	IT
DEGEN2	100	100	24	1349.5880	8.8406×10^{-8}	9.0911	IT
DEGEN3	100	100	16	17.7945	4.5670×10^{-7}	4.2674	IT
E226	100	100	0	6.3332×10^{-8}	1.3172×10^{-7}	7.9141×10^{-14}	IT
FFFFFF800	100	100	0	0.0060419	9.0141×10^{-8}	5.0878×10^{-7}	IT
ISRAEL	100	100	0	0.046959	0.020743	0.24648	IT
LOTFI	24	24	0	1.1102×10^{-16}	4.1512×10^{-12}	3.9280×10^{-10}	SO
MAROS-R7	-	-	-	-	-	-	MEM
QAP8	14	14	14	9.3365×10^{-16}	2.8730×10^{-13}	9.4983×10^{-10}	SO
QAP12	-	-	-	-	-	-	MEM
QAP15	-	-	-	-	-	-	MEM
SC105	13	13	0	4.9913×10^{-17}	5.6560×10^{-16}	1.3489×10^{-10}	SO
SC205	17	17	0	3.7777×10^{-17}	2.2624×10^{-15}	9.6628×10^{-9}	SO
SC50A	14	14	0	2.7171×10^{-17}	2.7682×10^{-16}	3.0852×10^{-9}	SO
SC50B	11	11	0	5.1878×10^{-17}	1.8884×10^{-16}	4.0194×10^{-9}	SO
SCAGR25	49	49	0	1.3717×10^{-15}	1.0543×10^{-15}	9.6253×10^{-10}	SO
SCAGR7	8	8	0	9.1986×10^{-16}	2.6358×10^{-16}	4.4944×10^{-9}	SO
SCFXM1	100	100	0	2.6009×10^{-7}	1.5149×10^{-15}	2.9628×10^{-15}	IT
SCFXM2	100	100	0	6.9358×10^{-7}	2.5249×10^{-15}	1.1114×10^{-14}	IT
SCFXM3	100	100	0	5.2018×10^{-7}	1.9568×10^{-15}	1.3120×10^{-14}	IT
SCORPION	23	23	23	4.7220×10^{-9}	5.5229×10^{-12}	1.8785×10^{-12}	SO
SCRS8	100	100	0	1.0783×10^{-9}	1.0779×10^{-7}	1.5629×10^{-12}	IT
SCSD1	14	14	0	2.9605×10^{-16}	3.9230×10^{-17}	1.9343×10^{-9}	SO
SCSD6	20	20	0	5.6808×10^{-16}	3.3306×10^{-16}	2.1887×10^{-9}	SO
SCSD8	14	14	0	1.8203×10^{-15}	1.7763×10^{-15}	1.6321×10^{-9}	SO
SCTAP1	23	23	0	8.7721×10^{-16}	7.1376×10^{-16}	1.0118×10^{-10}	SO
SCTAP2	22	22	0	4.3860×10^{-16}	4.8315×10^{-16}	1.0322×10^{-9}	SO

Problema	IT	ITB	NCJ	ρ_d	ρ_p	ρ_g	Motivo
SCTAP3	23	23	0	3.5088×10^{-16}	5.7970×10^{-16}	4.8991×10^{-9}	SO
SHARE1B	100	100	0	0.43592	0.028425	0.53769	IT
SHARE2B	15	15	0	9.5294×10^{-15}	1.7925×10^{-14}	1.0682×10^{-11}	SO
SHIP04L	11	11	11	6.4726×10^{-11}	6.2384×10^{-13}	8.3663×10^{-12}	SO
SHIP04S	10	10	10	6.4726×10^{-11}	5.0908×10^{-10}	3.2161×10^{-9}	SO
SHIP08L	13	13	13	5.8781×10^{-11}	4.0833×10^{-13}	1.5739×10^{-12}	SO
SHIP08S	10	10	10	5.8781×10^{-11}	2.2773×10^{-13}	1.2729×10^{-11}	SO
SHIP12L	15	15	15	1.6939×10^{-10}	7.9047×10^{-12}	2.0333×10^{-9}	SO
SHIP12S	13	13	13	1.6939×10^{-10}	4.0321×10^{-13}	4.7454×10^{-12}	SO
STOCFOR1	8	8	0	4.2186×10^{-16}	1.4437×10^{-14}	1.5246×10^{-10}	SO
STOCFOR2	24	24	0	5.2451×10^{-15}	3.2484×10^{-14}	4.4747×10^{-12}	SO
STOCFOR3	-	-	-	-	-	-	MEM
TRUSS	-	-	-	-	-	-	MEM
WOOD1P	100	100	0	4967.6649	1.8889e-6	3349.9197	IT
WOODW	-	-	-	-	-	-	MEM

Tabela 4.5: Resultados dos experimentos numéricos em problemas da NETLIB para o Algoritmo 9.

