

Grasielli Gava

*Alternativas para Resolução de Sistemas
Não-Lineares Especiais*

Florianópolis

Março de 2007

Grasielli Gava

*Alternativas para Resolução de Sistemas
Não-Lineares Especiais*

Dissertação apresentada ao Curso de Pós-Graduação em Matemática e Computação Científica, do Centro de Ciências Físicas e Matemáticas da Universidade Federal de Santa Catarina, para obtenção do grau de Mestre em Matemática, com Área de Concentração em Matemática Aplicada.

Orientador:

Prof. Dr. Mário César Zambaldi

UNIVERSIDADE FEDERAL DE SANTA CATARINA

Florianópolis

Março de 2007

Dissertação de Mestrado sob o título *Alternativas para Resolução de Sistemas Não-Lineares Especiais*, defendida por Grasielli Gava e aprovada em 30 de Março de 2007, em Florianópolis, Estado de Santa Catarina, pela comissão examinadora constituída pelos doutores:



Prof. Dr. Mário César Zambaldi
Orientador



Prof. Dr. Clóvis Caesar Gonzaga
Universidade Federal de Santa Catarina



Prof. Dr. Juliano de Bem Francisco
Universidade Federal de Santa Catarina



Profª. Dra. Marcia Aparecida Gomes Ruggiero
Universidade Estadual de Campinas

Aos meus pais,
aos meus irmãos,
aos amigos que estiveram ao meu lado
e ao Rodrigo.
Agradeço o apoio financeiro do CNPq.

Resumo

Neste trabalho fazemos um estudo e implementação computacional de alguns métodos numéricos para resolução de sistemas não-lineares especiais. Consideramos sistemas com restrições de caixas abordados por meio de técnicas de região de confiança específicas. Resolvemos também o problema de quadrados mínimos não-lineares por meio de técnicas de penalização para o sistema não-linear associado.

Abstract

In this work we make a study and computational implementation of some numerical methods for the solution of special nonlinear systems of equations. We solve some systems with bound constraints by using trust region techniques. We also solve nonlinear least squares problems by using penalization techniques designed to nonlinear systems.

Sumário

Introdução	p. 8
1 Elementos de Otimização para Sistemas de Equações Não-lineares	p. 11
1.1 Método de Newton para sistemas de equações não-lineares	p. 12
1.2 Método de Broyden	p. 12
1.3 Métodos de busca linear	p. 14
1.4 Métodos de região de confiança	p. 16
1.5 Problema de quadrados mínimos não-lineares	p. 19
1.5.1 Método de Gauss-Newton	p. 19
1.5.2 Método Levenberg-Marquardt	p. 20
2 Sistemas Não-lineares com Restrições de Caixa	p. 22
2.1 Introdução	p. 22
2.2 O algoritmo e sua implementação	p. 23
2.3 Análise de convergência	p. 29
2.4 Resultados numéricos	p. 33
3 O Problema de Quadrados Mínimos Não-lineares Restrito via Penalização	p. 37
3.1 Método Newton-Lagrange	p. 37
3.2 Penalização	p. 39
3.2.1 Análise de convergência	p. 42
3.3 Resultados numéricos	p. 45
3.3.1 Fluxo de potência	p. 45

3.3.1.1	Modelagem do problema	p. 47
3.3.1.2	O problema sem solução	p. 49
3.3.1.3	Resultados	p. 50
3.3.2	Outras aplicações	p. 52
Conclusão		p. 54
Referências		p. 56
Apêndice A - Problemas		p. 58
A.1	Problemas referentes ao capítulo 2	p. 58
A.2	Problemas referentes ao capítulo 3	p. 78

Introdução

Sistemas de equações não-lineares aparecem em grande parte dos modelos matemáticos provenientes de problemas da vida real. É o caso de sistemas de potências em engenharia elétrica, equações diferenciais parciais não-lineares em várias áreas, entre outras aplicações.

Em geral, os métodos para resolução desses modelos são baseados em linearizações sucessivas entre outras técnicas específicas. Usualmente métodos numericamente eficientes devem ser adequados às características dos sistemas, como por exemplo, o grau de não-linearidade ou o número de variáveis.

Um sistema não-linear pode ser representado por:

$$F(x) = 0, \quad F : \mathbb{R}^n \rightarrow \mathbb{R}^n. \quad (1)$$

O que referenciamos no título deste trabalho de sistemas não-lineares especiais, são sistemas mais específicos que (1). Neste trabalho consideramos dois modelos centrais. O primeiro deles consiste em limitar o domínio à restrições de caixas. Neste caso o sistema é descrito por:

$$\begin{aligned} F(x) &= 0, \\ l &\leq x \leq u. \end{aligned} \quad (2)$$

A introdução dessas restrições merece uma técnica específica para sua resolução, como veremos na seqüência do trabalho, e que não é usual para a resolução de (1). Para obter um algoritmo robusto, devemos incorporar técnicas que o fazem independente do ponto inicial. Uma forma clássica e conhecida é a estratégia de busca linear usando como função de mérito a soma de quadrados das componentes de F . Alternativamente, a técnica de região de confiança para o sistema com restrições de caixa. Ela se adequa precisamente ao uso das restrições por meio do uso do escalamento da região de confiança.

O método apresentado aqui, descrito em [1], utiliza regiões de confiança elipsoidais, definidas por Coleman e Li [2], originalmente descritas por Dikin [3]. Em cada passo do método, uma função quadrática é minimizada em uma região de confiança elíptica que depende dos limitantes l e u . Assim, evita-se a minimização da função quadrática na caixa.

Outro problema interessante e relacionado à (1) consiste em considerar algumas equações para as quais não existe solução. Pode ser representado assim:

$$F(x) = \begin{bmatrix} f(x) \\ c(x) \end{bmatrix},$$

em que $c(x) = 0$, e em geral $f(x) \neq 0$. Assim o problema de interesse fica

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & \frac{1}{2} \|f(x)\|_2^2 \\ \text{s.a.} & c(x) = 0. \end{aligned} \quad (3)$$

Em geral, problemas como estes são abordados por métodos de quadrados mínimos não-lineares. No problema de quadrados mínimos não-lineares em que a restrição $c(x) = 0$ não aparece, o mais comum é utilizarmos métodos como Gauss-Newton e Levenberg-Marquardt. Como em (3) aparecem restrições, ele pode ser resolvido utilizando-se a função Lagrangiano.

A motivação principal para resolução deste modelo é o fato de que estávamos investigando as soluções do problema de fluxo de potências [4], um sistema não-linear do tipo (1). Uma forma mais elaborada do modelo, considerando as especificidades do problema prático em engenharia elétrica, é relaxar algumas equações de modo que não tenham raízes. Economicamente este fato é muito importante, significando por exemplo, permitir que certas regiões estejam muito mais sujeitas à falta de energia que outras, evitando carregamento ou colapso do sistema elétrico em horários de pico. Então, como algumas equações não apresentam zeros, a idéia é tratá-las de modo distinto das outras. Isto, como veremos mais adiante, permite trabalhar apenas com informações de primeira ordem ao contrário das técnicas usuais para o problema de quadrados mínimos não-lineares.

Neste trabalho fazemos um estudo e implementação computacional de alguns métodos numéricos para resolução de sistemas não-lineares especiais. Consideramos sistemas com restrições de caixas abordados por meio de técnicas de região de confiança específicas. Damos ênfase ao problema de fluxo de potências sem solução real. Resolvemos também o problema de quadrados mínimos não-lineares por meio de técnicas de penalização para o sistema não-linear associado.

O trabalho está organizado como se segue. No Capítulo 1 apresentamos os tópicos de otimização não-linear clássicos, que serão referenciados durante o trabalho. No Capítulo 2 tratamos de métodos para a resolução de (2). No Capítulo 3 descrevemos o modelo das equações da rede elétrica que dão origem ao problema mais relevante do tipo (3); estudamos também

métodos de resolução. Conseqüências, conclusões e futuras linhas de seqüência do trabalho fecham este trabalho.

1 *Elementos de Otimização para Sistemas de Equações Não-lineares*

Neste capítulo apresentamos alguns conceitos e resultados que vamos explorar nos capítulos que se seguem. Esta exposição resumida tem apenas o propósito de referência usada neste trabalho. Detalhes podem ser encontrados em textos clássicos como [5] e [6].

Vamos considerar dois problemas centrais em otimização não linear. Um sistema de equações não-lineares consiste em obter $x \in \mathbb{R}^n$ tal que

$$F(x) = 0 \tag{1.1}$$

em que $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é continuamente diferenciável.

Uma forma de resolver (1.1) é exigir a redução de $\|F(x)\|_2$ se possível para 0. Nosso problema pode ser reescrito como

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|F(x)\|_2^2, \tag{1.2}$$

onde o quadrado da norma tem a vantagem de deixar a função objetivo continuamente diferenciável. Embora nem toda solução de (1.2) seja solução de (1.1) poderíamos pensar em utilizar um método de minimização para resolver o sistema não-linear. Neste caso seria importante usar a estrutura do problema (1.1). A direção de Newton de (1.1) é uma direção de descida para (1.2). Outro ponto de conexão é que estratégias globais para (1.1) estão baseadas em (1.2). Isso veremos mais adiante.

No que se segue, descrevemos resumidamente os métodos de Newton e Broyden para sistemas não-lineares e examinaremos suas propriedades de convergência local. Para obter a convergência global, apresentaremos os métodos de busca linear e métodos de região de confiança. Apresentaremos também uma introdução sobre o problema de quadrados mínimos não-lineares.

1.1 Método de Newton para sistemas de equações não-lineares

O método de Newton para sistemas de equações não-lineares consiste em resolver uma seqüência de problemas aproximados ao nosso problema original, com o objetivo de nos aproximar da solução de (1.1).

A cada iteração k , constrói-se um modelo M_k aproximado de F e encontra-se a solução x_{k+1} de $M_k(p) = 0$.

Sabemos que

$$F(x + p) = F(x) + \int_0^1 F'(x + tp)p dt \quad (1.3)$$

em que F' representa o Jacobiano de F .

Obtemos o modelo M_k , aproximando a integral em (1.3) pelo termo linear $F'(x_k)p$, obtendo

$$M_k(p) = F(x_k) + F'(x_k)p.$$

Assim, $x_{k+1} = x_k + p_k$ em que p_k é a solução do sistema linear

$$F'(x_k)p_k = -F(x_k). \quad (1.4)$$

Se a aproximação x_k está próxima de uma solução x^* e o Jacobiano F' é uma função contínua de x , o método de Newton possui convergência local superlinear, e, se a aproximação x_k está próxima de uma solução x^* e o Jacobiano F' é Lipschitz contínuo, então o método de Newton possui convergência local quadrática.

1.2 Método de Broyden

Como visto acima, o método de Newton possui boa convergência local, mas possui o inconveniente de calcular o Jacobiano de F em cada iteração. As direções quase-Newton são uma alternativa atraente, já que não requerem o cálculo do Jacobiano e ainda alcançam a razão de convergência superlinear. Em sistemas grandes, estes métodos têm a vantagem de trabalhar diretamente com formas não fatoradas das aproximações para a inversa do Jacobiano, com custo inferior à resolução do sistema não-linear.

Os métodos quase-Newton constróem, em cada iteração, uma aproximação para o Jacobiano de F . O Jacobiano aproximado $B(x_k)$ é usado no lugar de $F'(x_k)$ em (1.4). Assumindo que $B(x_k)$ seja não-singular, o passo para a próxima iteração será a solução de

$$B(x_k)p_k = -F(x_k).$$

As matrizes de aproximação geradas pelos métodos quase-Newton devem ser uma boa aproximação do Jacobiano ao longo da direção p_k , para isso exige-se que M_{k+1} seja igual à F ao menos nas duas últimas aproximações, x_k e x_{k+1} . Assim, como $M_{k+1}(0) = F(x_{k+1})$, basta exigirmos

$$M_{k+1}(-p_k) = F(x_{k+1}) - B(x_{k+1})p_k = F(x_k),$$

ou seja,

$$B(x_{k+1})p_k = F(x_{k+1}) - F(x_k).$$

Considerando $s_k = x_{k+1} - x_k$ e $y_k = F(x_{k+1}) - F(x_k)$ obtemos a *equação secante*:

$$y_k = B(x_{k+1})s_k, \tag{1.5}$$

que garante a boa concordância de $B(x_{k+1})$ e $F'(x_{k+1})$ ao longo de p_k .

No método de Broyden a matriz $B_{k+1} = B(x_{k+1})$ é atualizada da seguinte forma

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)s_k^T}{s_k^T s_k}. \tag{1.6}$$

Segundo [6, Lema 11.4], a atualização de Broyden é a atualização que faz as menores mudanças possíveis na matriz B_k (medida pela norma Euclideana $\|B_k - B_{k+1}\|$), dentre as que satisfazem (1.5).

O desempenho do método depende muito da escolha de $B_0 = B(x_0)$, algumas implementações recomendam escolher B_0 como sendo $F'(x_0)$ ou alguma aproximação por diferenças finitas desta matriz.

1.3 Métodos de busca linear

Podemos obter algoritmos com convergência global, aplicando os métodos de busca linear para a função $f(x) = \frac{1}{2}\|F(x)\|_2^2$. Na estratégia de busca linear, após escolhermos a direção de busca p_k , buscamos ao longo desta direção, a partir da aproximação x_k , uma nova aproximação x_{k+1} com um valor de função menor, ou seja, $f(x_{k+1}) < f(x_k)$.

A nova aproximação será dada por:

$$x_{k+1} = x_k + \alpha_k p_k,$$

em que o escalar positivo α_k é chamado de comprimento do passo, e pode ser encontrado resolvendo-se aproximadamente o seguinte problema de minimização unidimensional

$$\min_{\alpha > 0} \phi(\alpha_k),$$

sendo $\phi(\alpha_k) = f(x_k + \alpha_k p_k)$.

A condição de que $f(x_{k+1}) = f(x_k + \alpha_k p_k) < f(x_k)$ não garante que $\{x_k\}$ convirja para um minimizador de f . Podemos ter reduções muito pequenas no valor de f em relação ao tamanho dos passos e também podemos tomar passos muito pequenos em relação à taxa de decréscimo inicial de f .

Para evitar que a função tenha um decréscimo muito pequeno em relação ao seu valor anterior, podemos exigir que a redução em f seja proporcional ao tamanho do passo α_k e à sua derivada direcional na direção p_k , $\nabla f(x_k)^T p_k$, que representa a taxa de decréscimo da função nesta direção. Assim queremos encontrar α_k entre os α que satisfazem

$$f(x_k + \alpha p_k) \leq f(x_k) + c_1 \alpha \nabla f(x_k)^T p_k, \quad (1.7)$$

com $c_1 \in (0, 1)$. Esta desigualdade é chamada *Condição de Armijo* [6].

Para evitarmos passos muito pequenos, inclui-se uma condição que teste a inclinação $\phi'(\alpha) = \nabla f(x_k + \alpha p_k)^T p_k$ decorrente da escolha de α :

$$\nabla f(x_k + \alpha p_k)^T p_k \geq c_2 \nabla f(x_k)^T p_k,$$

com $c_2 \in (c_1, 1)$, c_1 escolhido na condição (1.7). Essa condição, chamada de *Condição de Cur-*

vatura, exige que a inclinação de ϕ em α_k seja maior do que a inclinação de ϕ em 0. Lembrando que $\phi(0) = f(x_k)$, esta exigência garante que α_k não será escolhido pequeno demais.

As condições de Armijo e de curvatura juntas são chamadas de *Condições de Wolfe* [6].

A convergência global depende não somente de uma boa escolha de α_k , mas também de uma boa escolha para a direção de busca. Considerando θ_k como o ângulo entre p_k e a direção de descida $-\nabla f(x_k) = -F'(x_k)^T F(x_k)$, a direção p_k deverá satisfazer

$$\cos \theta_k = -\frac{p_k^T \nabla f(x_k)}{\|p_k\| \|\nabla f(x_k)\|} > 0,$$

ou seja, p_k deverá ser uma direção de descida para F .

Segundo [6, Teorema 11.6], no caso da busca linear que escolhe α_k satisfazendo as condições de Wolfe, se garantirmos que

$$\cos \theta_k \geq \delta, \quad \text{para algum } \delta \in (0, 1) \text{ e para todo } k \text{ suficientemente grande,} \quad (1.8)$$

então $F'(x_k)^T F(x_k) \rightarrow 0$. Além disso, quando $\|F'(x_k)^{-1}\|$ é limitada em uma vizinhança \mathcal{D} do conjunto de nível $\mathcal{L} = \{x : f(x) \leq f(x_0)\}$ temos que $F(x_k) \rightarrow 0$, ou seja, que a aproximação x_k deverá se aproximar de um ponto limite x^* que resolve o problema de equações não-lineares $F(x) = 0$.

O passo de Newton (1.4) é uma direção de descida para f , e assim, podemos ver que

$$\cos \theta_k = -\frac{p_k^T \nabla f(x_k)}{\|p_k\| \|\nabla f(x_k)\|} = \frac{\|F(x_k)\|^2}{\|F'(x_k)^{-1} F(x_k)\| \|F'(x_k)^T F(x_k)\|} \geq \frac{1}{\|F'(x_k)^T\| \|F'(x_k)^{-1}\|} = \frac{1}{\kappa(F'(x_k))}$$

em que $\kappa(F'(x_k))$ é o número de condição de $F'(x_k)$. Assim se $\kappa(F'(x_k))$ é uniformemente limitado para todo k , então $\cos \theta_k$ é limitado no estilo de (1.8).

O mesmo acontece em um algoritmo de busca linear quase-Newton, com $p_k = -B_k^{-1} \nabla f(x_k)$, pois tudo o que precisamos é que os números de condições de $\{B_k\}$ sejam uniformemente limitados. O seguinte teorema, devido a Dennis e Moré, mostra que se a direção de busca de um método Quase-Newton aproxima muito bem a direção de Newton, então o comprimento de passo unitário irá satisfazer as condições de Wolfe assim como as aproximações convergem para a solução.

Teorema 1 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^2(D)$, em que D é um conjunto aberto, convexo de \mathbb{R} e $\nabla^2 f$*

é Lipschitz contínua em D . Considere uma seqüência $\{x_k\}$ gerada por $x_{k+1} = x_k + \alpha_k p_k$, em que $\nabla f(x_k)^T p_k < 0$ para todo k e α_k é escolhido para satisfazer as condições de Wolfe com $c_1 < \frac{1}{2}$. Se $x_k \rightarrow x^* \in D$, em que $\nabla^2 f(x^*)$ é positiva definida, e se

$$\lim_{k \rightarrow \infty} \frac{\|\nabla f(x_k) + \nabla^2 f(x_k) p_k\|}{\|p_k\|} = 0 \quad (1.9)$$

então existe um índice $k_0 \geq 0$ tal que para todo $k \geq k_0$, $\alpha_k = 1$ é admissível. Além disso, $\nabla f(x^*) = 0$, e se $\alpha_k = 1$ para todo $k \geq k_0$, então $\{x_k\}$ converge superlinearmente para x^* .

A condição (1.9) é necessária e suficiente para a taxa de convergência superlinear dos métodos quase-Newton.

Assim, se utilizarmos a estratégia de busca linear, os métodos de direção de descida convergem, desde que a escolha do passo satisfaça as condições de Wolfe, ou alguma condição equivalente.

1.4 Métodos de região de confiança

No método de busca linear descrito acima, tratamos exclusivamente com o problema de encontrar um comprimento de passo aceitável em uma direção de busca dada. Se a direção de busca não for satisfatória, significa que o modelo utilizado não está representando adequadamente a função f . Caso isso aconteça, os algoritmos de busca linear mantém essa mesma direção de busca e calculam um novo comprimento de passo.

