XXI ENMC
ENCONTRO NACIONAL DE MODELAGEM COMPUTACIONAL
IX ECTM
ENCONTRO DE CIÊNCIA E TECNOLOGIA DE MATERIAIS

08 a 11 de Outubro de 2018
Instituto Federal Fluminense
Búzios - RJ

Estudo sobre a utilização dos métodos de Luus-Jaakola e PSO em relação ao método de Newton para resolução de sistemas de equações não lineares

Rennan Mendes de Moraes dos Santos Dias[1] - rennandias2@hotmail.com
Lucas de Souza Almeida[1] - lucassouza.almeida15@gmail.com
Bárbara Lucena Gomides Ferreira [1] - barbaralucena@id.uff.br
Ramon Uicker Bertussi[1] - rmuicker@gmail.com
Rosilene Abreu Portella Corrêa[1] - rosileneportella@id.uff.br
[1]Universidade Federal Fluminense, INFES - Santo Antônio de Pádua, RJ, Brazil

Resumo. *No presente trabalho será abordada a viabilização da aplicação de métodos de otimização estocásticos, Luss-Jaakola e Método de Otimização por Enxame de Particulas (PSO) para resolução de sistemas de equações não lineares no lugar dos determinísticos. Para efeito de uma análise comparativa, focou-se no método de Newton, como determinístico, levando-se em consideração que é o mais amplamente estudado e conhecido quando busca-se resolver esse tipo de sistemas. A escolha dos métodos estocásticos é justificada pela baixa dependência do chute inicial e do não uso de derivadas.*

Palavras-chave. *Sistemas de equações não lineares, Luss-Jaakola, PSO e Newton.*

1. INTRODUÇÃO

Os sistemas de equações não lineares são capazes de modelar inúmeros problemas práticos. Segundo Santos (2016), A necessidade de resolver um sistema de equações não lineares emerge naturalmente durante a modelagem de problemas práticos relacionados às mais diferentes áreas. Existem muitas aplicações que abrangem os sistemas não lineares, e elas incluem otimização estrutural, design de engenharia, problemas de alocação e locação, tarefas quadráticas e vários outros problemas de otimização assim como programação inteira e problemas gráficos. Na resolução por métodos iterativos de um sistema de equações, é possível transformá-lo em um problema de otimização, submetido à minimização da função objetivo equivalente às soluções do sistema original.

Neste trabalho foram usados dois sistemas de equações não lineares. A partir dos mesmos será feita uma análise da viabilização dos métodos estocásticos, Luus-Jaakola e PSO, no lugar do método de Newton. Com isso, prentende-se mostrar que o PSO e o Luus-Jaakola também encontram soluções de sistemas não lineares e que além disso, não precisam do uso de derivadas e, dessa forma, são mais simples e práticos de serem implementados.

Para efeito de um estudo comparativo utilizou-se o trabalho de SILVA (2017) para a obtenção de todas as raízes de cada problema escolhido.

2. SISTEMAS DE EQUAÇÕES NÃO LINEARES

Seja,

$$\begin{cases} f_1(x_1, x_2, \ldots, x_n) = 0 \\ f_2(x_1, x_2, \ldots, x_n) = 0 \\ \quad\vdots \quad\quad \vdots \quad\quad \vdots \\ f_n(x_1, x_2, \ldots, x_n) = 0 \end{cases} \tag{1}$$

a representação de um sistema de equações não lineares. Também pode-se escrever cada equação representada pelas f_n na forma vetorial, assim:

$$f = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{2}$$

Para transformar uma sistema de equações não lineares em um problema de minimização não linear sem restrições, com intuito de facilitar e ampliar a aplicação de métodos computacionais, necessita-se da criação de uma função objetivo $f_0(x)$ para obtenção do ponto ótimo, sendo a mesma gerada a partir da função vetorial (2). Esse processo é realizado pela norma L_p de Holder, definida como:

$$f_0 \equiv L_p = \left\{ \sum_{k=1}^{n} |f_k(x)|^p \right\}^{\frac{1}{p}} \tag{3}$$

onde $p \in \mathbb{N}$ e n é a dimensão do espaço vetorial.

