

SISTEMAS NÃO LINEARES VIA REGIÃO DE CONFIANÇA

John Lenon C. Gardenghi
Profa. Dra. Sandra Augusta Santos

infor.john@gmail.com
sandra@ime.unicamp.br



Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq

Palavras-chave: Sistemas Não Lineares – Levenberg-Marquardt – Região de Confiança.

Introdução

Dada $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$, F suave, consideramos o problema de quadrados mínimos:

$$\begin{aligned} \min \frac{1}{2} F(x)^T F(x) &\equiv \varphi(x) \\ \text{s.a. } x &\in \mathbb{R}^n \end{aligned} \tag{1}$$

Aplicamos o método de Levenberg-Marquardt (LM) a (1), seguindo as sugestões de Moré [1].

Metodologia

Pesquisamos a melhor maneira para se resolver os dois problemas centrais da estratégia, implementamos o método no CAS Maxima e testamos a implementação com problemas clássicos apresentados em [2], validando a implementação feita, os critérios adotados e a robustez, eficiência e confiabilidade da estratégia abordada.

LM e Região de Confiança

O algoritmo de LM trabalha com o modelo linear da função residual $F(x)$ em torno do ponto corrente x_c . Deve-se resolver, a cada iteração, o problema de quadrados mínimos linear:

$$\begin{aligned} \text{minimizar } \frac{1}{2} \|J_c p + F_c\|^2 \\ \text{s.a. } \|p\| \leq \Delta_c, \end{aligned} \tag{2}$$

onde J_c é o Jacobiano de F em x_c e $\Delta_c > 0$ o raio da região de confiança, que nos dá o subproblema de região de confiança.

Solução do Subproblema

As condições de otimalidade de primeira ordem do problema (2) são:

$$(J_c^T J_c + \lambda^* I) p^* = -J_c^T F_c \tag{3a}$$

$$\lambda^* (\|p^*\| - \Delta_c) = 0 \tag{3b}$$

$$\|p^*\| \leq \Delta_c, \quad \lambda^* \geq 0, \tag{3c}$$

que, aliada à estrutura de (2), nos dão condições necessárias e suficientes para o ótimo.

As condições (3) propõem a solução p^* em função do multiplicador de Lagrange λ^* . De (3b):

- Se $\lambda^* = 0$ e $\|p^*\| \leq 0$, então a solução procurada será dada pela solução do sistema $(J_c^T J_c) p^* = -J_c^T F_c$;
- Se $\lambda^* > 0$, então a solução procurada será $p^* = -(J_c^T J_c + \lambda^* I)^{-1} J_c^T F_c$ tal que $\|p^*\| - \Delta_c = 0$.

A conexão do método de LM com a estratégia de região de confiança está no tratamento de λ .

Problemas Centrais

O problema central pode ser instanciado como segue:

- Resolver $\|p^*\| - \Delta_c = 0$;
- Controlar o decrescimento da função objetivo (1) para cada ponto candidato $x_n = x_c + p^*$ e o tamanho do raio Δ_c .

Aspectos Teóricos Explorados

Os principais aspectos teóricos são:

- Ao invés de resolver o problema $\|p\| - \Delta_c = 0$, vamos resolver a equação secular

$$\phi(\lambda) \stackrel{def}{=} \frac{1}{\|p\|} - \frac{1}{\Delta_c} = 0; \tag{4}$$

- A equação (3a) pode ser vista como um sistema de equação normais para

$$\begin{pmatrix} J_c \\ \sqrt{\lambda} I \end{pmatrix} p \cong - \begin{pmatrix} F_c \\ 0 \end{pmatrix},$$

reduzindo assim pela metade o número de operações necessárias para resolver o problema envolvendo o produto $J_c^T J_c$;

- Calcular a decomposição QR da matriz J_c e aplicar:

$$\begin{pmatrix} Q^T & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} J_c \\ \sqrt{\lambda} I \end{pmatrix} = \begin{pmatrix} R_c \\ 0 \\ \sqrt{\lambda} I \end{pmatrix}$$

nos dá uma matriz com estrutura muito característica, que pode ser triagularizada por $n(n+1)/2$ rotações de Givens;

- A triangularização por rotações de Givens resulta em uma matriz triangular superior R_λ ; essa matriz é o fator de *Cholesky* para a matriz $(J_c^T J_c + \lambda I)$ e será base para resolver o sistema (3a);
- Para resolver (4), adotamos o método de Newton com algumas salvaguardas para que o algoritmo nao convirja para um zero que não seja λ^* .

