

Análise dos métodos Steepest Ascent/Descente e evolução diferenciada para a caixa preta.

Pimentel, U.F.^{1*}; Machado, G. A. G.²; Da Silva, W. B.

¹ Programa Graduação em Engenharia Química, Universidade Federal do Espírito Santo, Alegre, ES, Brasil.

² Programa de Pós-graduação em Engenharia Química, Universidade Federal do Espírito Santo, Alegre, ES, Brasil.

* e-mail: uillapimenteleng@gmail.com

Resumo

Desde que foram propostos como método de otimização, os algoritmos têm sido usados com sucesso para resolver problemas complexos nas mais diversas áreas. O presente artigo teve como finalidade testar e comparar diferentes métodos de otimização não linear, onde foi desenvolvido um código computacional que integra o algoritmo baseado no método heurístico Evolução Diferenciada (ED) e o método Steepest Descent (SD). A fim de desenvolver a minimização da função objetivo (caixa preta) e identificar as características de cada método. Os resultados obtidos são consistentes e mostram que existe diferença entre os valores mínimos encontrados e o custo operacional necessário.

Abstract

Since they were proposed as optimization method, the algorithms have been successfully used to solve complex problems in several areas. This article was designed to test and compare different nonlinear optimization methods, which was developed a computer code that integrates the algorithm based on heuristic method Evolution Differentiated (ED) and the Steepest Descent method (SD). In order to develop the minimization of the objective function (black box) and identify the characteristics of each method. The results obtained are consistent and show that there is a difference between the minimum values found and the required operating cost.

Keywords (Palavras chaves): otimização não linear, evolução diferenciada, steepest descent.

1. Introdução

Entende-se por otimização a forma de descobrir mínimos ou máximos de determinada função com diversas variáveis presentes em uma determinada região multidimensional, que tem como objetivo tratar um conjunto de procedimentos capazes de definir as melhores configurações admissíveis para o funcionamento de sistemas de interesse para o indivíduo, isto é, aumentar a produtividade, diminuir gastos, elevar a concorrência, estes problemas, entre outros, após ser formulados matematicamente, apresentam precisamente a mesma estrutura, e sua solução advém principalmente por meio do uso do mesmo conjunto de técnicas, sendo esse característica do problema de otimização (TAKAHASHI, 2007).

Secchi (2015), aborda que os problemas presentes no contexto da otimização, são tratados pelos seguintes conceitos: função objetivo, que é quando almeja-se otimizar a função matemática de duas ou mais

variáveis, procurando definir mínimo ou máximo; variáveis de decisão, que são compreendidas como as variáveis independentes que surgem na função objetivo, em número podemos dizer que é o excesso de variáveis em relação ao número de equações, isto é, o grau de liberdade do sistema; restrições, sendo os limites atribuídos ao sistemas ou formados pelas leis naturais que gerenciam o comportamento do sistema, que sujeitam-se as variáveis de decisão, também compreendida como região viável, se define como região do espaço determinado pelas variáveis de decisão, demarcada pelas restrições, em cujo interior ou fronteira localizam o ótimo da função objetivo; ponto ótimo, se define como ponto desenvolvido pelas variáveis de projetos, onde extremizam a função objetivo e atendem as restrições; e valor ótimo, sendo esse o valor da função objetivo no ponto ótimo.

Nos métodos de identificação caixa-preta, não há nenhuma informação sobre o sistema, além dos dados, ou quando disponível não é utilizada no procedimento de obtenção do modelo. Nesta situação, serão usados

apenas dados de entrada e saída durante a identificação. A escolha da representação e de sua estrutura são realizadas por meio do método empírico nos casos mais simples, onde nos mais complexos, esta escolha se torna crítica o que explica o uso de método mais sofisticado para preferência da estrutura do modelo. Para estimativa de parâmetro usa-se procedimentos de otimização que não possuem restrições. Na identificação da caixa-preta não há qualquer relação óbvia entre a estrutura e seus parâmetros com feições do sistema sendo identificado (BISCAIA JUNIOR, 2014).

1.1 Classificação dos Métodos de Otimização não Linear

No que se refere aos métodos de programação matemática classificam-se como programação linear, programação não linear e métodos fundamentados em teoria de aproximações (SILVA, 2016).

