



COMBINAÇÃO DE METAHEURÍSTICA E BUSCA LOCAL DETERMINÍSTICA INDEPENDENTE DE DERIVADAS

Viviane de J. Galvão

vivianegalvao@ice.ufjf.br

Universidade Federal de Juiz de Fora

Campus Universitário, Bairro Martelos, 36036-330 Juiz de Fora, MG, Brazil

Helio J. C. Barbosa

hcbm@lncc.br

Laboratório Nacional de Computação Científica

Av. Getúlio Vargas 333, 25651-075, Petrópolis, RJ, Brasil

Universidade Federal de Juiz de Fora

Campus Universitário, Bairro Martelos, 36036-330 Juiz de Fora, MG, Brazil

Resumo. Apesar de serem de fácil compreensão e implementação, os métodos determinísticos independentes de derivadas possuem certas limitações devido à sua taxa de convergência lenta e em relação ao quão grande possa ser a dimensão do problema. Já as metaheurísticas, apesar de sua robustez, demandam demasiado número de avaliações de função para convergirem. Este trabalho busca combinar o popular método de Otimização por Enxame de Partículas com o método de Busca Padrão, avaliar seus resultados e, ainda, verificar em quais condições, este tipo de combinação, gera boas respostas. Um dos principais objetivos deste trabalho é atingir bons resultados reduzindo o número de cálculos da função objetivo. Para isto, no método de Otimização por Enxame de Partículas é inserida uma estratégia de população dinâmica que depende da quantidade de acessos à Busca Padrão. Além disso, propõem-se melhorar a Busca Padrão a fim de diminuir a quantidade de avaliações de função desnecessárias.

Palavras-Chave: Otimização sem Derivadas, Busca Padrão, Metaheurística, Otimização por Enxame de Partículas

1 INTRODUÇÃO

Frequentemente, nas Ciências Exatas e nas Engenharias, são encontrados problemas de otimização com restrições, em espaço contínuo, em que o uso de derivadas se inviabiliza seja pela instabilidade no seu cálculo ou pela sua inacessibilidade como em problemas conhecidos como “caixa preta”, onde o cálculo da função objetivo e/ou das restrições envolve o uso de um simulador. Além disso, esta classe de problemas exige frequentemente avaliações com alto custo computacional. Tais características dificultam a utilização de aproximação numérica das derivadas por diferenças finitas e a de métodos baseados nas técnicas de Newton e quasi-Newton por envolver alto custo computacional e possível ruído contido na função objetivo ou restrições do problema. Estas dificuldades induzem o uso de uma classe de métodos apropriada: os métodos de Otimização Sem Derivadas (OSD).

Desde a década de 1960, a demanda de problemas motivou o desenvolvimento dos métodos da classe OSD que, até os anos 90, não tinham muita teoria relacionada à garantia de convergência. Atualmente, a ênfase tem sido no entendimento dos métodos já existentes com interesse na sua convergência global. O interesse nesses métodos tem crescido em razão da facilidade de implementação e paralelização, e grande aplicabilidade. Esses métodos podem ser aplicados, por exemplo, ao problema de projeto de pás do rotor de helicópteros e o objetivo poderia ser minimizar a vibração transmitida ao seu ponto central. Neste exemplo, a função objetivo é uma simulação que inclui dinâmica das estruturas e aerodinâmica e cada avaliação de função pode requerer minutos ou até mesmo dias em tempo de CPU. Pode-se encontrar um denso estudo destes métodos em Conn et al. (2009), incluindo análises de convergência e algumas aplicações práticas.

Dentre os métodos da classe OSD, pode-se citar os Métodos de Busca Direta, que utilizam apenas os valores da função objetivo para convergirem até um ponto de ótimo local. Mais detalhes sobre os Métodos de Busca Direta, como uma abordagem histórica, novas estratégias e um panorama sobre o tratamento de restrições laterais, lineares e não-lineares pode ser encontrado em Kolda et al. (2003). O método de Busca Padrão também é descrito por Kolda et al. (2003) e é uma estratégia que opera com movimentos exploratórios, além de ter uma estrutura flexível que permite sua combinação com outras heurísticas.

