

Acoplamento de algoritmos de otimização e agrupamento para resolução de problemas globalmente multimodais

Deise Mara B. de Almeida, Gustavo M. Platt, Francisco D. Moura Neto,

UERJ - Instituto Politécnico

28630-050, Nova Friburgo, RJ

E-mail: deisemalmeida@gmail.com, gmplatt@iprj.uerj.br, fdmoura@iprj.uerj.br.

Resumo: *Um problema de otimização globalmente multimodal é aquele que possui mais de um ponto de ótimo global. Algoritmos determinísticos e evolucionários são métodos comumente utilizados para resolução de problemas de otimização, mas enfrentam dificuldades para obtenção das soluções de um problema globalmente multimodal. Os métodos de agrupamento utilizam relações de similaridade entre os dados para encontrar padrões dentro do conjunto e realizar o agrupamento dos dados. O objetivo deste trabalho é empregar conjuntamente algoritmos de otimização e técnicas de agrupamento para obter simultaneamente todas as soluções de um problema globalmente multimodal. Para tal, utiliza-se inicialmente o algoritmo evolucionário, denominado Busca Harmônica, para identificar pequenos grupos em torno de pontos de ótimo global do problema. Em seguida, usa-se o algoritmo de agrupamento espectral para mapear os grupos; então, os pontos são classificados conforme o grupo ao qual pertencem e retorna-se uma estimativa de cada uma das soluções, utilizando o método k-means. Por fim, com o auxílio de um método determinístico, calcula-se cada uma das soluções. A metodologia proposta foi aplicada a duas funções teste, ambas bidimensionais. Nota-se que para as funções utilizadas a metodologia obteve êxito, ao encontrar todas as soluções globais e com pequeno erro relativo.*

Palavras-chave: *Otimização, Agrupamento, Problemas multimodais*

1 Introdução

Dado um problema de otimização este pode ser caracterizado pela existência de duas ou mais soluções ótimas. Neste caso diz-se que o problema é multimodal. Ainda é possível que existam mais de um solução ótima global, então tem-se um problema globalmente multimodal. Os algoritmos determinísticos foram os primeiros métodos utilizados para resolução de problemas de otimização. Estes métodos dependem da qualidade da estimativa inicial, que influencia fortemente no resultado e pode levar à identificação de um ponto de ótimo local. Os algoritmos evolucionários são alternativas para escapar de pontos de ótimo local, pois frequentemente encontram soluções globais. Por outro lado, no decorrer do processo evolutivo destes métodos, a população (no caso de algoritmos populacionais) perde sua diversidade deparando-se com dificuldades para resolver problemas multimodais, pois não encontra todos os pontos de ótimo global.

Um conjunto de dados pode ser subdividido conforme suas características. Os métodos de agrupamento utilizam relações de similaridade entre os dados para encontrar padrões dentro de um conjunto e realizar o agrupamento dos dados. Como determinados algoritmos evolucionários – antes do fim do processo evolutivo – são capazes de formar pequenos grupos em torno de cada ponto de ótimo global do problema, aliar métodos de otimização a métodos de agrupamento mostra-se uma maneira de resolver problemas globalmente multimodais.

O objetivo deste trabalho é unir algoritmos de otimização e algoritmos de agrupamento para obter simultaneamente todos os pontos de ótimo global de um dado problema de otimização.

Para tal, utiliza-se o algoritmo evolucionário Busca Harmônica [4] para a geração de pequenos grupos em torno de pontos de ótimo global. Após a obtenção dos pontos utiliza-se o agrupamento espectral para fazer o mapeamento dos grupos. Em seguida, usa-se o algoritmo k-means para classificar os pontos conforme o grupo ao qual pertencem e, então, obter uma estimativa de cada um dos pontos de ótimo global. Por fim, executa-se o algoritmo determinístico de Hooke-Jeeves [3] com cada uma das estimativas para obter todos os pontos ótimo global.