Denomina-se um método de otimização de determinístico, caso seja possível prever todo o caminho, distinguindo seu ponto de partida. Contrário a este método, apresenta-se os métodos estocástico ou aleatórios, em que o caráter aleatório de diversos processos é simulado. Nestes métodos, diversas escolhas são realizadas com base de números aleatórios, que são sorteados na hora da execução do código. Como a cada execução do código os números que serão sorteados são distintos, um método aleatório não realizará a mesma sequência de operações em duas execuções consecutivas. Partindo do mesmo ponto inicial, cada execução do código adotará o seu caminho, e provavelmente apresentará a uma resposta final desigual (BISCAIA JUNIOR, 2014).

1.2. Evolução Diferenciada

Trata-se de um método de busca direta estocástica que surgiu de tentativas de resolver de ajuste polinomial de Chebychev, Kenedy Price introduziu a ideia de usar diferenças de vetores para perturbar a população de vetores (indivíduos) resultando em um método que requer poucas variáveis de controle, é de rápida convergência, fácil de usar e robusto (STORN 2006).

Os Algoritmos Genéticos, são conhecidos como métodos recentes que não utilizam qualquer informação de derivada e, assim, possuem chances consideráveis de não serem aprisionados em ótimos locais. Provas concretas de sua convergência não forma apresentadas ainda, todavia, sua aplicação em problemas práticos normalmente leva para o ótimo global ou ao menos para soluções mais aceitáveis do que as obtidas por outros métodos. Tal método baseia-se no processo de seleção natural e evolução da espécie, especialmente no paradigma de sobrevivência dos mais aptos. Os seres que demonstrarem melhor adequação, ou seja, menor valor da função objetivo, apresentam melhores possibilidades de se reproduzirem e transferirem seus genes para gerações futuras. Todavia, todos os seres estão passíveis a mutações aleatórias em todos os tempos. Em sua

implementação, necessitam a definições de diversos parâmetros que afetam o seu desempenho do algoritmo de diversos meios. O tamanho da população deve ser satisfatoriamente para assegurar diversidade suficiente e dessa forma cobrir bem o espaço que se refere a soluções. Diferentes parâmetros como a probabilidade de cruzamento, mecanismo e taxa de mutações, comprometem o método de forma menos significativa. Não existe valores de parâmetros reconhecidamente ótimos, somente faixas sugeridas de trabalho. Ainda que seja muito mais veloz que métodos tipo busca exaustiva, são métodos bastante lentos se checados com métodos do tipo gradiente, uma vez que não aplicam qualquer informação no que se refere à derivada da função objetivo (BISCAIA JUNIOR, 2014).

1.3. Steepest Descent

O método "Gradient Descent" ou "Steepest Descent", Abud Filho (2010), cita que é conhecido como um dos mais antigos, foi proposto em 1847 por Cauchy e é o mais simples na família dos métodos de "otimização por gradiente" são os que empregam a derivada de uma função para encontrar seu ótimo. Este método é determinístico, ou seja, sempre chegará à mesma resposta se saírem do mesmo ponto inicial. É considerado um algoritmo de otimização de primeira ordem, no qual a alternativa da direção do mínimo da função 'f' é a direção adversa ao gradiente.

Com o decorrer do processo iterativo, o valor do passo de procura pode ser alterado (aumentado) para admitir uma convergência mais veloz, garantido que a função converge a cada iteração; caso adverso, o valor do mesmo deve ser diminuído para que o processo possa estabilizar (ABUD FILHO, 2010).

Dessa forma, destaca que o método steepest descent possui a vantagem de não necessitar calcular segundas derivadas, e nem resolver sistemas de equações, entretanto, a sua convergência se faz muito lenta, tendo em vista que a progressão no sentido ao ótimo ser efetuada em diversas mudanças de direção (AZEVEDO, 1995).

2. Objetivo

Neste trabalho é apresentado um estudo comparativo de desempenho dos modelos implementados pelos métodos de otimização: Steepest Ascent/Descent e Evolução diferenciada, afim de minimizar a função objetivo (caixa preta), ou seja, encontrar o ponto ótimo.