Na estratégia de região de confiança, em cada iteração é construído um modelo quadrático de f e uma região em que considera-se que o modelo represente a função adequadamente e minimiza-se este modelo na região, daí o nome região de confiança, ou seja, primeiro escolhemos o comprimento máximo para o passo e depois, utilizando o modelo, escolhemos a direção do passo. Dependendo do resultado obtido com esta minimização, algumas decisões são tomadas: aceitar ou não o minimizador como um novo ponto (uma nova aproximação); aumentar, reduzir, ou deixar inalterada a região de confiança

Considerando então o modelo

$$m_k(p) = f_k + \nabla f_k^T p + \frac{1}{2} p^T H_k p,$$

em que $f_k = f(x_k)$, $\nabla f_k = \nabla f(x_k)$ e H_k é a hessiana $\nabla^2 f(x_k)$ ou uma aproximação dela, precisa-

mos resolver o seguinte problema

$$\min_{p \in \mathbb{R}^n} m_k(p) \text{ sujeito a } \|p\| \leq \Delta_k, \quad (1.10)$$

em que $\Delta_k > 0$ é o raio da região de confiança. Para a convergência global, é suficiente encontrar uma solução aproximada para o problema (1.10), mais adiante apresentaremos uma das formas de encontrar essa solução. Outras formas de encontrar solução aproximada ou até mesmo soluções exatas podem ser obtidas em [6] e [3].

A escolha do raio de região de confiança em cada iteração dependerá do desempenho da direção na iteração anterior, precisamos saber se a redução obtida no modelo foi tão boa quanto a redução da função. Para isso, dada uma direção p_k , vamos nos basear na seguinte razão

$$\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)},$$

ou seja, a razão entre a redução real da função e a redução prevista pelo modelo.

Se ρ_k está próximo de 0 ou é negativo, significa que a concordância entre o modelo e a função sobre este passo é ruim, assim diminui-se o raio da região de confiança e minimiza-se o mesmo modelo quadrático sob esse raio menor. Se ρ_k está próximo de 1, então a concordância entre m_k e f está muito boa e podemos aumentar o raio para a próxima iteração. Se ρ_k é positivo mas não está tão próximo de 1, o raio não é alterado.

Decidido o tamanho da região de confiança, precisamos agora resolver (1.10). Vamos encontrar uma solução aproximada p_k que está dentro da região de confiança e dê uma redução suficiente no modelo. A redução suficiente será quantificada em termos do ponto de Cauchy, que iremos denotar por p_k^C . O ponto de Cauchy é o minimizador do modelo $m_k(\cdot)$ ao longo da direção de máximo declive, restrito à bola de raio Δ_k em torno de x_k , e pode ser calculado, segundo [6], da seguinte forma:

$$p_k^C = -\tau_k \frac{\Delta_k}{\|\nabla f_k\|} \nabla f_k$$

em que

$$\tau_k = \begin{cases} 1, & \text{se } \nabla f_k^T H_k \nabla f_k \leq 0; \\ \min(\|\nabla f_k\|^3 / (\Delta_k \nabla f_k^T H_k \nabla f_k), 1), & \text{caso contrário.} \end{cases}$$

Vários algoritmos que geram soluções aproximadas p_k para o subproblema (1.10) começam

calculando o ponto de Cauchy e então tentam melhorá-lo. Essa é a idéia do método dogleg que vamos descrever abaixo.

Quando H_k é positiva definida, o minimizador irrestrito de m_k é dado pela direção de Newton $p_k^N = -H_k^{-1} \nabla f(x_k)$. Se este for um ponto viável para (1.10), ou seja, se $\|p_k^N\| \leq \Delta$, ele será uma solução e aí,

$$p_k = p_k^N.$$

Quando Δ_k for muito pequeno, a solução de (1.10) será dada pelo ponto de Cauchy, assim

$$p_k = -\Delta \frac{\nabla f_k}{\|\nabla f_k\|}$$

Para valores intermediários de Δ_k a solução exata do problema de região de confiança segue uma trajetória curva como a que está apresentada pela chamada trajetória ótima na figura 1.

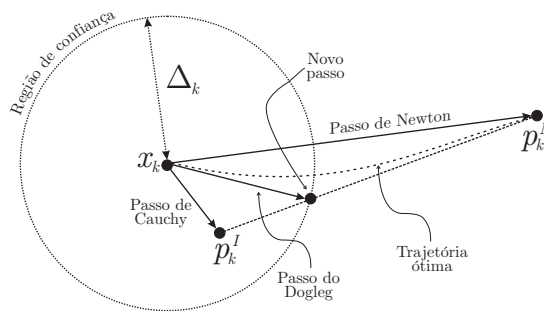


Figura 1: O passo do método dogleg.

O método dogleg encontra uma solução aproximada trocando a trajetória ótima por um caminho com dois segmentos de reta. O primeiro segmento vai da aproximação x_k ao minimizador irrestrito ao longo da direção de máxima descida definido por

$$p_k^I = -\frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T B_k \nabla f_k} \nabla f_k,$$

enquanto o segundo vai de p_k^I a p_k^N . Esta trajetória pode ser escrita como

$$p_k = \begin{cases} \mu p_k^I, & 0 \leq \mu \leq 1, \\ p_k^I + (\mu - 1)(p_k^N - p_k^I), & 1 \leq \mu \leq 2. \end{cases}$$

Assim, no caso em que p_k^N está fora da região de confiança, o passo dogleg será a intersecção da trajetória definida acima com o raio da região de confiança, este ponto de intersecção pode

ser calculado analiticamente, como pode ser visto em [6].

1.5 Problema de quadrados mínimos não-lineares

O problema de quadrados mínimos não-lineares é mais freqüentemente encontrado no contexto de ajuste de curvas, mas pode aparecer também quando um sistema de equações não-lineares tem mais equações do que incógnitas. É um caso especial de minimização irrestrita, e, por isso, aproveitando as vantagens da estrutura do problema pode-se modificar as técnicas de minimização irrestrita e criar algoritmos mais adequados ao problema.

Dada a função $r : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq n$, o problema de quadrados mínimos não-lineares consiste em encontrar $x \in \mathbb{R}^n$ para o qual $G(x) = \sum_{i=1}^m (r_i(x))^2$ é minimizado, em que $r_i(x)$ denota a i -ésima componente da função r .

Embora a minimização possa ser feita por um método de minimização irrestrita padrão, em muitos casos é vantajoso levar em conta as propriedades da função G . As derivadas de $G(x)$ podem ser expressas em termos do Jacobiano J de r da seguinte forma:

$$\nabla G(x) = \sum_{j=1}^m r_j(x) \nabla r_j(x) = J(x)^T r(x)$$

e

$$\begin{aligned} \nabla^2 G(x) &= \sum_{j=1}^m \nabla r_j(x) \nabla r_j(x)^T + \sum_{j=1}^m r_j(x) \nabla^2 r_j(x) \\ &= J(x)^T J(x) + \sum_{j=1}^m r_j(x) \nabla^2 r_j(x). \end{aligned} \quad (1.11)$$

Podemos então observar que a Hessiana da função objetivo $G(x)$ consiste de uma combinação das informações de primeira e segunda ordem. Muitos algoritmos para quadrados mínimos não-lineares exploram a estrutura da Hessiana considerando que para o resíduo $\|r_j(x)\|$ pequeno, o termo $J(x)^T J(x)$ dominará o segundo termo de (1.11).

1.5.1 Método de Gauss-Newton

O método Gauss-Newton [7] pode ser visto como uma modificação do método de Newton para minimização irrestrita de G com busca linear. Aqui, o segundo termo de (1.11) é excluído

e obtemos a direção de busca p_k , resolvendo

$$J(x_k)^T J(x_k) p_k = -J(x_k)^T r(x_k). \quad (1.12)$$

Note que o sistema (1.12) envolve apenas as derivadas de primeira ordem de r . A solução de (1.12) é uma solução do problema de quadrados mínimos não-lineares

$$\min_{p \in \mathbb{R}^n} \frac{1}{2} \|J(x_k)p + r(x_k)\|_2^2, \quad (1.13)$$

e é única se $J(x_k)$ tem posto completo. O vetor que resolve (1.13) é chamado de direção de Gauss-Newton e será denotado por p^{GN} .

Uma vantagem do método Gauss-Newton é que quando $J(x_k)$ tem posto completo, a direção p^{GN} é uma direção de descida para G . Assim, é uma direção adequada para a busca linear e portanto localmente convergente para quase todos os problemas de quadrados mínimos não-lineares.

Uma alternativa para o método Gauss-Newton é o método de Levenberg-Marquardt.

1.5.2 Método Levenberg-Marquardt

O método Levenberg-Marquardt [7] é uma modificação do método Gauss-Newton, no sentido que usa a estratégia de região de confiança ao invés da estratégia de busca linear. A direção de busca do método Levenberg-Marquardt é definida como a solução das equações

$$(J(x_k)^T J(x_k) + \eta_k I) p_k^{LM} = -J(x_k)^T r(x_k) \quad (1.14)$$

em que η_k é um escalar não-negativo.

Pode-se mostrar que, dado $\eta_k > 0$ existe $\Delta = \Delta(\eta_k) > 0$ tal que o vetor p_k^{LM} é a solução do subproblema restrito

$$\min_{p \in \mathbb{R}^n} \frac{1}{2} \|J(x_k)p + r(x_k)\|_2^2 \quad \text{s.a.} \quad \|p\|_2 \leq \Delta.$$

O uso de técnicas baseadas em região de confiança evita uma das fragilidades do método Gauss-Newton, isto é, é adequado quando o jacobiano $J(x)$ tem posto deficiente, ou está próximo disto. Outra vantagem é que quando o passo de Gauss-Newton é muito longo, o passo de

Levenberg-Marquardt está próximo de estar em uma direção de descida .

2 *Sistemas Não-lineares com Restrições de Caixa*

2.1 Introdução

Seja $F : X \rightarrow \mathbb{R}^n$, $F \in C^2(X)$, em que $X \subseteq \mathbb{R}^n$ é um conjunto aberto contendo a caixa n -dimensional

$$\Omega = \{x \in \mathbb{R}^n; l \leq x \leq u\}.$$

Os vetores $l \in (\mathbb{R} \cup -\infty)^n$ e $u \in (\mathbb{R} \cup \infty)^n$ são cotas inferiores e superiores para as variáveis de forma que Ω tenha interior não vazio.

O problema é encontrar um vetor $x \in \mathbb{R}^n$ satisfazendo

$$F(x) = 0, \quad x \in \Omega. \quad (2.1)$$

Problemas constituídos de sistemas não-lineares $F(x) = 0$ com restrições de desigualdades e possivelmente com restrições de caixa, podem ser colocados na forma (2.1) através da introdução das variáveis de folga. Uma observação importante é que a formulação (2.1) permite remover as descontinuidades quando a função F não está definida em todo o espaço \mathbb{R}^n .

Métodos globalmente convergentes para o problema irrestrito $F(x) = 0$ podem ser impróprios para resolver (2.1). Esses métodos, como por exemplo o método de Newton globalizado, podem gerar soluções falsas (vetores que satisfazem $F(x) = 0$ mas que não estão em Ω) tão rápido quanto soluções significativas. Mesmo que uma solução significativa esteja no interior do conjunto viável, tomar um ponto inicial numa vizinhança dessa solução não garante que as soluções falsas sejam evitadas.

Segundo [1], uma forma simples de resolver sistemas de equações não-lineares com restrições de caixa é transformar (2.1) em um problema de programação não-linear usando a norma

Euclidiana de F como função objetivo. Considerando as diferenças importantes entre equações não-lineares e o problema de minimização, é essencial estudar algoritmos específicos para resolver (2.1) em sua forma original, explorando suas principais características.

O método apresentado a seguir combina idéias do método clássico de região de confiança para resolver o sistema de equações não-lineares irrestrito $F(x) = 0, x \in \mathbb{R}^n$, e a abordagem afim escala para a solução de problemas de otimização restritos dada por [2]. Uma propriedade importante do método é a exigência de que todas as aproximações estejam no interior estrito de Ω .

2.2 O algoritmo e sua implementação

Para definir um processo iterativo robusto, o método de Newton pode ser incorporado em um esquema de região de confiança globalmente convergente. No método clássico de região de confiança, uma região em torno da aproximação corrente x_k é definida e dentro dessa região o seguinte modelo quadrático

$$m_k(p) = \frac{1}{2}\|F'_k p + F_k\|^2 = \frac{1}{2}\|F_k\|^2 + F_k^T F'_k p + \frac{1}{2}p^T F_k'^T F'_k p \quad (2.2)$$

é confiável para ser uma representação adequada da função de mérito

$$f(x) = \frac{1}{2}\|F(x)\|^2.$$

Portanto, a direção de busca p_k é o vetor solução do subproblema

$$\min_p \{m_k(p); \|p\| \leq \Delta_k\}, \quad (2.3)$$

para um raio da região de confiança Δ_k dado. Quando o sistema não-linear é restrito, temos que levar em conta que a exigência da viabilidade estrita pode levar a problemas que impedem a convergência da seqüência $\{x_k\}$ para uma solução de (2.1). Para prevenir estes problemas, Bellavia, Macconi e Morini [1] usaram a estratégia afim escala proposta por Coleman e Li [2]. Considere o gradiente $F'^T(x)F(x)$ da função de mérito f e tome $v(x)$ como a função vetorial com componentes $v_i(x)$, $i = 1, 2, \dots, n$, dadas por

$$\begin{aligned}
v_i(x) &= x_i - u_i, & \text{se } (F'^T(x)F(x))_i < 0 \text{ e } u_i < \infty; \\
v_i(x) &= x_i - l_i, & \text{se } (F'^T(x)F(x))_i \geq 0 \text{ e } l_i > -\infty; \\
v_i(x) &= -1, & \text{se } (F'^T(x)F(x))_i < 0 \text{ e } u_i = \infty; \\
v_i(x) &= 1, & \text{se } (F'^T(x)F(x))_i \geq 0 \text{ e } l_i = -\infty.
\end{aligned} \tag{2.4}$$

Tomando $D(x)$ como a matriz diagonal de escala tal que

$$D(x) = \text{diag}(|v_1(x)|^{-1/2}, |v_2(x)|^{-1/2}, \dots, |v_n(x)|^{-1/2}),$$

podemos considerar a região de confiança elíptica

$$\|D_k p\| \leq \Delta_k.$$

Assim, a fim de determinar x_{k+1} , um passo p_k é calculado resolvendo-se o seguinte subproblema de região de confiança elíptico

$$\min_p \{m_k(p); \|D_k p\| \leq \Delta_k\}. \tag{2.5}$$

No contexto de sistemas não-lineares irrestritos, se x_k é uma boa aproximação da solução, o método de Newton pode ser aplicado e o passo p_k^N dado por

$$F'_k p_k^N = -F_k$$

resolve (2.5) se Δ_k é grande o suficiente para que $\|D_k p_k^N\| \leq \Delta_k$ seja satisfeito.

Por outro lado, se $\|D_k p_k^N\| > \Delta_k$, calculamos uma solução aproximada de (2.5) reescalando a variável p de maneira que a região de confiança seja esférica na variável escalada. Definindo $\tilde{p} = D_k p$ e substituindo em (2.5), obtemos:

$$\begin{cases} \min_{\tilde{p}} \tilde{m}_k(\tilde{p}) = f_k + \nabla f_k^T D_k^{-1} \tilde{p} + \frac{1}{2} \tilde{p}^T (D_k^{-1} F'^T F'_k D_k^{-1}) \tilde{p} \\ \text{s.a} \quad \|\tilde{p}\| \leq \Delta_k \end{cases} \tag{2.6}$$

Se $\|\tilde{p}_k^N\| = \|D_k p_k^N\| > \Delta_k$, calculamos uma solução aproximada de (2.6) usando o método dogleg [6]. De acordo com o capítulo 1, a trajetória curva ótima é aproximada por um caminho constituído de dois segmentos, o primeiro segmento une a aproximação x_k ao minimizador

irrestrito \tilde{p}_k^I de $\tilde{m}_k(\tilde{p})$ ao longo da direção de máxima descida $D_k^{-1}\nabla f_k$:

$$\tilde{p}_k^I = -\frac{\|D_k^{-1}\nabla f_k\|^2}{\|F_k' D_k^{-2}\nabla f_k\|^2} D_k^{-1}\nabla f_k, \quad (2.7)$$

enquanto o segundo segmento conecta \tilde{p}_k^I à \tilde{p}_k^N . O método dogleg aproxima a solução \tilde{p}_k de (2.6) calculando o minimizador do modelo \tilde{m}_k ao longo desse caminho, ou seja,

$$\tilde{p}_k = \begin{cases} \Delta_k D_k^{-1}\nabla f_k / \|D_k^{-1}\nabla f_k\|, & \text{se } \|\tilde{p}_k^I\| \geq \Delta_k \\ \tilde{p}_k^I + (1 - \mu)(\tilde{p}_k^N - \tilde{p}_k^I), & \text{caso contrário} \end{cases} \quad (2.8)$$

em que μ é a solução positiva da seguinte equação

$$\|\tilde{p}_k^I + (1 - \mu)(\tilde{p}_k^N - \tilde{p}_k^I)\|^2 = \Delta_k^2. \quad (2.9)$$

Por fim, para voltar ao espaço original e calcular uma solução aproximada p_k de (2.5), simplesmente fazemos $p_k = D_k^{-1}\tilde{p}_k$. Note que para pequenos valores de Δ_k , a direção de p_k é a direção de máxima descida escalada d_k dada por

$$d_k = -D_k^{-2}\nabla f_k. \quad (2.10)$$

A discussão acima leva ao seguinte algoritmo:

Algoritmo 1: Método Dogleg

Entrada: p_k^N , ∇f_k , D_k e Δ_k

1 se $\|D_k p_k^N\| \leq \Delta_k$ então

2 $p_k = p_k^N$ e pare.

3 **fim**

4 Calcule \tilde{p}_k^I usando (2.7).

5 se $\|\tilde{p}_k^I\| \geq \Delta_k$ então

6 $\tilde{p}_k = \Delta_k D_k^{-1}\nabla f_k / \|D_k^{-1}\nabla f_k\|$

7 **senão**

8 $\tilde{p}_k^N = D_k p_k^N$

9 Calcule μ resolvendo (2.9) e tome $\tilde{p}_k = \tilde{p}_k^I + (1 - \mu)(\tilde{p}_k^N - \tilde{p}_k^I)$.

10 **fim**

11 Tome $p_k = D_k^{-1}\tilde{p}_k$.

Para manter a viabilidade estrita das iterações, são feitas, quando necessário, restrições adequadas sobre a escolha do passo. Assim, o método pode lidar com problemas em que a função F não está definida fora de Ω .

A fim de garantir que x_{k+1} esteja no interior de Ω , calcula-se o tamanho do passo ao longo de p_k até a fronteira, isto é

$$\lambda(p_k) = \begin{cases} \infty, & \text{se } \Omega = \mathbb{R}^n; \\ \min_{1 \leq i \leq n} \Lambda_i(p_k), & \text{se } \Omega \subset \mathbb{R}^n; \end{cases}$$

em que

$$\Lambda_i(p_k) = \begin{cases} \max \left\{ \frac{l_i - (x_k)_i}{(p_k)_i}, \frac{u_i - (x_k)_i}{(p_k)_i} \right\}, & \text{se } (p_k)_i \neq 0; \\ \infty, & \text{se } (p_k)_i = 0. \end{cases}$$

Note que se $\lambda(p_k) > 1$, então $x_k + p_k$ está no interior de Ω ; caso contrário, um passo para trás ao longo de p_k será necessário para ficar dentro de Ω .

Basta considerar agora

$$x_{k+1} = x_k + \alpha(p_k),$$

em que o passo $\alpha(p_k)$ é definido por

$$\alpha(p_k) = \begin{cases} p_k, & \text{se } \lambda(p_k) > 1, \\ \max\{\theta, 1 - \|p_k\|\} \lambda(p_k) p_k, & \text{caso contrário,} \end{cases} \quad (2.11)$$

e $\theta \in (0, 1)$ é uma constante fixa independente de k .

O objetivo do método de região de confiança é obter uma nova aproximação que permaneça no interior da região de confiança e que resulte numa boa redução no modelo. A redução suficiente é quantificada em termos do minimizador de m_k ao longo da direção de máxima descida d_k dada por (2.10), sujeito aos limites da região de confiança, isto é, em termos do ponto de Cauchy:

$$p_k^C = \tau_k d_k = \tau_k D_k^{-2} \nabla f_k, \quad (2.12)$$

em que $\tau_k = \arg \min_{\tau > 0} \{m_k(\tau d_k) : \|\tau D_k d_k\| \leq \Delta_k\}$ (ver [6]).

