

Sobre a convergência do método de Levenberg-Marquardt com *scaling* singular para problemas de mínimos quadrados com resíduo não nulo

Rafaela Filippozzi, Everton Boos,
Douglas S. Gonçalves, Fermín S. V. Bazán

UFSC - Departamento de Matemática

Campus Trindade

88040-900, Florianópolis, SC

E-mail: rafaela.filippozzi@posgrad.ufsc.br, everton.boos@ufsc.br,
douglas.goncalves@ufsc.br, fermin.bazan@ufsc.br

Resumo: Este trabalho aborda a análise de convergência local e global de uma variante do método de Levenberg-Marquardt (LMM), projetada para problemas de mínimos quadrados não lineares com resíduo não nulo. Tal variante, chamada LMM com *scaling* singular (LMMSS), permite que a chamada matriz de *scaling* do LMM seja singular, o que pode ser útil em certas aplicações. Para lidar com o resíduo não nulo enquanto preservamos a convergência local, uma escolha criteriosa do parâmetro do LMM é feita com base no erro de linearização do gradiente, que é ditado pela não linearidade e pelo tamanho do resíduo. As principais contribuições deste trabalho estão relacionadas à convergência local e global do LMMSS neste cenário. Mais especificamente, demonstramos que a sequência de direções d_k gerada pelo LMMSS é *gradient-related* e então provamos que pontos limite de uma versão com busca linear do LMMSS são estacionários para a função de mínimos quadrados. Além disso, a análise local é ainda demonstrada sob uma condição de *error bound* no gradiente e para diferentes hipóteses sobre o erro de linearização.

Palavras-chave: *Levenberg-Marquardt, Matriz de Scaling Singular, Análise de convergência, Resíduo não nulo*

1 Introdução

Neste artigo, investigamos as propriedades de convergência global e local do método de Levenberg-Marquardt aplicado a um problema específico de mínimos quadrados não lineares (NLS):

$$\min_{x \in \mathbb{R}^n} \quad \frac{1}{2} \|F(x)\|^2 := \phi(x), \quad (1)$$

em que $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é duas vezes continuamente diferenciável. Em particular, focamos no problema *sobredeterminado*, em que $m \geq n$.

Diferentemente de trabalhos anteriores, que focaram predominantemente em casos de resíduo nulo ou assumiram posto completo da matriz Jacobiana [6, 5, 12, 9, 10], aqui assumimos que o resíduo pode ser não nulo em uma solução e não temos informações sobre a Jacobiana em tal ponto, podendo esta, portanto, ter posto não completo.

Definimos o problema (1) como um problema NLS com resíduo não nulo se x^* é um mínimo global de ϕ e $F(x^*) \neq 0$. Consequentemente, nosso interesse está em identificar os pontos estacionários de ϕ , i.e., o conjunto $X^* = \{x \in \mathbb{R}^n \mid J(x)^T F(x) = 0\}$, em que $J(x) \in \mathbb{R}^{m \times n}$ é a Jacobiana de F em x , sob a hipótese de que $X^* \neq \emptyset$.

Embora trabalhos como [1, 2] explorem problemas com resíduo não nulo, apresentando convergência local do LMM, a matriz de *scaling* permanece sendo a matriz identidade. Neste trabalho, além de considerar resíduo não nulo, consideramos que o método é construído através da seguinte iteração:

$$(J_k^T J_k + \lambda_k L^T L) d_k = -J_k^T F_k \quad (2)$$

$$x_{k+1} = x_k + \alpha_k d_k, \quad \forall k \geq 0, \quad (3)$$

em que $F_k := F(x_k)$, $J_k := J(x_k)$, $\{\lambda_k\}$ é um escalar positivo, $\{\alpha_k\}$ é o tamanho do passo, e $L^T L$ é chamada *matriz de scaling*, que é permitida ser singular.

Referimo-nos à iteração (2)–(3) como o método de Levenberg-Marquardt com *Scaling* Singular (LMMSS). A escolha da matriz de *scaling* como $L^T L$, permitida ser singular, difere da abordagem clássica do LMM. Esta estratégia é motivada por aplicações em problemas inversos, e se distancia dos resultados explorados para problemas com resíduo nulo em [4].