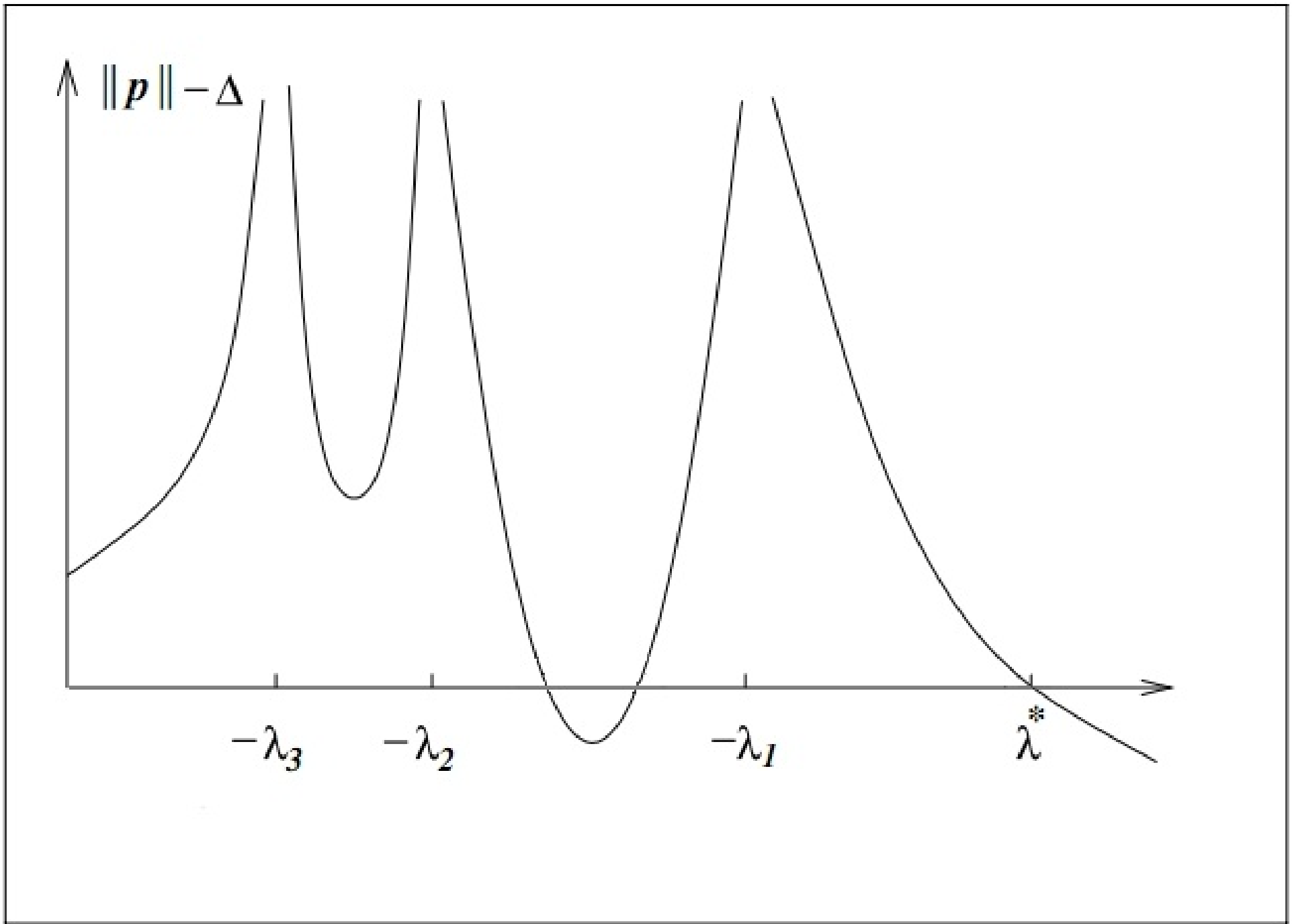


Gráfico da função $\|p\| - \Delta_c$.