Apesar da sua aplicabilidade, os métodos de otimização sem derivadas têm taxa de convergência relativamente baixa e não atingem bons resultados para altas dimensões, diferentemente dos métodos baseados em derivadas. Pode-se esperar o sucesso desses métodos em problemas que não tenham mais do que uma centena de variáveis, que possuam custo muito alto de avaliação e nos quais uma rápida taxa de convergência não é o objetivo primordial. Alguns trabalhos tentam melhorar a eficiência e robustez desses métodos, como no trabalho de Custódio (2007) que aplica a teoria dos gradientes simpléticos no método de busca padrão e em Vaz and Vicente (2007) que combina a técnica de otimização por enxame de partículas com a busca padrão.

Por outro lado, existem os métodos de natureza estocástica que são populares devido à sua robustez, mas que possuem alto custo computacional. Dentre eles, pode-se citar a Otimização por Enxame de Partículas (PSO), um método populacional inspirado em mecanismos de inteligência coletiva. Um resumo de sua dinâmica, suas variações e aplicações pode ser encontrado em Polli et al. (2007).

A proposta deste trabalho consiste em combinar uma Busca Padrão com PSO para resolver problemas de otimização em espaço de busca contínuo com restrições laterais a fim de melhorar

o desempenho do PSO. Além disso, pretende-se modificar a busca padrão inicialmente adotada com o intuito de realizar menos cálculos da função objetivo. Experimentos computacionais com problemas geralmente usados na literatura foram realizados para analisar a técnica proposta.

2 BUSCA PADRÃO

Métodos de busca padrão (Kolda et al., 2003) podem ser vistos como métodos de busca direta que usam um conjunto de direções positivamente geradoras para se guiarem na exploração do espaço de busca. Mais informações sobre bases positivamente geradoras podem ser obtidas em Conn et al. (2009). Pode-se citar, como exemplos de busca padrão, a busca padrão original (Hooke and Jeeves, 1961), com movimentos exploratórios, e a busca multidirecional (Dennis et al., 1991).

A busca padrão adotada aqui (chamada de Busca Direta Direcional por Conn et al. (2009)) considera dois passos em sua estrutura: o passo de procura e o passo de avaliação. No passo de procura é feita uma sondagem dentro de um conjunto finito de pontos A onde, dada uma iteração k , uma função objetivo f e x_{k-1} o melhor ponto da iteração $k - 1$, se existe um ponto $x \in A$ tal que $f(x) < f(x_{k-1})$, então $x_k = x$. Caso não haja sucesso no passo de procura, o passo de avaliação é iniciado. Considera-se o conjunto G de bases positivamente geradoras, um passo α_{k-1} e o conjunto $P_k = \{x_{k-1} + \alpha d\}$ tal que $d \in G$. Assim, se existe um ponto $x \in P_k$ tal que $f(x) < f(x_{k-1})$, então $x_k = x$ e $\alpha_k = \gamma \alpha_{k-1}$ para $\gamma \geq 1$; caso contrário, $x_k = x_{k-1}$ e $\alpha_k = \beta \alpha_{k-1}$ para $\beta < 1$. O conjunto de bases considerado é composto pelos vetores que formam a base canônica e seus negativos.

Para o caso de problemas com restrições laterais a convergência dos métodos de busca padrão é verificada em Lewis and Torczon (1999) sob a condição da função objetivo ser diferenciável e o conjunto de direções ser apropriadamente escolhido. Neste caso, a busca padrão procede da mesma forma, mas avaliando se o ponto analisado é factível em relação às restrições.

3 OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS

A otimização por enxame de partículas (PSO) (Kennedy and Eberhart, 1995) foi inspirada no comportamento social de grupos de pássaros, simulando mecanismos de inteligência coletiva. Nesta técnica, as soluções candidatas são representadas por uma população de partículas, onde cada uma tem sua velocidade e carrega a informação do melhor ponto visitado por ela e pelo enxame.

Denotando por t a iteração atual e por n a dimensão do problema, enquanto um critério de parada não for alcançado, a velocidade (ν) e a posição (x) de cada partícula é atualizada via:

$$\nu_i^{t+1} = \delta \nu_i^t + \eta \omega_{1i}^t (x_i^t(\text{melhorLocal}) - x_i^t) + \psi \omega_{2i}^t (x_i^t(\text{melhorGlobal}) - x_i^t) \quad (1)$$

$$x_i^{t+1} = x_i^t + \nu_i^{t+1} \quad (2)$$

onde δ é um fator de inércia da partícula, $\eta > 0$ é um fator de aprendizagem, $\psi > 0$ é um fator social, ω_{1i} e ω_{2i} são números aleatórios no intervalo $(0, 1)$ e $i = 1, \dots, n$.

