

Resolução de sistemas não lineares restritos utilizando a Inicialização Global Topográfica

Janaína I. da Costa¹
 IPRJ, Nova Friburgo, RJ
 Marroni de Sá Rêgo²
 IFPA, Óbidos, PA
 Luiz N. H. G. de Oliveira³
 IPRJ, Nova Friburgo, RJ
 Adriana da Rocha Silva⁴
 IPRJ, Nova Friburgo, RJ
 Narcisa C. da Silva⁵
 IPRJ, Nova Friburgo, RJ
 Sérgio S. de Sousa⁶
 IPRJ, Nova Friburgo, RJ

Resumo. Em geral, os métodos clássicos para resolver sistemas de equações não lineares são conhecidos por sua eficiência. Entretanto, dependem fortemente da localização dos pontos de partida. Neste trabalho, utilizamos a Inicialização Global Topográfica para gerar bons pontos iniciais para o método de busca local utilizado na resolução de problemas restritos de minimização global, cujas soluções são raízes de sistemas não lineares associados. Para realizar as tarefas de busca local, usamos o Algoritmo de Direções Viáveis e Pontos Interiores (FDIPA). Em seguida, utilizamos quatro problemas descritos na literatura para avaliar a eficácia da nossa metodologia. Os resultados indicaram que a presente abordagem é uma estratégia poderosa para encontrar todas as raízes de sistemas de equações não lineares.

Palavras-chave. FDIPA, Sistemas de Equações, Inicialização Topográfica

1 Introdução

Neste trabalho lidamos com o problema de encontrar todas as soluções de sistemas de equações não lineares com restrições de desigualdade. Tal problema pode ser expresso como,

$$F(x) = 0, x \in \Omega \subset \mathbb{R}^n, \quad (1)$$

onde $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é uma aplicação diferenciável e Ω é o conjunto de restrições dado por,

$$\Omega = \{x \in S \subset \mathbb{R}^n; g_i(x) \leq 0, \forall i = 1, \dots, l\}. \quad (2)$$

¹janainaimbiriba@gmail.com

²marronidesa2@gmail.com

³neliohenderson@gmail.com

⁴adrianarochanf@gmail.com

⁵narcisa.correa@gmail.com

⁶pfsergiosousa@yahoo.com.br

Aqui, g_i , $i = 1, \dots, l$ são por hipótese funções suaves e diferenciáveis, mas não necessariamente convexas. O subconjunto S é um hiper-cubo definido por $S = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$, onde $[a_j, b_j]$ é o intervalo real no qual a variável x_j está definida e “ \times ” denota o produto cartesiano.

O problema descrito na equação (1), pode ser reescrito como um problema de otimização. Para isso, utiliza-se a chamada função de mérito dada por,

$$f(x) = \sum_{j=1}^m [WF_j(x)F_j(x)]. \quad (3)$$

sendo W as constantes de penalidade. Note que $x \in \Omega$ é solução de (1) se, e somente se, $x \in \Omega$ é uma solução global do seguinte problema de minimização restrita,

$$\text{minimize } f(x) \text{ sujeito a } g_i(x) \leq 0, i = 1, \dots, l. \quad (4)$$

No problema (4), assumimos que o conjunto viável Ω possui o interior não vazio. Além disso, por definição, tal conjunto é limitado e fechado, garantindo a existência de soluções do problema de minimização restrita, ver [3].

Os métodos clássicos de otimização usados para resolver o problema (4), como os métodos tipo Newton, a cada iteração realizam uma busca local ao redor de um determinado ponto inicial [5]. Entretanto, devido as características locais da teoria matemática utilizada no desenvolvimento de tais métodos, a menos que os pontos iniciais sejam próximos às soluções de $F(x) = 0$, não é dada preferência às soluções globais de (4). Sendo esta uma severa limitação ao uso direto de tais métodos de otimização para solução do problema (4). Além disso, em geral, os pontos iniciais apropriados não estão disponíveis na formulação de problemas práticos da engenharia.

Para contornar esta dificuldade recorreu-se ao uso da Inicialização Global Topográfica, proposta em [7]. Tal algoritmo de agrupamento, utiliza uma abordagem baseada em conceitos elementares da teoria dos grafos, a fim de gerar bons pontos de partida para os métodos de busca local, a partir de pontos distribuídos de modo uniforme no interior da região viável. Conforme o que foi descrito em [7], essa estratégia de inicialização está dividida em três etapas: (i) a obtenção de N pontos amostrais aleatórios distribuídos uniformemente no espaço de busca, (ii) a construção do grafo (topografo), sendo este um grafo com arcos direcionados conectando a cada ponto amostral os seus k vizinhos mais próximos e (iii) a seleção dos mínimos do topografo que serão os pontos iniciais para o método de busca local.