Em problemas que envolvem esse tipo de processo tem-se três casos mais clássicos: $p = 1$, $p = 2$ e $p \to +\infty$. Para todos os métodos estocásticos estudados nesse trabalho utilizou-se $p = 2$ e omitiu-se a raiz, por conseguinte, será minimizada a soma dos quadrados dos componentes. Para maiores informações sobre a utilização do norma de Holder em problemas de otimização, veja VIANA (2017). Agora, a equação (3) é denotada por

$$f_0 \equiv L_2 = \sum_{k=1}^{n} |f_k(x)|^2 \tag{4}$$

3. Métodos utilizados na resolução dos sistemas não lineares

3.1 Método de Otimização por Enxame de Particulas - PSO

O método de otimização PSO (Particle Swarm Optimization), desenvolvido por Kennedy e Eberhart, é inspirado em um modelo de interação social onde cada indivíduo, ou partícula, na posição $\mathbf{p}$, é uma solução candidata, a qual se move no espaço de busca com velocidade $\mathbf{v}$, dinamicamente ajustada de acordo com sua própria experiência de movimento e pela experiência de movimento do grupo (KENNEDY & EBERHARD, 1995).

Em cada geração t, cada uma das Q partículas são manipuladas de acordo com as seguintes relações,

$$\mathbf{v}_i(t+1) = w\mathbf{v}_i(t) + r_1c_1(\mathbf{b}_i(t) - \mathbf{p}_i(t)) + r_2c_2(\mathbf{b}_g(t) - \mathbf{p}_i(t)); \tag{5}$$

$$\mathbf{p}_i(t+1) = \mathbf{p}_i(t) + \mathbf{v}_i(t+1), \tag{6}$$

onde w é denominado termo de inércia, c_1 e c_2 são coeficientes de aceleração; r_1 e r_2 são valores randômicos no intervalo [0,1]; $\mathbf{b}_i$ é a melhor posição ocupada pela partícula i, e $\mathbf{b}_g$ é a melhor posição encontrada por todo o grupo.

3.2 Método de Luus-Jaakola

O método estocástico Luus-Jaakola, criado por Luss e Jaakola em 1973, define-se por sua busca aleatória, utilizando somente as informações da função objetivo e ao longo de suas iterações, ocorre uma contração gradativa do seu raio de busca, fazendo com que o intervalo no qual se encontra a solução seja reduzido o máximo possível (LUUS & JAAKOLA, 1973). Ao contrário do método populacional PSO, LJ gera apenas uma estimativa inicial, que ao decorrer de cada iteração é melhorada tendendo a convergir para o ótimo da função (SOUZA & OLIVEIRA, 2016).

Para iniciar o método, escolhe-se o tamanho de busca inicial $\mathbf{r}^0$; o número de laços internos (n_{int}) e externos (n_{ext}); o coeficiente de contração c e a solução inicial $x^* = x_0$.

Para cada iteração j do laço interno, uma nova candidata a solução do problema será dada por,

$$x^j = x^* + \mathbf{R}^j\mathbf{r}^{(j-1)} \tag{7}$$

onde $\mathbf{R}$ é uma matriz diagonal formada por números aleatórios entre -0,5 e 0,5. Para cada nova solução candidata, o valor do funcional é avaliado e, caso um menor valor seja encontrado (em um problema de minimização) faz-se,

$$x^* = x^j. \tag{8}$$

Ao final de cada laço interno, enquanto o número de iteração i for menor que n_{ext}, o raio de busca deve ser contraído da seguinte forma,

$$\mathbf{r}^i = (1-c)\mathbf{r}^{(i-1)}. \tag{9}$$

3.3 Método de Newton para sitemas de equações não lineares

O método de Newton consiste em aproximar uma função não linear por meio de um modelo linear local, tendo em vista que essa aproximação é feita em uma vizinhança de um ponto x_n no qual procura-se encontrar uma solução para o sitema de equações não lineares.

Quando é aplicado em funções de uma variável real, esse modelo é gerado a partir da reta tangente da função no ponto. Entretanto, expandido esse conceito para funções com mais de uma variável, necessita-se do cálculo de derivadas parciais e, por consequência, da matriz jacobiana. Matemáticamente, tem-se a seguinte equação que resume o método de Newton,

$$x_{n+1} = x_n - J^{-1}(x_n)F(x_n) \tag{10}$$

Levando em consideração que,

$$J = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix} \tag{11}$$

representa a matriz jacobiana, então J^{-1} corresponde a sua inversa.