2 Métodos de otimização

Um método é considerado determinístico caso seja possível prever todas as etapas de seu desenvolvimento a partir do conhecimento de seu ponto de partida. Um método determinístico sempre chega a mesma solução quando é fornecida a mesma estimativa inicial. O algoritmo de Hooke-Jeeves é um método determinístico de busca direta. Conforme os próprios autores desta metodologia [3], o método de Hooke-Jeeves baseia-se em uma sequência de movimentos exploratórios que se inicia a partir de um ponto base e um incremento pré-definido. Dado um ponto base, adiciona-se a uma coordenada deste ponto um incremento, caso o ponto com a coordenada incrementada tenha menor valor de função este é aceito como ponto teste. Caso contrário, subtrai-se o incremento da coordenada do ponto base, se o novo ponto obtiver menor valor na função torna-se o ponto teste. Este procedimento é repetido para todas as coordenadas do ponto base. Se ao final do processo o ponto teste apresentar menor valor na função objetivo torna-se o ponto base, caso contrário, reduz-se o incremento e recomeça-se o processo de busca por um novo ponto teste. O procedimento é interrompido ao atingir-se uma tolerância estabelecida *a priori*.

Os métodos estocásticos são técnicas de otimização baseadas em processos aleatórios e fundamentados por algumas heurísticas. Algoritmos evolucionários são métodos baseados muitas vezes em processos de evolução natural. Estes métodos visam encontrar pontos de ótimo global em diferentes problemas (embora não haja garantias de localização de ótimos globais, uma vez que tais algoritmos dependem fundamentalmente de alguns parâmetros de controle). O algoritmo de Busca Harmônica é um método evolucionário que mimetiza o processo desenvolvido por um músico para chegar a uma música com notas harmônicas perfeitas. Conforme [4], durante o processo de improvisação, o músico utiliza-se de três opções: usa uma nota famosa de uma música de sua memória, toca algo parecido com algo já conhecido ou compõe novas notas, que podem ser aleatórias. Estas opções são denominadas no processo de otimização respectivamente como: uso da memória harmônica, ajuste do tom e aleatoriedade. O processo evolutivo inicia-se a partir da geração da memória harmônica inicial, então avança-se para as três etapas do processo. O uso da memória harmônica é feito conforme uma taxa de aceitação e o ajuste do tom é determinado por uma variável que controla o tamanho máximo da variação de tom. Caso as etapas anteriores não obtenham resultados satisfatórios, calcula-se aleatoriamente uma nova nota. A nova harmonia (ou seja, um novo ponto na matriz que contém a memória harmônica) é aceita caso seja melhor do que a pior existente na memória harmônica (a pior é então removida da memória harmônica); então, repete-se o processo até que a solução seja encontrada. A descrição detalhada do algoritmo – que foge ao escopo deste trabalho – encontra-se em [4].

3 Métodos de agrupamento

Os dados de um conjunto apresentam certas características que podem ser similares para alguns dados. Dessa forma, um conjunto de dados pode ser subdividido conforme as características semelhantes que os dados possuam. Os métodos de agrupamento utilizam relações de similaridade entre os dados para encontrar padrões dentro de um conjunto e realizar o agrupamento dos dados.

O *k-means* é um método de agrupamento antigo e amplamente utilizado [5]. Trata-se de

um algoritmo de agrupamento simples que objetiva encontrar k grupos não sobrepostos. Os grupos do *k-means* são representados por centroides e estes comumente são dados pela média dos pontos pertencentes ao grupo o qual ele representa. O processo de agrupamento inicia-se a partir da seleção dos k centroides, sendo k o número de grupos que se deseja obter. Em seguida, cada um dos pontos são comparados a cada um dos centroides e, então, os grupos são formados de acordo com a proximidade dos pontos com o centroide. O próximo passo é atualizar os centroides calculando-se a média dos pontos do grupo e repetir o processo de seleção dos pontos para os novos centroides. O processo chega ao fim quando não houver mais migração de pontos para outro grupo. A desvantagem do *k-means* é que a divisão dos grupos depende dos centroides escolhidos inicialmente, que pode fazer com que os grupos encontrados não representem adequadamente a similaridade entre os pontos.