Vale ressaltar que a escolha do parâmetro k tem um papel fundamental no desempenho da Inicialização Global Topográfica. Por esse motivo, em [1] os autores desenvolveram uma fórmula para estimar k , que utiliza apenas informações sobre os possíveis topografos associados aos pontos de amostragem.

No presente trabalho, usamos a inicialização global topográfica revisada em [1], para obter estimativas iniciais adequadas para o método de busca local FDIPA (Algoritmo de Direções Viáveis e Pontos Interiores), utilizado aqui para encontrar as soluções globais do problema (4).

2 Inicialização Global Topográfica

Seja Ω , um subconjunto limitado do espaço Euclidiano n -dimensional \mathbb{R}^n , cujo interior é um conjunto não vazio, e $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função real definida em Ω . A Inicialização Global Topográfica é um método que foi desenvolvido com o objetivo de selecionar pontos iniciais apropriados para métodos de minimização baseados na busca local.

Inicialmente, são escolhidos uniformemente N pontos amostrais em $\Omega \subset \mathbb{R}^n$, denotados por P_i , $i = 1, \dots, N$. Para cada ponto P_i , é construída uma lista de referência (uma lista de índices dos pontos). Isto é feito para organizar os outros $(N - 1)$ pontos, através da chamada ordem do

vizinho mais próximo. Assim, para P_i , o j -ésimo elemento desta lista é o ponto amostral j mais próximo de P_i . Essa lista é ainda complementada pela atribuição de um sinal para cada índice j , do seguinte modo:

$$j = \begin{cases} +j, & \text{se } f(P_j) \geq f(P_i) \\ -j, & \text{se } f(P_j) < f(P_i) \end{cases} \quad (5)$$

As N listas de referências constituem uma matriz $N \times (N - 1)$, que é chamada de matriz topográfica (t -matriz) da função f associada aos pontos de amostragem. A matriz topográfica pode ser interpretada por um grafo orientado, onde os sinais indicam as direções dos arcos no grafo. O sinal positivo representa a “ponta final” do arco e o sinal negativo a “ponta inicial” do arco. Este grafo associado é chamado de grafo topográfico (ou simplesmente topografo) de f associado com a referida amostra.

Dado um inteiro $1 \leq k \leq (N - 1)$, a submatriz $N \times k$ obtida considerando apenas os k vizinhos mais próximos é chamada de k - t -matriz, cujo grafo correspondente é o k^+ -topografo. O conjunto de todos os k^+ -topografos de f define a k^+ -topografia de f .

Seja P_i um ponto amostral, se a i -ésima linha da k - t -matriz é uma linha positiva, tal ponto é dito um minimizador local de f no k^+ -topografo. Aqui, todos os minimizadores locais de f no k^+ -topografo são selecionados para a busca local. Entretanto, de acordo com observações feitas por [7], uma vez que o número de minimizadores locais de f no k^+ -topografo aumenta conforme o valor de k cresce, este procedimento pode levar a um grande custo computacional no passo de busca local. Por esse motivo, o parâmetro k deve ser escolhido de forma adequada. Nesse contexto, em [1] foi proposta uma fórmula para o cálculo de k . Utilizando esta fórmula os autores obtiveram, na maioria dos seus experimentos, $k = 4$. Por esse motivo, no presente trabalho utilizamos apenas os 4 vizinhos mais próximos.

Para exemplificar o uso da Inicialização Global Topográfica, consideremos o seguinte problema: obtenha $x \in \Omega$, onde $\Omega = \{x \in \mathbb{R}; -1 \leq x \leq 0\}$, tal que,

$$F(x) = \text{sen}(x) + x^4 + x^2 = 0. \quad (6)$$

Definindo $W = 10$, a função de mérito é dada por,

$$f(x) = 10(\text{sen}(x) + x^4 + x^2)^2. \quad (7)$$

Observe que Ω é um conjunto com restrições de caixa, que são desigualdades simples. Além disso, nesse conjunto a função f possui dois minimizadores globais, a saber: $-0,652593$ e 0 . Assim, o objetivo é determinar tais minimizadores.

Para isso, são gerados $N = 10$ pontos amostrais P_i em Ω , $i = 1, \dots, N$. Tais pontos são descritos na Tabela 1 com os respectivos valores de f .