Esse método é determinístico e comumente utilizado na resolução de sistemas de equações não lineares, pois possui uma boa precisão. Em contrapartida é definido por meio de derivadas, ou seja, é necessário que a função seja diferenciável para o funcionamento do método (OLIVEIRA & MARTINEZ, 2016).

4. PROBLEMAS PROPOSTOS

Neste artigo serão apresentados dois problemas envolvendo sistemas de equações não lineares.

O primeiro é um problema com equações trigonométricas, usado para testar o algoritmo C-GRASP, por HIRSCH et. al (2006), para resolver sistemas não-lineares. Este problema possui 13 soluções. Abaixo segue as equações que representam o sistema,

$$\begin{cases} -\sin(x_1)\cos(x_2) - 2cos(x_1)\sin(x_2) = 0 \\ \cos(x_1)\sin(x_2) - 2\sin(x_1)\cos(x_2) = 0 \\ \qquad 0 \leq x \leq 2\pi \end{cases} \tag{12}$$

Já o segundo problema é de Bini e Mourrain de 2004

$$\begin{cases} -x_2^2x_3^2 - x_2^2 + 24x_2x_3 - x_3^2 - 13 = 0 \\ -x_1^2x_3^2 - x_1^2 + 24x_1x_3 - x_3^2 - 13 = 0 \\ -x_1^2x_2^2 - x_1^2 + 24x_1x_2 - x_2^2 - 13 = 0 \\ \qquad 0 \leq x_1, x_2, x_3 \leq 20 \end{cases} \tag{13}$$

Todos os dois exemplos dispostos acima foram tirados da tese de SILVA (2017) que por sua vez serviu como mecanismo de comparação no momento da obtenção dos resultados.

5. RESULTADOS

Os sistemas apresentados na seção anterior foram escolhidos para servirem de exemplo pois apresentam grande número de soluções e, além disso, o segundo deles necessita de 9 derivadas para ser executado no método de Newton. Para obter os resultados foi utilizado um computador com processador AMD Phenom II X4 830 e memória RAM de 4,0GB. Assim, pretendemos testar a eficiência dos métodos estocásticos propostos na busca por soluções de sistemas não-lineares.

5.1 Método de Otimização Luus-Jaakola

Resultados com critério de parada para o valor da função objetivo de 10^{-7}, para laços externos igual a 100, laços internos igual a 200 e coeficiente de contração igual a 0,2, após 200 execuções.

Tabela 1- Tabela referente as raízes do primeiro sistema de equações não lineares encontradas pelo Luus-Jaakola

Tempo de execução para 100 execuções = 29,361285 s.			
Raízes	x_1	x_2	$f(x_1, x_2)$
1	3.1415469	6.2831483	0.0000000
2	1.5707688	1.5707218	0.0000000
3	1.5709282	4.7123458	0.0000000
4	0.0000318	0.0000213	0.0000000
5	3.1415944	3.1416633	0.0000000
6	4.7124805	4.7124738	0.0000000
7	0.000074	3.1415138	0.0000000
8	4.7125049	1.5708138	0.0000000
9	6.2832568	3.1416041	0.0000000
10	3.1416704	0.0000577	0.0000000
11	0.0000647	6.2832023	0.0000000
12	6.2830759	6.2832553	0.0000000
13	6.2832661	0.0001085	0.0000000

Para o sistema de Bini e Mourrain teve-se os resultados com critério de parada para o valor da função objetivo de 10^{-7}, para laços externos igual a 100, laços internos igual a 200 e coeficiente de contração igual a 0,2, após 200 execuções onde 50 destas não convergiram suficientemente. Segue a tabela que mostra todas as raízes encontradas

Tabela 2- Tabela referente as raízes do segundo sistema de equações não lineares encontradas pelo Luus-Jaakola

Tempo de execução para 100 execuções = 209.89969 s.				
Raízes	x_1	x_2	x_3	$f(x_1, x_2, x_3)$
1	10.857714	0.779552	0.7795505	0.0000000
2	0.7795515	0.7795495	10.857721	0.0000000
3	0.7795527	10.85773	0.7795517	0.0000000
4	0.7795503	0.7795547	0.7795535	0.0000000
5	4.6251797	4.6251822	4.6251817	0.0000000

5.2 Método de otimização por enxame de partículas-PSO

Resultados obtidos com critério de parada para o valor da função objetivo de 10^{-7}, para 150 indivíduos, coeficiente do resultado individual igual a 0,7, coeficiente do resultado de grupo igual a 0,7 e coeficiente da velocidade igual a 0,7, após 200 execuções.