O agrupamento espectral é conhecido por identificar agrupamentos naturais dentro de um conjunto de dados baseando-se na similaridade entre estes dados [2]. Para seu desenvolvimento é necessário a representação dos dados como grafos, então, para cada grafo pode-se associar uma matriz Laplaciana, ou operador Laplaciano. Este operador pode ser definido utilizando a subtração da matriz de grau D pela matriz de similaridade W , ambas relacionadas ao grafo. A matriz de similaridade é uma matriz quadrada cuja ordem corresponde ao número de dados e a matriz de graus é diagonal cujos elementos podem ser obtidos pela soma das linhas da matriz de similaridade. O Laplaciano possui ainda uma versão normalizada que pode ser definida da seguinte forma: $L_n = D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}$, que é capaz de obter agrupamentos mais precisos quando existe um ou mais dados sem semelhanças ou com similaridade relativamente grande ou muito pequena. O agrupamento dos dados é representado pelos autovetores do Laplaciano, ou do Laplaciano normalizado, a partir do autovetor associado ao segundo menor autovalor. O autovetores responsáveis pelo agrupamento são também conhecidos como vetores de Fiedler. Estes vetores fazem o mapeamento dos grupos, que podem ser visto através de um gráfico de dispersão. Além disso, este mapeamento pode ser observado em dimensões menores que a original.

4 Metodologia

Um problema de otimização que possui mais de um ponto de ótimo global é denominado globalmente multimodal. Tradicionais métodos de otimização frequentemente enfrentam dificuldades para encontrar todas as soluções globais. Os algoritmos determinísticos foram os primeiros métodos utilizados na resolução de problemas de otimização. Estes métodos são dependentes de estimativas iniciais que nem sempre são de boa qualidade e influenciam na solução, podendo levar a soluções locais. Além disso, para encontrar mais um ponto de ótimo global seria necessário um maior número de estimativas iniciais e que estas estejam próximas de cada um dos pontos de ótimo global do problema (o que nem sempre é uma tarefa fácil de ser conduzida). Já os algoritmos evolucionários – que representam uma alternativa aos algoritmos determinísticos pois frequentemente encontram soluções globais – também encontram dificuldades com problemas multimodais. Isto porque, no decorrer do processo evolutivo, estes métodos podem perder diversidade e não encontrar todos os pontos de ótimo global. Entretanto, durante o processo evolutivo parte destes métodos são capazes de formar pequenos grupos de pontos em torno de cada um dos pontos ótimos globais e como algoritmos de agrupamento são capazes de identificar grupos dentro de um conjunto de pontos, uma opção para resolução de problemas multimodais pode ser a junção de métodos de otimização e agrupamento.

A ideia é executar sequencialmente métodos de otimização e agrupamento para que ao final sejam encontrados pontos ótimos globais, possivelmente todos eles (embora, naturalmente, não haja garantias de localização de todos os ótimos globais). Assim, a metodologia empregada pode ser resumida nos seguintes passos:

1. Inicialmente, utilizou-se o algoritmo evolucionário Busca Harmônica [4] para gerar gru-

pos de pontos nas vizinhanças dos ótimos globais do problema. O algoritmo de Busca Harmônica é então interrompido, após um pequeno número de iterações;

2. Após a obtenção dos grupos de pontos gerados pelo algoritmo de Busca Harmônica [4], utiliza-se a técnica de agrupamento espectral para mapear estes grupos e exibir o resultado em dimensão reduzida, caso seja possível. Para classificar os grupos utiliza-se o método *k-means* e um ponto de cada grupo é escolhido para servir como estimativa inicial na próxima etapa do processo, que é a execução do método determinístico de Hooke-Jeeves [3];
3. O último passo da metodologia é o refinamento das soluções (com vistas a aumentar a acurácia dos resultados) através de um método determinístico. Conforme já ressaltado, o algoritmo de Hooke-Jeeves [3] foi empregado para esta finalidade. Este método é executado para todas as estimativas retornando um ponto de ótimo global para cada estimativa fornecida. Assim, após identificar os grupos e tomar como estimativa um ponto de cada grupo considera-se que o método de Hooke-Jeeves deve retornar soluções com boa acurácia.