Para cada partícula, verifica-se se a nova posição é melhor do que a posição atual; se sim, o melhor ponto da partícula é atualizado e é feita a comparação com o melhor ponto global. A iteração tem sucesso quando o melhor ponto global for melhorado.

Uma técnica para estimar o coeficiente de inércia é inicializá-lo com um valor relativamente alto e reduzi-lo gradativamente para um valor muito menor. Isto porque o coeficiente de inércia pode ser interpretado com a fluidez do meio no qual se encontram as partículas. Para valores altos o meio é pouco viscoso, o que proporciona maior exploração. Já para valores baixos, a viscosidade é alta e a busca se concentra próxima ao ótimo local.

Quando são considerados problemas com restrições laterais de limites inferiores e superiores para cada variável, o problema é tratado com a projeção dos pontos para dentro do domínio factível. Neste caso, o valor limite admissível é atribuído à posição violada.

4 A TÉCNICA HÍBRIDA

A estrutura da técnica híbrida proposta consiste na execução do PSO no passo de procura da Busca Padrão e, caso necessário, do passo de avaliação em seguida. A técnica híbrida é detalhada no Algoritmo 1.

Uma adaptação, baseada na quantidade de sucessos que cada partícula tem, foi incorporada ao PSO; se a quantidade de sucessos de uma partícula superar a quantidade de acessos à busca local então esta é desativada, caso contrário é ativada. Outra adaptação foi incorporada, onde as partículas mais próximas da melhor partícula global são desativadas.

Um grande problema é que a Busca Padrão, no seu passo de avaliação, chega a consumir $2n$ avaliações de função objetivo em cada acesso, o que pode acarretar num desperdício de avaliações de função, sendo n a dimensão do problema. Para diminuir o número de avaliações de função na Busca Padrão, foi feita a memorização da última coordenada na qual se obteve sucesso e, na sua próxima execução, a busca começa pela coordenada posterior.

5 EXPERIMENTOS COMPUTACIONAIS

Foram realizados experimentos computacionais utilizando a técnica híbrida aqui proposta e os resultados encontrados foram comparados com aqueles obtidos pelo PSwarm, um *software* que considera uma implementação de um PSO com uma Busca Padrão, de Vaz and Vicente (2007). Nenhum ajuste fino dos parâmetros foi realizado, de modo que os valores adotados aqui são os mesmos usados por Vaz and Vicente (2007). Para a Busca Padrão tem-se $\gamma = 2.0$ e $\beta = 0.5$; para o PSO tem-se $\eta = \psi = 0.5$ e $\delta(t) = 0.9 - (0.5)(t_{max}/t)$, onde t é a iteração corrente e $t_{max} = 2000$ é o máximo de iterações. São 2000 cálculos de função permitidos e foram realizadas 30 execuções independentes.

Foram usados 10 problemas comumente encontrados na literatura (Bernardino et al., 2011). Os problemas f_{01} , f_{02} , f_{06} e f_{07} são unimodais, enquanto os problemas f_{03} , f_{04} , f_{05} e f_{10} são multimodais. O problema f_{08} é uma função degrau e possui apenas um mínimo e f_{09} é um polinômio de quarto grau com um ruído sendo incorporado via um valor aleatório $U(0, 1)$ uniformemente distribuído no intervalo $[0; 1)$. As formulações dos problemas-teste usados aqui e suas restrições de limites (mínimo e máximo) admissíveis podem ser encontradas na Tabela 1. É importante destacar que as funções possuem dimensão n ajustável e os experimentos foram conduzidos assumindo $n = 30$.

Algoritmo 1: Otimização por Enxame de Partículas com Busca Padrão.

