

## O Recozimento Simulado como Ferramenta de Otimização Global

*Vanderlei de Campos Bueno\**  
*Francisco José da Cunha Pires Soeiro\*\**

*Existe uma útil e profunda relação entre a mecânica estatística (características de sistemas com muitos graus de liberdade em equilíbrio térmico a uma dada temperatura) e a otimização global (processo de encontrar o ponto mínimo de uma função dada, sendo a mesma dependente de muitos parâmetros). Uma analogia detalhada com o recozimento em sólidos resulta em uma ferramenta para otimização das propriedades de sistemas muito grandes e complexos que compõe essencialmente um algoritmo de otimização. Esse algoritmo resultante denominado Recozimento Simulado (Simulated Annealing), é um método de busca randômica sendo dessa forma classificado como método estocástico, que apresenta como critério de convergência princípios termodinâmicos.*

### 1 – INTRODUÇÃO

**R**ecozimento Simulado (RS) é uma técnica simples que se baseia em idéias da mecânica estatística (Janes 1951) e no algoritmo de simulação proposto por Metropolis (1953). Essa técnica

de otimização foi apresentada inicialmente por Kirkpatrick (1983), que utilizou o método no projeto de circuitos eletrônicos. Além de Kirkpatrick muitos outros pesquisadores implementaram o RS com bons resultados nas mais diversas áreas de aplicações possíveis: projetos de redes neurais, reconstrução de estruturas policristalinas, processamento de imagens e outros. Neste artigo serão revistos brevemente conceitos de otimização global e da mecânica estatística apontando as similaridades entre os dois campos. E também como o

\*Graduação em Engenharia Mecânica pela Escola de Engenharia de Piracicaba EEP (1996). Atualmente aluno de Mestrado de Engenharia Mecânica do Instituto Militar de Engenharia – IME.

\*\*Curso de Material Bélico da AMAN (1971). Graduação em Engenharia Mecânica e de Automóveis do IME (1979). Mestrado em Engenharia Mecânica do IME (1986). Doutorado em Engenharia Mecânica da Universidade da Flórida. (EUA). Atualmente é professor do Departamento de Engenharia Mecânica da UERJ.

algoritmo de Metropolis, que foi desenvolvido para simulação numérica aproximada das características de muitos sistemas de corpos a uma dada temperatura, resulta em uma ferramenta natural para otimização global.

## 2 – OTIMIZAÇÃO GLOBAL

A otimização global consiste na técnica de encontrar máximos ou mínimos valores de uma função com muitas variáveis independentes. Essa função normalmente denominada como função custo ou objetivo, pode apresentar vários mínimos locais. A função que é apresentada ao final deste artigo, por exemplo, é a chamada função “Camelback” e possui 15 mínimos locais. Um mínimo global, a não ser que a função seja convexa, não tem nenhuma condição matemática a caracterizá-lo, diferentemente dos mínimos locais que são caracterizados pelo comportamento local das funções (gradientes e matrizes Hessianas). Portanto, o problema de se encontrar um mínimo global permanece como um desafio do ponto de vista matemático e computacional.

Métodos de otimização global podem ser divididos basicamente em duas categorias: determinísticos e estocásticos. Métodos determinísticos são basicamente métodos heurísticos que podem ou não conduzir ao ponto de ótimo global. Já os métodos estocásticos são métodos randômicos que apresentam ou não um critério de convergência baseado em algum acontecimento físico, como por exemplo o RS que apresenta como critério de convergência princípios termodinâmicos.

## 3 – MECÂNICA ESTATÍSTICA

Mecânica estatística é a disciplina central da física da matéria condensada, um conjunto de métodos para análise de propriedades agregadas do grande número de átomos encontrados em amostras de matérias líquidas ou sólidas. Devido ao número de átomos ser da ordem de  $10^{23}$  por centímetro cúbico, apenas as mais prováveis características do sistema em equilíbrio térmico em uma dada temperatura são observadas em experimentos. O sistema pode ser caracterizado por suas propriedades médias e pequenas flutuações sobre as mesmas. Essas propriedades são levantadas considerando um universo de sistemas idênticos. Cada configuração é definida pelo campo de posições atômicas  $\{r_i\}$ , ponderado pelo fator de probabilidade de Boltzmann,  $\exp(-E(\{r_i\})/k_B T)$ , onde  $E(\{r_i\})$  é a energia da configuração,  $k_B$  é a constante de Boltzmann, e  $T$  é a temperatura.