Na Tabela 4.5, podemos perceber que foram resolvidos 31 problemas, cerca de 60% dos problemas testados. Além disso, o limite de iterações foi atingido em 15 dos problemas, e 6 problemas não foram resolvidos por falta de memória. Ou seja, não houve mudança com relação ao Algoritmo 7 Alternativo nesses aspectos. O número de iterações também se manteve entre 10 e 30, sendo o maior diferencial entre os dois métodos o número de iterações de Broyden executadas. Os problemas em que o critério de parada mais se aproximou de ser satisfeito foram BANDM, BNL2, E226, SCFX1, SCFX2, SCFX3 e SCRS8.

Na Tabela 4.6, realizamos uma comparação lado a lado do número de iterações gasto por cada algoritmo para resolver os problemas selecionados. Nesta tabela, os menores valores obtidos para cada problema foram marcados com um asterisco.

Problema	Algoritmo 7	Tulip	Algoritmo 7 alternativo	Algoritmo 9
25FV47	100	21*	100	100

Problema	Algoritmo 7	Tulip	Algoritmo 7 alternativo	Algoritmo 9
ADLITTLE	12	7*	7*	7*
AFIRO	28	10	7*	7*
AGG	4*	11	5	5
AGG2	4*	17	6	6
AGG3	3*	11	6	6
BANDM	100*	100*	100*	100*
BEACONFD	11*	11*	12	12
BNL1	41*	100	41*	41*
BNL2	100*	100*	100*	100*
D2Q06C	100*	100*	100*	100*
DEGEN2	100	14*	100	100
DEGEN3	100	19*	100	100
E226	100	66*	100	100
FFFFFF800	46*	100	100	100
ISRAEL	81	33*	100	100
LOTFI	21*	23	24	24
MAROS-R7	-	13*	-	-
QAP8	33	19	14*	14*
QAP12	-	50*	-	-
QAP15	-	38*	-	-
SC105	17	22	13*	13*
SC205	12*	26	17	17
SC50A	23	15	14*	14*
SC50B	26	17	11*	11*
SCAGR25	32	26*	49	49
SCAGR7	16	19	8*	8*
SCFXM1	100*	100*	100*	100*
SCFXM2	100*	100*	100*	100*
SCFXM3	100*	100*	100*	100*
SCORPION	28	57	23*	23*
SCRS8	100	19*	100	100
SCSD1	37	12*	14	14
SCSD6	37	13*	20	20

Problema	Algoritmo 7	Tulip	Algoritmo 7 alternativo	Algoritmo 9
SCSD8	32	9*	14	14
SCTAP1	30	13*	23	23
SCTAP2	33	12*	22	22
SCTAP3	32	14*	23	23
SHARE1B	100	55*	100	100
SHARE2B	28	19	15*	15*
SHIP04L	20	26	11*	11*
SHIP04S	20	22	10*	10*
SHIP08L	21	23	13*	13*
SHIP08S	20	18	10*	10*
SHIP12L	21	17	15*	15*
SHIP12S	21	15	13*	13*
STOCFOR1	22	11	8*	8*
STOCFOR2	29	16*	24	24
STOCFOR3	-	100*	-	-
TRUSS	-	14*	-	-
WOOD1P	100*	100*	100*	100*
WOODW	-	100*	-	-

Tabela 4.6: Comparação lado a lado do número de iterações de cada algoritmo.

Como evidencia a Tabela 4.6, o Algoritmo 7 obteve o menor número de iterações 15 vezes, enquanto que o Tulip gastou o menor número de iterações 30 vezes. Ambos os Algoritmos 7 alternativo e 9 obtiveram o mesmo número de iterações em todos os problemas, e proporcionaram o menor número de iterações 24 vezes.