Embora a teoria clássica do método de Levenberg-Marquardt (LMM) seja bem estabelecida, a consideração de uma matriz de *scaling* singular introduz desafios únicos na análise de convergência, exigindo novas ferramentas teóricas e suposições adicionais. O estudo visa demonstrar que a convergência do LMM com matrizes de *scaling* singulares é alcançável sob condições razoáveis, enfatizando o impacto de tais matrizes na qualidade das soluções aproximadas obtidas, quando devidamente orientadas ao problema em questão. Esta abordagem, combinada com a consideração de resíduo não nulo, diverge da literatura existente, fornecendo uma nova perspectiva sobre convergência local e global para o LMM na presença de matrizes de *scaling* singulares.

Nosso estudo está organizado da seguinte forma. Na Seção 2, começamos com algumas suposições que serão consideradas ao longo do manuscrito e também o contexto matemático e resultados preliminares necessários para as seções seguintes. Comentários sobre a análise de convergência local e resultados obtidos são apresentados na Seção 3, que são categorizados em dois casos baseados no comportamento do posto da Jacobiana em torno de uma solução (posto constante e posto decrescente). Finalmente, a Seção 4 prova que, para o LMMSS com um esquema de busca linear de Armijo, todo ponto limite da sequência gerada é um ponto estacionário da função de soma de quadrados, estabelecendo assim a convergência global deste método. Nossas considerações finais são fornecidas na Seção 5.

Enfatizamos que este trabalho é apenas uma breve apresentação do artigo completo [7], no qual o leitor pode encontrar uma abordagem mais aprofundada dos conceitos discutidos aqui, juntamente com provas e exemplos ilustrativos.

2 Hipóteses e ferramentas auxiliares

Nossas principais hipóteses são apresentadas a seguir.

Hipótese 1. A matriz $L \in \mathbb{R}^{p \times n}$ tem posto completo, em que $m \geq n \geq p$, e existe $\gamma > 0$ tal que, para todo $x \in \mathbb{R}^n$

$$\|J(x)v\|^2 + \|Lv\|^2 \geq \gamma \|v\|^2, \quad \forall v \in \mathbb{R}^n. \quad (4)$$

Hipótese 2. Para qualquer $x^* \in X^*$, existe uma constante $\delta \in (0, 1)$ e $L_0 > 0$ tal que

$$\|J(x) - J(y)\| \leq L_0 \|x - y\|, \quad (5)$$

para todo $x, y \in B(x^*, \delta)$.

A Hipótese 2 exige que a Jacobiana seja localmente Lipschitz e implica:

$$\|J(y)(x - y) - (F(x) - F(y))\| \leq L_1 \|x - y\|^2, \quad \forall x, y \in B(x^*, \delta) \quad (6)$$

em que $L_1 = L_0/2$, isto é, o erro na aproximação linear de $F(x)$ em torno de y é $O(\|x - y\|^2)$, para x e y em uma vizinhança de x^* .

Devido à compacidade da bola $B(x^*, \delta)$, existem constantes positivas L_2 e β tais que $\|J(x)\| \leq L_2$ e $\|F(x)\| \leq \beta$ para todo $x \in B(x^*, \delta)$. Portanto, como $\|J(x)\|$ é limitado em $B(x^*, \delta)$, aplicando a desigualdade do valor médio, podemos inferir que $\|F(x) - F(y)\| \leq L_2\|x - y\|$, para todo $x, y \in B(x^*, \delta)$. Além disso, o gradiente $\nabla\phi(x) = J(x)^T F(x)$ é Lipschitz em $B(x^*, \delta)$:

$$\|J(x)^T F(x) - J(y)^T F(y)\| \leq L_3\|x - y\|, \quad \forall x, y \in B(x^*, \delta), \quad (7)$$

em que $L_3 = L_2^2 + \beta L_0$.

Além disso, observe que para $z \in X^* \cap B(x^*, r)$ e $x, y \in B(x^*, r)$, temos

$$\begin{aligned} \|(J(x) - J(y))^T F(y)\| &= \|(J(x) - J(z) + J(z) - J(y))^T F(y)\| \\ &\leq \|(J(x) - J(z))^T F(y)\| + \|(J(z) - J(y))^T F(y)\| \\ &\leq L_0 L_2 \|x - z\| \|y - z\| + \|J(x)^T F(z)\| \\ &\quad + L_0 L_2 \|y - z\|^2 + \|J(y)^T F(z)\|. \end{aligned} \quad (8)$$

Lema 2.1. [1, Lema 2.1] *Se a Hipótese 2 é satisfeita, então existe $L_4 > 0$ tal que*

$$\|\nabla\phi(y) - \nabla\phi(x) - J(x)^T J(x)(y - x)\| \leq L_4 \|x - y\|^2 + \|(J(x) - J(y))^T F(y)\|, \quad (9)$$

para todo $x, y \in B(x^*, \delta)$.