Uma questão fundamental na mecânica estatística é determinar o que acontece ao sistema no limite inferior de temperatura. Por exemplo, se os materiais permanecem fluidos ou sólidos, e no caso de se solidificarem, se eles formam um sólido cristalino ou não. Isto está relacionado com nível interno de energia do material. No entanto, baixa temperatura não é uma condição suficiente para encontrar estados baixos de energia na matéria. O estado de temperatura inferior de um material pode ser determinado, por exemplo, obtendo-se um cris-

tal simples de um material fundido através de um cuidadoso recozimento, ou seja, após o material fundido resfriar-se lentamente até a proximidade, se não o próprio ponto de temperatura inferior. Caso isso não ocorra o cristal obtido pode apresentar muitos defeitos, ou seja, pode-se encontrar um estado de energia inferior “local” e não o estado de menor energia do sistema.

Pode-se dizer que de uma maneira análoga, utiliza-se a filosofia desse processo na otimização global, onde o valor da função objetivo será a energia do sistema que se deseja otimizar. Entretanto o conceito da temperatura de um sistema físico não possui nenhum equivalente claro na otimização, mas sua introdução na otimização é de suma importância, visto que a mesma será o parâmetro que controla o processo de otimização utilizando o RS.

Melhoramentos sucessivos, ou seja, a aceitação apenas dos projetos que diminuem a energia em problemas de otimização, são muito comuns como no processo de rearranjo microscópico modelado pela mecânica estatística. Entretanto, aceitando apenas os rearranjos que diminuem a Energia (função objetivo) do sistema, o processo converge rapidamente de altas temperaturas para  $T = 0$ , podendo assim, encontrar estados cristalinos com defeitos (mínimos locais em otimização).

O procedimento de Metropolis da mecânica estatística fornece uma generalização do melhoramento sucessivo em que soluções intermediárias piores também são aceitas, ou seja, os rearranjos que aumentam a energia também podem ser aceitos.

Metropolis nos primeiros passos da computação científica introduziu um algoritmo simples que pode ser usado para realizar uma simulação eficiente de um sistema de átomos em equilíbrio a uma dada temperatura. Em cada passo desse algoritmo, a um átomo é dado um pequeno deslocamento randômico e a variação da energia  $\Delta E$  ( $\Delta E = E_i - E_{i+1}$ ) é calculada. Se  $\Delta E \leq 0$ , o deslocamento é aceito, e a configuração com o átomo deslocado é usado como novo ponto de partida para o próximo deslocamento. Caso porém,  $\Delta E > 0$ , então trata-se probabilisticamente o sistema, ou seja, a probabilidade dessa configuração ser aceita é dada por:  $P(\Delta E) = \exp(-\Delta E / k_B T)$ . Um número randômico uniformemente distribuído no intervalo  $(0,1)$  é gerado, dando então um significado randômico conveniente ao algoritmo. Esse número gerado randomicamente é então comparado com  $P(\Delta E)$ . Se o número randômico for maior que  $P(\Delta E)$ , essa configuração é aceita, caso contrário, essa configuração é descartada e a configuração anterior é novamente utilizada como ponto de partida. Repetindo-se os passos básicos muitas vezes, temos então a simulação do movimento térmico dos átomos em um caminho do resfriamento da temperatura  $T$ . A escolha dessa probabilidade se baseia na distribuição de Boltzmann.

Usando o valor da função objetivo no lugar da energia e substituindo a configuração pelo campo de variáveis de projeto do sistema  $\{x_i\}$ , pode-se então adaptar o critério de Metropolis para o problema de otimização global. A temperatura é simplesmente um parâmetro de controle, que deve estar na mesma ordem de grandeza da função objetivo. O processo do RS consiste de se partir temperatura inicial alta. Então lentamente a temperatura, passando por estágios intermediários, diminui até que não haja mais mudanças significativas no sistema. A seqüência de temperaturas e o número de rearranjos de  $\{x_i\}$  fornecidos para alcançar o equilíbrio em cada temperatura pode ser considerado um processo de recozimento.

Recozimento como implementado por Metropolis, difere dos processos de melhoramentos iterativos como já visto anteriormente, pois ele pode aceitar estados piores, e isso faz com que ele consiga superar os pontos de ótimos locais chegando no ponto de ótimo global. É importante frisar também que no início do processo ele aceita pontos piores com mais facilidade que no fim do processo. Isso faz com que em altas temperaturas apenas as características “grossas” sejam levadas em consideração, porém já em baixas temperaturas ele leva em conta detalhes “finos” do sistema, tornando o algoritmo RS altamente robusto.