Considerando apenas os $k = 4$ vizinhos mais próximos, a 4 - t -matriz é dada por,

$$4-t-matriz = \begin{bmatrix} -8 & -9 & -4 & -5 \\ 7 & -5 & 10 & 4 \\ 10 & 7 & -6 & -2 \\ 9 & 8 & -5 & -2 \\ 4 & 2 & 9 & 8 \\ 3 & 10 & 7 & 2 \\ -10 & -3 & -2 & -6 \\ -9 & -4 & 1 & -5 \\ 8 & -4 & -5 & 1 \\ 7 & -3 & -6 & -2 \end{bmatrix} \begin{matrix} \leftarrow 1 \\ \leftarrow 2 \\ \leftarrow 3 \\ \leftarrow 4 \\ \leftarrow 5 \\ \leftarrow 6 \\ \leftarrow 7 \\ \leftarrow 8 \\ \leftarrow 9 \\ \leftarrow 10 \end{matrix}. \quad (8)$$

Tabela 1: Pontos amostrais P_i com os respectivos valores de f .

i	P_i	$f(P_i)$
1	-1	13,4219
2	-0,5	0,278641
3	-0,25	0,327602
4	-0,75	0,389145
5	-0,625	0,017543
6	-0,125	0,118387
7	-0,375	0,423833
8	-0,875	3,41363
9	-0,8125	1,36866
10	-0,3125	0,400983

Note que na $4 - t -$ matriz apenas as linhas 5 e 6, correspondentes aos pontos P_5 e P_6 , são positivas. Assim, os pontos P_5 e P_6 são os minimizadores locais de f no $4^+ -$ topografo e são as estimativas iniciais para o método de busca local.

Em [7], os autores inferem experimentalmente que a Inicialização Global Topográfica deve ser usada em conjunto com pontos amostrais suficientes e distribuídos uniformemente em Ω . Este fato também foi descrito por [1]. Aqui, utilizamos o gerador de amostras uniformes chamado de sequência de Sobol, ver [6]. Observe que inicialmente são gerados N pontos amostrais dentro de uma caixa (restrições simples) e em seguida é verificado quais desses pontos satisfazem (caso existam) as demais condições de desigualdade.

3 Algoritmo de Direções Viáveis e Pontos Interiores

Dada uma estimativa inicial viável adequada, o método FDIPA proposto por Herskovits em [2], gera sequências de pontos em Ω com valores decrescentes da função de mérito f , que converge para o conjunto de pontos KKT do problema (4). Assim, em cada iteração, este método de pontos interiores realiza uma busca linear em uma direção de descida de f , que é também uma direção viável do problema (4).

Se $x \in \Omega$ é um minimizador local de f em Ω , as condições de otimalidade KKT do problema são dadas por:

$$\nabla f(x) + \nabla g(x)\mu = 0, \tag{9}$$

$$G(x)\mu = 0, \tag{10}$$

$$g(x) \leq 0, \tag{11}$$

$$\mu \geq 0, \tag{12}$$

onde $G(x) \in \mathbb{R}^{m \times m}$ denota a matriz diagonal definida como $G_{ii}(x) = g_i(x)$, para todo $i = 1, \dots, m$.

As equações (9) e (10) descrevem um sistema não linear em (x, μ) . Dados $\mu > 0$ e $x \in \Omega^0$, onde Ω^0 denota o interior de Ω , ou seja, $\Omega^0 = \{x \in \mathbb{R}^n; g(x) < 0\}$. Aplicando o método de Newton para resolver este sistema não linear, com o ponto inicial (x, μ) , obtemos (d_α, μ_α) que é solução do seguinte sistema linear,

$$\begin{bmatrix} B(x, \mu) & \nabla g(x) \\ M\nabla g(x)^T & G(x) \end{bmatrix} \begin{bmatrix} d_\alpha \\ \mu_\alpha \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ 0 \end{bmatrix}. \tag{13}$$

Aqui, $M \in \mathbb{R}^{m \times m}$ é a matriz diagonal definida por $M_{ii} = \mu_i \neq 0, i = 1, \dots, m$. A matriz $B(x, \mu)$ é igual a hessiana $H(x, \mu) = \nabla^2 L_{xx}(x, \mu) = \nabla^2 f(x) + \sum_{i=1}^m \nabla^2 g_i(x) \mu_i$, ou ainda, uma aproximação quasi-Newton de $H(x, \mu)$.