Em seguida, apresentamos a hipótese de *error bound* assumida ao longo deste trabalho.

Hipótese 3 (*Error bound*). *Para qualquer $x^* \in X^*$, $\|J(x)^T F(x)\|$ fornece um limitante de erro (error bound) em $B(x^*, \delta)$, i.e., existe $\omega \in (0, \infty)$ tal que*

$$\omega \text{dist}(x, X^*) \leq \|J(x)^T F(x)\|, \quad \forall x \in B(x^*, \delta), \quad (10)$$

em que $\text{dist}(x, X^*) = \inf_{z \in X^*} \|x - z\|$. Ao longo do texto, dado x , denotaremos por $\bar{x}$ um elemento de X^* tal que $\|x - \bar{x}\| = \text{dist}(x, X^*)$.

Da Hipótese 3 e da equação (7), obtemos

$$\omega \text{dist}(x_k, X^*) \leq \|J_k^T F_k\| \leq L_3 \text{dist}(x_k, X^*). \quad (11)$$

As hipóteses restantes focam, como delineado em (8), nos termos

$$\|J(x)^T F(z)\| \text{ e } \|J(y)^T F(z)\|.$$

Esses termos desempenham um papel crucial no controle do erro, expresso em (9), da “linearização incompleta” do gradiente. Vale notar que cada uma dessas hipóteses, motivadas por [1], dá origem a taxas de convergência distintas e análises correspondentes.

Hipótese 4. *Para todo $x \in B(x^*, \delta)$ e todo $z \in X^* \cap B(x^*, \delta)$, a seguinte desigualdade vale:*

$$\|(J(x) - J(z))^T F(z)\| \leq \theta \|x - z\|,$$

com $0 \leq \theta < \bar{\theta}(\omega, L_3, \lambda^*)$, em que $\theta(\omega, L_3, \lambda^*)$ é uma constante positiva dependendo de ω de (10), L_3 de (7), e o menor autovalor positivo de $J_*^T J_*$, denotado por λ^* .

Hipótese 5. *Para todo $x \in B(x^*, \delta)$ e todo $z \in X^* \cap B(x^*, \delta)$, a seguinte desigualdade vale:*

$$\|(J(x) - J(z))^T F(z)\| \leq C \|x - z\|^{1+r}, \quad (12)$$

com $r \in]0, 1[$ e $C \geq 0$.

Hipótese 6. Para todo $x \in B(x^*, \delta)$ e todo $z \in X^* \cap B(x^*, \delta)$, a seguinte desigualdade vale:

$$\|(J(x) - J(z))^T F(z)\| \leq K \|x - z\|^2, \quad (13)$$

com $K \geq 0$.

Em [1], alguns exemplos simples com duas variáveis e resíduo não nulos são apresentados, ilustrando a motivação para tais hipóteses. Além disso, é evidente que a Hipótese 6 implica a Hipótese 5, que por sua vez implica a Hipótese 4.

Além das hipóteses acima, algumas ferramentas auxiliares são necessárias, a principal sendo a Decomposição em Valores Singulares Generalizada (GSVD) [11]. Uma versão particularizada da GSVD para o contexto deste trabalho pode ser encontrada em [8, p. 22]. Esta é uma ferramenta importante na análise teórica porque permite a conexão, em uma única decomposição, de um par de matrizes não relacionadas, como é o caso da Jacobiana de F em x_k e L .

Os seguintes resultados estabelecem um limite no tamanho do passo com base na distância da iteração atual ao conjunto solução, considerando dois casos distintos. O Lema 2.2 aborda o cenário em que o posto da Jacobiana em torno da solução é constante, enquanto o Lema 2.3 trata do caso de posto decrescente, dependendo da definição da sequência λ_k .

Lema 2.2. Suponha que as Hipóteses 1-3 são válidas, e $\text{rank}(J(x)^T J(x)) = \text{rank}(J(x^*)^T J(x^*)) = q \geq 1$ para todo $x \in B(x^*, \delta)$. Se $x_k \in B(x^*, \delta)$ e $\lambda_k := \lambda(x_k) > 0$, então existe $c_1 > 0$ tal que

$$\|d_k\| \leq c_1 \text{dist}(x_k, X^*). \quad (14)$$

Lema 2.3. Suponha que as Hipóteses 1-2 são válidas. Para $x_k \in B(x^*, \delta)$ e $\text{rank}(J_k) = \ell \geq \text{rank}(J_*) = q \geq 1$, e

(a) Hipótese 5 é satisfeita com $\lambda_k = \|J_k^T F_k\|^r$, e $r \in]0, 1[$,

(b) Hipótese 6 é satisfeita com $\lambda_k = \|J_k^T F_k\|$,

então existe $c_1 > 0$ tal que $\|d_k\| \leq c_1 \text{dist}(x_k, X^*)$.