CONCLUSÕES

Neste trabalho, foram estudados os métodos de pontos interiores primais-duais em problemas de programação linear. Foram estudados os métodos seguidores de caminho, contudo o foco do trabalho foram os métodos preditores-corretores. Os métodos preditores-corretores clássicos utilizam o método de Newton para calcular seus passos preditor e corretor. Nossa proposta foi alterar o passo corretor para que o cálculo das direções fosse realizado a partir de um método Quasi-Newton, mais especificamente, o método de Broyden. Foram encontrados alguns estudos que utilizaram abordagens Quasi-Newton em métodos de pontos interiores, no entanto não foi encontrado um estudo envolvendo métodos preditores-corretores.

Os estudos encontrados tiveram papel fundamental na motivação de nossa proposta. O sucesso de abordagens Quasi-Newton em trabalhos envolvendo métodos de pontos interiores e também um trabalho sugerindo a possibilidade do método de Broyden atuar como um corretor foram algumas das motivações que levaram a nossa proposta. Deste modo, o método de Broyden foi estudado, assim como sua teoria de convergência local. Em seguida, foi elaborado um algoritmo preditor-corretor no qual o passo preditor realiza um passo de Newton, enquanto que o passo corretor executa algumas iterações do método de Broyden. Estudar a teoria dos métodos de pontos interiores e a teoria de convergência local do método de Broyden foi essencial para que pudéssemos desenvolver uma teoria de convergência local para o algoritmo elaborado. Uma implementação deste algoritmo foi realizada na linguagem Julia, e testes foram feitos em problemas pequenos e depois em problemas maiores provenientes da biblioteca NETLIB.

Os experimentos em problemas pequenos mostraram excelentes resultados, onde o algoritmo encontrou a solução esperada em todos os casos. Já nos problemas maiores, enfrentou-se algumas dificuldades, como por exemplo, a não convergência do método de

Broyden, ou a presença de Jacobianas não inversíveis por serem mal condicionadas. Com essas adversidades observadas na prática, percebeu-se a necessidade de realizar algumas alterações no algoritmo para poder contornar esses problemas. Entre essas alterações, foi proposto um método preditor-corretor Quasi-Newton alternativo para ser utilizado em caso de não convergência do método de Broyden. Também foi realizada uma correção para melhorar o condicionamento da Jacobiana e tornar possível a sua inversão.

Após essas modificações, a implementação foi capaz de resolver a maioria dos problemas testados. Contudo, foi observado um alto número de iterações de Broyden, que ocorreu por conta dos testes que acontecem após cada tentativa falha do Passo 4 do Algoritmo 7. Fazendo mais uma modificação no algoritmo foi possível reduzir consideravelmente o número de iterações de Broyden e reduzir ligeiramente o número de iterações do algoritmo. No entanto, o número de problemas resolvidos com essa modificação foi um pouco menor. Observou-se que o método preditor-corretor Quasi-Newton alternativo foi executado com sucesso muitas vezes, o que nos motivou a testá-lo de forma isolada.

Ao testar o método alternativo de forma isolada, não foram observadas mudanças significativas nos resultados obtidos anteriormente, no entanto o número de iterações de Broyden reduziu drasticamente. O contraste observado no mal funcionamento do algoritmo proposto em comparação com o bom funcionamento do método alternativo nos leva a crer que o método de Broyden puro pode não ser um bom corretor, pois em nossos testes ele divergiu frequentemente. No entanto, ao controlar os passos de Broyden por um tamanho de passo seguindo o mesmo raciocínio do método de Newton em pontos interiores, onde procura-se manter a positividade estrita, como foi feito no método alternativo, observou-se uma melhora no comportamento do método de Broyden, que certamente acarretou nos bons resultados do método alternativo. Isto sugere que essa variação do método de Broyden pode funcionar melhor como um corretor do que o método de Broyden puro.

Com respeito as implementações, é importante mencionar que algumas limitações afetaram o seu bom funcionamento. Por exemplo, faltou um pouco de experiência em programação relacionada a otimização do código para fazer um uso eficiente de memória, o que levou a não resolução de alguns dos problemas. Além disso, certamente o código pode ser otimizado para a realização de alguns cálculos de maneira mais eficiente. Trabalhos futuros podem tentar estudar a complexidade de métodos preditores-corretores com

corretor Quasi-Newton ou até explorar o caso de métodos que não requerem a viabilidade do ponto inicial. Ainda, problemas quadráticos podem ser considerados.