Testamos assim, se o passo $\alpha(p_k)$ satisfaz a seguinte condição:

$$\rho_k^C(p_k) = \frac{m_k(0) - m_k(\alpha(p_k))}{m_k(0) - m_k(\alpha(p_k^C))} \geq \beta_1, \quad (2.13)$$

em que $\beta_1 \in (0, 1)$ é uma constante dada.

A solução do subproblema de região de confiança (2.5), satisfaz (2.13), mas se um passo para trás ao longo de p_k for necessário, isto é, se $\alpha(p_k) \neq p_k$, então a condição (2.13) não é necessariamente satisfeita.

Para que exista uma boa concordância entre o modelo m_k e a função objetivo f , exigimos que $\alpha(p_k)$ satisfaça a condição padrão:

$$\rho_k^f(p_k) = \frac{f(x_k) - f(x_k + \alpha(p_k))}{m_k(0) - m_k(\alpha(p_k))} \geq \beta_2, \quad (2.14)$$

em que $\beta_2 \in (0, 1]$ é uma constante dada.

As considerações acima nos levam ao seguinte método:

Algoritmo 2: Método de Região de Confiança Escalado

Entrada: $x_0 \in \text{int}(\Omega)$, $\Delta_0 > 0$

Escolha $\theta \in (0, 1)$, $\beta_1 \in (0, 1]$, β_2 e β_3 tais que $0 < \beta_2 < \beta_3 < 1$, δ_1 e δ_2 tais que $0 < \delta_1 < 1 < \delta_2$.

- 1 $k = 0$
 - 2 Calcule a matriz D_k .
 - 3 Tome $\rho_k^f(p_k) = 0$.
 - 4 **enquanto** $\rho_k^f(p_k) < \beta_2$ **faça**
 - 5 Encontre $p_k = \arg \min_{\|D_k p\| \leq \Delta_k} m_k(p)$.
 - 6 Calcule o ponto de Cauchy p_k^C usando (2.12).
 - 7 Calcule $\alpha(p_k)$ e $\alpha(p_k^C)$ usando (2.11).
 - 8 Calcule $\rho_k^c(p_k)$ usando (2.13).
 - 9 **se** $\rho_k^c(p_k) \leq \beta_1$ **então**
 - 10 $p_k = p_k^C$
 - 11 **fim**
 - 12 Tome $\Delta_k^* = \Delta_k$ e $\Delta_k = \delta_1 \Delta_k$.
 - 13 Calcule $\rho_k^f(p_k)$ usando (2.14).
 - 14 **fim**
 - 15 Tome $x_{k+1} = x_k + \alpha(p_k)$ e $\Delta_k = \Delta_k^*$.
 - 16 **se** $\rho_k^f(p_k) \geq \beta_3$ **então**
 - 17 $\Delta_{k+1} = \delta_2 \Delta_k$
 - 18 **senão**
 - 19 $\Delta_{k+1} = \Delta_k$
 - 20 **fim**
 - 21 $k = k + 1$
-

Como podemos observar na linha 9 do algoritmo, se a condição (2.13) não é válida, ou seja, se o passo encontrado não é tão bom quanto o passo de Cauchy, tomamos $p_k = p_k^C$.

O passo é analisado também pela condição (2.14) e com base nela, na linha 4 é que o passo é aceito ou rejeitado. Se a condição (2.14) é válida, a linha 14 é finalizada com sucesso e $x_k + \alpha(p_k)$ é a próxima iteração. Caso contrário, na linha 12 diminui-se o tamanho da região de confiança, tomando

$$\Delta_k = \min\{\alpha_1 \Delta_k, \alpha_2 \|D_k \alpha(p_k)\|\},$$

para α_1, α_2 tais que $0 < \alpha_1 \leq \alpha_2 < 1$ e um novo passo deve ser calculado na linha 5.

Se o tamanho da região de confiança é muito pequeno quando comparado com a concordância entre o modelo e a função objetivo, o método permite, na linha 17, que o tamanho da região de confiança seja aumentado, tomando

$$\Delta_{k+1} = \max\{\Delta_k, 2\|D_k \alpha(p_k)\|\}.$$

Em cada iteração, resolve-se aumentar Δ_k se a condição (2.14) é satisfeita para uma constante $\beta_3 \in (0, 1)$ tal que $\beta_3 > \beta_2$.

2.3 Análise de convergência

As propriedades de convergência do método descrito na seção anterior foram investigadas em [1]. Vamos apresentar algumas dessas propriedades.

Sejam $\{x_k\}$ a seqüência gerada pelo método proposto, $r > 0$ e $L = \bigcup_{k=0}^{\infty} \{x \in \Omega / \|x - x_k\| \leq r\}$.

Suponha que

(H1) F' é Lipschitz contínua em L , com constante de Lipschitz igual a $2\gamma_L$;

(H2) $\|F'(x)\|$ é limitada superiormente em L .

Para os resultados de convergência do método, vamos assumir que a seqüência $\{x_k\}$ é limitada. Assim, evitamos as falhas devido à falta de pontos limite.

Note que se $\{x_k\}$ é limitada, então existe uma constante $\chi_D > 0$ tal que

$$\|D^{-1}(x)\| < \chi_D, \quad x \in L. \quad (2.15)$$

Ainda, se a hipótese (H1) é satisfeita, $F'(x)^T F(x)$ é Lipschitz contínua em L com constante $2\gamma_L \tilde{\gamma} + \tilde{\gamma}^2$ em que $\tilde{\gamma} = \max\{\sup_{x \in L} f(x), \sup_{x \in L} \|F'(x)\|\}$ e se a hipótese (H2) é satisfeita existe um escalar positivo χ_g tal que o gradiente $F_k'^T F_k$ da função de mérito f satisfaz

$$\|F_k'^T F_k\|_{\infty} < \chi_g, \quad (2.16)$$

para todo $x \in L$.

O teorema 2, que será provado por contradição, juntamente com o Lema 1, nos diz que se a seqüência $\{x_k\}$ é limitada então todos os seus pontos limites são pontos estacionários para o problema $\min_{\Omega} f$.

Lema 1 [8, Lema 2.1] *Seja $x_k \in \Omega$. Então $D_k^{-1}F_k'^T F_k = 0$ se e somente se x_k é um ponto estacionário para o problema $\min_{\Omega} f$.*

Prova: Ver [2]. ■

Teorema 2 [1, Teorema 3.1] *Suponha que (H1) e (H2) são satisfeitas. Se $\{x_k\}$ é limitada, então*

$$\lim_{k \rightarrow \infty} \|D_k^{-1}F_k'^T F_k\| = 0.$$

Prova: Pela hipótese (H2), segue que existe uma constante positiva $\chi_B > 0$ tal que $\|F_k'^T F_k'\| < \chi_B$, $\forall x_k \in L$. Então, por [2, Teorema 3.4] temos que

$$\liminf_{k \rightarrow \infty} \|D_k^{-1}F_k'^T F_k\| = 0. \quad (2.17)$$

De (2.15) e (2.16) existe uma constante $\chi_f > 0$ tal que

$$\|D(x)^{-1}F'(x)^T F'(x)D(x)^{-1}\| < \chi_f, \quad x \in L.$$

Sendo $\text{pred}(p_k) = m_k(0) - m_k(\alpha(p_k))$ e usando [1, Lema 3.3] obtemos

$$\text{pred}(p_k) \geq \frac{1}{2}\beta_1 \|D_k^{-1}F_k'^T F_k\| \min \left\{ \Delta_k, \frac{\|D_k^{-1}F_k'^T F_k\|}{\chi_f}, \frac{\theta \|D_k^{-1}F_k'^T F_k\|}{\chi_g} \right\},$$

e (2.14) fornece

$$\begin{aligned} f(x_k) - f(x_k + \alpha(p_k)) &\geq \beta_2 \text{pred}(p_k) \\ &\geq \frac{1}{2}\beta_1\beta_2 \|D_k^{-1}F_k'^T F_k\| \min \left\{ \Delta_k, \frac{\|D_k^{-1}F_k'^T F_k\|}{\chi_f}, \frac{\theta \|D_k^{-1}F_k'^T F_k\|}{\chi_g} \right\} \end{aligned} \quad (2.18)$$

Agora vamos supor que exista uma seqüência $\{m_i\}$ tal que $\|D_{m_i}^{-1}F_{m_i}'^T F_{m_i}\| \geq \varepsilon_1$ para algum $\varepsilon_1 \in (0, 1)$. Usando (2.17) podemos afirmar que para qualquer $\varepsilon_2 \in (0, \varepsilon_1)$ existe uma subseqüência de $\{m_i\}$, sem perda de generalidade vamos assumir que é a seqüência inteira, e uma seqüência $\{l_i\}$ tal que

$$\|D_k^{-1}F_k'^T F_k\| \geq \varepsilon_2, \quad m_i \leq k < l_i \quad \|D_{l_i}^{-1}F_{l_i}'^T F_{l_i}\| < \varepsilon_2. \quad (2.19)$$

Então de (2.18)

$$f(x_k) - f(x_{k+1}) \geq \frac{1}{2}\beta_1\beta_2\varepsilon_2 \min \left\{ \Delta_k, \frac{\varepsilon_2}{\chi_f}, \frac{\theta\varepsilon_2}{\chi_g} \right\}, \quad m_i \leq k < l_i.$$

Agora, de (2.15) temos que $\|x_{k+1} - x_k\| = \|\alpha(p_k)\| \leq \|D_k^{-1}\|\Delta_k \leq \chi_D$ e aí

$$f(x_k) - f(x_{k+1}) \geq \frac{1}{2}\beta_1\beta_2\varepsilon_2 \min \left\{ \frac{\|x_{k+1} - x_k\|}{\chi_D}, \frac{\varepsilon_2}{\chi_f}, \frac{\theta\varepsilon_2}{\chi_g} \right\}, \quad m_i \leq k < l_i. \quad (2.20)$$

A seqüência $\{f(x_k)\}$ converge, pois é não-crescente e limitada inferiormente por zero. Portanto, $f(x_k) - f(x_{k+1})$ tende à zero.

De (2.20) segue que

$$f(x_k) - f(x_{k+1}) \geq \varepsilon_3\|x_{k+1} - x_k\|, \quad m_i \leq k < l_i,$$

para i suficientemente grande e $\varepsilon_3 = \frac{1}{2}\beta_1\beta_2\varepsilon_2/\chi_D$. Então, usando a desigualdade triangular obtemos

$$f(x_{m_i}) - f(x_{k_i}) \geq \varepsilon_3\|x_{m_i} - x_{k_i}\|, \quad m_i \leq k_i < l_i, \quad (2.21)$$

e podemos concluir que $\|x_{m_i} - x_{k_i}\|$ tende à zero. Ainda, pela continuidade Lipschitz de $F'^T F$ e do fato que $\|x_{m_i} - x_{k_i}\|$ tende à zero, segue que

$$\|F_{m_i}'^T F_{m_i} - F_{k_i}'^T F_{k_i}\| \leq \varepsilon_2, \quad (2.22)$$

para i suficientemente grande.

Sem perda de generalidade, assuma que toda a seqüência x_{l_i} converge para um ponto, digamos x^* . De (2.21) temos que $\{x_{m_i}\}$ converge para x^* também.

Se $(F'(x^*)^T F(x^*))_j \neq 0$ para algum $1 \leq j \leq n$, então (2.4) implica $|(v_{m_i})_j - (v_{l_i})_j| \leq |(x_{m_i})_j - (x_{l_i})_j|$ para i suficientemente grande. Conseqüentemente $\|(D_{m_i}^{-1} - D_{l_i}^{-1})F_{l_i}'^T F_{l_i}\| \rightarrow 0$ e portanto

$$\|(D_{m_i}^{-1} - D_{l_i}^{-1})F_{l_i}'^T F_{l_i}\| \leq \varepsilon_2, \quad (2.23)$$

para i suficientemente grande. Finalmente, de $\|D_{m_i}^{-1}F_{m_i}'^T F_{m_i}\| \geq \varepsilon_1$, (2.19), (2.22), (2.23) e

$$\|D_{m_i}^{-1}F_{m_i}'^T F_{m_i}\| \leq \|D_{m_i}^{-1}\| \cdot \|F_{m_i}'^T F_{m_i} - F_{l_i}'^T F_{l_i}\| + \|(D_{m_i}^{-1} - D_{l_i}^{-1})F_{l_i}'^T F_{l_i}\| + \|D_{l_i}^{-1}F_{l_i}'^T F_{l_i}\|,$$

temos

$$\varepsilon_1 \leq (\chi_D + 2)\varepsilon_2,$$

isto é, uma contradição desde que $\varepsilon_2 \in (0, \varepsilon_1)$ pode ser arbitrariamente pequeno. ■

Observe que pode acontecer de nenhum ponto limite de $\{x_k\}$ satisfazer $F(x) = 0$. Neste caso, a matriz $D(x)^{-1}F'(x)^T$ é singular em cada ponto limite x^* . O fato de $D(x^*)^{-1}F'(x^*)^T$ ser singular ocorre ou se $F'(x^*)$ é singular ou se $D(x^*)^{-1}$ é singular. Este último caso ocorre quando x^* está na fronteira de Ω .

No Teorema 3, encontramos as propriedades de convergência do método quando $\{x_k\}$ tem pelo menos um ponto de acumulação x^* tal que $F(x^*) = 0$ e $F'(x^*)$ é invertível.

Teorema 3 *Suponha que (H1) e (H2) são satisfeitas. Se $\{x_k\}$ é limitada e existe um ponto limite isolado x^* tal que $F'(x^*)$ é invertível e $F(x^*) = 0$, então*

1. $\|F_k\| \rightarrow 0$ e $x_k \rightarrow x^*$;
2. se o ponto limite $x^* \in \text{int}(\Omega)$, então $x_k \rightarrow x^*$ q -quadraticamente.

Prova: Ver [1]. ■

O Teorema 3 diz que se x^* é um ponto limite da seqüência $\{x_k\}$ e se $F'(x^*)$ é não-singular, então $\{x_k\}$ converge para x^* . Além disso, quando $x^* \in \text{int}(\Omega)$ a convergência é quadrática. Se não supormos a posição de x^* , as propriedades de convergência local do método podem ser vistas no seguinte resultado.

Teorema 4 *Sejam $\{x_k\}$ a seqüência gerada pelo método proposto, e ζ_k tal que $\alpha(p_k^N) = \zeta_k p_k^N$. Suponha que exista uma solução x^* de (2.1) tal que $F'(x^*)$ é não-singular. Se a seqüência $\{x_k\}$ converge para x^* e eventualmente p_k^N resolve (2.5) e satisfaz (2.13) e (2.14), então*

1. Se ζ_k é limitada fora do 0, $\|F(x_k)\| \rightarrow 0$ q -linearmente.

2. Se $\zeta_k \rightarrow 1$, então $x_k \rightarrow x^*$ *q-superlinearmente*.
3. Se eventualmente $\zeta_k = 1$, então $x_k \rightarrow x^*$ *q-quadraticamente*.

Prova: Ver [1]. ■

Deste resultado podemos observar que a convergência local quadrática para uma solução na fronteira da caixa não é garantida. Isto deve-se ao fato de que nem sempre o passo de Newton p_k^N resolve o subproblema de região de confiança (2.5). A exigência da viabilidade estrita das iterações pode inibir a convergência local rápida já que, eventualmente, se a iteração não for viável o escalamento produzirá um passo $\alpha(p_k^N) = \zeta_k p_k^N$ tal que $\zeta_k \rightarrow 1$. Neste caso, a taxa de convergência será superlinear.

2.4 Resultados numéricos

Nesta seção apresentamos os resultados da implementação computacional do algoritmo descrito na seção 2.2. O Algoritmo foi implementado utilizando o software Matlab versão R2006a, e os parâmetros utilizados foram: $\theta = 0,99995$, $\beta_1 = 0,1$, $\beta_2 = 0,25$, $\beta_3 = 0,75$, $\alpha_1 = 0,25$ e $\alpha_2 = 0,5$.

Foram extraídos 31 problemas testes da página eletrônica [http://www. polymath-software. com/library](http://www.polymath-software.com/library), a qual possui uma série de problemas que permite ao usuário comprovar o desempenho de seus algoritmos. Todos esses problemas provém de modelos matemáticos relacionados à fenômenos físicos, e, os aqui apresentados possuem de 2 à 14 variáveis. A maioria dos problemas possuem mais de um ponto inicial. Como o método exige viabilidade estrita, foram testados apenas aqueles em que o ponto inicial estivesse no interior da região viável Ω , totalizando assim, 107 testes.

A tabela 1 mostra os resultados dos problemas, em que no modelo quadrático (2.2) foi usado exatamente $F'(x)$, ou seja, a direção de Newton. Em comparação com os resultados apresentados em [9], listamos na tabela 1 o número de testes, NT, o número de testes bem sucedidos, NS, a média de iterações, MIT e a média de avaliações da função F , MAF, para dois diferentes valores do raio inicial da região de confiança ($\Delta_0 = 1$ e $\Delta_0 = \|D_0^{-1}\nabla f_0\|$). Observamos que dos 107 testes, 74 foram bem sucedidos quando escolhemos o raio da região de confiança igual a 1 e 77 com raio igual a $\|D_0^{-1}\nabla f_0\|$. O desempenho do método foi ligeiramente afetado quando mudamos o raio de região de confiança, podemos destacar os problemas: Threeq3, Threeq5, Threeq6, Sixeq4b, Seveq1, Teneq1a, 14eq1. Em todos esses problemas o número

médio de iterações e de avaliações de F foi menor com $\Delta_0 = \|D_0^{-1}\nabla f_0\|$, mas novos testes devem ser realizados para estabelecer a escolha adequada do raio inicial de região de confiança.

O método não obteve convergência nos problemas Sixeq1 e Sixeq4a com nenhuma das estimativas iniciais. Em algumas, o número máximo de avaliações da função F foi atingido e em outras o raio da região de confiança tornou-se muito pequeno. Esses problemas são caracterizados naquela coleção, como problemas de alta dificuldade, o primeiro pode possuir solução fora do conjunto viável e a função F do segundo problema possui uma descontinuidade dentro do conjunto viável.

Tabela 1: Direções obtidas combinadas por Newton.

Problema	Raio=1				Raio= $\ D_0^{-1}\nabla f_0\ $			
	NT	NS	MIT	MAF	NT	NS	MIT	MAF
Twoeq2	4	2	8	9	4	2	5	6
Twoeq3	5	3	11	12	5	4	8	9
Twoeq4a	3	3	4	5	3	3	5	6
Twoeq4b	3	3	4	5	3	3	6	7
Twoeq5a	4	4	4	5	4	4	5	7
Twoeq5b	4	4	6	7	4	4	8	9
Twoeq6	4	4	9	15	4	4	8	12
Twoeq7	4	4	7	8	4	4	8	10
Twoeq8	4	2	4	5	4	2	4	5
Twoeq9	4	4	307	343	4	4	310	348
Twoeq10	4	4	6	7	4	4	8	10
Threeq1	4	4	21	27	4	4	27	34
Threeq2	4	1	5	6	4	1	5	6
Threeq3	4	4	14	15	4	4	5	6
Threeq4a	4	1	5	6	4	1	5	6
Threeq4b	4	4	7	8	4	4	6	8
Threeq5	4	1	20	28	4	3	6	7
Threeq6	4	4	103	144	4	4	74	109
Threeq8	1	1	6	7	1	1	5	6
Fiveq1	4	2	9	10	4	2	13	14
Sixeq1	4	0	-	-	4	0	-	-
Sixeq2a	2	1	3	4	2	1	3	4
Sixeq2b	2	1	4	5	2	1	4	5
Sixeq2c	2	1	3	4	2	1	3	4
Sixeq3	4	2	9	11	4	3	7	9
Sixeq4a	3	0	-	-	3	0	-	-
Sixeq4b	4	3	343	385	4	3	7	8
Sevneq1	3	2	22	31	3	2	12	16
Sevneq2a	3	1	5	6	3	1	4	5
Teneq1a	2	2	13	14	2	2	8	9
14eq1	2	2	10	11	2	1	5	7

Na tabela 2 apresentamos os resultados de mais 6 testes feitos com problemas encontrados em [10], os problemas possuem de 2 à 10 variáveis.