#### 4 – ALGORITMO

O fluxograma do algoritmo RS implementado por Goffe (1992), utilizado nesse artigo, foi elaborado por Corana (1987) e como pode ser verificado na figura a seguir, foi escolhido por reunir dentre vários artigos sobre RS as características de maior robustez e simplicidade de implementação:

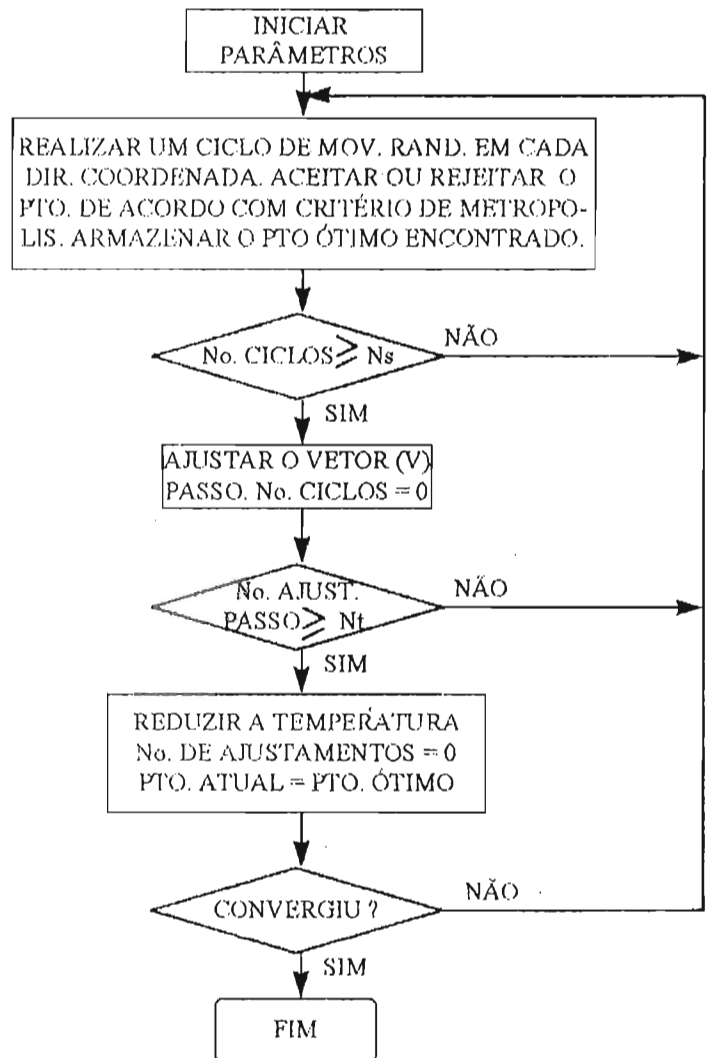


FIGURA 01. Algoritmo de otimização RS

**Inicialização**

Escolher

- Um ponto de partida  $x_0$ .
- Um passo de partida do vetor  $v_0$ .
- Uma temperatura de partida  $T_0$ .
- Um critério de converg.  $\hat{N}$  e um número de redução de temperatura para testar o término no  $\hat{N}$ .
- Um teste para redução de temperatura  $N_1$  e um coeficiente de redução  $r_1$ .

Fazer  $i, j, m, k$  igual a 0 (zero).  $i$  é o índice que denota pontos sucessivos,  $j$  indica ciclos sucessivos ao longo de todas direções,  $m$  descreve sucessivos ajustamentos do passo, e  $k$  indica reduções de temperaturas sucessivas.

Fazer  $h = 1$ ,  $h$  é o índice que indica a direção ao longo do qual um novo ponto é gerado, partindo sempre do último ponto aceito.

Calcular  $f_0 = f(x_0)$ .

Fazer  $x_{opt} = x_0, f_{opt} = f_0$ .

Fazer  $n_u = 0, u = 1, \dots, n$ .

Fazer  $F_u^* = f_0, u = 0, -1, \dots, -\hat{N} + 1$ .

**Primeiro passo:**

Partindo de um ponto  $x_j$ , gerar um ponto randômico  $x'$  na direção  $h$ :

$$x' = x_j + r v_{mh} e_h$$

onde  $r$  é um número randômico gerado no campo de  $[-1,1]$  por um gerador de número pseudorandômico;  $e_h$  é o vetor da  $h$ (th) direção coordenada; e  $v_{mh}$  é a componente do vetor passo  $v_m$  na mesma direção.