Observação 2.1. Para o cenário em que o posto da matriz Jacobiana diminui, sob a Hipótese 4, o resultado sobre um limite superior no tamanho do passo em termos da distância da iteração atual ao conjunto solução permanece em aberto.

3 Convergência local

Os Lemas 2.2 e 2.3 mostram que, sob hipóteses específicas, $\|d_k\| \leq c_1 \|x_k - \bar{x}_k\|$. Tal desigualdade é chave para a análise de convergência local sob a condição de *error bound* (Hipótese 3). Nesta seção, consideramos a iteração LMMSS “pura”, i.e., (2)–(3) com $\alpha_k = 1$ para todo k .

A abordagem para obter resultados sobre a convergência local do método proposto é separada em dois casos de acordo com se o posto da Jacobiana próximo ao conjunto solução é constante ou não e de acordo com a hipótese em consideração (Hipótese 4, 5 ou 6). Devido à complexidade da explicação e para encurtar a exposição, mencionamos que a convergência quadrática local pode ser estabelecida sob a Hipótese 6, com o posto da Jacobiana sendo constante ou decrescente. Adicionalmente, a convergência superlinear local pode ser obtida sob a Hipótese 5, novamente considerando posto constante na Jacobiana ou não. Sob a Hipótese 4, entretanto, a convergência linear local é provada apenas para o caso em que o posto da matriz Jacobiana permanece constante. Supondo que a Hipótese 4 é satisfeita quando o posto está decrescendo, i.e., $\text{rank}(J_k) = \ell \geq \text{rank}(J_*) = q \geq 1$, a prova de convergência local ainda está em aberto.

Novamente, para uma explicação completa e provas destes resultados, o leitor é direcionado a [7].

4 Convergência global

Para a análise de convergência global, consideramos o Algoritmo 1, uma versão do LMMSS com busca linear para mínimos quadrados não lineares com resíduo não nulo. Vale ressaltar que o Algoritmo 1 difere do algoritmo proposto em [4] na escolha do parâmetro λ_k bem como no critério de aceitação do passo completo ($\alpha_k = 1$).

Algoritmo 1: LMMSS para problemas com resíduo não nulo com busca linear

```

1 Entrada:  $\nu, \zeta, \vartheta \in (0, 1)$ ,  $F$ ,  $J$ ,  $L$ , e  $x_0 \in \mathbb{R}^n$ 
2 Defina  $\lambda_0 = \|\nabla\phi(x_0)\|$ 
3 para  $k = 0, 1, 2, \dots$  faça
4   se  $\|\nabla\phi(x_k)\| = 0$  então
5     Pare com  $\bar{x} = x_k$ 
6   fim
7   Calcule  $d_k = -(J_k^T J_k + \lambda_k L^T L)^{-1} J_k^T F_k$ 
8   se  $\|\nabla\phi(x_k + d_k)\| \leq \vartheta \|\nabla\phi(x_k)\|$  então
9      $\alpha_k = 1$ 
10  senão
11    Escolha  $m$  como o menor inteiro não negativo tal que
                                
$$\phi(x_k + \zeta^m d_k) - \phi(x_k) \leq \nu \zeta^m \nabla\phi(x_k)^T d_k \quad (15)$$

    Defina  $\alpha_k = \zeta^m$ 
12  fim
13  Atualize  $x_{k+1} = x_k + \alpha_k d_k$ 
14   $\lambda_{k+1} = \|\nabla\phi(x_{k+1})\|$ 
15 fim

```

Inicialmente, demonstraremos que a sequência de direções d_k gerada pelo Algoritmo 1 é *gradient-related* (ver [3, Proposição 1.2.1]). Subsequentemente, estabeceremos que qualquer ponto limite da sequência produzida por este algoritmo é um ponto estacionário para (1), independentemente do ponto inicial. Isso serve para provar o teorema central, de convergência global. Além disso, nesta seção, assumimos que a Hipótese 6 é satisfeita para simplificar a notação e as provas.