Como podemos ver, o método obteve convergência para o problema Test1 e para o problema Test25 a convergência foi obtida somente quando considerado o raio 1, no caso do raio $\|D_0^{-1}\nabla f_0\|$, o método acusou erro 3, o que significa que o tamanho da região de confiança tornou-se muito pequeno. O método acusou erro 6 no problema Test3, significando que as iterações se aproximaram muito da fronteira da caixa e portanto a matriz de escala D não pode ser calculada.

Tabela 2: Problemas diversos.

Problema	Raio=1		Raio= $\ D_0^{-1}\nabla f_0\ $	
	MIT	MAF	MIT	MAF
Test1	14	19	16	26
Test25	11	13	erro 3	
Test3	erro 6		erro 6	
Test5	não convergiu		6	7
Test38	não convergiu		não convergiu	
Test110	4	5	3	5

O método poderia ter acusado 4 outros erros: erro 1, indicando que o número máximo de iterações (1000) foi atingido, erro 2, indicando que o número máximo de avaliações da função F (1000) foi atingido, erro 4, significando que nenhuma melhora para o resíduo não-linear foi obtida, e, por fim, erro 5, indicando que a norma do gradiente escalado da função de mérito se tornou muito pequena.

Nas tabelas 3 e 4 foram feitos testes com os mesmos problemas citados acima, só que no modelo quadrático (2.2) é usado B_k em lugar de $F'(x)$, em que B_k é dada pela matriz de correção da fórmula de Broyden (1.6). Usamos como aproximação inicial B_0 para o método de Broyden a matriz $F'(x_0)$ calculada por diferenças finitas.

Observamos que em vários casos, o método convergiu para as mesmas soluções obtidas quando consideramos a direção de Newton. Acreditamos que os casos em que o método não obteve convergência, deva-se ao fato de que a direção obtida pelo método de Broyden nem sempre é uma direção de descida. Entretanto, estes testes servem para avaliar o comportamento de um método Quasenewtoniano padrão com um esquema de região de confiança para sistemas não-lineares, o que não vimos até agora na literatura. Como obtivemos a convergência para vários casos, não podemos considerar que o método não seja efetivo, dado que seu custo por iteração é muito inferior ao método de Newton para problemas com muitas variáveis, o que pode torná-lo, eventualmente, eficiente do ponto de vista computacional.

Tabela 3: Direções obtidas combinadas por Broyden.

Problema	Raio=1				Raio= $\ D_0^{-1}\nabla f_0\ $			
	NT	NS	MIT	MAF	NT	NS	MIT	MAF
Twoeq2	4	1	11	14	4	1	10	12
Twoeq3	5	2	22	28	5	3	10	12
Twoeq4a	3	3	6	7	3	3	11	8
Twoeq4b	3	2	7	8	3	3	9	10
Twoeq5a	4	4	8	11	4	3	8	9
Twoeq5b	4	2	8	9	4	3	12	14
Twoeq6	4	3	36	56	4	3	39	58
Twoeq7	4	1	8	9	4	1	8	9
Twoeq8	4	1	5	6	4	1	5	6
Twoeq9	4	0	-	-	4	0	-	-
Twoeq10	4	2	7	8	4	3	10	11
Threeq1	4	0	-	-	4	1	24	31
Threeq2	4	1	19	26	4	1	19	26
Threeq3	4	4	15	16	4	4	6	7
Threeq4a	4	1	8	9	4	1	8	9
Threeq4b	4	2	10	12	4	2	10	12
Threeq5	4	0	-	-	4	0	-	-
Threeq6	4	0	-	-	4	1	74	103
Threeq8	1	1	10	11	1	1	8	9
Fiveq1	4	2	42	55	4	2	20	25
Sixeq1	4	0	-	-	4	0	-	-
Sixeq2a	2	1	7	9	2	1	7	9
Sixeq2b	2	1	11	15	2	1	38	58
Sixeq2c	2	0	-	-	2	0	-	-
Sixeq3	4	0	-	-	4	2	23	28
Sixeq4a	3	0	-	-	3	0	-	-
Sixeq4b	4	0	-	-	4	0	-	-
Seven1	3	0	-	-	3	0	-	-
Seven2a	3	0	-	-	3	1	9	10
Teneq1a	2	0	-	-	2	0	-	-
14eq1	2	1	204	278	2	0	-	-

Tabela 4: Problemas diversos.

Problema	Raio=1		Raio= $\ D_0^{-1}\nabla f_0\ $	
	MIT	MAF	MIT	MAF
Test1	24	34	28	40
Test25	erro 3		erro 3	
Test3	erro 6		erro 6	
Test5	não convergiu		erro 3	
Test38	erro 3		erro 3	
Test110	5	7	5	7

3 *O Problema de Quadrados Mínimos Não-lineares Restrito via Penalização*

Neste capítulo trataremos de sistemas do tipo (1.1) que não possuem solução, ou seja, problemas em que não existe $x \in \mathbb{R}^n$ tal que $F(x) = 0$. Dentre os vários problemas práticos que apresentam este tipo de modelagem, podemos citar o importante problema de encontrar o mínimo corte de carga em um sistema elétrico, que aparece em sistemas de potência. Este problema será abordado na seção 3.3.1 deste trabalho.

Embora o sistema não-linear não tenha solução, em muitas aplicações é conveniente impor que algumas equações não-lineares não admitam resíduo. Desse modo, definindo $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $c : \mathbb{R}^n \rightarrow \mathbb{R}^p$, este problema, tema central deste capítulo, pode ser transformado em um problema de quadrados mínimos não-lineares restrito:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|h(x)\|_2^2 \quad \text{sujeito a } c(x) = 0. \quad (3.1)$$

Neste caso,

$$F(x) = \begin{bmatrix} h(x) \\ c(x) \end{bmatrix}.$$

Apresentaremos dois métodos de resolver (3.1): o primeiro trabalha com as condições de KKT [6] do problema, e o segundo, descrito na seção 3.2, substitui o problema original restrito por uma seqüência de problemas irrestritos.

3.1 Método Newton-Lagrange

Uma forma de resolver (3.1) é aplicar o método de Newton para as condições de otimalidade de KKT do problema.

Considerando a função Lagrangeana

$$\mathcal{L}(x, \lambda) = \frac{1}{2}h(x)^T h(x) - \lambda^T c(x)$$

e $A(x) \in \mathbb{R}^{p \times n}$ como a matriz Jacobiana das restrições, ou seja,

$$A(x)^T = [\nabla c_1(x), \nabla c_2(x), \dots, \nabla c_p(x)],$$

temos que as condições de primeira ordem para o caso de igualdade restrita formam o seguinte sistema de $n + p$ equações em $n + p$ incógnitas x e λ :

$$G(x, \lambda) = \begin{pmatrix} R(x)^T h(x) - A(x)^T \lambda \\ c(x) \end{pmatrix} = 0,$$

em que $R(x)^T = [\nabla h_1(x), \nabla h_2(x), \dots, \nabla h_m(x)]$.

Neste momento temos então um sistema de equações não-lineares dado pelo gradiente do Lagrangiano. Podemos assim, utilizar o método de Newton para obter a solução (x^*, λ^*) .

O Jacobiano de $G(x, \lambda)$ é dado por

$$W(x, \lambda) = \begin{pmatrix} R(x)^T R(x) + \sum_{i=1}^m h_i(x) \nabla^2 h_i(x) - \sum_{i=1}^p \lambda_i \nabla^2 c_i(x) & -A(x)^T \\ A(x) & 0 \end{pmatrix}.$$

Na k -ésima iteração do método de Newton temos que resolver o sistema linear,

$$W(x_k, \lambda_k) \begin{pmatrix} \Delta x_k \\ \Delta \lambda_k \end{pmatrix} = \begin{pmatrix} -R(x_k)^T h(x_k) + A(x_k)^T \lambda_k \\ -c(x_k) \end{pmatrix}. \quad (3.2)$$

Assim, sendo α_k a correção do passo de Newton obtido pela busca linear na direção obtida por (3.2), obtemos,

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k \Delta x_k; \\ \lambda_{k+1} &= \lambda_k + \alpha_k \Delta \lambda_k. \end{aligned}$$

A busca linear é um dos possíveis métodos de globalização do método de Newton. Neste

trabalho o parâmetro α_k é obtido pela minimização unidimensional da seguinte função de mérito:

$$M(x, \lambda) = \frac{1}{2} \nabla \mathcal{L}(x, \lambda)^T \nabla \mathcal{L}(x, \lambda).$$

Como é conhecido, o processo de minimização unidimensional não deve ser exato e, portanto, várias estratégias inexatas de busca podem ser empregadas.

3.2 Penalização

Uma importante classe de métodos para otimização restrita procura converter problemas complexos em uma subsequência de problemas irrestritos. Um importante e bem conhecido procedimento é o chamado método de penalidade, em que a não satisfação de uma restrição é sancionada com o acréscimo de um parâmetro de penalização para cada restrição na função objetivo, de maneira que as restrições sejam incorporadas na função objetivo a qual queremos minimizar.

Nesta seção discutiremos a penalidade externa, muitas vezes chamada simplesmente de penalidade. Na penalidade externa o custo do parâmetro acrescentado na função objetivo aumenta proporcionalmente à medida que as restrições são violadas. A solução de um problema penalizado externamente geralmente está fora do conjunto viável, mas tende à viabilidade à medida que o parâmetro de penalidade cresce.

O princípio da penalidade externa é a utilização de uma função contínua que se anula no conjunto a ser penalizado e é positiva fora dele. Assim, definindo $\Omega_1 = \{x \in \mathbb{R}^n; c(x) = 0\}$, para penalizar o problema (3.1) basta escolher $P : \mathbb{R}^n \rightarrow \mathbb{R}$, $P \in C^0(\mathbb{R}^n)$ tal que

$$P(x) \begin{cases} = 0, & \text{se } x \in \Omega_1; \\ > 0, & \text{se } x \notin \Omega_1. \end{cases}$$

Dessa maneira, a função objetivo para o problema penalizado $\mathcal{P}_\rho(x)$ associado à (3.1) é

$$\mathcal{P}_\rho(x) = \frac{1}{2} \|h(x)\|_2^2 + \rho P(x),$$

em que $\rho > 0$ é o parâmetro de penalidade. Portanto, o problema penalizado associado à (3.1) é

$$\min_{x \in \mathbb{R}^n} \mathcal{P}_\rho(x). \quad (3.3)$$

Note que, uma vez fixado $\rho \in \mathbb{R}$, a resolução de (3.3) produz uma solução $x(\rho)$, que, como veremos mais adiante, para um aumento gradativo do parâmetro de penalização ρ e uma escolha propícia para $P(x)$, geramos uma seqüência que converge para a solução do problema original (3.1).

Poderíamos considerar as funções de penalidade exatas, em que a solução do problema penalizado (3.3) é exatamente a solução do problema original (3.1), para um valor finito do parâmetro de penalização. A vantagem é que com estas funções, não seria preciso resolver uma seqüência infinita de subproblemas.

Entretanto, uma desvantagem é que a maioria das funções de penalidade exatas são não-diferenciáveis na solução. Uma dessas funções bem conhecida é a baseada na $\|\cdot\|_1$ que para o conjunto Ω_1 é da forma

$$P(x) = \sum_{i=1}^p |c_i(x)| = \|c(x)\|_1. \quad (3.4)$$

Propriedades de convergência dos subproblemas penalizados associados à (3.1), para um parâmetro ρ finito e para a função (3.4) são citadas em [11]. Dificuldades podem surgir se o parâmetro ρ não for escolhido adequadamente, pois a convergência da solução do subproblema para a solução do problema original depende muito da escolha de ρ . Se ρ for muito pequeno, a função penalizada pode ser inferiormente ilimitada. Por outro lado, se ρ for muito grande, surgem os problemas de mal condicionamento.

Neste trabalho, vamos considerar a seguinte função de penalidade

$$P(x) = \sum_{i=1}^p c_i^2(x) = \|c(x)\|_2^2,$$

também conhecida como função de penalidade quadrática. Assim, dado $\rho > 0$, definimos

$$Q_\rho(x) = \frac{1}{2} \|h(x)\|_2^2 + \frac{\rho}{2} \|c(x)\|_2^2 \quad (3.5)$$

e resolvemos, para uma seqüência $\{\rho_k\}_{k \in \mathbb{N}}$, com $\rho_k \rightarrow \infty$, o problema irrestrito

$$\min_{x \in \mathbb{R}^n} Q_{\rho_k}(x).$$

As considerações acima levam ao seguinte algoritmo

Algoritmo 3: Penalidade Externa

Entrada: $\rho_1 \geq 0$, $x_0 \in \mathbb{R}^n$, $\epsilon_0 > 0$

1 $k = 1$

2 Calcular $x_k = x(\rho_k)$ como solução aproximada de

$$\min_{x \in \mathbb{R}^n} Q_{\rho_k}(x) \quad (3.6)$$

3 e terminar quando $\|\nabla Q_{\rho_k}(x)\| \leq \epsilon_k$.

4 Escolher $\rho_{k+1} > \rho_k$

5 $k = k + 1$ e voltar ao passo 2.

Segundo [11] e [6], os parâmetros de penalização podem ser atualizados da seguinte forma: $\rho_1 = 1$, $\rho_k = 10\rho_{k-1}$, e ainda aconselha-nos a usar x_{k-1} como ponto inicial para resolver o subproblema (3.6).

Note que, resolver (3.6) equivale a resolver o seguinte problema

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \left\| \begin{array}{c} h(x) \\ \sqrt{\rho_k} c(x) \end{array} \right\|_2^2, \quad (3.7)$$

o qual pode ser resolvido, por exemplo, pelo método de Levenberg-Marquardt da seção 1.5.

Assim, dado $x_k \in \mathbb{R}^n$, considere

$$F_{\rho_k}(x_k) = \begin{pmatrix} h(x_k) \\ \sqrt{\rho_k} c(x_k) \end{pmatrix}$$

e defina $J_{\rho_k}(x_k)$ o Jacobiano de F_{ρ_k} em x_k . Portanto, um passo do método Levenberg-Marquardt a partir de x_k para o problema (3.7) consiste em resolver o sistema linear

$$(J_{\rho_k}(x_k)^T J_{\rho_k}(x_k) + \eta_k I) p_k = -J_{\rho_k}(x_k)^T F_{\rho_k}(x_k), \quad (3.8)$$

para algum $\eta_k > 0$ devidamente escolhido, chamado neste trabalho de parâmetro de Levenberg-Marquardt.

Note que resolver o sistema (3.8) é equivalente a resolver o subproblema

$$\min_{p \in \mathbb{R}^n} \|J_{\rho_k}(x_k)p + F_{\rho_k}(x_k)\|_2^2 + \eta_k \|p\|_2^2,$$

ou ainda,

$$\min_{p \in \mathbb{R}^n} \left\| \begin{pmatrix} J_{\rho_k}(x_k) \\ \sqrt{\eta_k} I \end{pmatrix} p + \begin{pmatrix} F_{\rho_k}(x_k) \\ 0 \end{pmatrix} \right\|_2^2 \quad (3.9)$$

que é um problema de quadrados mínimos linear.

Uma outra alternativa para resolver o problema (3.7) é usar um método diretamente para $F_{\rho_k}(x) = 0$, por exemplo o método de Newton. Esta estratégia será empregada nos testes numéricos.

3.2.1 Análise de convergência

Para a análise de convergência do método de penalidade quadrática vamos definir $f(x) = \frac{1}{2} \|h(x)\|_2^2$ em (3.5).

Teorema 5 [6, pag. 496] *Suponha que cada x_k é o minimizador global de $Q_{\rho_k}(x)$ no Algoritmo 3 e que $\rho_k \rightarrow \infty$. Então todo ponto limite x^* da seqüência $\{x_k\}$ é uma solução do problema (3.1).*

Prova: Seja \tilde{x} uma solução global de (3.1), isto é

$$f(\tilde{x}) \leq f(x), \quad \text{para todo } x \text{ com } c_i(x) = 0, i \in \mathcal{E},$$

em que \mathcal{E} representa um conjunto finito de índices e $[c_i(x)]_{i \in \mathcal{E}} = c(x)$.

Como x_k minimiza $Q_{\rho_k}(\cdot)$ para cada k , nós temos que $Q_{\rho_k}(x_k) \leq Q_{\rho_k}(\tilde{x})$, assim obtemos a desigualdade

$$f(x_k) + \frac{\rho_k}{2} \sum_{i \in \mathcal{E}} c_i^2(x_k) \leq f(\tilde{x}) + \frac{\rho_k}{2} \sum_{i \in \mathcal{E}} c_i^2(\tilde{x}) = f(\tilde{x}). \quad (3.10)$$

Organizando a equação obtemos

$$\sum_{i \in \mathcal{E}} c_i^2(x_k) \leq 2 \frac{[f(\tilde{x}) - f(x_k)]}{\rho_k}. \quad (3.11)$$

Suponha que x^* é um ponto limite de $\{x_k\}$, de modo que exista uma subsequência \mathcal{K} tal que

$$\lim_{k \in \mathcal{K}} x_k = x^*.$$

Tomando o limite com $k \rightarrow \infty, k \in \mathcal{K}$, em ambos os lados de (3.11), nós obtemos

$$\sum_{i \in \mathcal{E}} c_i^2(x^*) = \lim_{k \in \mathcal{K}} \sum_{i \in \mathcal{E}} c_i^2(x_k) \leq \lim_{k \in \mathcal{K}} 2 \frac{[f(\bar{x}) - f(x_k)]}{\rho_k} = 0,$$

em que a última desigualdade segue de $\rho_k \rightarrow \infty$. Agora, x^* é viável, pois $c_i(x^*) = 0$ para todo $i \in \mathcal{E}$. Além disso, tomando o limite com $k \rightarrow \infty, k \in \mathcal{K}$ em (3.10), temos pela não-negatividade de ρ_k e de cada $c_i^2(x_k)$ que

$$f(x^*) \leq f(x^*) + \lim_{k \in \mathcal{K}} \frac{\rho_k}{2} \sum_{i \in \mathcal{E}} c_i^2(x_k) \leq f(\bar{x}),$$

Como x^* é um ponto viável e $f(x^*) \leq f(\bar{x})$, concluímos que x^* é também um minimizador global. ■

Para que um ponto limite da seqüência x_k seja solução do problema (3.1), o teorema apresentado acima exige que resolvamos o passo 2 do Algoritmo 3 exatamente. Como nem sempre isso é possível, o próximo teorema apresenta-nos propriedades de convergência da seqüência x_k quando diminuimos a tolerância, ou seja, quando permitimos a minimização inexata de Q_{ρ_k} .

Teorema 6 [6, pag. 497] *Se a tolerância ϵ_k no Algoritmo 3 satisfaz*

$$\lim_{k \rightarrow \infty} \epsilon_k = 0$$

e o parâmetro de penalidade satisfaz $\rho_k \rightarrow \infty$, então para todos os pontos limites x^ da seqüência $\{x_k\}$ no qual os gradientes das restrições $\nabla c_i(x^*)$ são linearmente independentes, temos que x^* é um ponto KKT para o problema (3.1). Além disso para uma subsequência infinita \mathcal{K} tal que $\lim_{k \in \mathcal{K}} x_k = x^*$ temos que*

$$\lim_{k \in \mathcal{K}} -\rho_k c_i(x_k) = \lambda_i^*, \text{ para todo } i \in \mathcal{E},$$

em que λ^ é o vetor contendo os multiplicadores de Lagrange que satisfaz as condições KKT.*

Prova: Diferenciando $Q_{\rho_k}(x)$, obtemos

$$\nabla Q_{\rho_k}(x_k) = \nabla f(x_k) + \sum_{i \in \mathcal{E}} \rho_k c_i(x_k) \nabla c_i(x_k).$$

Aplicando o critério de parada do Algoritmo 3, temos que

$$\left\| \nabla f(x_k) + \sum_{i \in \mathcal{E}} \rho_k c_i(x_k) \nabla c_i(x_k) \right\| \leq \epsilon_k. \quad (3.12)$$

Reagrupando esta expressão (em particular usando a desigualdade $\|a\| - \|b\| \leq \|a + b\|$), obtemos

$$\left\| \sum_{i \in \mathcal{E}} c_i(x_k) \nabla c_i(x_k) \right\| \leq \frac{1}{\rho_k} [\epsilon_k + \|\nabla f(x_k)\|]. \quad (3.13)$$

Quando tomamos o limite com $k \rightarrow \infty$ para $k \in \mathcal{K}$ o termo entre colchetes do lado direito tende à $\|\nabla f(x^*)\|$ e, como $\rho_k \rightarrow \infty$, o lado direito tende a zero. Assim, tomando o limite $k \rightarrow \infty$ em (3.13), obtemos

$$\sum_{i \in \mathcal{E}} c_i(x^*) \nabla c_i(x^*) = 0.$$

Como por hipótese os gradientes das restrições $\nabla c_i(x^*)$ são linearmente independentes, temos que $c_i(x^*) = 0$ para todo $i \in \mathcal{E}$ e portanto x^* é viável.