5 Resultados e discussão

A metodologia proposta foi aplicada a duas funções-teste, que caracterizam problemas de minimização, ambas escolhidas bidimensionais para que todas as etapas da metodologia pudessem ser observadas graficamente. As funções foram retiradas de uma página de internet que contém uma série de *benchmarks* para testar algoritmos de otimização global, [1]. A primeira função teste possui quatro pontos de mínimo global e é conhecida como *cross in tray*, Eq.(1).

$$f(x) = -0,0001 \left(\left| e^{\left| 100 - \frac{\sqrt{x_1^2 + x_2^2}}{\pi} \right|} \left| \sin(x_1) \sin(x_2) \right| + 1 \right)^{0,1} \quad (1)$$

A segunda função teste utilizada é conhecida pelo nome de seu autor (*Parsopoulos*) e possui doze pontos de mínimo global. Esta função é definida como a seguir

$$g(x) = \cos(x_1)^2 + \sin(x_2)^2. \quad (2)$$

A representação gráfica destas funções pode ser observada na Figura 1. Na Figura 1(a) tem-se o gráfico da primeira função-teste (*cross in tray*) e na Figura 1(b) da segunda função-teste (*Parsopoulos*).

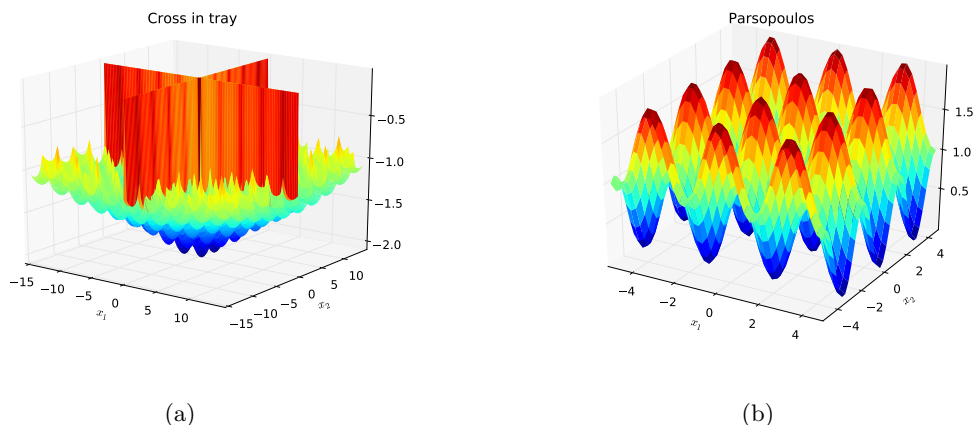


Figura 1: Gráficos das funções-testes: (a) *cross in tray function* e (b) *Parsopoulos function*

Para a execução do Busca Harmônica aplicado à primeira função foram utilizados 100 pontos e 300 iterações do método. Já para a segunda função foram usados 400 pontos e 500 iterações. O uso de um maior número de pontos para a segunda função se justifica pelo número de pontos de ótimo global da função. Com um número maior de pontos de ótimo, são necessárias mais iterações para formar os grupos. O domínio utilizado para as funções foi o mesmo, $x_1, x_2 \in [-5, 5]$. A Figura 2 exibe as curvas de nível das funções analisadas, no domínio dado, e a distribuição dos pontos ao final da execução do método evolucionário. Nas Figuras 2(a) e 2(b) pode-se notar que ao final do processo evolutivo os pontos formam grupos ao redor de cada um dos pontos de mínimo global.