3. Metodologia

Para otimizar a caixa preta utilizou-se dois métodos: Steepest Descent e Evolução diferenciada. Os códigos computacionais foram Implementados no software SCILAB.

Utilizou-se como entrada duas variáveis utilizando um intervalo de $[-600 \ 600]$ para cada variável. O procedimento para a realização do método Steepest Descent apresenta-se na figura 1.

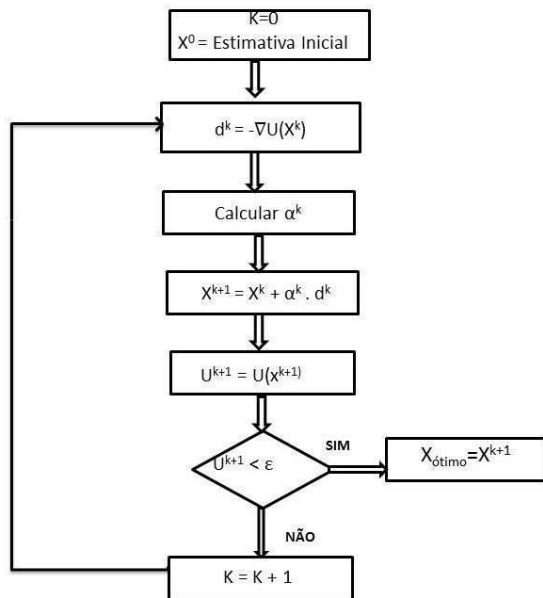


Figura 1: Fluxograma de descrição para o método Steepest Descent (minimização).

Em que:

$$\nabla = \frac{\partial U}{\partial x_1} \mathbf{i} + \frac{\partial U}{\partial x_2} \mathbf{j}, \text{ Gradiente de } U; \quad (1)$$

U= Função Objetiva (caixa preta); α - passo de procura; x - vetor de variáveis; d - direção de procura; k - contador de iterações; ϵ - erro.

Os procedimentos para o Evolução diferenciada esta apresentada na figura 2.

Pelo fluxograma observa-se que, primeiramente para a utilização deste método deve-se gerar uma população inicial de variáveis (n_p), em que normalmente escolhe-se n_p 10 vezes maior que o número de variável. Dessa forma, utilizou-se $n_p=20$, pois 2 é o número de variáveis do problema.

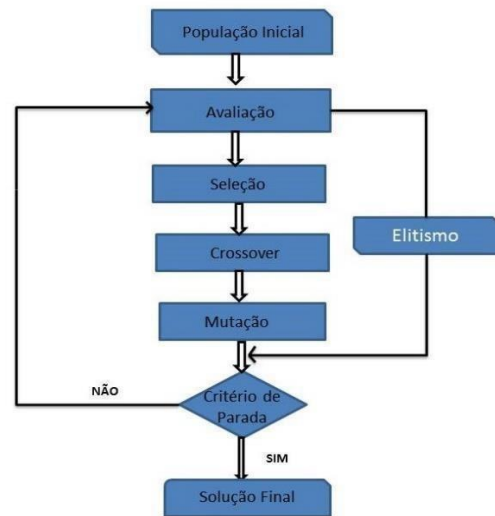


Figura 2: Fluxograma de descrição do método Evolução Diferenciada.

Em sequência gera-se uma região limitada de interesse de onde será obtido o valor ótimo, sendo esta região denominada população que é ilustrada na figura 3.

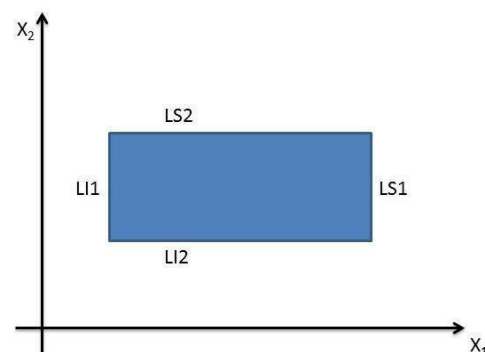


Figura 3: Demonstração da Região limitada.