Usando $A(x)$ para denotar a matriz dos gradientes das restrições, isto é,

$$A(x)^T = [\nabla c_i(x)]_{i \in \mathcal{E}},$$

e λ_k para denotar o vetor $-\rho_k c(x_k)$, temos como em (3.12) que

$$A(x_k)^T \lambda_k = \nabla f(x_k) - \nabla Q_{\rho_k}(x_k), \quad \|\nabla Q_{\rho_k}(x_k)\| \leq \epsilon_k.$$

Para todo $k \in \mathcal{K}$ suficientemente grande, a matriz $A(x_k)$ tem posto completo, e portanto $A(x_k)A(x_k)^T$ é não-singular. Multiplicando a última expressão por $A(x_k)$ e reagrupando os termos obtemos

$$\lambda_k = [A(x_k)A(x_k)^T]^{-1} A(x_k) [\nabla f(x_k) - \nabla Q_{\rho_k}(x_k)].$$

Assim, tomando o limite quando $k \xrightarrow{k \in \mathcal{K}} \infty$, encontramos

$$\lim_{k \in \mathcal{K}} \lambda_k = \lambda^* = [A(x^*)A(x^*)^T]^{-1}A(x^*)\nabla f(x^*).$$

Tomando o limite em (3.12), concluimos que

$$\nabla f(x^*) - A(x^*)^T \lambda^* = 0$$

Assim, x^* é um ponto KKT, com um único vetor multiplicador de Lagrange. ■

Com isso, usando os teoremas anteriores garantimos convergência para pontos estacionários do problema (3.1), desde que para um dado parâmetro de penalidade ρ_k sejam encontrados pontos estacionários do subproblema irrestrito (3.7).

3.3 Resultados numéricos

O método de penalização descrito na seção 3.2 e o método Newton-Lagrange, da seção 3.1 serão agora aplicados ao problema de fluxo de carga em redes de energia elétrica [12] e à problemas diversos da literatura encontrados em [10] e [13]. Daremos nessa primeira seção uma breve explicação sobre a formulação do problema de fluxo de carga.

3.3.1 Fluxo de potência

O cálculo de fluxo de potência (ou fluxo de carga) em uma rede de energia elétrica consiste, essencialmente, na determinação do estado da rede, da distribuição dos fluxos e de outras grandezas de interesse. Nesse tipo de problema, a rede é representada por um conjunto de equações e inequações algébricas.

Classificamos dois grupos que compõem um sistema de energia elétrica: as barras, que compreendem os geradores, cargas, reatores e capacitores, e os circuitos, que compreendem as linhas de transmissão, transformadores e defasadores (elementos que interligam as barras).

As equações algébricas básicas do fluxo de carga são obtidas impondo-se a conservação das potências ativa e reativa em cada nó da rede, ou seja, a potência líquida injetada deve ser igual a soma das potências que fluem pelos componentes internos das barras.

A cada barra são definidas 4 variáveis correspondentes à tensão nodal e à injeção de potên-

cia na barra:

- V_k - magnitude da tensão nodal;
- θ_k - ângulo da tensão nodal;
- P_k - injeção líquida de potência ativa;
- Q_k - injeção líquida de potência reativa.

Duas dessas variáveis são dados do problema e as duas outras devem ser calculadas. Dependendo de quais são dadas e quais são incógnitas, definem-se três tipos básicos de barras:

Tabela 5: Tipos de barras.

Tipo	Dados	Incógnitas	Características
PQ	P_k, Q_k	V_k, θ_k	Representam as barras de carga
PV	P_k, V_k	Q_k, θ_k	Representam as barras de geração
Folga	V_k, θ_k	P_k, Q_k	Fornecem a referência angular e fecham o balanço de potência do sistema

O conjunto de equações do problema do fluxo de carga é formado por duas equações para cada barra

$$P_k = \sum_{m \in \Omega_k} P_{km}(V_k, V_m, \theta_k, \theta_m) \quad \text{e} \quad Q_k = \sum_{m \in \Omega_k} Q_{km}(V_k, V_m, \theta_k, \theta_m)$$

em que $k = 1, \dots, nb$, sendo nb o número de barras e Ω_k é o conjunto de barras vizinhas à barra k (duas barras são vizinhas quando existe um circuito interligando-as).

Seja I_{km} a corrente em uma linha de transmissão (que liga a barra k à barra m). A injeção líquida de corrente na barra k é obtida aplicando-se a primeira lei de Kirchhoff:

$$I_k + I_k^{sh} = \sum_{m \in \Omega_k} I_{km}, \quad k = 1, \dots, nb.$$

Esta expressão, utilizando as equações nodais da rede elétrica, pode ser escrita na seguinte forma matricial:

$$I = YE,$$

em que I é o vetor das injeções de corrente, cujas componentes são I_k , $k = 1, \dots, nb$; o vetor E representa as tensões nodais cujas componentes são $E_k = V_k e^{j\theta_k}$ e a matriz $Y = G + jB$ é denominada matriz de admitância nodal, sendo G a matriz de condutância e B a matriz de susceptância. Os elementos da matriz Y são obtidos da seguinte maneira:

$$Y_{km} = -y_{km} \quad \text{e} \quad Y_{kk} = jb_k^{sh} + \sum_{m \in \Omega_k} (jb_{km}^{sh} + y_{km}),$$

em que b_k^{sh} corresponde à susceptância de equipamentos reativos conectados à barra k , b_{km}^{sh} é a metade da susceptância shunt do circuito conectando as barras k e m ; e y_{km} é a admitância série do circuito conectando as barras k e m . Vale a pena observar que, em geral, a matriz é esparsa, pois $Y_{km} = 0$ sempre que entre as barras k e m não existirem circuitos conectando-as.

As equações de potência ativa e reativa são deduzidas aplicando as leis de Kirchhoff, e são dadas por:

$$P_k^{cal} = G_{kk} V_k^2 + V_k \sum_{m \in \Omega_k} V_m [G_{km} \cos(\theta_k - \theta_m) + B_{km} \sin(\theta_k - \theta_m)] \quad (3.14)$$

$$Q_k^{cal} = -B_{kk} V_k^2 + V_k \sum_{m \in \Omega_k} V_m [G_{km} \cos(\theta_k - \theta_m) + B_{km} \sin(\theta_k - \theta_m)] \quad (3.15)$$

em que $k = 1, \dots, nb$ e Ω_k é o conjunto de índices das barras vizinhas à barra k , excluindo a própria barra k .

3.3.1.1 Modelagem do problema

As equações básicas do fluxo de carga representadas por (3.14) e (3.15), formam um sistema de $2nb$ equações com $4nb$ incógnitas, com as seguintes características:

- nb equações do tipo (3.14);
- nb equações do tipo (3.15);
- $4nb$ variáveis (para cada barra k associam-se V_k , θ_k , P_k , Q_k).

O fluxo de carga básico em redes de energia consiste em resolver este problema definindo parte das variáveis ($2nb$ variáveis) e calculando as demais. Consideremos o problema em que são dados P_k e Q_k para as barras PQ , P_k e V_k para as barras PV e V_k e θ_k para a barra $V\theta$ (folga)

e pede-se para calcular V_k e θ_k nas barras PQ , θ_k nas barras PV e P_k e Q_k na barra de folga (consideramos apenas uma barra de folga).

Sejam npq e npv , respectivamente, o número de barras PQ e PV da rede. Podemos decompor o problema descrito acima, em dois subsistemas de equações algébricas:

Subsistema 1: Neste subproblema são dados P_k e Q_k nas barras PQ e P_k e V_k nas barras PV . Pretende-se calcular V_k e θ_k nas barras PQ e θ_k nas barras PV , ou seja, trata-se de um sistema de $(2npq + npv)$ equações algébricas não-lineares com o mesmo número de incógnitas:

$$\begin{aligned} P_k^{dado} - P_k^{cal} &= 0 && \text{para as barras } PQ \text{ e } PV; \\ Q_k^{dado} - Q_k^{cal} &= 0 && \text{para as barras } PQ. \end{aligned}$$

Resolvido o Subsistema 1 será conhecido o estado (V_k, θ_k) para todas as barras da rede.

Subsistema 2: Conhecidos (V_k, θ_k) , deseja-se calcular P_k e Q_k na barra de folga, e Q_k nas barras PV . Trata-se de um sistema com $npv+2$ equações algébricas não-lineares com o mesmo número de incógnitas.

As incógnitas do Subsistema 1 podem ser agrupadas no vetor x dado a seguir:

$$x = \begin{bmatrix} V \\ \theta \end{bmatrix}$$

em que V é o vetor de magnitude das tensões das barras PQ e θ é o vetor dos ângulos das tensões das barras PQ e PV . As expressões que formam o Subsistema 1, podem ser reescritas do seguinte modo:

$$\begin{aligned} \Delta P_k &= P_k^{dado} - P_k^{cal} = 0 && \text{para as barras } PQ \text{ e } PV; \\ \Delta Q_k &= Q_k^{dado} - Q_k^{cal} = 0 && \text{para as barras } PQ. \end{aligned}$$

em que ΔP_k e ΔQ_k são respectivamente os balanços de potência ativa e reativa na barra k .

As funções ΔP_k e ΔQ_k podem ser colocadas na forma vetorial

$$\begin{aligned} \Delta P &= P^{dado} - P^{cal}(V, \theta) \\ \Delta Q &= Q^{dado} - Q^{cal}(V, \theta) \end{aligned}$$

em que $P^{cal}(V, \theta)$ é o vetor das injeções de potência ativa nas barras PQ e PV , e $Q^{cal}(V, \theta)$ o das injeções de potência reativa nas barras PQ .

Considerando a função vetorial dada por

$$F(x) = \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix}$$

o Subsistema 1 pode ser colocado na forma

$$F(x) = 0.$$

3.3.1.2 O problema sem solução

Pode-se dividir o estado da rede em três níveis: região de operação, região de emergência e região sem solução real.

A região de operação caracteriza-se por apresentar pontos em que as equações estáticas do fluxo de carga possuem solução real $F(x) = 0$ e não há violação dos limites operacionais. É nessa região que se deseja que os sistemas de energia elétrica operem.

Na região de emergência, as equações estáticas do fluxo de carga apresentam solução real, porém com violações de um ou mais limites operacionais. A princípio, é possível operar nesta região por um intervalo de tempo limitado, mas o importante é, a partir de um ponto de operação nessa região, utilizar mecanismos que possibilitem a migração desse ponto para a região de operação.

A região sem solução é aquela onde as equações estáticas do fluxo de carga não apresentam solução real. Qualquer tentativa de operar no sistema nesta região pode causar instabilidade na tensão e até mesmo o colapso da tensão. Esse fenômeno pode ser observado em duas situações: quando o sistema elétrico torna-se carregado devido ao aumento de demanda e/ou quando o sistema sofre uma contingência severa.

Vamos utilizar-nos do procedimento de penalização para resolver o problema de fluxo de carga, visando encontrar a partir de um ponto na região sem solução, um ponto na região de emergência. Será considerada a existência de barras de injeção nula, nas quais não pode haver resíduos de potências. Assim, ao transformar uma barra PV (digamos a barra k) em uma barra PQ , exige-se que $\Delta P_k = 0$ e $\Delta Q_k = 0$, ou seja, nessa barra não pode haver resíduos de potências ativa e reativa. Isto significa que as barras PV transformadas em PQ passam a ser tratadas como barras de injeção nula.

Como em [14] vamos considerar o seguinte problema

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|h(x)\|_2^2$$

s.a $c(x) = 0$

sendo que no vetor $c(x) : \mathbb{R}^n \rightarrow \mathbb{R}^p$ estão contidas as equações dos balanços de potências $(\Delta P, \Delta Q)$ para as barras de injeção nula, e em $h(x)$ estão as demais equações de $F(x)$. O valor de p depende do sistema elétrico em questão, sendo que em problemas reais, p encontra-se entre 10% a 15% do número de barras.

Obtemos então um problema na forma de (3.1), que será resolvido através da teoria descrita nas seções anteriores.

3.3.1.3 Resultados

Os métodos Newton-Lagrange e de Penalização foram implementados utilizando o software Matlab versão R2006a. Fizemos os testes para o problema de fluxo de potência em que as estimativas iniciais utilizadas foram $V_0 = [1; 1; \dots; 1]$, $\theta_0 = [0; 0; \dots; 0]$. Para o método de penalidade $\rho_1 = 1$ e $\rho_{k+1} = 10 \cdot \rho_k$ e para o método de Newton-Lagrange $\lambda_0 = [0; 0; \dots; 0]$.

Os resultados referente ao método Newton-Lagrangeano para os sistemas de 6, 30 e 57 barras são apresentados nas tabelas 6, 7 e 8 respectivamente, que mostram a evolução da convergência do gradiente do Lagrangeano em cada iteração.

Tabela 6: Sistema de 6 barras

Iteração	Grad. do Lagrangeano
1	4.7098e+00
2	1.3514e+00
3	3.0659e-01
4	1.1519e-01
5	2.7824e-02
6	1.5370e-03
7	5.3066e-06

Tabela 7: Sistema de 30 barras

Iteração	Grad. do Lagrangeano
1	1.0494e+01
2	1.0141e+00
3	3.4897e-01
4	1.1138e-01
5	2.9974e-02
6	5.6653e-03
7	1.1400e-03
8	6.4668e-05

Tabela 8: Sistema de 57 barras

Iteração	Grad. do Lagrangeano
1	2.1407e+02
2	7.5749e+01
3	1.9875e+01
4	1.3723e+01
5	7.4064e+00
6	4.5129e+00
7	2.4919e+00
8	1.3279e+00
9	5.9365e-01
10	6.9662e-01
11	2.2048e-01
12	3.0670e-02
13	1.6098e-02
14	7.8667e-03
15	2.3935e-03
16	2.6047e-04

Podemos observar que o processo de convergência fica gradativamente mais lento a medida que a dimensão aumenta (7 iterações em 6 barras e 16 iterações em 57 barras).

Um fato relevante a ser destacado é que, para as dimensões acima reportadas, o método de Newton não exigiu o procedimento de busca, ou seja, $\alpha_k = 1$ em todas as iterações. Isto geralmente não ocorre com dimensões maiores, em que o procedimento de busca é frequentemente requisitado para obter um decréscimo suficiente.

Apesar do passo completo de Newton ser aceito, esperar-se-ia que o processo de convergência apresentasse ordem quadrática [5]. Este fato tem repercussões tanto teóricas quanto práticas. Do ponto de vista teórico, este comportamento linear da convergência do método de Newton nos leva a supor que a Hessiana do Lagrangiano é singular na solução. Este fato também nos conduz a cogitar como estaria se comportando o ângulo das direções de Newton com o gradiente do Lagrangiano, dado que o avanço em cada iteração é relativamente pequeno.

Do ponto de vista prático, as observações acima induzem a um estudo numérico e também da estrutura do problema. Na prática, pelo comportamento dos resultados, devemos verificar um mau condicionamento da matriz Jacobiana, em paralelo à sua singularidade teórica.

O emprego de um método iterativo linear, considerando futuras abordagens ao problema com dimensão maior, poderia ser aconselhável. Entretanto, os resultados acima mostram, em princípio, que esta estratégia pode apresentar resultados indesejáveis, já que um preconditionador baseado numa fatoração incompleta de uma matriz mal condicionada dificilmente apresentará bons resultados. Uma observação na estrutura da matriz é necessária para afirmar com

mais clareza sobre usar ou não os métodos iterativos.

Os mesmos 3 sistemas foram testados com o método de penalização. A tentativa de resolver $F_{\rho_k}(x) = 0$ pelo método de Newton puro gerou um mal condicionamento no Jacobiano do sistema. Este fato parece ter uma relação com a convergência linear obtida pelo método Newton-Lagrange. Assim, resolvemos o problema (3.7) pelo método de Levenberg-Marquardt da seção 1.5.2 com a atualização $\eta_k = \|F_{\rho_k}\|_2^2$ e utilizando a função `\` do Matlab para o problema (3.9). No problema do fluxo de potência, o subsistema 1 possui a mesma quantidade de incógnitas e de equações, desta forma a função `\` do Matlab resolverá o problema (3.9) através da eliminação gaussiana.

O número de iterações referente ao método de penalização descrito na tabela 9 refere-se à soma do número de iterações para resolver (3.9) para cada k .

Tabela 9: Número de iterações.

Problema	Iterações	
	Penalização	Newton-Lagrange
sistema de 6 barras	17	7
sistema de 30 barras	20	8
sistema de 57 barras	41	16

Quando utilizamos o método de penalização, precisamos resolver vários problemas do tipo (3.9) de dimensão n (n colunas), enquanto que na utilização do método Newton-Lagrange o sistema (3.2) tem dimensão $n + p$. A tabela 10 apresenta as dimensões dos sistemas de cada método, relacionados aos problemas de 6, 30 e 57 barras. Podemos observar, conforme a tabela 9, que apesar de o método de penalização levar mais iterações para alcançar a solução, os sistemas relacionados ao método são menores. Além disso, os sistemas relacionados ao método de penalização trabalham apenas com o Jacobiano da função F_{ρ_k} , enquanto que os relacionados ao método Newton-Lagrange utilizam a Hessiana da função f e o Jacobiano das restrições.

Tabela 10: Dimensões dos sistemas.

Problema	Penalização	Newton-Lagrange
sistema de 6 barras	9	11
sistema de 30 barras	53	65
sistema de 57 barras	106	138

3.3.2 Outras aplicações

Para analisar o desempenho do método de penalização, outros testes foram feitos com problemas de [10] e [13]. O método convergiu para a solução em 13 problemas dos 17 testados.

Nesses testes o problema (3.7) foi resolvido através do método de Newton puro para sistemas não-lineares, ou seja, para cada ρ_k encontramos a solução de

$$F_{\rho_k}(x) = 0. \quad (3.16)$$

Ao resolver o problema (3.16), se atingirmos o número máximo de iterações, ou seja, se o critério de convergência não for satisfeito, o valor obtido como ponto inicial para a iteração seguinte do Algoritmo 3 pode não ser um valor satisfatório. Acreditamos que o mal desempenho do método nos problemas Test27, Test42 e Test316 se deva ao fato de usarmos o método de Newton puro, pois, para todo ρ_k , ao resolvermos o sistema (3.16) atingimos o número máximo de iterações. Isso indica que um algoritmo mais adequado deve ser usado na resolução dos subproblemas (3.7).

Problema	Resultado
Test26	não convergiu
Test27	não convergiu
Test28	convergiu
Test42	não convergiu
Test48	convergiu
Test49	convergiu
Test50	convergiu
Test51	convergiu
Test52	convergiu
Test77	convergiu
Test79	convergiu
Test216	convergiu
Test269	convergiu
Test316	não convergiu
Test344	convergiu
Test345	convergiu
Test373	convergiu

Conclusão

Neste trabalho estudamos e implementamos métodos alternativos para resolução de alguns sistemas não-lineares especiais.

Os sistemas não-lineares com restrições de caixas, foram abordados com técnicas de região de confiança adequadas às restrições seguindo as linhas gerais de [1]. Implementamos o método dogleg com combinações de direções de Cauchy-Newton e Cauchy-Broyden. Este estudo nos facilitou a compreensão da importância do escalamento das regiões de confiança para as restrições de caixa, assim como a robustez da direção de Newton para o bom desempenho do método. Ressaltamos que o método Broyden por ter obtido a convergência para a mesma solução de Newton para vários problemas, pode ser uma alternativa para sistemas grandes, em que o custo do cálculo da direção é essencial. Uma alternativa para estes casos seria o método de região de confiança do gradiente conjugado de Steihaug [6]. Nele a direção inicial de Cauchy vai sendo atualizada passo a passo pelas fórmulas CG, parando em geral muito antes do passo final, que seria Newton.