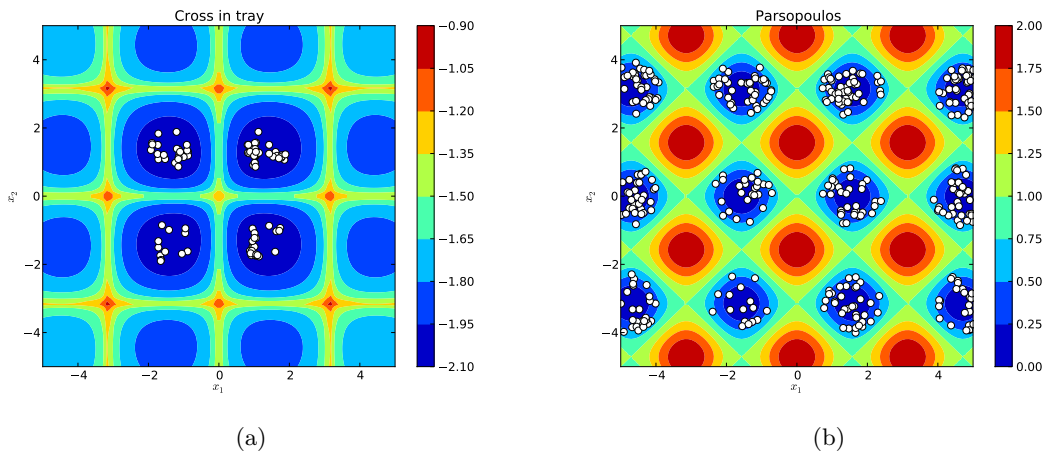


Figura 2: Curvas de nível e distribuição final dos pontos (a) *cross in tray function* e (b) *Parsopoulos function*.

A identificação dos grupos é feita pelo agrupamento espectral, utilizando o Laplaciano normalizado, e com o auxílio do gráfico de dispersão, do primeiro vetor de Fiedler pelo segundo, é possível observar o mapeamento dos pontos, exibido na Figura 3(a). Observa-se que o método identificou perfeitamente os quatro grupos distintos. A classificação dos pontos realizada pelo k-means é ilustrada na Figura 3(b), em que pontos de um mesmo grupo possuem a mesma cor.

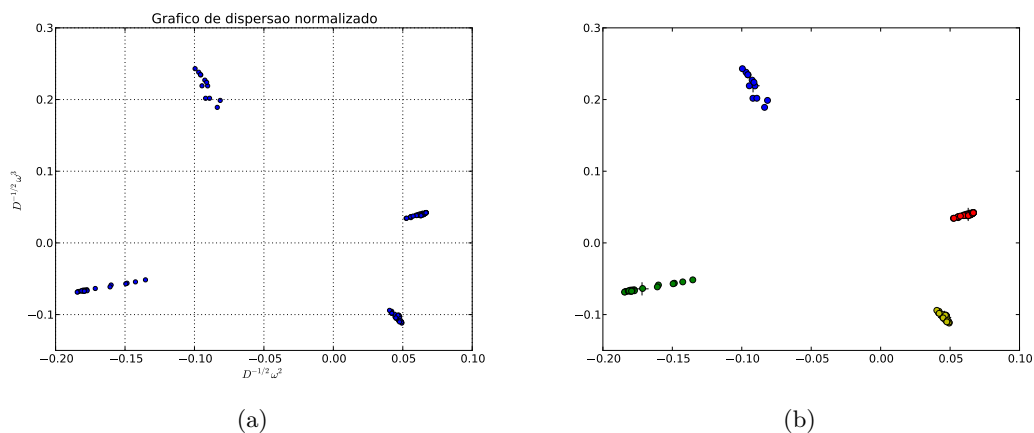


Figura 3: (a) Mapeamento dos grupos com o agrupamento espectral. (b) Classificação dos pontos com o k-means.

Conforme já ressaltado, após a classificação dos pontos sorteia-se um integrante de cada grupo para servir como estimativa inicial para o algoritmo de Hooke-Jeeves. Portanto, o método determinístico foi executado quatro vezes para a obtenção dos quatro pontos de mínimo global da função. Os resultados encontram-se na Tabela 1, juntamente com as soluções da referência e

o erro relativo. Nota-se que as soluções obtidas estão próximas das retiradas da referência.