De acordo com [2], o vetor d_α pode não ser uma direção viável para o problema de minimização (4). Para contornar esta dificuldade, em [2] o autor propõe a solução de um sistema linear adicional nas novas variáveis d_β e μ_β , dado por,

$$B(x, \mu)d_\beta + \nabla g(x)\mu_\beta = 0, \tag{14}$$

$$M\nabla g(x)^T d_\beta + G(x)\mu_\beta = -\mu. \tag{15}$$

Assim, a nova direção de busca é dada por,

$$d = d_\alpha + \rho d_\beta, \tag{16}$$

onde $\rho > 0$ é tal que, se $\nabla f(x)^T d_\beta > 0$, então

$$\rho \leq \frac{(\xi - 1)\nabla f(x)^T d_\alpha}{\nabla f(x)^T d_\beta}, \quad \xi \in (0, 1). \tag{17}$$

Note que a direção de busca d é computada em dois passos. Primeiramente, resolvemos o sistema linear (13) a fim de obter (d_α, μ_α) . O segundo passo é resolver o sistema linear dado pelas equações (14) e (15). Finalmente, a direção de busca deste método de pontos interiores é determinado pela combinação linear descrita na equação (16). Observa-se que estes dois sistemas lineares possuem a mesma matriz dos coeficientes.

Obtida a direção de busca d , realiza-se uma busca linear inexata que obedece a condição de Armijo. Assim, o novo ponto é determinado por $\hat{x} = x + td$, onde t é o comprimento do passo. A matriz B (simétrica e definida positiva) é atualizada usando a fórmula de atualização quasi-Newton *BFGS* com a modificação de Powell, ver [2]. Assim, a cada iteração obtemos uma aproximação B da matriz hessiana da função Lagrangiana do problema (4).

4 Resultados

Para avaliar a eficácia da metodologia aqui apresentada, utilizamos quatro exemplos descritos na literatura, ver [4] e [8]. Os problemas 1 e 2 possuem apenas restrições de caixa, onde o teste 1 apresenta duas soluções. Os problemas 3 e 4 além das restrições de caixa, apresentam também restrições de desigualdade de outros tipos.

Os resultados numéricos foram obtidos utilizando um computador equipado com o processador Intel® Core™ i7-4510U, 8 GB de memória e sistema operacional Ubuntu 15.04. O código foi desenvolvido em linguagem C.

Problema 1:

$$F(x) = \begin{bmatrix} 2x_1 + x_2 + x_3 + x_4 + x_5 - 6 \\ x_1 + 2x_2 + x_3 + x_4 + x_5 - 6 \\ x_1 + x_2 + 2x_3 + x_4 + x_5 - 6 \\ x_1 + x_2 + x_3 + 2x_4 + x_5 - 6 \\ x_1 x_2 x_3 x_4 x_5 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{18}$$

onde $-2 \leq x_i \leq 2, i = 1, \dots, 5$.

Problema 2:

$$F(x) = \begin{bmatrix} x_1 x_2 + x_1 - 3x_5 \\ 2x_1 x_2 + x_1 + 3nx_2 x_2 + x_2 x_3 x_3 + dx_2 x_3 + mx_2 x_4 + ex_2 - ax_5 \\ 2x_2 x_3 x_3 + dx_2 x_3 + 2bx_3 x_3 + cx_3 - 8x_5 \\ mx_2 x_4 + 2x_4 x_4 - 4ax_5 \\ x_2(x_1 + x_3^2 + dx_3 + mx_4 + e) + x_1 + nx_2^2 + bx_3^2 + cx_3 + x_4^2 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{19}$$

com $a = 10$; $b = 0,193$; $c = 4,106 \times 10^{-4}$; $d = 5,451 \times 10^{-4}$; $e = 4,497 \times 10^{-7}$; $m = 3,407 \times 10^{-5}$; $n = 9,615 \times 10^{-7}$ e $0,0001 \leq x_i \leq 100, i = 1, \dots, 5$.

Problema 3:

$$F(x) = \begin{bmatrix} x_1 + x_1^3 - 10 \\ x_2 - x_3 + x_2^3 + 1 \\ x_2 + x_3 + 2x_3^3 - 3 \\ 2x_4^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (20)$$

onde

$$x_1 + x_2 + x_3 + x_4 - 3 \leq 0, \quad (21)$$

com $0 \leq x_i \leq 3, i = 1, \dots, 4$.

Problema 4:

$$F(x) = \begin{bmatrix} (2 + 0,5x_3)x_1 + x_4 - 4 \\ (2 + 2x_3)x_2 - 2x_4 - 2 \\ x_1 - 2x_2 + 1 \\ x_3(0,25x_1^2 + x_2^2 - 1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (22)$$

onde

$$0,25x_1^2 + x_2^2 - 1 \leq 0, \quad (23)$$

com $0 \leq x_i \leq 2, i = 1, \dots, 4$.