Em que LI1 e LS1 representam respectivamente, o limite inferior e superior, para X_1 . E, LI2 e LS2 representam respectivamente, o limite inferior e superior, para $X_2=Y$.

A partir desta região buscou-se escolher aleatoriamente os membros da população:

1. Escolheu-se uma mãe de acordo com a equação (2) e determinou-se a mutação.

$$\text{Mãe} = \alpha + F(\beta - \gamma) \quad (2)$$

Em que α , β , γ -Representam os membros aleatórios da População.

F- Fator que define a mutação. Utilizou-se $F=0.8$.

2. Escolheu-se o pai.

3. Realizou-se o cruzamento entre a mãe escolhida com os pais até obter o n_p de filhos gerando a

população de uma geração. No nascimento tem-se o crossover que é um número randômico entre 0,5 e 1. Usou-se CR=0.9 que participa do cruzamento, que juntamente com a mutação garante variabilidade dos filhos.

4. Assim, para cada filho obtido por cada pai foi realizado uma comparação, se o filho for mais forte, ele assume o lugar do pai na próxima geração. Caso seja mais fraco, ele morrerá e o pai continua na próxima geração.

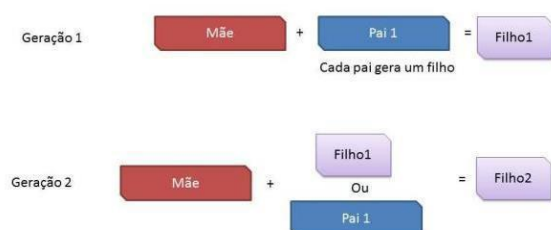


Figura 4: Representação dos itens 3 e 4.

Definiu-se um total de 200 gerações. Cada cruzamentos realizados tem um filho ótimo que pode substituir o pai ou não, isso ocorre para cada geração

4. Resultados e Discussões

Os resultados obtidos pelo Método Steepest Descent estão apresentados na tabela 1 e pelas figuras 5 e 6.

Tabela 1 - Resultados do Método Steepest Descent

Condições Iniciais	Tempo	Ponto Ótimo	Valor Ótimo
[-10 10]	0.877912 s	[-21.9802 13.3153]	0.1652
[-200 100]	0.719926 s	[-197.8210 102.0836]	12.3961
[-600 600]	0.822016 s	[-599.7018 599.1295]	179.8083

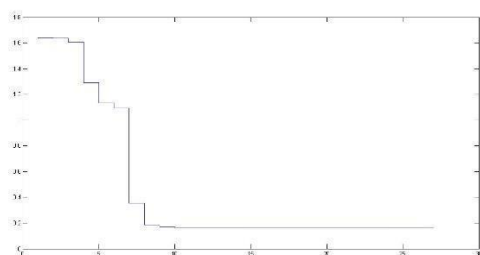


Figura 5: Comportamento da função objetiva com a execução do algoritmo Steepest Descent com a condição inicial $[x1 \ y1] = [-10 \ 10]$.

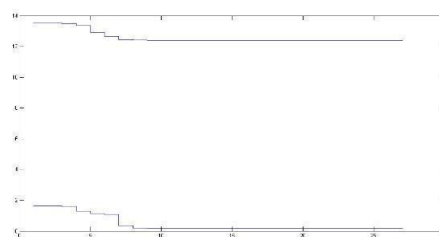


Figura 6 – Comportamento da função objetiva com a execução do algoritmo Steepest Descent com a condições iniciais: $[x1 \ y1] = [-10 \ 10]$ (Curva inferior) e $[x1 \ y1] = [-200 \ 100]$ (Curva superior)

A figura 5 demonstra o comportamento de uma função objetiva desconhecida (caixa preta) de acordo com o algoritmo implementado, sendo uma função decrescente que converge a um valor ótimo (menor valor). O mesmo é mostrado na figura 6 em que se realiza uma comparação entre duas condições iniciais diferentes. Em que se percebe que para cada condição tem-se um comportamento diferente, isso mostra que o método Steepest Descent não fornece valores absolutos apenas máximos e mínimos locais.