Os problemas de quadrados mínimos não-lineares são muito representativos. O caso das equações da rede elétrica é um modelo relevante em engenharia. Neste caso, apresentamos este modelo com detalhes como parte do nosso estudo. O problema clássico nessa aplicação típica é o caso particular denominado Fluxo de Potências, em que o modelo é um sistema não-linear usual. A opção de incorporar equações que não apresentam zeros permite tornar o modelo mais representativo em termos da engenharia elétrica. Para estes problemas apresentamos uma idéia nova: resolvê-lo usando a estrutura mais próxima possível do sistema não-linear associado. Mais precisamente, ao invés do sistema não-linear com penalização diretamente, usamos o método de Levenberg-Marquardt. Esta é uma alternativa interessante ao método do Lagrangeano que usa sistemas lineares maiores além de informações de segunda ordem. Embora tenhamos verificado que o número de iterações lineares aumenta, o importante aqui é ressaltar a convergência do método à mesma solução. Isso nos permite cogitar futuras pesquisas usando estas idéias com um conjunto de testes mais representativos. Em alguns testes menores, extraídos de [10] e [13], verificamos que trabalhar com o sistema não-linear penalizado diretamente resulta em convergência à mesma solução da formulação de quadrados mínimos. Finalmente ressaltamos que o contexto deste trabalho foi empregar métodos numéricos baseados em mo-

delos mais simples para modelos mais elaborados. Este estudo, assim como a implementação computacional nos permitiram aprender conceitos básicos porém fundamentais em otimização numérica.

Referências

- [1] BELLAVIA, S.; MACCONI, M.; MORINI, B. An affine scaling trust-region approach to bound-constrained nonlinear systems. *Applied Numerical Mathematics*, v. 44, p. 257–280, 2003.
- [2] COLEMAN, T. F.; LI, Y. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, v. 6, p. 418–445, 1996.
- [3] CONN, A. R.; GOULD, N. I. M.; TOINT, P. L. *Trust-region methods*. Philadelphia: Society for Industrial and Applied Mathematics, 2000. 959 p. (MPS-SIAM series on optimization). ISBN 0-89871-460-5.
- [4] BARBOZA, L. V. *Análise do Máximo Carregamento de Sistemas de Potência via Métodos de Pontos Interiores*. 208 f. Dissertação (Pós-Graduação em Engenharia Elétrica) — Universidade Federal de Santa Catarina, Santa Catarina, 1997.
- [5] DENNIS JR., J. E.; SCHNABEL, R. B. *Numerical methods for unconstrained optimization and nonlinear equations*. Philadelphia: Society for Industrial and Applied Mathematics, 1996. 378 p. (Classics in applied mathematics, 16). ISBN 0-89871-364-1.
- [6] NOCEDAL, J.; WRIGHT, S. J. *Numerical optimization*. New York: Springer-Verlag New York, Inc., 1999. 636 p. (Springer Series in Operations Research). ISBN 0-387-98793-2.
- [7] GILL, P. E.; MURRAY, W.; WRIGHT, M. H. *Practical optimization*. London: Academic Press, 1981. 401 p. ISBN 0-12-283952-8.
- [8] FRANCISCO, J. B.; KREJIĆ, N.; MARTÍNEZ, J. M. An interior-point method for solving box-constrained underdetermined nonlinear systems. *Journal of Computational and Applied Mathematics*, v. 177, p. 67–88, 2005.
- [9] BELLAVIA, S.; MACCONI, M.; MORINI, B. STRSCNE: A scaled trust-region solver for constrained nonlinear equations. *Computational Optimization and Applications*, v. 28, p. 31–50, 2004.
- [10] HOCK, W.; SCHITTKOWSKI, K. *Test Examples for Nonlinear Programming Codes*. Berlin, Heidelberg, New York: Springer, 1981. 178 p. (Lecture Notes in Economics and Mathematical Systems, 187).
- [11] MARTÍNEZ, J. M.; SANTOS, S. A. *Métodos Computacionais de Otimização*. Rio de Janeiro: IMPA, 1995. 256 p. (20º Colóquio Brasileiro de Matemática). ISBN 8524400927.
- [12] MONTICELLI, A. J. *Fluxo de carga em redes de energia elétrica*. São Paulo: Editora Edgar Blücher Ltda, 1983. 164 p.

- [13] SCHITTKOWSKI, K. *More Test Examples for Nonlinear Programming Codes*. Berlin, Heidelberg, New York: Springer Verlag, 1987. 261 p. (Lecture Notes in Economics and Mathematical Systems, 282). ISBN 3-540-17182-7.
- [14] FRANCISCO, J. B. *Métodos Numéricos Aplicados à Resolução das Equações da Rede Elétrica*. 71 f. Dissertação (Mestrado em Matemática e Computação Científica) — Universidade Federal de Santa Catarina, Santa Catarina, 2002.

APÊNDICE A - Problemas

Abaixo apresentamos a formulação dos problemas que foram utilizados nos testes dos capítulos 2 e 3. Esses problemas foram extraídos da página eletrônica <http://www.polymath-software.com/library>, de [10] e de [13].

A.1 Problemas referentes ao capítulo 2

1. Twoeq2

$$F(x) = \begin{bmatrix} 120x_1 - 75k(1 - x_1) \\ -x_1(873 - x_2) + 11(x_2 - 300) \end{bmatrix}$$

em que $k = 0, 12e^{12581 \frac{x_2 - 298}{298x_2}}$.

Limitantes: $l = [-0, 01; -\infty]$ e $u = [1, 1; +\infty]$.

Estimativas iniciais: $x = [1; 400]$, $x = [0; 300]$, $x = [0, 5; 320]$ e $x = [0; 350]$.

Solução: $[0, 9638680512795; 346, 16369814640]$.

2. Twoeq3

$$F(x) = \begin{bmatrix} k \sqrt{1 - x_1} \left(\frac{0,91 - 0,5x_1}{9,1 - 0,5x_1} - \frac{x_1^2}{(1 - x_1)^2 Kp} \right) \\ x_2(1, 84x_1 + 77, 3) - 43260x_1 - 105128 \end{bmatrix}$$

em que

- $k = e^{\frac{-149750}{x_2} + 92,5}$;
- $Kp = e^{\frac{42300}{x_2} - 24,2 + 0,17 \ln(x_2)}$.

Limitantes: $l = [0; -\infty]$ e $u = [1; +\infty]$.

Estimativas iniciais: $x = [0, 5; 1700]$, $x = [0; 1600]$, $x = [0; 1650]$, $x = [0, 9; 1600]$ e $x = [0, 9; 1700]$.

Solução: $[0, 5333728995523; 1637, 70322946500]$.

3. Twoeq4a

$$F(x) = \begin{bmatrix} \ln(\gamma_2) + \ln(t_2) + v_1 \frac{x_2 \cdot t_2 - x_1 \cdot t_1}{t_1 \cdot t_2} \\ \ln(\gamma_1) + \ln(t_1) - v_2 \frac{x_2 \cdot t_2 - x_1 \cdot t_1}{t_1 \cdot t_2} \end{bmatrix}$$

em que

- $t = 58, 7;$
- $pw1 = 21;$
- $v_1 = \frac{pw1/46,07}{pw1/46,07+100-pw1/86,18};$
- $P_2 = 10^{6,87776-1171,53/(224,366+t)};$
- $P_1 = 10^{8,04494-1554,3/(222,65+t)};$
- $\gamma_2 = \frac{760}{P_2};$
- $\gamma_1 = \frac{760}{P_1};$
- $v_2 = 1 - v_1;$
- $t_1 = v_1 + v_2 \cdot x_2;$
- $t_2 = v_2 + v_1 \cdot x_1;$
- $g2calc = e^{-\ln(t_2)-v_1 \cdot (x_2 \cdot t_2 - x_1 \cdot t_1)/(t_1 \cdot t_2)};$
- $g1calc = e^{-\ln(t_1)+v_2 \cdot (x_2 \cdot t_2 - x_1 \cdot t_1)/(t_1 \cdot t_2)}.$

Limitantes: $l = [0; 0]$ e $u = [+ \infty; + \infty]$.

Estimativas iniciais: $x = [0, 1; 0, 1]$, $x = [0, 5; 0, 5]$ e $x = [0, 8; 0, 8]$.

Solução: $[0, 0785888476379; 0, 3017535930592]$.

4. Twoeq4b

$$F(x) = \begin{bmatrix} t_1 \cdot t_2 \cdot \ln(\gamma_2) + \ln(t_2) + v_1(x_2 \cdot t_2 - x_1 \cdot t_1) \\ t_1 \cdot t_2 \cdot \ln(\gamma_1) + \ln(t_1) - v_2(x_2 \cdot t_2 - x_1 \cdot t_1) \end{bmatrix}$$

com os mesmos parâmetros, limitantes e estimativas iniciais do problema Twoeq4a.

Solução: [0, 0785888934885; 0, 3017535528355].

5. Twoeq5a

$$F(x) = \begin{bmatrix} \log(\gamma_1) - x_1 \frac{v_2^2}{(x_1 \cdot v_1/x_2 + v_2)^2} \\ \log(\gamma_2) - x_2 \frac{v_1^2}{(v_1 + x_2 \cdot v_2/x_1)^2} \end{bmatrix}$$

em que

- $t = 70,9$;
- $pw1 = 49$;
- $v1 = \frac{pw1/46,07}{pw1/46,07 + (100 - pw1)/100,2}$;
- $P2 = 10^{6,9024 - 1268,115/(216,9+t)}$;
- $P1 = 10^{8,04494 - 1554,3/(222,65+t)}$;
- $\gamma_1 = \frac{760}{P1}$;
- $v2 = 1 - v1$;
- $\gamma_2 = \frac{760}{P2}$;
- $g1calc = 10^{x_1 \cdot v_2^2 / (x_1 \cdot v_1/x_2 + v_2)^2}$;
- $g2calc = 10^{x_2 \cdot v_1^2 / (v_1 + x_2 \cdot v_2/x_1)^2}$.

Limitantes: $l = [0; 0]$ e $u = [+∞; +∞]$.

Estimativas iniciais: $x = [0, 5; 0, 5]$, $x = [1; 1]$, $x = [5; 5]$ e $x = [8; 2]$.

Solução: [0, 7580768059470; 1, 1249034445330].

6. Twoeq5b

$$F(x) = \begin{bmatrix} \log(\gamma_1) \cdot (x_1 \cdot \frac{v_1}{x_2} + v_2)^2 - x_1 \cdot v_2^2 \\ \log(\gamma_2) \cdot (v_1 + x_2 \cdot \frac{v_2}{x_1})^2 - x_2 \cdot v_1^2 \end{bmatrix}$$

com os mesmos parâmetros, limitantes e estimativas iniciais do problema Twoeq5a.

Solução: [0, 7580768059470; 1, 1249034445330].

7. **Twoeq6**

$$F(x) = \begin{bmatrix} \frac{x_1}{1-x_1} - 5 \ln\left(0,4 \frac{1-x_1}{x_2}\right) + 4,45977 \\ x_2 - (0,4 - 0,5 \cdot x_1) \end{bmatrix}$$

Limitantes: $l = [0; -\infty]$ e $u = [1; +\infty]$.

Estimativas iniciais: $x = [0,9; 0,5]$, $x = [0,5; 0,5]$, $x = [0,4; 0,5]$ e $x = [0,6; 0,1]$.

Solução: $[0,7573962468236; 0,0213018765882]$.

Solução não viável: $[1,0989839337750; -0,1494919668876]$.

8. **Twoeq7**

$$F(x) = \begin{bmatrix} a - \frac{(c+2x_1)^2 \cdot (a+b+c-2x_1)^2}{x_2} - x_1 \\ x_2 - kp \cdot P^2 \cdot (b - 3x_1)^3 \end{bmatrix}$$

em que

- $a = 0,5$;
- $b = 0,8$;
- $c = 0,3$;
- $kp = 604500$;
- $P = 0,00243$.

Limitantes: $l = [0; -\infty]$ e $u = [1; +\infty]$.

Estimativas iniciais: $x = [0; -1]$, $x = [0; 1]$, $x = [0,5; 0,1]$ e $x = [0,5; -0,1]$.

Solução 1: $[0,6003231171445; -3,57990244801]$.

Solução 2: $[0,0586545710394; 0,86743788245]$.

9. **Twoeq8**

$$F(x) = \begin{bmatrix} x_1 - 0,327x_2^{0,804} e^{\left(\frac{-5230}{1,987(373+184000x_1)}\right)} \\ x_2 - (0,06 - 161x_1) \end{bmatrix}$$

Limitantes: $l = [0; -\infty]$ e $u = [+ \infty; + \infty]$.

Estimativas iniciais: $x = [0,0001; 0,01]$, $x = [0,001; 0,01]$, $x = [0,0001; 0,1]$ e $x = [0,5; 0,5]$.

Solução: [0,0003406054400; 0,0051625241669].

10. Twoeq9

$$F(x) = \left[\begin{array}{c} x_1 - \frac{1}{(2,28-4 \log(eps/D+4,67/(Re \cdot x_1^{0,5})))^2} \\ 133,7 - \frac{2x_1 \cdot \rho \cdot x_2^2 \cdot L/D + \rho \cdot g \cdot 200}{g \cdot 144} \end{array} \right]$$

em que

- $L = 6000$;
- $D = 0,505$;
- $\rho = 53$;
- $g = 32,2$;
- $eps = 0,00015$;
- $Re = \frac{\rho \cdot D \cdot x_2}{13,2 \cdot 0,000672}$.

Limitantes: $l = [0; 0]$ e $u = [+∞; +∞]$.

Estimativas iniciais: $x = [0, 1; 10]$, $x = [1; 10]$, $x = [0, 1; 1]$ e $x = [0, 1; 0, 1]$.

Solução: [0,006874616348157; 5,6728221306731].

11. Twoeq10

$$F(x) = \left[\begin{array}{c} \frac{1}{v1 \cdot v2} - 2x_2 \frac{c2}{(v1+v2 \cdot c1)^3} - 2x_1 \frac{c4}{(v2+v1 \cdot c3)^3} \\ \frac{v1-v2}{(v1 \cdot v2)^2} + 6x_2 \cdot c2 \frac{1-c1}{(v1+v2 \cdot c1)^4} + 6x_1 \cdot c4 \frac{c3-1}{(v2+v1 \cdot c3)^4} \end{array} \right]$$

em que

- $c1 = e^{-\alpha \cdot x_2}$;
- $c2 = e^{-2\alpha \cdot x_2}$;
- $c3 = e^{-\alpha \cdot x_1}$;
- $c4 = e^{-2\alpha \cdot x_1}$;
- $v2 = 1 - v1$;
- $\alpha = 0,4$;
- $v1 = 0,5$.

Limitantes: $l = [-\infty; -\infty]$ e $u = [+ \infty; + \infty]$.

Estimativas iniciais: $x = [0, 1; 0, 1]$, $x = [1; 1]$, $x = [10; 10]$ e $x = [15; 15]$.

Solução 1: [1, 6043843214350; 1, 6043843214350].

Solução 2: [2, 9353711137400; 2, 9353711137400].

12. Threeq1

$$F(x) = \begin{bmatrix} x_2 + x_3 - 1 \\ x_2 - \frac{y_1}{k_1} \\ x_3 - \frac{y_2}{k_2} \end{bmatrix}$$

em que

- $y_2 = 0, 8;$
- $y_1 = 0, 2;$
- $p_1 = 10^{7,62231 - \frac{1417,9}{191,15+x_1}};$
- $p_2 = 10^{8,10765 - \frac{1750,29}{235+x_1}};$
- $B = 0, 7;$
- $A = 1, 7;$
- $\gamma_2 = 10^{\frac{B \cdot x_2^2}{(x_2 + B \cdot x_3 / A)^2}};$
- $\gamma_1 = 10^{\frac{A \cdot x_3^2}{(A \cdot x_2 / B + x_3)^2}};$
- $k_2 = \gamma_2 \frac{p_2}{760};$
- $k_1 = \gamma_1 \frac{p_1}{760}.$

Limitantes: $l = [-\infty; 0; 0]$ e $u = [+ \infty; 1; 1]$.

Estimativas iniciais: $x = [100; 0, 2; 0, 8]$, $x = [70; 0, 5; 0, 5]$, $x = [80; 0, 2; 0, 8]$ e $x = [80; 0, 5; 0, 5]$.

Solução: [93, 96706523770; 0, 0078754574659; 0, 9921245425339].

13. Threeq2

$$F(x) = \begin{bmatrix} x_1 - \frac{z_1}{1+x_3(k_1-1)} \\ x_2 - \frac{z_2}{1+x_3(k_2-1)} \\ x_1 + x_2 - (y_1 + y_2) \end{bmatrix}$$

em que

- $p1 = 10^{7,62231 - \frac{1417,9}{191,15+t}}$;
- $p2 = 10^{8,10765 - \frac{1750,29}{235+t}}$;
- $\gamma_2 = 10^{\frac{B \cdot x_1^2}{(x_1 + Bx_2/A)^2}}$;
- $\gamma_1 = 10^{\frac{A \cdot x_2^2}{(Ax_1/B + x_2)^2}}$;
- $k1 = \gamma_1 \frac{p1}{760}$;
- $k2 = \gamma_2 \frac{p2}{760}$;
- $y1 = k1 \cdot x_1$;
- $y2 = k2 \cdot x_2$;
- $t = 88,538$;
- $B = 0,7$;
- $A = 1,7$;
- $z1 = 0,2$;
- $z2 = 0,8$.

Limitantes: $l = [0; 0; 0]$ e $u = [1; 1; 1]$.

Estimativas iniciais: $x = [0; 1; 0,5]$, $x = [0,5; 0,5; 0,9]$, $x = [0,4; 0,6; 0,9]$ e $x = [0,1; 0,9; 0,5]$.

Solução: $[0,0226974766367; 0,9773025233633; 0,5322677863643]$.

Solução não viável: $[0,6867568052506; 0,3132431762583; 1,4708209249600]$.

Solução falsa: $[-7,97906491100 \cdot 10^{-13}; 2,54516136300 \cdot 10^{-10}; -9,14388809400 \cdot 10^9]$.

14. Threeq3

$$F(x) = \begin{bmatrix} F \cdot \frac{T0-x_1}{V} + 30000 \cdot k \cdot \frac{x_2}{rhocp} - U \cdot A \cdot \frac{x_1-x_3}{rhocp \cdot V} \\ F \cdot \frac{Ca0-x_2}{V} - kx_2 \\ Fj \cdot \frac{Tj0-x_3}{3,85} + U \cdot A \cdot \frac{x_1-x_3}{62,3 \cdot 1,0 \cdot 3,85} \end{bmatrix}$$

em que

- $k = 7,08 \cdot 10^{10} e^{\frac{-30000/1,9872}{x_1}}$;
- $rhocp = 50 \cdot 0,75$;

- $F = 40$;
- $T0 = 530$;
- $V = 48$;
- $U = 150$;
- $A = 250$;
- $Ca0 = 0,55$;
- $Fj = 49,9$;
- $Tj0 = 530$.

Limitantes: $l = [-\infty; 0; -\infty]$ e $u = [\infty; \infty; \infty]$.

Estimativas iniciais: $x = [100; 0,2; 0,8]$, $x = [70; 0,5; 0,5]$, $x = [80; 0,2; 0,8]$ e $x = [80; 0,5; 0,5]$.

Solução 1: [671, 27832050250; 0,0354195308679; 660,46287831030].

Solução 2: [590, 34979512380; 0,3301868979161; 585,72976766210].

Solução 3: [537, 85475413080; 0,5213904932394; 537,25344007970].

15. Threeq4a

$$F(x) = \begin{bmatrix} CC \frac{x_1}{CA \cdot CB} - KC1 \\ x_2 \frac{CY}{CB \cdot CC} - KC2 \\ \frac{x_3}{CA \cdot x_2} - KC3 \end{bmatrix}$$

em que

- $CY = x_2 + x_3$;
- $CC = x_1 - CY$;
- $CA = CA0 - x_1 - x_3$;
- $CB = CB0 - x_1 - CY$;
- $KC1 = 1,06$;
- $KC2 = 2,63$;
- $KC3 = 5$;
- $CA0 = 1,5$;
- $CB0 = 1,5$.