**Segundo passo:**

Se a  $h$ (th) coordenada de  $x'$  estiver fora do domínio definido de  $f$ , ou seja, se  $x'_h < a_h$  ou  $x'_h > b_h$ , então retorne ao primeiro passo.

**Terceiro passo:**

Calcular  $f' = f(x')$ .

Se  $f' \leq f_j$ , então aceitar o novo ponto.

Fazer  $x_{i+1} = x'$ ,

Fazer  $f_{i+1} = f'$ ,

Adicionar 1 a variável  $i$ ,

Adicionar 1 ao  $n_h$ ;

Se  $f' < f_{opt}$ , então fazer

$$x_{\text{opt}} = x',$$

$$f_{\text{opt}} = f'.$$

Senão ( $f' > f_i$ ) aceitar ou rejeitar o ponto, de acordo com a probabilidade (critério de Metropolis):

$$P(\Delta E) = \exp\left(\frac{f_i - f'}{T_k}\right)$$

Na prática, um número pseudorandômico  $P'$  é gerado no campo  $[0,1]$  e é comparado com  $P(\Delta E)$ . Se  $P' < P(\Delta E)$ , o ponto é aceito, caso contrário é rejeitado.

Caso o ponto seja aceito:

$$\text{Fazer } x_{i+1} = x',$$

$$\text{Fazer } f_{i+1} = f',$$

Adicionar 1 ao  $i$ ,

Adicionar 1 ao  $n_i$ .

#### Quarto passo:

Adicionar 1 ao  $h$ .

Se  $h \leq n$ , então vá para o primeiro passo;

Senão fazer  $h = 1$  e adicionar 1 ao  $j$ .

#### Quinto passo:

Se  $j < N_s$ , então vá para o primeiro passo:

Senão adaptar o vetor passo  $v_m$ :

Para cada direção  $u$  uma nova componente do vetor passo  $v_u'$  é

$$v_u' = v_{mu} = \left(1 + c_u \frac{n_u/N_s - 0.6}{0.4}\right) \text{ se } n_u > 0.6N_s,$$

$$v_u' = \frac{v_{mu}}{0.4 - n_u/N_s} \text{ se } n_u < 0.4N_s,$$

$$1 + c_u \frac{\quad}{0.4}$$

$$v_u' = v_{mu} \quad \text{caso contrário}$$

$$\text{Fazer } V_{m+1} = v',$$

Fazer  $j = 0$ ,  
 Fazer  $n_u = 0$ ,  $u = 1, \dots, n$ ,  
 Adicionar 1 ao  $m$ .

O propósito dessa variação no comprimento dos passos é manter a porcentagem média de movimentos aceitos em torno da metade do número total de movimentos. O parâmetro  $c_u$  controla o passo de variação ao longo de cada direção  $u$ (th).

#### Sexto passo:

Se  $m < N_t$ , então vá para o primeiro passo.  
 Senão, reduzir a temperatura  $T_k$ :  
 Fazer  $T_{k+1} = r_1 \cdot T_k$ ,  
 Fazer  $f_k^* = f_i$ ,  
 Adicionar 1 ao  $k$ ,  
 Fazer  $m = 0$ .

Vale ressaltar que a redução de temperatura ocorre em todo  $N_s \cdot N_t$  ciclos de movimentos ao longo de toda direção e depois de  $N_t$  ajustamentos de passo.

#### Sétimo passo:

Se:  
 $|f_k^* - f_{k-u}^*| \leq \epsilon$ ,  $u = 1, \dots, N_s$   
 $f_k^* - f_{opt} \leq \epsilon$   
 Então parar a busca;

Senão:

Adicionar 1 a  $i$ ;  
 Fazer  $x_i = x_{opt}$ ,  
 Fazer  $f_i = f_{opt}$ .  
 Vá para o primeiro passo.