BIBLIOGRAFIA

- [1] WRIGHT, S. Primal-dual Interior-Point Methods. Society for Industrial and Applied Mathematics, 1997. (Other Titles in Applied Mathematics). ISBN 9781611971453. Disponível em: <<https://books.google.com.br/books?id=oQdBzXhZeUkC>>.
- [2] NOCEDAL, J.; WRIGHT, S. Numerical Optimization. Springer New York, 2006. (Springer Series in Operations Research and Financial Engineering). ISBN 9780387400655. Disponível em: <<https://books.google.com.br/books?id=VbHYoSyeIFcC>>.
- [3] KLEE, V.; MINTY, G. J. How good is the simplex algorithm. Inequalities, New York, v. 3, n. 3, p. 159–175, 1972.
- [4] KARMARKAR, N. A new polynomial-time algorithm for linear programming. In: Proceedings of the sixteenth annual ACM symposium on Theory of computing. [S.l.: s.n.], 1984. p. 302–311.
- [5] GONDZIO, J.; SOBRAL, F. Quasi-Newton approaches to interior point methods for quadratic problems. Computational Optimization and Applications, Springer, v. 74, n. 1, p. 93–120, 2019.
- [6] GONDZIO, J.; SOBRAL, F. N. Polynomial worst-case iteration complexity of quasi-Newton primal-dual interior point algorithms for linear programming. arXiv preprint arXiv:2208.08771, 2022.
- [7] MOSCHEN, L. Application of Newton’s and Broyden’s methods in linear programming. Programa de Pós-Graduação em Matemática da Universidade Estadual de Maringá, 2022.
- [8] MEHROTRA, S. On the implementation of a primal-dual interior point method. SIAM Journal on optimization, SIAM, v. 2, n. 4, p. 575–601, 1992.

- [9] GONDZIO, J. Multiple centrality corrections in a primal-dual method for linear programming. Computational optimization and applications, Springer, v. 6, n. 2, p. 137–156, 1996.
- [10] DENNIS, J.; SCHNABEL, R. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Society for Industrial and Applied Mathematics, 1996. (Classics in Applied Mathematics). ISBN 9780898713640. Disponível em: <<https://books.google.com.br/books?id=4y1IDOXNOqMC>>.
- [11] ANJOS, M. F.; LODI, A.; TANNEAU, M. Tulip.jl: an open-source interior-point linear optimization solver with abstract linear algebra. [S.l.], 2019. Disponível em: <<https://www.gerad.ca/fr/papers/G-2019-36>>.
- [12] COLOMBO, M.; GONDZIO, J. Further development of multiple centrality correctors for interior point methods. Computational Optimization and Applications, Springer, v. 41, p. 277–305, 2008.
- [13] BAZARAA, M. S.; JARVIS, J. J.; SHERALI, H. D. Linear programming and network flows. [S.l.]: John Wiley & Sons, 2011.
- [14] FRIEDLANDER, A. Elementos de programação não-linear. Livro-texto. UNICAMP, 2012.
- [15] ANDRADE, D. O Teorema da Função Inversa e da Função Implícita. 2018. Disponível em: <<http://www.dma.uem.br/kit/arquivos/arquivospdf/inversa.pdf>>. Acesso em: 25 de janeiro de 2024.
- [16] MUNKRES, J. R. Analysis on manifolds. [S.l.]: CRC Press, 2018.
- [17] MIZUNO, S.; TODD, M. J.; YE, Y. On adaptive-step primal-dual interior-point algorithms for linear programming. Mathematics of Operations research, INFORMS, v. 18, n. 4, p. 964–981, 1993.
- [18] VITOR, F.; EASTON, T. Projected orthogonal vectors in two-dimensional search interior point algorithms for linear programming. Computational Optimization and Applications, Springer, v. 83, n. 1, p. 211–246, 2022.

-
- [19] DENNIS, J. E.; MORÉ, J. J. A characterization of superlinear convergence and its application to quasi-Newton methods. Mathematics of computation, v. 28, n. 126, p. 549–560, 1974.
- [20] KINCAID, D.; CHENEY, W. Numerical analysis, brooks. Cole Publishing Company, v. 20, p. 10–13, 1991.
- [21] DONGARRA, J. J.; GROSSE, E. Distribution of mathematical software via electronic mail. Communications of the ACM, ACM New York, NY, USA, v. 30, n. 5, p. 403–407, 1987.
- [22] TANNEAU, M.; ORBAN, D.; contributors. QPSReader.jl: A Reader for Linear and Quadratic Optimization Problems in MPS, QPS, and SIF Formats. 2020. Disponível em: <<https://github.com/JuliaSmoothOptimizers/QPSReader.jl>>.