Limitantes: $l = [0; 0; 0]$ e $u = [+∞; +∞; +∞]$.

Estimativas iniciais: $x = [0, 7; 0, 2; 0, 4]$, $x = [0; 0, 1; 0]$, $x = [1; 1; 1]$ e $x = [10; 10; 10]$.

Solução: $[0, 7053344059695; 0, 1777924200537; 0, 3739765850146]$.

16. Threeq4b

$$F(x) = \begin{bmatrix} CC \cdot x_1 - KC1 \cdot CA \cdot CB \\ x_2 \cdot CY - KC2 \cdot CB \cdot CC \\ x_3 - KC3 \cdot CA \cdot x_2 \end{bmatrix}$$

com os mesmos parâmetros, limitantes e estimativas iniciais do problema Threeq4a.

Solução: $[0, 7053344059695; 0, 1777924200537; 0, 3739765850146]$.

Solução não viável 1: $[0, 0555561340633; 0, 5972196082270; 1, 0820734849200]$.

Solução não viável 2: $[1, 0701041278950; -0, 3227156095103; 1, 1305335070740]$.

17. Threeq5

$$F(x) = \begin{bmatrix} -0, 16 \cdot x_1 \cdot \frac{F0}{x_3} + k1(1 - x_1) - k2 \cdot x_1 \\ 0, 16 \cdot F0 \cdot \frac{T0}{x_3} - 0, 16 \cdot x_2 \cdot \frac{F0}{x_3} + 5(k1(1 - x_1) - k2 \cdot x_1) \\ 0, 16 \cdot F0 - 0, 4 \sqrt{x_3} \end{bmatrix}$$

em que

- $k1 = 300000e^{-5000/x_2}$;
- $k2 = 60000000e^{-7500/x_2}$;
- $F0 = 1$;
- $T0 = 300$.

Limitantes: $l = [0; -∞; 0]$ e $u = [1; +∞; +∞]$.

Estimativas iniciais: $x = [0, 5; 500; 0, 5]$, $x = [0, 5; 200; 0, 1]$, $x = [0, 7; 700; 0, 2]$ e $x = [0, 001; 400; 0, 01]$.

Solução: $[0, 0171035426725; 300, 08551771340; 0, 16]$.

18. Threeq6

$$F(x) = \begin{bmatrix} 0,1(1-x_1) - k_1 \cdot x_1^2 \\ -0,1x_2 + k_1 \cdot x_1^2 - k_2 \cdot x_2 \\ 0,1(25-x_3) - 418 \cdot k_1 \cdot x_1^2 - 418 \cdot k_2 \cdot x_2 + Q \cdot 0.00001 \end{bmatrix}$$

em que

- $k_1 = 11e^{\frac{-4180}{8,314(x_3+273,16)}}$;
- $k_2 = 172,2e^{\frac{-34833}{8,314(x_3+273,16)}}$;
- $Q = 5100000$.

Limitantes: $l = [0; 0; -273.16]$ e $u = [1; 1; +\infty]$.

Estimativas iniciais: $x = [0,5; 0,5; 500]$, $x = [0,1; 0,2; 700]$, $x = [0,9; 0,8; 200]$ e $x = [0,01; 0,01; 500]$.

Solução: $[0,1578109142617; 0,77071354919; 153,09]$.

19. Threeq8

$$F(x) = \begin{bmatrix} 70 \cdot 32,3 - x_1(144 \cdot \frac{32,2}{62,35}) + k \cdot 45 \cdot (x_2 + x_3)^2 \\ (x_1 - p_2) \cdot (144 \cdot \frac{32,2}{62,35}) + k \cdot 24 \cdot x_2^2 \\ (x_1 - p_3) \cdot (144 \cdot \frac{32,2}{62,35}) + k \cdot 34 \cdot x_3^2 \end{bmatrix}$$

em que

- $p_2 = 156,6 - 0,00752x_2^2$;
- $p_3 = 117,1 - 0,00427x_3^2$;
- $D_{34} = \frac{2,067}{12}$;
- $D_{24} = \frac{1,278}{12}$;
- $D_{45} = \frac{2,469}{12}$;
- $fF = 0,015$;
- $\pi = 3,1416$;
- $k_{24} = 2 \cdot fF \cdot \frac{125}{(60 \cdot 7,48)^2 \cdot (\pi \cdot D_{24}^2 / 4)^2 \cdot D_{24}}$;
- $k_{34} = 2 \cdot fF \cdot \frac{125}{(60 \cdot 7,48)^2 \cdot (\pi \cdot D_{34}^2 / 4)^2 \cdot D_{34}}$;
- $k_{45} = 2 \cdot fF \cdot \frac{145}{(60 \cdot 7,48)^2 \cdot (\pi \cdot D_{45}^2 / 4)^2 \cdot D_{45}}$;

Limitantes: $l = [0; 0; 0]$ e $u = [+∞; +∞; +∞]$.

Estimativa inicial: $x = [50; 100; 100]$.

Solução não viável 1: $[33, 85347741000; -57, 48733479292; 109, 47131250920]$.

Solução não viável 2: $[57, 12556053860; -51, 75154635713; -92, 91811127982]$.

Solução não viável 3: $[33, 85347740990; 57, 48733479339; -109, 47131250920]$.

20. Fiveq1

$$F(x) = \begin{bmatrix} F \cdot \frac{2,88-x_1}{V} - k \cdot x_1^2 \\ F \cdot \frac{66-x_2}{V} - dhr \cdot k \cdot \frac{x_1^2}{\rho \cdot cp} - u \cdot a \cdot \frac{x_2-x_3}{V \cdot \rho \cdot cp} \\ u \cdot a \cdot \frac{x_2-x_3}{1,82 \cdot 1000 \cdot 4184} - fc \cdot \frac{x_3-27}{1,82} \\ \frac{(x_2-80)/20-x_4}{20} \\ \frac{m-x_5}{\tau_i} \end{bmatrix}$$

em que

- $k = 0,0744 \cdot e^{-1,182 \cdot 10^7 / (8314,39 \cdot (x_2 + 273,16))}$;
- $fc = 0,02 \cdot 50^{-m}$;
- $m = x_5 + kc \cdot \left(\frac{10}{20} - x_4\right)$;
- $F = 0,0075$;
- $V = 7,08$;
- $dhr = -9,86 \cdot 10^7$;
- $\rho = 19,2$;
- $cp = 1,815 \cdot 10^5$;
- $u = 3550$;
- $a = 5,4$;
- $\tau_i = 600$;
- $kc = 1$.

Limitantes: $l = [0; -∞; -∞; 0; 0]$ e $u = [+∞; +∞; +∞; 1; 1]$.

Estimativas iniciais: $x = [1; 100; 50; 0,4; 0,25]$, $x = [0,5; 50; 25; 0,1; 0,1]$,

$x = [0,2; 20; 10; 0,01; 0,01]$ e $x = [2; 200; 150; 0,8; 0,8]$.

Solução: $[1, 1206138931808; 90; 54, 8512245178517; 0,5; 0,3172119884117]$.

21. Sixeq1

$$F(x) = \begin{bmatrix} x_1 + x_2 + x_4 - 0,001 \\ x_5 + x_6 - 55 \\ x_1 + x_2 + x_3 + 2x_5 + x_6 - 110,001 \\ x_1 - 0,1x_2 \\ x_1 - 1 \cdot 10^4 \cdot x_3x_4 \\ x_5 - 55 \cdot 10^{14} \cdot x_3x_6 \end{bmatrix}$$

Limitantes: $l = [0; 0; 0; 0; 0; 0]$ e $u = [+∞; +∞; +∞; +∞; +∞; +∞]$.

Estimativas iniciais: $x = [10; 10; 10; 10; 10; 10]$, $x = [1; 1; 1; 1; 1; 1]$, $x = [0; 0; 0; 0; 0; 0]$ e $x = [1 \cdot 10^{-4}; 1 \cdot 10^{-3}; 0; 1 \cdot 10^{-4}; 55; 1 \cdot 10^{-4}]$.

Solução: $[8,2645 \cdot 10^{-5}; 8,2645 \cdot 10^{-4}; 9,0909 \cdot 10^{-5}; 9,0909 \cdot 10^{-5}; 54,9999; 1,1 \cdot 10^{-10}]$.

Solução não viável 1: $[-1 \cdot 10^{-10}; -1 \cdot 10^{-9}; -1 \cdot 10^{-11}; 0,001; 55,001; -0,001]$.

22. Sixeq2a

$$F(x) = \begin{bmatrix} 1 - x_1 - k1 \cdot x_1 \cdot x_6 + kr1 \cdot x_4 \\ 1 - x_2 - k2 \cdot x_2 \cdot x_6 + kr2 \cdot x_5 \\ -x_3 + 2 \cdot k3 \cdot x_4 \cdot x_5 \\ k1 \cdot x_1 \cdot x_6 - kr1 \cdot x_4 - k3 \cdot x_4 \cdot x_5 \\ 1,5(k2 \cdot x_2 \cdot x_6 - kr2 \cdot x_5) - k3 \cdot x_4 \cdot x_5 \\ 1 - x_4 - x_5 - x_6 \end{bmatrix}$$

em que

- $k1 = 31,24;$
- $k2 = 2,062;$
- $kr1 = 0,272;$
- $kr2 = 0,02;$
- $k3 = 303,03;$

Limitantes: $l = [0; 0; 0; 0; 0; 0]$ e $u = [+∞; +∞; +∞; +∞; +∞; +∞]$.

Estimativas iniciais: $x = [0,99; 0,05; 0,05; 0,99; 0,05; 0]$ e $x = [0,05; 0,99; 0,05; 0,05; 0,99; 0]$.

Solução: [0, 970; 0, 980; 0, 0598; 0, 990; $9.9751 \cdot 10^{-5}$; 0, 0098].

Solução não viável: [1, 0333; 1, 0222; -0, 0666; $-1.0980 \cdot 10^{-4}$; 1, 001; -0, 001].

23. Sixeq2b

$$F(x) = \begin{bmatrix} 1 - x_1 - k_1 \cdot x_1 \cdot x_6 + kr_1 \cdot x_4 \\ 1 - x_2 - k_2 \cdot x_2 \cdot x_6 + kr_2 \cdot x_5 \\ -x_3 + 2 \cdot k_3 \cdot x_4 \cdot x_5 \\ k_1 \cdot x_1 \cdot x_6 - kr_1 \cdot x_4 - k_3 \cdot x_4 \cdot x_5 \\ 1,5(k_2 \cdot x_2 \cdot x_6 - kr_2 \cdot x_5) - k_3 \cdot x_4 \cdot x_5 \\ 1 - x_4 - x_5 - x_6 \end{bmatrix}$$

em que

- $k_1 = 17,721$;
- $k_2 = 3,483$;
- $kr_1 = 0,118$;
- $kr_2 = 0,033$;
- $k_3 = 505,051$.

com os mesmos limitantes e estimativas iniciais do problema Sixeq2a.

Solução: [0, 9499; 0, 9666; 0, 1001; 0, 9899; $1.0012 \cdot 10^{-4}$; 0, 0099].

Solução não viável: [1, 0698; 1, 0465; -0, 1396; $-1.3775 \cdot 10^{-4}$; 1, 0038; -0, 0036].

24. Sixeq2c

$$F(x) = \begin{bmatrix} 1 - x_1 - k_1 \cdot x_1 \cdot x_6 + kr_1 \cdot x_4 \\ 1 - x_2 - k_2 \cdot x_2 \cdot x_6 + kr_2 \cdot x_5 \\ -x_3 + 2 \cdot k_3 \cdot x_4 \cdot x_5 \\ k_1 \cdot x_1 \cdot x_6 - kr_1 \cdot x_4 - k_3 \cdot x_4 \cdot x_5 \\ 1,5(k_2 \cdot x_2 \cdot x_6 - kr_2 \cdot x_5) - k_3 \cdot x_4 \cdot x_5 \\ 1 - x_4 - x_5 - x_6 \end{bmatrix}$$

em que

- $k_1 = 17,721$;
- $k_2 = 6,966$;

- $kr1 = 0,118$;
- $kr2 = 333,333$;
- $k3 = 505,051$.

com os mesmos limitantes e estimativas iniciais do problema Sixeq2a.

Solução: $[0,9499; 0,9666; 0,1001; 0,9899; 1.0013 \cdot 10^{-4}; 0,0099]$.

25. Sixeq3

$$F(x) = \begin{bmatrix} x_1 - \frac{0,2}{x_6 + (1-x_6) \cdot k11/k12} \\ x_2 - x_1 \frac{k11}{k12} \\ x_3 - \frac{0,8}{x_6 + (1-x_6) \cdot k21/k22} \\ x_4 - x_3 \frac{k21}{k22} \\ x_1(1 - k11) + x_3(1 - k21) \\ (x_1 - x_2) + (x_3 - x_4) \end{bmatrix}$$

em que

- $p1 = 10^{7,62231 - \frac{1417,9}{191,15+x_5}}$;
- $p2 = 10^{8,10765 - \frac{1750,29}{235+x_5}}$;
- $A = 1,7$;
- $B = 0,7$;
- $\gamma_{11} = 10^{\frac{A \cdot x_3^2}{(A \cdot x_1/B + x_3)^2}}$;
- $\gamma_{21} = 10^{\frac{B \cdot x_1^2}{(x_1 + B \cdot x_3/A)^2}}$;
- $\gamma_{12} = 10^{\frac{A \cdot x_4^2}{(A \cdot x_2/B + x_4)^2}}$;
- $\gamma_{22} = 10^{\frac{B \cdot x_2^2}{(x_2 + B \cdot x_4/A)^2}}$;
- $k11 = \gamma_{11} \cdot \frac{p1}{760}$;
- $k21 = \gamma_{21} \cdot \frac{p2}{760}$;
- $k12 = \gamma_{12} \cdot \frac{p1}{760}$;
- $k22 = \gamma_{22} \cdot \frac{p2}{760}$;

Limitantes: $l = [0; 0; 0; 0; -\infty; 0]$ e $u = [1; 1; 1; 1; +\infty; 1]$.

Estimativas iniciais: $x = [0; 1; 1; 0; 100; 0, 8]$, $x = [0, 05; 0, 95; 1; 0; 100; 0, 8]$,

$x = [0, 1; 0, 9; 1; 0; 100; 0, 8]$ e $x = [0; 1, 0; 0, 3; 0, 7; 100; 0, 8]$.

Solução: $[0, 0226; 0, 6867; 0, 9773; 0, 3132; 88, 5378; 0, 7329]$.

Solução falsa: $[-3, 607 \cdot 10^{-12}; 2, 210 \cdot 10^{-10}; 2, 255 \cdot 10^{-10}; 9, 117 \cdot 10^{-13}; 80, 859; 6, 610 \cdot 10^9]$.

26. Sixeq4a

$$F(x) = \begin{bmatrix} V - vO \cdot \frac{CAO - x_1}{-rA} \\ V - vO \cdot \frac{CBO - x_2}{-rB} \\ V - vO \cdot \frac{x_3}{rC} \\ V - vO \cdot \frac{x_4}{rD} \\ V - vO \cdot \frac{x_5}{rE} \\ 5000 \cdot (350 - x_6) - 25 \cdot (20 + 40) \cdot (x_6 - 300) + V \cdot SRH \end{bmatrix}$$

em que

- $rA = 2 \cdot r1B$;
- $rB = r1B + 2 \cdot r2C$;
- $rC = -3 \cdot r1B + r2C$;
- $rD = -r3E - r2C$;
- $rE = r3E$;
- $r1B = -k1B \cdot CA \cdot x_2$;
- $r2C = -k2C \cdot x_3 \cdot x_2^2$;
- $r3E = k3E \cdot x_4$;
- $k1B = 0, 4e^{(20000/R) \cdot (1/300 - 1/x_6)}$;
- $k2C = 10e^{(5000/R) \cdot (1/310 - 1/x_6)}$;
- $k3E = 10e^{(10000/R) \cdot (1/320 - 1/x_6)}$;
- $SRH = -rA \cdot 20000 + 2 \cdot r2C \cdot 10000 + 5000 \cdot r3E$;
- $R = 1, 987$;
- $V = 500$;
- $vO = \frac{75}{3,3}$;
- $CAO = \frac{25}{vO}$;

$$\bullet CBO = \frac{50}{vo};$$

Limitantes: $l = [0; 0; 0; 0; 0; -\infty]$ e $u = [+∞; +∞; +∞; +∞; +∞; +∞]$.

Estimativas iniciais: $x = [0, 5; 0, 01; 1; 0, 01; 1; 420]$, $x = [0, 05; 0, 001; 1; 0, 05; 1; 400]$ e $x = [0, 1; 0, 2; 0, 5; 0, 1; 0, 7; 350]$.

Solução: $[0, 00266; 0, 03346; 0, 83706; 0, 00039; 0, 80853; 372, 76458]$.

27. Sixeq4b

$$F(x) = \begin{bmatrix} V \cdot (-rA) - vo \cdot (CAO - x_1) \\ V \cdot (-rB) - vo \cdot (CBO - x_2) \\ V \cdot rC - vo \cdot x_3 \\ V \cdot rD - vo \cdot x_4 \\ V \cdot rE - vo \cdot x_5 \\ 5000 \cdot (350 - x_6) - 25 \cdot (20 + 40) \cdot (x_6 - 300) + V \cdot SRH \end{bmatrix}$$

com os mesmos parâmetros e limitantes do problema Sixeq4a.

Estimativas iniciais: $x = [0, 5; 0, 01; 1; 0, 01; 1; 420]$, $x = [0, 05; 0, 001; 1; 0, 05; 1; 400]$, $x = [0, 1; 0, 2; 0, 5; 0, 1; 0, 7; 350]$ e $x = [0, 1; 0, 2; 0, 5; 0, 1; 0, 7; 380]$.

Solução: $[0, 00266; 0, 03346; 0, 8370; 0, 00039; 0, 8085; 372, 76458]$.

Solução não viável: $[-0, 00266; -0, 03507; 0, 8123; 0, 00043; 0, 8414; 371, 4219]$.

28. Seveq1

$$F(x) = \begin{bmatrix} 0,5x_1 + x_2 + 0,5x_3 - \frac{x_6}{x_7} \\ x_3 + x_4 + 2x_5 - \frac{2}{x_7} \\ x_1 + x_2 + x_5 - \frac{1}{x_7} \\ -28837x_1 - 139009x_2 - 78213x_3 + 18927x_4 + 8427x_5 + \frac{13492}{x_7} - 10690\frac{x_6}{x_7} \\ x_1 + x_2 + x_3 + x_4 + x_5 - 1 \\ 400 \cdot x_1 \cdot x_4^3 - 1,7837 \cdot 10^5 \cdot x_3 \cdot x_5 \\ x_1 \cdot x_3 - 2,6058 \cdot x_2 \cdot x_4 \end{bmatrix}$$

Limitantes: $l = [0; 0; 0; 0; 0; 0; 0]$ e $u = [+∞; +∞; +∞; +∞; +∞; +∞; +∞]$.

Estimativas iniciais: $x = [0, 5; 0; 0; 0, 5; 0; 0, 5; 2]$, $x = [0, 2; 0, 2; 0, 2; 0, 2; 0, 2; 0, 5; 0, 2]$ e $x = [0, 22; 0, 075; 0, 001; 0, 58; 0, 125; 0, 435; 2, 35]$.

Solução: [0, 322; 0, 0092; 0, 046; 0, 618; 0, 0037; 0, 576; 2, 977].

Solução não viável: [0, 4569; -0, 0004; -0, 0021; 0, 9151; -0, 3695; 2, 6105; 11, 5004].