Para a segunda função-teste utilizou-se um maior número de pontos para que os grupos pudessem ter um maior número de representantes e, assim, melhorar sua identificação. O mapeamento dos grupos da segunda função pode ser visto na Figura 4(a) onde observa-se que doze grupos foram identificados.

Solução	Referência	Metodologia	Erro relativo
1	(-1,34940661; -1,34940661)	(-1,3494371; -1,34945128)	$2,83 \times 10^{-5}$
2	(1,34940661; 1,34940661)	(1,34944656; 1,34940232)	$2,11 \times 10^{-5}$
3	(-1,34940661; 1,34940661)	(-1,34944357; 1,34946129)	$3,46 \times 10^{-5}$
4	(1,34940661; -1,34940661)	(1,34937644; -1,34937625)	$2,24 \times 10^{-5}$

Tabela 1: Pontos de ótimo global da função *cross in tray*

O agrupamento espectral apenas mapeia os grupos, a classificação dos pontos é feita pelo k-means. Esta classificação pode ser vista na Figura. 3(b), entretanto, foram necessárias algumas tentativas para que todos os grupos fossem classificados corretamente. O classificação feita pelo k-means depende fortemente dos centroides escolhidos inicialmente; caso estes não sejam tomados de forma adequada podem levar a classificações indesejadas.

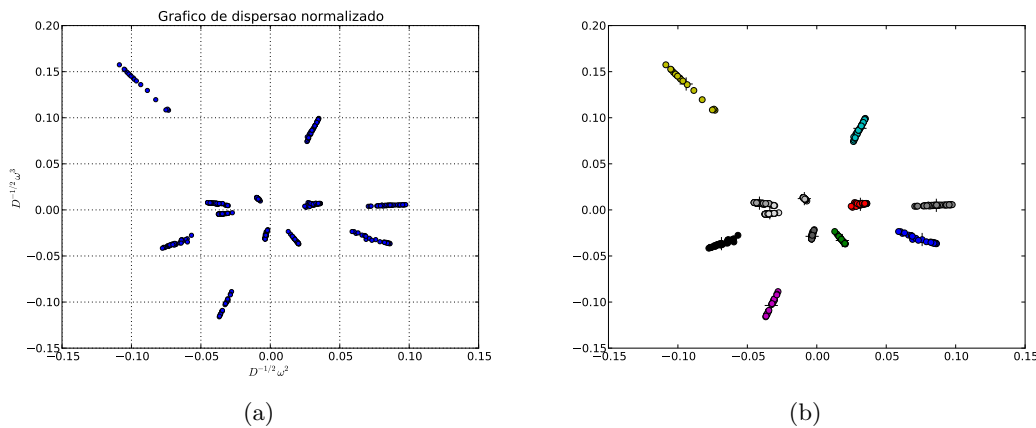


Figura 4: (a) Mapeamento dos grupos com o agrupamento espectral. (b) Classificação dos pontos com o k-means.

Novamente, após a classificação dos grupos, foram retirados de cada grupo um representante. Este foram tomados como estimativas iniciais para o método Hooke-Jeeves. Na Tabela 2 tem-se os resultados encontrados, os valores de referência e o erro relativo.

Como os métodos de agrupamento identificaram todos os grupos, então todas as doze soluções foram encontradas. Observa-se que apenas as soluções 10 e 11 obtiveram maiores erros relativos. Entretanto, os resultados são considerados satisfatórios, visto que todas as soluções foram encontradas e com pequenos erros. Ademais, a qualidade final das soluções pode ser controlada pelo critério de parada do método de Hooke-Jeeves.