Na Tabela 2, são apresentadas as soluções obtidas utilizando a metodologia aqui apresentada. Para cada problema teste, apresentamos os minimizadores globais (denotados por x^*) com os respectivos valores da função de mérito f (denotados por f^*), o número de pontos amostrais gerados (indicado por N) e o número de estimativas iniciais (denotado por \bar{N}).

Tabela 2: Soluções obtidas para os problemas 1-4 utilizando a metodologia aqui apresentada.

Problema	N	\bar{N}	x^*	f^*
1	100	10	(0,91635; 0,91635; 0,91635; 0,91635; 1,4183) (1; 1; 1; 1)	$5,23 \times 10^{-11}$ $1,17 \times 10^{-10}$
2	100	14	(0,003373; 31,852; 0,06779; 0,8595; 0,03696)	$6,73 \times 10^{-9}$
3	150	1	(2; $3,62 \times 10^{-8}$; 1; $1,81 \times 10^{-8}$)	$2,65 \times 10^{-13}$
4	150	5	(0,82286; 0,91143; 1,8467; 1,5945)	$2,32 \times 10^{-9}$

Observe que o número N aumenta conforme são adicionadas restrições além das restrições de caixa. De fato, quanto mais rígidas as restrições de desigualdade, maior o valor de N de forma a obter pontos no interior da região viável.

Na Tabela 3 são apresentados os números de avaliações (NAF) da função de mérito f necessárias para determinar as soluções de cada sistema. Também são mostrados os tempos (em segundos) necessários para obter tais soluções utilizando a presente metodologia.

Tabela 3: Tempo e o número de avaliações da função de mérito necessários para obter as soluções descritas na Tabela 2.

Problema	Tempo (s)	NAF
1	0,031	27314
2	0,038	31500
3	0,005	420
4	0,018	33375

5 Conclusões

No presente trabalho, testamos a eficácia da Inicialização Global Topográfica na resolução de sistemas de equações não lineares, equipados com restrições de desigualdade. Para isso, tal problema foi reformulado como um problema de otimização global com restrições, cujas soluções são as raízes dos sistemas não lineares. A Inicialização Global Topográfica foi então empregada para gerar pontos iniciais apropriados para o método de busca local. Tal busca foi realizada usando o Algoritmo de Direções Viáveis e Pontos Interiores (FDIPA).

Experimentos computacionais foram realizados usando diferentes problemas de teste previamente estudados na literatura, com o objetivo de verificar se a metodologia aqui apresentada é capaz de obter as soluções dos sistemas. Os resultados obtidos mostram que a presente abordagem é uma poderosa ferramenta numérica para encontrar as raízes de sistemas não lineares restritos.

Agradecimentos

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

Os autores também agradecem o apoio financeiro da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro - FAPERJ e do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq.

Referências

- [1] Henderson, N., Sá Rêgo, M., Sacco, W. F. e Rodrigues, R. A. A new look at the topographical global optimization method and its application to the phase stability analysis of mixtures, *Chemical Engineering Science*, 127:151-174, 2015. DOI:10.1016/j.ces.2015.01.029.
- [2] Hershkovits, J. Feasible Direction Interior-Point Technique for Nonlinear Optimization, *Journal of Optimization Theory and Applications*, 99:121-146, 1998. DOI:10.1023/A:1021752227797.
- [3] Ismailov, A., Solodov, M. V. *Otimização: Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade. 2a edição.* IMPA, Rio de Janeiro, 2009.
- [4] Maranas, C.D., Floudas, C.A. Finding all solutions of nonlinearly constrained systems of equations, *Journal of Global Optimization*, 7:143-182, 1995. DOI:10.1007/BF01097059.
- [5] Nocedal, J., Wright, S.J. *Numerical Optimization, 2a. edição.* Springer, New York, 2006.
- [6] Sobol, I. M. On the distribution of points in a cube and the approximate evaluation of integrals, *USSR Computational Mathematics and Mathematical Physics*, 7:86-112, 1967. DOI:10.1016/0041-5553(67)90144-9.
- [7] Törn, A. e Viitanen, S.V. Topographical global optimization. In *Recent Advances in Global Optimization*. Princeton University Press, 1992. ISBN: 9780691631875.
- [8] Yang, L., Chen, Y., Tong, X. e Deng, C. A new smoothing Newton method for solving constrained nonlinear equations, *Applied Mathematics and Computation*, 217:9855-9863, 2011. DOI:10.1016/j.amc.2011.04.045.