29. Sevepeq2a

$F(x) =$

$$\left[\begin{array}{c} 5,5 \cdot \frac{x_4 - x_2}{x_1 - x_2 - x_3} - \frac{x_2}{x_2 + x_3} \\ 3,5 \cdot \frac{0,888x_1 - x_3 - x_4}{x_1 - x_2 - x_3} - \frac{x_3}{x_2 + x_3} \\ 5,5 \cdot \frac{0,0986x_1 - 0,01(1,098x_1 - x_2 - 9x_4 - x_5 - x_6 + x_7) - (x_2 + x_5)}{1,098x_1 - x_2 - 9x_4 - x_5 - x_6 + x_7} - \frac{x_5}{x_5 + x_6} \\ 3,5 \cdot \frac{0,986x_1 - 9x_4 - x_6 - 0,0986x_1 + 0,01(1,098x_1 - x_2 - 9x_4 - x_5 - x_6 + x_7)}{1,098x_1 - x_2 - 9x_4 - x_5 - x_6 + x_7} - \frac{x_6}{x_5 + x_6} \\ \frac{150,2 \cdot 133,7 \cdot 0,32^2}{x_1 - x_2 - x_3} - x_4 \cdot \frac{x_1 - x_2 - x_3}{(0,0986 \cdot x_1 - x_4)^2} \\ \frac{446,8 \cdot 133,7 \cdot 0,0032^2}{1,098x_1 - x_2 - 9x_4 - x_5 - x_6 + x_7} - \frac{x_1 \cdot 0,0986 - x_4}{1,098x_1 - x_2 - 9x_4 - x_5 - x_6 + x_7} + 0,01 \\ 0,04 \cdot \frac{74,12 \cdot (0,986x_1 - 10x_4) + 222,24 \cdot (0,0986x_1 - x_4) + 18 \cdot (x_4 - x_2) + 278,84x_4 + 98,09 \cdot 0,0136 \cdot x_1}{98,01} - 0,0136x_1 - x_7 \end{array} \right]$$

Limitantes: $l = [0; 0; 0; 0; 0; 0; 0]$ e $u = [+∞; +∞; +∞; +∞; +∞; +∞; +∞]$.

Estimativas iniciais: $x = [80; 7; 60; 7; 0,5; 4; 0,05]$, $x = [100; 10; 50; 10; 0,5; 4; 0,05]$ e $x = [50; 5; 25; 5; 0,5; 4; 0,05]$.

Solução: [80, 5340; 6, 9725; 61, 1190; 7, 2042; 0, 5476; 3, 6532; 0, 05970].

30. Teneq1a

$$F(x) = \left[\begin{array}{c} x_1 + x_4 - 3 \\ 2x_1 + x_2 + x_4 + x_7 + x_8 + x_9 + 2x_{10} - R \\ 2x_2 + 2x_5 + x_6 + x_7 - 8 \\ 2x_3 + x_5 - 4 \cdot R \\ x_1 \cdot x_5 - 0,193 \cdot x_2 \cdot x_4 \\ x_6 \cdot \sqrt{x_2} - 0,002597 \cdot \sqrt{x_2 \cdot x_4 \cdot x_5} \\ x_7 \cdot \sqrt{x_4} - 0,003448 \cdot \sqrt{x_1 \cdot x_4 \cdot x_5} \\ x_8 \cdot x_4 - 1,799 \cdot 10^{-5} \cdot x_2 \cdot x_5 \\ x_9 \cdot x_4 - 0,0002155 \cdot x_1 \cdot \sqrt{x_3 \cdot x_5} \\ x_{10} \cdot x_4^2 - 3,846 \cdot 10^{-5} \cdot x_4^2 \cdot x_5 \end{array} \right]$$

em que

- $R = 10$;
- $x_5 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10}$.

Limitantes: $l_i = 0$ e $u_i = +\infty$, $\forall i$.

Estimativas iniciais: $x = [1; 1; 10; 1; 1; 1; 0; 0; 0; 0]$ e $x = [2; 2; 10; 1; 1; 2; 0; 0; 0; 0]$.

Solução: $[2, 8801; 3, 9506; 19, 9841; 0, 1198; 0, 0317; 0, 0046; 0, 0304; 0, 01608; 0, 1205; 0, 001]$.

31. 14eq1

$$F(x) = \begin{bmatrix} -((x_{12} - L0) \cdot k11 + L1) \cdot x_1 + x_{13} \cdot k12 \cdot x_2 \\ L1 \cdot x_1 - (x_{13} \cdot k12 + L2) \cdot x_2 + x_{14} \cdot k13 \cdot x_3 + z1 \cdot F \\ L2 \cdot x_2 - (x_{14} \cdot k13 + B) \cdot x_3 \\ -((x_{12} - L0) \cdot k21 + L1) \cdot x_4 + x_{13} \cdot k22 \cdot x_5 \\ L1 \cdot x_4 - (x_{13} \cdot k22 + L2) \cdot x_5 + x_{14} \cdot k23 \cdot x_6 + z2 \cdot F \\ L2 \cdot x_5 - (x_{14} \cdot k23 + B) \cdot x_6 \\ k11 \cdot x_1 + k21 \cdot x_4 - 1 \\ k12 \cdot x_2 + k22 \cdot x_5 - 1 \\ k13 \cdot x_3 + k23 \cdot x_6 - 1 \\ k1f \cdot z1 + k2f \cdot z2 - 1 \\ k10 \cdot k11 \cdot x_1 + k20 \cdot k21 \cdot x_4 - 1 \\ -x_{12} \cdot hv1 + x_{13} \cdot hv2 - L1 \cdot hl1 + L0 \cdot h0 \\ -x_{13} \cdot hv2 + x_{14} \cdot hv3 + hf + L1 \cdot hl1 - L2 \cdot hl2 \\ -x_{14} \cdot hv3 + Q + L2 \cdot hl2 - L3 \cdot hl3 \end{bmatrix}$$

em que

- $L0 = x_{12} - D$;
- $L1 = x_{13} - D$;
- $L2 = x_{14} + F - D$;
- $L3 = B$;
- $hl1 = x_7 \cdot (29,6 + 0,04x_7) \cdot x_1 + x_7 \cdot (38,5 + 0,025x_7) \cdot x_4$;
- $hv1 = (8003 + x_7 \cdot (43,8 - 0,04x_7)) \cdot k11 \cdot x_1 + (12004 + x_7 \cdot (31,7 + 0,007x_7)) \cdot k21 \cdot x_4$;
- $hl2 = x_8 \cdot (29,6 + 0,04x_8) \cdot x_2 + x_8 \cdot (38,5 + 0,025x_8) \cdot x_5$;
- $hv2 = (8003 + x_8 \cdot (43,8 - 0,04x_8)) \cdot k12 \cdot x_2 + (12004 + x_8 \cdot (31,7 + 0,007x_8)) \cdot k22 \cdot x_5$;
- $hl3 = x_9 \cdot (29,6 + 0,04x_9) \cdot x_3 + x_9 \cdot (38,5 + 0,025x_9) \cdot x_6$;
- $hv3 = (8003 + x_9 \cdot (43,8 - 0,04x_9)) \cdot k13 \cdot x_3 + (12004 + x_9 \cdot (31,7 + 0,007x_9)) \cdot k23 \cdot x_6$;

- $hf = x_{10} \cdot (29,6 + 0,04x_{10}) \cdot z1 + x_{10} \cdot (38,5 + 0,025x_{10}) \cdot z2;$
- $h0 = x_{11} \cdot (29,6 + 0,04x_{11}) \cdot k10 \cdot k11 \cdot x_1 + x_{11} \cdot (38,5 + 0,025 \cdot x_{11}) \cdot k20 \cdot k21 \cdot x_4;$
- $k11 = (10^{6,80776-935,77/((x_7-32) \cdot 5/9+238,789)})/P;$
- $k12 = (10^{6,80776-935,77/((x_8-32) \cdot 5/9+238,789)})/P;$
- $k13 = (10^{6,80776-935,77/((x_9-32) \cdot 5/9+238,789)})/P;$
- $k21 = (10^{6,85296-1064,84/((x_7-32) \cdot 5/9+232,012)})/P;$
- $k22 = (10^{6,85296-1064,84/((x_8-32) \cdot 5/9+232,012)})/P;$
- $k23 = (10^{6,85296-1064,84/((x_9-32) \cdot 5/9+232,012)})/P;$
- $k1f = (10^{6,80776-935,77/((x_{10}-32) \cdot 5/9+238,789)})/P;$
- $k2f = (10^{6,85296-1064,84/((x_{10}-32) \cdot 5/9+232,012)})/P;$
- $k10 = (10^{6,80776-935,77/((x_{11}-32) \cdot 5/9+238,789)})/P;$
- $k20 = (10^{6,85296-1064,84/((x_{10}-32) \cdot 5/9+232,012)})/P;$
- $F = 1;$
- $z1 = 0,40;$
- $z2 = 1 - z1;$
- $B = 0,75;$
- $D = 0,25;$
- $P = 760 \cdot \frac{120}{14,7};$
- $Q = 10000;$
- $y11 = k11 \cdot x_1;$
- $y12 = k12 \cdot x_2;$
- $rec = y11 \cdot \frac{D}{z1 \cdot F};$

Limitantes: $l = [0; 0; 0; 0; 0; 0; -\infty; 0; 0; 0; 0; 0; 0; 0] e u_i = +\infty, \forall i.$

Estimativas iniciais: $x = [0, 1; 0, 1; 0, 1; 0, 1; 0, 1; 0, 1; 100; 100; 100; 100; 100; 5; 5; 5]$ e $x = [0, 5; 0, 4; 0, 3; 0, 3; 0, 4; 0, 5; 145; 190; 210; 200; 200; 1; 1; 1].$

Solução: $[0, 579; 0, 3957; 0, 2718; 0, 4209, 0, 6043; 0, 7281; 186, 3785, 200, 5268; 211, 486; 200, 1675, 169, 064; 1, 081; 1, 0668, 1, 0495].$

32. **Test1**

$$F(x) = \begin{bmatrix} 10(x_2 - x_1^2) \\ 1 - x_1 \end{bmatrix}$$

Limitantes: $l = [-\infty; -1, 5]$ e $u = [+ \infty; + \infty]$.

Estimativa inicial: $x = [-2; 1]$.

Solução: $[1; 1]$.

33. **Test25**

$$F(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_{99}(x) \end{bmatrix}$$

em que $f_i(x) = -0,011 + e^{-\frac{1}{x_1}(u_i - x_2)^{x_3}}$ e $u_i = 25 + (-50 \ln(0,011))^{2/3}$, $i = 1, \dots, 99$.

Limitantes: $l = [-1; 0; 0]$ e $u = [100; 25, 6; 5]$.

Estimativa inicial: $x = [100; 12, 5; 3]$.

Solução: $[50; 25; 1, 5]$.

34. **Test3**

$$F(x) = \begin{bmatrix} -2 \cdot 10^{-5} \cdot (x_2 - x_1) \\ 1 + 2 \cdot 10^{-5} \cdot (x_2 - x_1) \end{bmatrix}$$

Limitantes: $l = [-\infty; 0]$ e $u = [+ \infty; + \infty]$.

Estimativa inicial: $x = [10; 1]$.

Solução: $[0; 0]$.

35. **Test5**

$$F(x) = \begin{bmatrix} \cos(x_1 + x_2) + 2 \cdot (x_1 - x_2) - 1,5 \\ \cos(x_1 + x_2) - 2 \cdot (x_1 - x_2) + 2,5 \end{bmatrix}$$

Limitantes: $l = [-1, 5; -3]$ e $u = [4; 3]$.

Estimativa inicial: $x = [0; 0]$.

Solução: $[-\pi/3 + 1/2; -\pi/3 - 1/2]$.

36. Test38

$$F(x) = \begin{bmatrix} -400 \cdot x_1 \cdot (x_2 - x_1^2) - 2(1 - x_1) \\ 200(x_2 - x_1^2) + 20, 2(x_2 - 1) + 19, 8(x_4 - 1) \\ -360 \cdot x_3 \cdot (x_4 - x_3^2) - 2(1 - x_3) \\ 180(x_4 - x_3^2) + 20, 2(x_4 - 1) + 19, 8(x_2 - 1) \end{bmatrix}$$

Limitantes: $l = [-10; -10; -10; -10]$ e $u = [10; 10; 10; 10]$.

Estimativa inicial: $x = [-3; -1; -3; -1]$.

Solução: $[1; 1; 1; 1]$.

37. Test110

$$F(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_{10}(x) \end{bmatrix}$$

em que $f_i(x) = 2 \frac{\log(x_i - 2)}{x_i - 2} - 2 \frac{\log(10 - x_i)}{10 - x_i} - 0, 2x_i^{-0,8} \left(\prod_{\substack{j=1 \\ j \neq i}}^{10} x_j \right)^{0,2}$, $i = 1, \dots, 10$.

Limitantes: $l_i = 2, 001$ e $u = 9, 999$, $\forall i$

Estimativa inicial: $x_i = 9$, $\forall i$.

Solução: $[9, 35025655; \dots; 9, 35025655]$.

A.2 Problemas referentes ao capítulo 3

1. Test26

$$F(x) = \begin{bmatrix} x_1 - x_2 \\ (x_2 - x_3)^2 \\ (1 + x_2^2) \cdot x_1 + x_3^4 - 3 \end{bmatrix}$$

Estimativa inicial: $x = [-2, 6; 2; 2]$.

Solução: $[1; 1; 1]$ e $[a; a; a]$ em que $a = \sqrt[3]{\alpha - \beta} - \sqrt[3]{\alpha + \beta} - 2/3$, $\alpha = \sqrt{139/108}$ e $\beta = 61/54$.

2. Test27

$$F(x) = \begin{bmatrix} 0, 1(x_1 - 1) \\ x_2 - x_1^2 \\ x_1 + x_3^2 + 1 \end{bmatrix}$$

Estimativa inicial: $x = [2; 2; 2]$.

Solução: $[-1; 1; 0]$.

3. Test28

$$F(x) = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + 2x_2 + 3x_3 - 1 \end{bmatrix}$$

Estimativa inicial: $x = [-4; 1; 1]$.

Solução: $[0, 5; -0, 5; 0, 5]$.

4. Test42

$$F(x) = \begin{bmatrix} x_1 - 1 \\ x_2 - 2 \\ x_3 - 3 \\ x_4 - 4 \\ x_1 - 2 \\ x_3^2 + x_4^2 - 2 \end{bmatrix}$$

Estimativa inicial: $x = [1; 1; 1; 1]$.

Solução: $[2; 2; 0,6\sqrt{2}; 0,8\sqrt{2}]$.

5. Test48

$$F(x) = \begin{bmatrix} x_1 - 1 \\ x_2 - x_3 \\ x_4 - x_5 \\ x_1 + x_2 + x_3 + x_4 + x_5 - 5 \\ x_3 - 2(x_4 + x_5) + 3 \end{bmatrix}$$

Estimativa inicial: $x = [3; 5; -3; 2; -2]$.

Solução: $[1; 1; 1; 1; 1]$.

6. Test49

$$F(x) = \begin{bmatrix} x_1 - x_2 \\ x_3 - 1 \\ (x_4 - 1)^2 \\ (x_5 - 1)^3 \\ x_1 + x_2 + x_3 + 4x_4 - 7 \\ x_3 + 5x_5 - 6 \end{bmatrix}$$

Estimativa inicial: $x = [10; 7; 2; -3; 0, 8]$.

Solução: $[1; 1; 1; 1; 1]$.

7. Test50

$$F(x) = \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \\ (x_3 - x_4)^2 \\ x_4 - x_5 \\ x_1 + 2x_2 + 3x_3 - 6 \\ x_2 + 2x_3 + 3x_4 - 6 \\ x_3 + 2x_4 + 3x_5 - 6 \end{bmatrix}$$

Estimativa inicial: $x = [35; -31; 11; 5; -5]$.

Solução: [1; 1; 1; 1; 1].

8. Test51

$$F(x) = \begin{bmatrix} x_1 - x_2 \\ x_2 + x_3 - 2 \\ x_4 - 1 \\ x_5 - 1 \\ x_1 + 3x_2 - 4 \\ x_3 + x_4 - 2x_5 \\ x_2 - x_5 \end{bmatrix}$$

Estimativa inicial: $x = [2, 5; 0, 5; 2; -1; 0, 5]$.

Solução: [1; 1; 1; 1; 1].

9. Test52

$$F(x) = \begin{bmatrix} 4x_1 - x_2 \\ x_2 + x_3 - 2 \\ x_4 - 1 \\ x_5 - 1 \\ x_1 + 3x_2 \\ x_3 + x_4 - 2x_5 \\ x_2 - x_5 \end{bmatrix}$$

Estimativa inicial: $x = [2; 2; 2; 2; 2]$.

Solução: $[-33; 11; 180; -158; 11]/349$.

10. **Test77**

$$F(x) = \begin{bmatrix} x_1 - 1 \\ x_1 - x_2 \\ x_3 - 1 \\ (x_4 - 1)^2 \\ (x_5 - 1)^3 \\ x_4 \cdot x_1^2 + \text{sen}(x_4 - x_5) - 2\sqrt{2} \\ x_2 + x_3^4 \cdot x_4^2 - 8 - \sqrt{2} \end{bmatrix}$$

Estimativa inicial: $x = [2; 2; 2; 2; 2]$.

Solução: $[1, 166172; 1, 182111; 1, 380257; 1, 506036; 0, 6109203]$.

11. **Test79**

$$F(x) = \begin{bmatrix} x_1 - 1 \\ x_1 - x_2 \\ x_2 - x_3 \\ (x_3 - x_4)^2 \\ (x_4 - x_5)^2 \\ x_1 + x_2^2 + x_3^3 - 2 - 3\sqrt{2} \\ x_2 - x_3^2 + x_4 + 2 - 2\sqrt{2} \\ x_1 \cdot x_5 - 2 \end{bmatrix}$$

Estimativa inicial: $x = [2; 2; 2; 2; 2]$.

Solução: $[1, 191127; 1, 362603; 1, 472818; 1, 635017; 1, 679081]$.

12. **Test216**

$$F(x) = \begin{bmatrix} 10(x_1^2 - x_2) \\ x_1 - 1 \\ x_1(x_1 - 4) - 2x_2 + 12 \end{bmatrix}$$

Estimativa inicial: $x = [-1, 2; 1]$.

Solução: $[2; 4]$.

13. **Test269**

$$F(x) = \begin{bmatrix} x_1 - x_2 \\ x_2 + x_3 - 2 \\ x_4 - 1 \\ x_5 - 1 \\ x_1 + 3x_2 \\ x_3 + x_4 - 2x_5 \\ x_2 - x_5 \end{bmatrix}$$

Estimativa inicial: $x = [2; 2; 2; 2; 2]$.

Solução: $[-0,7674; 0,2558; 0,6279; -0,1163; 0,2558]$.

14. **Test316**

$$F(x) = \begin{bmatrix} x_1 - 20 \\ x_2 + 20 \\ 0,01(x_1^2 + x_2^2) - 1 \end{bmatrix}$$

Estimativa inicial: $x = [0; 0]$.

Solução: $[7,071; -7,071]$.

15. **Test344**

$$F(x) = \begin{bmatrix} x_1 - 1 \\ x_1 - x_2 \\ (x_2 - x_3)^2 \\ x_1(1 + x_2^2) + x_3^4 - 4 - 3\sqrt{2} \end{bmatrix}$$

Estimativa inicial: $x = [2; 2; 2]$.

Solução: $[1,105; 1,197; 1,535]$.

16. Test345

$$F(x) = \begin{bmatrix} x_1 - 1 \\ x_1 - x_2 \\ (x_2 - x_3)^2 \\ x_1(1 + x_2^2) + x_3^4 - 4 - 3\sqrt{2} \end{bmatrix}$$

Estimativa inicial: $x = [0; 0; 0]$.

Solução: $[1, 105; 1, 197; 1, 535]$.

17. Test373

$$F(x) = \begin{bmatrix} x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_1 + x_2 e^{-5x_3} + x_4 - 127 \\ x_1 + x_2 e^{-3x_3} + x_5 - 151 \\ x_1 + x_2 e^{-x_3} + x_6 - 379 \\ x_1 + x_2 e^{x_3} + x_7 - 421 \\ x_1 + x_2 e^{3x_3} + x_8 - 460 \\ x_1 + x_2 e^{5x_3} + x_9 - 426 \end{bmatrix}$$

Estimativa inicial: $x = [300; -100; -0, 1997; -127; -151; 379; 421; 460; 426]$.

Solução: $[523, 31; -156, 95; -0, 2; 29, 61; -86, 62; 47, 33; 26, 24; 22, 92; -39, 47]$.