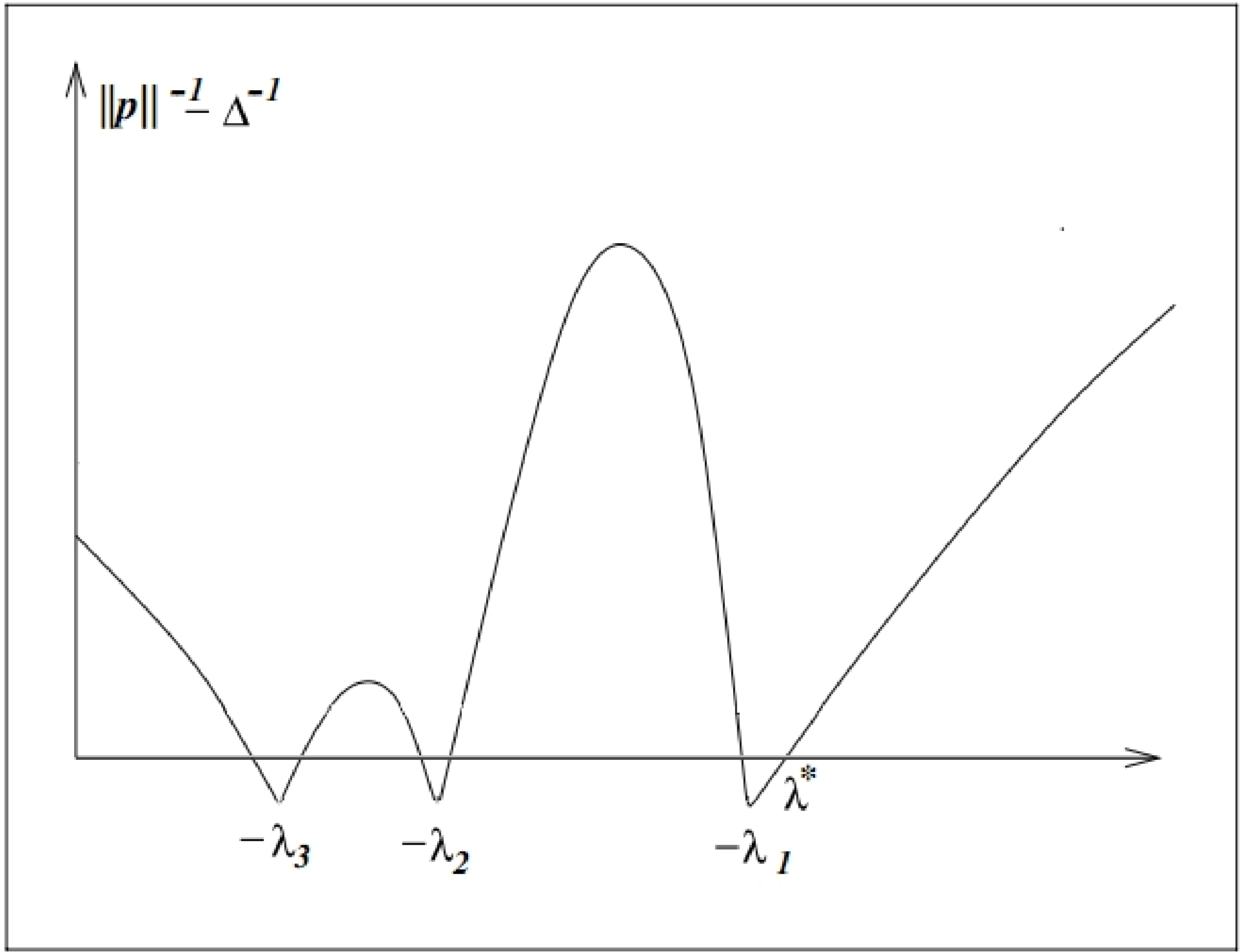


Gráfico da função secular (4).

Resultados e Discussão

Na primeira parte, tomamos as seguintes funções para o experimento, os pontos iniciais são os mesmos da literatura [2].

Função [2]	Sigla	n	m
1	ROS	2	2
4	BBS	2	3
5	BEA	2	3
13	POWS	4	4
14	WOOD	4	6
16	BDF	4	20
20	WAT	12	31

Disso, tivemos os seguintes resultados:

CP O critério de parada: $\begin{cases} 1 : \|F_c\| \leq \epsilon \\ 2 : \|J_c^T F_c\| \leq \epsilon, \end{cases}$