Os valores dos pontos ótimo e valor ótimo obtido estão presentes na tabela 01, sendo o menor valor obtido para a função foi de 0.1652 correspondendo a condição inicial $[x1 \ y1] = [-10 \ 10]$. Tem-se que o tempo de processamento foi praticamente igual para as três condições analisadas.

Os resultados obtidos utilizando o método Evolução Diferenciada encontra-se na figura 7, abaixo.

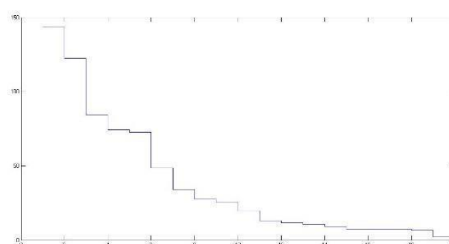


Figura 7 – Comportamento da função objetiva com a execução de Evolução diferenciada.

Na figura 7 tem-se o comportamento de uma função desconhecida (caixa preta) de acordo como o algoritmo implementado, sendo uma função decrescente que converge a um valor ótimo (menor valor= -143.7360) na posição $[64.5440 \ 318.3860]$ (ponto de valor mínimo) e tempo computacional pelo ponto 3.394382 s.

Não foi realizado comparações entre diferentes condições iniciais para esse método uma vez que ele tende a procurar o mínimo global da função otimizada.

Ao se comparar os tempos operacionais gastos para se obter o valor mínimo entre os dois métodos de otimização utilizados, tem-se que o tempo foi maior para o método de Evolução diferenciada, uma vez que este método é mais robusto.

Cabe ressaltar ainda que a implementação dos dois métodos ocorreu de forma gradual e conforme estudos e pesquisas bibliográficas foram sendo realizados, e para a confecção do algoritmo foi necessário conhecimento da linguagem do programa computacional SCILAB, sendo que muitas funções dos mesmos foram sendo aperfeiçoadas e substituídas por outras funções para garantir o melhor método de procura para os valores mínimos.

[5] STORN.; PRICE K. Home page of differential Evolution (DE) for Continuous Function Optimization. Disponível em: <http://www.icsi.berkeley.edu/~storn/code.html#matl>

[6] ABUD FILHO, E. Aplicação de Métodos Variacionais e Formulações Heurísticas para Análise e Síntese Numérica de Transformadores em Guia de Onda Retangulares. Disponível em: http://www.maxwell.vrac.puc-rio.br/16456/16456_9.PDF

[7] AZEVEDO, A. F. M. Otimização De Estruturas. Disponível em: <http://www.alvaroazevedo.com/publications/reports/1995/OE.pdf>

5. Conclusão

Com os dados e comportamentos obtidos por este trabalho foi possível perceber que o método de Evolução Diferenciada tende a buscar mínimos globais e não são afetados pelas condições iniciais, diferente do método Steepest Descent que apenas fornece máximos e mínimos locais. Além disso identificou-se que o gasto operacional do método de Evolução Diferenciada é maior que o método Steepest Descent. Outro aspecto observado, percebe que para a implementação do algoritmo se faz necessário conhecimento da linguagem do software utilizado e do método a ser otimizado para que se possa realizar passos e obter o mínimo com o menor custo operacional.

4. Referências

[1] TAKAHASHI, R. H. C. Otimização Escalar e Vetorial, Volume 3: Otimização Vetorial. Notas de aula. Departamento de Matemática. Universidade Federal de Minas Gerais, 2007.

[2] SECCHI, A. R. COQ-897 – Otimização de processos. Apostila do Curso de Engenharia Química da Universidade Federal do Rio de Janeiro – UFRJ. Rio de Janeiro, 2015.

[3] BISCAIA JUNIOR, E. C. Métodos não Determinísticos. Disponível em: http://www2.peq.coppe.ufrj.br/Pessoal/Professores/Evaristo/CO897_2014/M%E9todos%20N%E3o%20Determin%C3%94sticos/aula1.pdf

[4] SILVA, Emílio Carlos Nelli. Métodos Numéricos Aplicados em Problemas de Otimização de Engenharia.44-67. Disponível em: <http://sites.poli.usp.br/d/pmr5215/a2-5215.pdf>