Tabela 3- Tabela referente as raízes do primeiro sistema de equações não lineares encontradas pelo PSO

Tempo de execução para 100 execuções = 76,711476 s.			
Raízes	x_1	x_2	$f(x_1, x_2)$
1	3.1415467	3.1416388	0.0000000
2	3.1416986	6.2831252	0.0000000
3	6.2832119	3.1415799	0.0000000
4	6.2833104	0.0000331	0.0000000
5	4.7125202	4.7123837	0.0000000
6	4.7124517	1.5707809	0.0000000
7	1.5708929	1.5708979	0.0000000
8	1.570658	4.712367	0.0000000
9	0.0000242	6.283193	0.0000000
10	0.0000265	3.1415179	0.0000000
11	0.0000667	0.0000563	0.0000000
12	6.2832424	6.2832752	0.0000000
13	3.14156	0.000032	0.0000000

Para o problema de Bini e Mourrain foram encontrados os resultados com critério de parada para o valor da função objetivo de 10^{-7}, usando-se 150 indivíduos, coeficiente cognitivo individual igual a 0,7, coeficiente cognitivo de grupo igual a 0,7 e coeficiente da velocidade igual a 0,7, após 200 execuções onde 11 destas não convergiram suficientemente. Segue abaixo a tabela com as raízes obtidas

Tabela 4- Tabela referente as raízes do segundo sistema de equações não lineares encontradas pelo PSO

Tempo de execução para 100 execuções = 189,53589 s.				
Raízes	x_1	x_2	x_3	$f(x_1, x_2, x_3)$
1	0.7795491	10.857695	0.7795464	0.0000000
2	10.857724	0.7795547	0.7795521	0.0000000
3	0.7795361	0.7795425	0.7795558	0.0000000
4	4.6251675	4.6251985	0.3320731	0.0000000
5	4.6251452	0.3320742	4.625218	0.0000000
6	0.7795463	0.7795452	10.857704	0.0000000
7	0.3320724	4.6251896	4.6251763	0.0000000
8	4.6251828	4.6251835	4.6251802	0.0000000

5.3 Método de Newton

Resultados obtidos com critério de parada da norma infinita, entre o resultado anterior e o resultado atual, menor que 10^{-7}, após 200 execuções para o sistema trigonométrico. E os chutes inicias foram dados de forma aleatória dentro do intervalo de busca da solução do sistema.

Tabela 5- Tabela referente as raízes do primeiro sistema de equações não lineares encontradas pelo método de Newton

Tempo de exeução para 100 execuções = 0.2444796 s.			
Raízes	x_1	x_2	$f(x_1, x_2)$
1	1.5705963	3.1416388	0.0000000
2	3.1415927	3.1415927	0.0000000
3	6.2831853	3.1415927	0.0000000
4	3.1415927	0.0000000	0.0000000
5	4.712389	1.5705963	0.0000000
6	1.5705963	1.5705963	0.0000000
7	0	0	0.0000000
8	3.1415927	6.2831853	0.0000000
9	6.2831853	0	0.0000000
10	0	3.1415927	0.0000000
11	4.712389	4.712389	0.0000000
12	6.2831853	6.2831853	0.0000000
13	0	6.2831853	0.0000000

Para o problema de Bini e Mourrain os resultados obtidos com critério de parada da norma infinita, entre o resultado anterior e o resultado atual, menor que 10^{-7}, após 200 execuções foram os mostrados na tabela abaixo.