Relembramos a definição de direções *gradient-related* de [3, Eq. (1.13)]:

Definição 4.1 (*Gradient-related*). *Sejam $\{x_k\}$ e $\{d_k\}$ sequências em $\mathbb{R}^n$. Dizemos que a sequência $\{d_k\}$ é gradient-related de $\{x_k\}$ se, para cada subsequência $\{x_k\}_{k \in \mathcal{K}}$ (com $\mathcal{K} \subseteq \mathbb{N}$) convergindo para um ponto não estacionário de ϕ , a subsequência correspondente $\{d_k\}_{k \in \mathcal{K}}$ é limitada e satisfaz*

$$\limsup_{k \rightarrow \infty, k \in \mathcal{K}} \nabla\phi(x_k)^T d_k < 0. \quad (16)$$

Proposição 4.1. *Suponha que as Hipóteses 1-3, 6, e $\lambda_k = \|J_k^T F_k\|$ são satisfeitas. Sejam $\{d_k\}$, $\{x_k\}$ sequências geradas pelo Algoritmo 1. Então, $\{d_k\}$ é gradient-related.*

Esta proposição garante que o Algoritmo 1 gera uma sequência de direções $\{d_k\}$ que são *gradient-related*. Portanto, usando [3, Proposição 1.2.1], a convergência global pode ser estabelecida, como segue.

Teorema 4.1. *Seja $\{x_k\}$ uma sequência gerada pelo Algoritmo 1. Então, todo ponto limite $\hat{x}$ de $\{x_k\}$ é tal que $\nabla\phi(\hat{x}) = 0$.*

Demonstração. Seja $K_1 = \{k \in \mathbb{N} \mid \|J(x_k + d_k)^T F(x_k + d_k)\| \leq \vartheta \|J(x_k)^T F(x_k)\|\}$. Se K_1 é infinito, segue que $\|J(x_k)^T F(x_k)\| \rightarrow 0$, e portanto qualquer ponto limite $\hat{x}$ de $\{x_k\}$ é tal que

$J(\hat{x})^T F(\hat{x}) = 0$, logo $\nabla \phi(\hat{x}) = 0$. Caso contrário, se K_1 é finito, vamos assumir, sem perda de generalidade, que $\|J(x_k + d_k)^T F(x_k + d_k)\| > \vartheta \|J(x_k)^T F(x_k)\|$, para cada k , de modo que o tamanho do passo é escolhido para satisfazer a condição de Armijo. Como, pela Proposição 4.1, as direções do Algoritmo 1 são *gradient-related*, segue de [3, Proposição 1.2.1] que qualquer ponto limite $\hat{x}$ de $\{x_k\}$ é um ponto estacionário de ϕ . $\square$

Agora, para conectar convergência local e global, o próximo resultado estabelece que, sob certas condições, $\alpha_k = 1$ para todo k suficientemente grande e então podemos aplicar a teoria de convergência local para mostrar que $\{\text{dist}(x_k, X^*)\}$ converge para zero quadraticamente.

Teorema 4.2. *Suponha que as Hipóteses 1-3 e 6 são satisfeitas e $\lambda_k = \|J_k^T F_k\|$. Além disso, assuma que o conjunto de nível $C_n := \{x \in \mathbb{R}^n : \phi(x) \leq \phi(x_0)\}$ é compacto para algum x_0 . Seja $\{x_k\}$ gerada pelo Algoritmo 1, usando x_0 como ponto inicial. Então, para todo k suficientemente grande $\alpha_k = 1$ e a sequência $\{\text{dist}(x_k, X^*)\}$ converge para 0 quadraticamente.*

Observação 4.1. *Observamos que no caso de resíduo nulo, um resultado anterior em [12, Teorema 3.1] requer a suposição de que o ponto limite x^* é tal que $F(x^*) = 0$. Além disso, no caso de otimização irrestrita, para provar um resultado similar a suposição requerida é que a Hessiana em x^* é definida positiva; entretanto, tal suposição implicaria que x^* é um ponto estacionário isolado [3, Proposição 1.2.1]. Como queremos abordar o caso de resíduo não nulo e possivelmente pontos estacionários não isolados, consideramos a limitação do conjunto de nível em vez destas outras duas.*

Observação 4.2. *No cenário em que a Hipótese 5 é satisfeita com posto decrescente, é necessário ajustar o parâmetro λ_k do Algoritmo 1 para $\lambda_k = \|J_k^T F_k\|^r$. Além disso, quando esta hipótese é válida, os resultados acima seguem com pequenas alterações em algumas constantes, e podemos mostrar que existe k_0 tal que $\alpha_k = 1, \forall k \geq k_0$ e assim $\text{dist}(x_k, X^*)$ vai para zero superlinearmente. Quando apenas a Hipótese 4 é satisfeita, embora as provas da Proposição 4.1 e do Teorema 4.1 sejam análogas (quando consideramos posto constante), a prova de que o tamanho de passo completo $\alpha_k = 1$ é sempre aceito para k suficientemente grande permanece em aberto.*