Valores razoáveis, encontrados para os parâmetros de controle do recozimento simulado são:

$N_s = 20$   
 $N_t = 5$   
 $c_i = 2, i = 1, \dots, n$ .  
 $N_e = 4$   
 $r_1 = [0.5, 0.75]$

Esses valores foram obtidos empiricamente e propiciam ao algoritmo a robustez necessária para a otimização de diversos problemas. Porém é importante deixar claro que caberá ao usuário da técnica estabelecer valores apropriados para essas variáveis caso o algoritmo

não encontre bons resultados. Com relação à temperatura inicial, ela deve estar na ordem da grandeza da função objetivo. Por isso, para determinação da temperatura, deve-se inicialmente colocar um valor para T relativamente alto e um  $r_i$  bem baixo ( $r_i = 0.1$ ). Assim quando o método RS começar a determinar valores para variáveis de projeto, poderá ser monitorado o comportamento da função objetivo e ser ajustado a temperatura na mesma unidade da função objetivo. Feito isso ajusta-se a taxa de resfriamento ( $r_i$ ), lembrando que um resfriamento muito lento resultará em uma otimização muito lenta com um conseqüente aumento do número de avaliações da função objetivo, e um resfriamento muito rápido pode não conduzir à solução global.

**5 – EXEMPLO**

Como exemplo, utilizou-se a função abaixo:

$$f(x,y) = 4x^2 - 2.1x^4 + (x^6/3) + xy + 4y^4 - 4y^2$$

que é a chamada função “Camelback” e possui quinze mínimos, dos quais dois são globais. Os valores e o gráfico em curvas de nível (“Contour plot”) são apresentados abaixo. Utilizando-se o algoritmo descrito acima chegou-se ao resultado  $x = 0.0898$  ,  $y = 0.71269$  e  $f(x,y) = -1,0316$

num.	x	y	f(x,y)
1	0.0	0.0	0.0
2	0.0898	-0.7127	-1.0316 <sup>G</sup>
3	-1.1092	0.76827	0.5437
4	1.2302	0.1623	2.4963
5	1.2961	0.6051	2.2295
6	1.67071	0.5687	2.1043
7	1.6381	0.2287	2.2294
8	1.7036	-0.7961	-0.2155
9	-0.0898	0.7127	-1.0316 <sup>G</sup>
10	-1.1092	0.76827	0.5437
11	-1.2302	-0.1623	2.4963
12	-1.2961	-0.6051	2.2295
13	-1.67071	-0.5687	2.1043
14	-1.6381	-0.2287	2.2294
15	-1.7036	0.7961	-0.2155

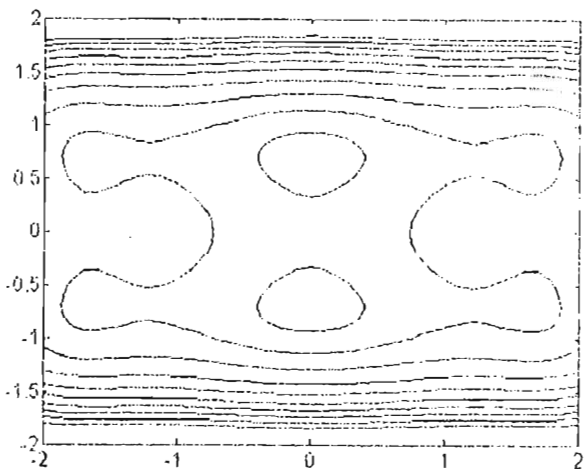


FIGURA 02. “Contour plot” da função.

<sup>G</sup> Ponto de mínimo global.



## 5 – CONCLUSÕES

Foi apresentado nesse trabalho um método de otimização global que é baseado no processo de recozimento, denominado Recozimento Simulado (Simulated Annealing). Esse método requer muitas avaliações da função objetivo requerendo um certo esforço computacional para otimização de funções, que aumenta linearmente de acordo com o número de dimensões do problema.

Trata-se de um método de otimização global de fácil implementação e muito robusto na solução dos mais diversos problemas, tornando-se extremamente interessante quando o número de mínimos for elevado.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Atiquilah, Mir M. e Rao, S.S.; 1995: Parallel Processing in Optimal Structural Design Using Simulated Annealing. *AIAA Journal* 33, núm. 12, 2386-2392.
- Corana, A.; Marchesi, M.; Martini, C. e Ridella, S.; 1987: Minimizing Multimodal Functions of Continuous Variables With the “Simulated Annealing Algorithm”. *ACM Transactions on Mathematical Software* 13, 262-280.
- Goffe, W. L.; Ferrier, G. D.; Rogers, J.; 1992: Global Optimization of Statistical Functions With Simulated Annealing. *Journal of Econometrics* 60, 65-99.
- Jaynes, E. T.; 1957: Information theory and statistical mechanics. *Phys. Rev.* 106, 620-630.
- Kirkpatrick, S.; Gelatt, Jr., C. D.; Vecchi, M. P.; 1983: Optimization by Simulated Annealing. *Science* 220, núm. 4598, 671-680.
- Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; 1953: Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* 21, 1087-1092.