$\|F_*\|$ O valor da função avaliada na solução,

IT Número de iterações do processo,

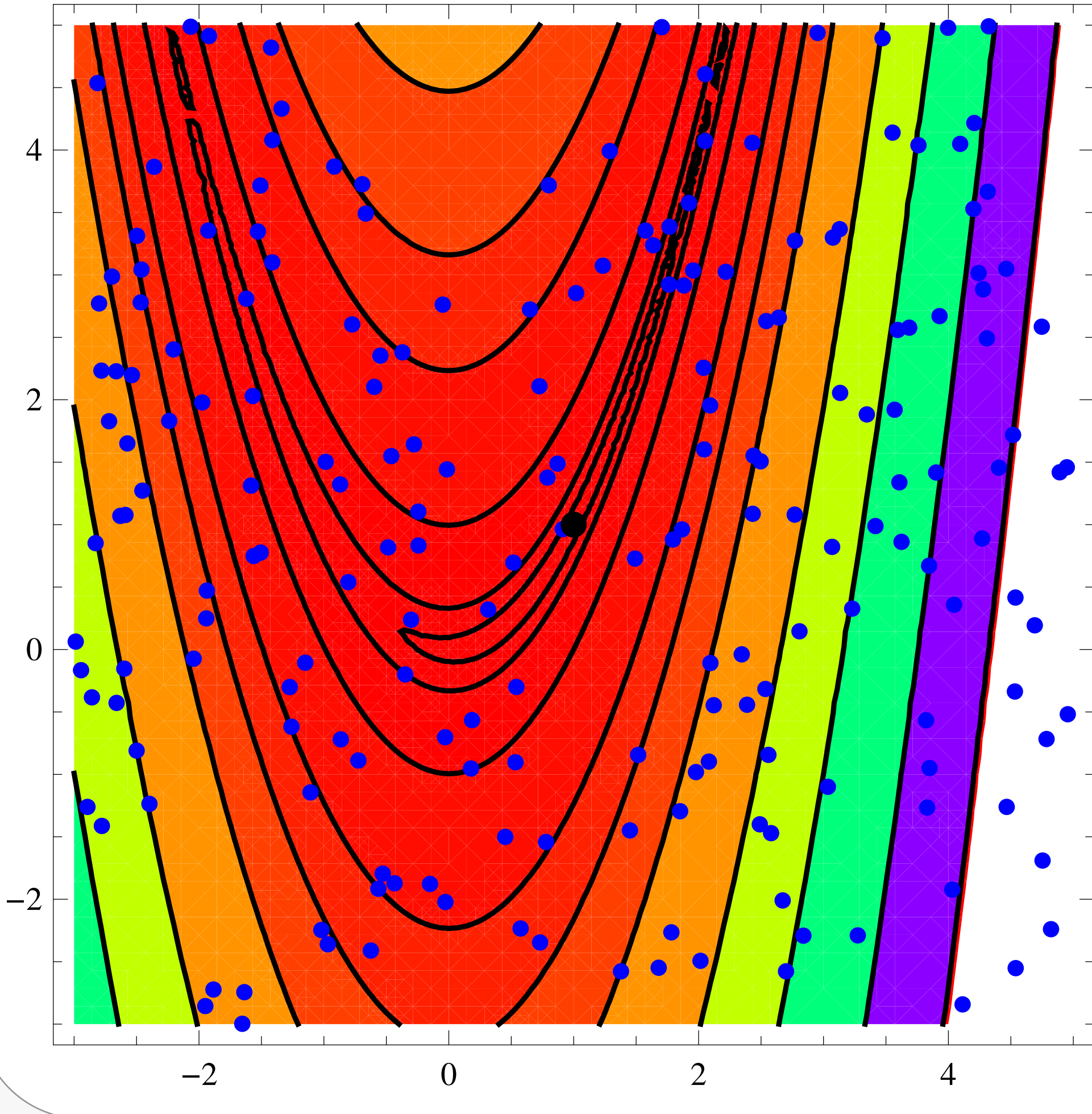
AF Avaliações da função efetuadas,

AJ Avaliações da jacobiana efetuadas,

TE Tempo total de execução (em segundos).

Função	CP	$\ F_*\ $	IT	AF	AJ	TE
ROS	1	0	18	19	14	0.16
BBS	1	$3E-13$	13	14	12	0.11
BEA	1	$6E-12$	5	6	6	0.16
POWS	2	$5E-5$	9	10	10	0.07
WOOD	1	$2E-9$	67	68	63	0.91
BDF	2	$3E2$	46	47	28	1.87
WAT	2	$2E-5$	5	6	6	1.01

Na segunda parte, tomamos a função *ROS* e testamos 200 pontos iniciais sorteados dentro da caixa $[-3, 5] \times [-3, 5]$, como ilustrado na figura abaixo. Em todos os pontos obtivemos convergência à solução (1,1), com médias de 7.735 iterações e 0.036s.



Conclusões

Neste trabalho, implementamos o método de Levenberg-Marquardt aliado às técnicas de região de confiança; os experimentos efetuados validaram o bom desempenho e a robustez do algoritmo proposto.

Referências

- [1] J.J. Moré, “The Levenberg-Marquardt algorithm: implementation and theory”, em: G.A. Watson, *Lecture Notes in Mathematics 630: Numerical Analysis*, Springer-Verlag, New York, 1978, pp. 105-116.
- [2] J.J. Moré, B.S. Garbow & K.E. Hillstrom, “Testing unconstrained optimization software”, *ACM Transactions on Mathematical Software*, Volume 7, pp. 17-41, 1981.