Solução	Referência	Metodologia	Erro relativo
1	(-1,57079632; -3,14159265)	(-1,57083632; -3,14164000)	$1,76 \times 10^{-5}$
2	(-1,57079632; 0)	(-1,57075712; $3,45746843 \times 10^{-5}$)	$3,33 \times 10^{-5}$
3	(-1,57079632; 3,14159265)	(-1,57082952; 3,14155946)	$1,34 \times 10^{-5}$
4	(-4,71238898; -3,14159265)	(-4,71242218; -3,14161208)	$6,8 \times 10^{-6}$
5	(-4,71238898; 0)	(-4,71234877; $-4,05957267 \times 10^{-5}$)	$1,21 \times 10^{-5}$
6	(-4,71238898; 3,14159265)	(-4,71242925; 3,14155446)	$9,8 \times 10^{-6}$
7	(1,57079632; -3,14159265)	(1,57083636; -3,14162646)	$1,49 \times 10^{-5}$
8	(1,57079632; 0)	(1,57076374; $-3,64151187 \times 10^{-5}$)	$3,11 \times 10^{-5}$
9	(1,57079632; 3,14159265)	(1,57083545; 3,14156221)	$1,41 \times 10^{-5}$
10	(4,71238898; -3,14159265)	(4,70305524; -3,14765344)	$1,97 \times 10^{-3}$
11	(4,71238898; 0)	(4,62997834; $-4,98087875 \times 10^{-2}$)	$2,04 \times 10^{-2}$
12	(4,71238898; 3,14159265)	(4,71235607; 3,14163917)	$1,01 \times 10^{-5}$

Tabela 2: Ponto de ótimo global da função *Parsopoulos*

6 Considerações Finais

A metodologia proposta objetivou resolver problemas de otimização multimodal utilizando métodos de otimização e agrupamento. Os algoritmos Busca Harmônica, agrupamento espectral, k-means e Hooke-Jeeves foram executados sequencialmente para obter todas as soluções de problemas globalmente multimodais. A metodologia foi aplicada às funções-teste *cross in tray* e *Parsopoulos*, que possuem quatro e doze pontos de ótimo global, respectivamente.

Como as funções escolhidas apresentam duas variáveis, todas as etapas da metodologia foram observadas graficamente. O método de Busca Harmônica obteve sucesso ao separar os pontos ao redor de cada uma das soluções e o agrupamento espectral identificou todos os grupos de forma adequada. O algoritmo k-means deparou-se com dificuldades na classificação dos doze grupos para a segunda função-teste, pois depende fortemente dos centroides escolhidos inicialmente. O método de Hooke-Jeeves obteve todas as soluções a partir das estimativas iniciais retiradas de cada um dos grupos. As soluções encontradas, para as duas funções teste, obtiveram erro relativo pequeno, sendo que para segunda função duas soluções apresentaram erros relativos maiores que as demais.

Além de mapear grupos dentro de um conjunto de dados baseando-se em relações de similaridade, o agrupamento espectral é capaz de proporcionar redução de dimensionalidade. Dessa forma, o uso deste método torna possível a visualização dos grupos para funções que não podem ser observadas graficamente. Ademais, a redução de dimensionalidade reduz o custo computacional do k-means. Assim, a metodologia proposta foi aplicada a funções bidimensionais para analisar sua eficácia mas objetiva a resolução de problemas n-dimensionais. Em trabalhos futuros, funções n-dimensionais, que representam problemas globalmente multimodais, serão testadas e espera-se encontrar todas as suas soluções ótimas simultaneamente.

Referências

- [1] Global Optimization Benchmarks and AMPGO. Disponível em: <http://infinity77.net/global_optimization/>. Acesso em: 8 de jan. de 2014.
- [2] D. J. Higham, G. Kalna, M. Kibble, “Spectral clustering and its use in bioinformatics”, *J. Computational and Appl. Mathematics*, v. 204, pp. 25-37, 2007.
- [3] R. Hooke, T. A. Jeeves, “Direct search solution of numerical and statistical problems”, *Journal of the Association for Computing Machinery*, v. 8, n. 2, pp. 212-229, 1961.
- [4] X. Yang, “Nature-Inspired Metaheuristic Algorithms”, 2 ed., Luniver Press, 2010.
- [5] J. Wu, “Advances in k-means clustering: a Data mining thinking”, Springer Teses: recognizing outstanding Ph.D. Research, 1 2012.