Tabela 6- Tabela referente as raízes do segundo sistema de equações não lineares encontradas pelo método de Newton

Tempo de execução para 100 execuções = 0,7375828 s.				
Raízes	x_1	x_2	x_3	$f(x_1, x_2, x_3)$
1	4.6251816	0.3320731	4.6251816	0.0000000
2	4.6251816	4.6251816	4.6251816	0.0000000
3	0.3320731	4.6251816	4.6251816	0.0000000
4	4.6251816	4.6251816	0.3320731	0.0000000
5	10.857704	0.779548	0.779548	0.0000000
6	0.779548	0.779548	0.779548	0.0000000
7	0.779548	0.779548	10.857704	0.0000000

6. CONCLUSÃO

A partir dos resultados obtidos, pode-se construir a seguinte tabela comparativa:

Tabela 7- Tabela referente aos resultados encontrados pelos métodos de LJ, PSO e Newton

Sistemas	Quantidade de raízes encontradas pra cada sistema			Número de soluções
	Luus-Jaakola	PSO	Newton	
Sistema Trigonométrico	13	13	13	13
Bini e Mourrain	5	8	7	8

Dessa forma, pode-se observar que os métodos estocásticos Luus-Jaakola e PSO foram capazes de encontrar as soluções dos dois sistemas. Como no primeiro sistema que foi possível determinar as 13 raízes do problema. Além disso, para o segundo sistema, somente o método estocástico PSO foi capaz de encontrar as 8 soluções, após 200 execuções. Vale atentar que, numa varredura completa (no lugar da aleatória) do intervalo de busca, é esperado que o método de Newton encontre, também, todas as soluções.

Percebe-se que o tempo de execução, dos métodos de otimização, são maiores em comparação ao de Newton, no entanto essa diferença é razoável tendo em vista que os métodos determinísticos utilizam derivadas para encontrar a solução do problema. Em compensação os métodos estocásticos não utlizam as derivadas e, dessa forma, a complexidade ou, possíveis problemas envolvendo a continuidade da função, não são relevantes. Isso é fácil de observar no sistema de Bini e Mourrain já que há a necessidade de calcular nove derivadas.

Sendo assim, confrontando a quantidade de soluções encontradas, o tempo de execução e a praticidade de implementação dos métodos, pode-se concluir que o método de otimização estocástico populacional PSO foi o mais viável para resolver os sistemas não-lineares propostos, quando os mesmos foram transformados em problemas de otimização.

Referências

CORRÊA, R. A. P. *Identificação de danos em estruturas bi-dimensionais via matriz de flexibilidade baseada em um modelo de dano contínuo*. Tese de Doutorado, UERJ, Nova Friburgo, RJ, 2013, 39-43.

SILVA, M. R. *Aplicação de um algoritmo híbrido estocástico-determinístico na solução de um problema de direção automotiva*. XXXVII Encontro nacional de Engenharia de Produção. 2017, 80-101.

SOUZA e OLIVEIRA, J. S. e M. B. *Resolução de Sistemas não Lineares através do Algoritmo Luus Jaakola*. CNAMAC 2016, Gramado, Rio Grande do Sul, Brasil.

OLIVEIRA & MARTINEZ, G. L. e A. L. M. *Um estudo introdutório do método de Newton para minimização funções de uma ou várias variáveis*. CMAC Sudeste 2013,Bauru, São Paulo, Brasil.

VIANA, P. E. S.*Método de Simulação Estocástica Aplicados em Problemas Não Lineares*. Dissertação de Pós-Graduação, UFF, Volta Redonda, RJ, 2017, 12-31.

SANTOS, T. M. dos. *Um estudo sobre a resolução de sistemas não lineares : perspectivas teóricas e aplicações*. Tese de Pós-Graduação, Unicamp, SP, 2016.

LUUS, R. e JAAKOLA, T.H.I. *Optimization by Direct Search and Systematic Reduction of the Size of Search Region*. AIChE Journal, 19, 1973, 760-766.

KENNEDY, James e EBERHART, Russell. *Particle Swarm Optimization*. Proceedings of the IEEE International Conference on Neural Networks, 1995.

HIRSCH M. J.; MENESES C. N.; PARDALOS P. M.; RESENDE M.G.C. *Global optimization by continuous grasp*. Mar. 8, 2006. Disponível em:<http://www.optimization-online.org/DB_FILE/2006/03/1344.pdf>.

Study on the use of Luus-Jaakola and PSO methods in relation to Newton's method for solving systems of non-linear equations

Resumo. *In the present work, the stochastic optimization method of the Luss-Jaakola and Particle Swarm Optimization Method (PSO) for solving systems of non-linear equations are studied. For the purpose of a comparative analysis, it was focused on Newton's method because it is widely used in solving this type of problem. The choice of stochastic methods is justified by the low dependence of initial points and non-use of derivatives.*