5 Conclusão e trabalhos futuros

Motivados pela abordagem do artigo [4], que emprega uma versão do método de Levenberg-Marquardt com uma matriz de *scaling* singular (denominada LMMSS) para problemas de mínimos quadrados não lineares com resíduo nulo, propomos um novo método LMMSS para problemas com resíduo não nulo. Nosso estudo investigou minuciosamente a convergência local e global do método proposto.

Para convergência local, mostramos que a convergência quadrática local de $\text{dist}(x_k, X^*)$ para zero é possível, mesmo na presença de um resíduo não nulo, com o posto da Jacobiana sendo constante ou decrescente. Quando o posto da Jacobiana é constante para $x \in B(x^*, \delta)$, mesmo considerando hipóteses mais fracas, conseguimos estabelecer convergência linear e superlinear local (ver Hipótese 4 e 5, respectivamente). Entretanto, se apenas a Hipótese 4 for válida no caso de posto decrescente, a convergência local permanece uma questão em aberto.

Além disso, demonstramos a natureza estacionária dos pontos limite na sequência gerada por uma versão do LMMSS com busca linear (Algoritmo 1). A convergência global foi estabelecida sob as Hipóteses 5 e 6. Infelizmente, quando apenas a Hipótese 4 é satisfeita, no caso de posto decrescente, a convergência global do algoritmo permanece inconclusiva.

Como trabalhos futuros, além de abordar essas questões em aberto, pretendemos realizar experimentos computacionais para validar nossos resultados teóricos, bem como testar o algoritmo em certos problemas inversos para avaliar seu comportamento prático.

Agradecimentos

Este trabalho foi apoiado pelas agências brasileiras FAPESC (Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina, número da concessão 2023TR000360). DG agradece o apoio do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), Brasil [número da concessão 305213/2021-0].

Referências

- [1] R. Behling, D. S. Gonçalves, S. A. Santos, Local convergence analysis of the Levenberg-Marquardt framework for nonzero-residue nonlinear least-squares problems under an error bound condition, *Journal of Optimization Theory and Applications*, 183.3 (2019) 1099-1122.
- [2] E. H. Bergou, Y. Diouane, V. Kungurtsev, Convergence and Complexity Analysis of a Levenberg-Marquardt Algorithm for Inverse Problems, *Journal of Optimization Theory and Applications*, 185 (2020) 927-944.
- [3] D. P. Bertsekas, “Nonlinear Programming”, Athena Scientific, 1999.
- [4] E. Boos, D.S. Gonçalves, F.S.V. Bazán, Levenberg-Marquardt method with singular scaling and applications, *Applied Mathematics and Computation* 474 (2024) 128688.
- [5] J. E. Dennis Jr., Nonlinear least squares and equations, em “The State of the Art in Numerical Analysis” (D. Jacobs, ed.), pp. 269-312, Academic Press, London, 1977.
- [6] J. E. Dennis Jr, R. B. Schnabel, “Numerical Methods for Unconstrained Optimization and Nonlinear Equations”, SIAM, Philadelphia, 1996.
- [7] R. Filippozzi, E. Boos, D.S. Gonçalves, F.S.V. Bazán, Convergence analysis of Levenberg-Marquardt method with Singular Scaling for nonzero residue nonlinear least-squares problems, *Numerical Algorithms*, 1 (2025), 1.
- [8] P. C. Hansen, “Rank-Deficient and Discrete Ill-Posed Problems”, SIAM, Philadelphia, 1998.
- [9] K. Levenberg, A Method for the Solution of Certain Non-Linear Problems in Least Squares, *Quarterly of Applied Mathematics*, 2 (1944), 164-168.
- [10] D. Marquardt, An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *SIAM Journal on Applied Mathematics*, 11.2 (1963), 431-441.
- [11]] C. F. Van Loan, Generalizing the Singular Value Decomposition, *SIAM Journal on Numerical Analysis* 13.1 (1976), 76-83.
- [12] N. Yamashita, M. Fukushima, On the rate of convergence of the Levenberg-Marquardt method, *Topics in Numerical Analysis*, (2001), 239-249.