Defina o tamanho p da população e um critério de parada;
 Inicialize as posições e velocidades de cada partícula $x_1^0, \dots, x_p^0$ aleatoriamente;
 Considere $x_i^0(local) = x_i^0, i = 1, \dots, p$;
 Seja $x_{global}^0 = \operatorname{argmin}_{y \in \{x_1^0, \dots, x_p^0\}} f(y)$. Declare $t = 0$;
enquanto critério de parada é falso **faça**
 sucesso = falso ; $t = t + 1$;
 para cada $i = 1, \dots, p$ **faça**
 atualizar a velocidade v_i^t e a posição x_i^t ;
 se $f(x_i^t) < f(x_i^{t-1}(local))$ **então**
 $x_i^t(local) = x_i^t$;
 se $f(x_i^t(local)) < f(x_{global}^{t-1})$ **então**
 $x_{global}^t = x_i^t(local)$;
 sucesso = verdadeiro;
 se (sucesso = falso) **então**
 se $\exists d \in G$ tal que $f(x_{global}^{t-1} + \alpha^{t-1}d) < f(x_{global}^{t-1})$ **então**
 $\alpha^t = \gamma \alpha^{t-1}$ (aumentar o passo);
 $x_{global}^t = x_{global}^{t-1} + \alpha^{t-1}d$;
 senão
 $\alpha^t = \beta \alpha^{t-1}$ (diminuir o passo);
 $x_{global}^t = x_{global}^{t-1}$;
fim enquanto

5.1 Resultados

Os resultados encontrados nos experimentos computacionais são apresentados nas Tabelas 2, 3, 4 e 5, apresentando, respectivamente, médias, desvios padrões, mínimos e máximos valores encontrados entre as soluções das execuções independentes. Os melhores resultados para cada problema são destacados em **negrito**. Os rótulos adotados para cada uma técnicas são: PSO refere-se ao PSO sem Busca Padrão, PSOBL ao PSO com Busca Padrão, PSwarm é o PSwarm original de Vaz and Vicente (2007), e PSwarm* e PSOBL* contêm a Busca Padrão aperfeiçoada. Os métodos são indicados na primeira linha de cada tabela.

Os resultados da Tabela 2 mostram melhora no desempenho do PSO com a inserção da Busca Padrão e a eficácia da proposta para diminuir o número de avaliações de função na Busca Padrão, indicando que boas respostas podem ser alcançadas com sua utilização. Ainda, nota-se que, dos 10 problemas testados, a técnica PSOBL consegue alcançar melhores respostas que o PSWARM em 6 problemas, o que significa que a técnica de população dinâmica atribuída ao PSO é eficaz. No entanto, da Tabela 3, o PSWARM se mostrou mais estável que PSOBL proposto aqui. É importante destacar que (i) todos os melhores resultados médios são obtidos em alguma das técnicas com a Busca Padrão aperfeiçoada, (ii) o PSWARM* obteve os melhores resultados médios em 5 dos 10 problemas, e (iii) o PSOBL* alcançou os melhores resultados médios nos outros 5 problemas.

Da Tabela 3 tem-se que o PSWARM* conseguiu resultados mais próximos da sua mé-

Tabela 1: Problemas

Funções	Restrições
$f_{01}(x) = \sum_{i=1}^n x_i^2$	$[-100, 100]$
$f_{02}(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2) + (x_i + 1)^2]$	$[-30, 30]$
$f_{03}(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	$[-5.12, 5.12]$
$f_{04}(x) = -20 \exp \left(\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \right) - \exp \left[\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i) \right] + 20 + e$	$[-32, 32]$
$f_{05}(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos \left(\frac{x_i}{\sqrt{i}} \right) + 1$	$[-600, 600]$
$f_{06}(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	$[-10, 10]$
$f_{07}(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2$	$[-100, 100]$
$f_{08}(x) = \sum_{i=1}^n [x_i + 0.5]$	$[-100, 100]$
$f_{09}(x) = \sum_{i=1}^n i x_i^4 + U(0, 1)$	$[-1.28, 1.28]$
$f_{10}(x) = \sum_{i=1}^n -x_i \sin(\sqrt{ x_i })$	$[-500, 500]$

dia que o PSWARM. Apesar de obter melhores resultados em média que o PSOBL, a técnica PSOBL* não é tão estável quando comparada tanto ao PSOBL quanto ao PSWARM*. Em termos globais, não é possível concluir qual das técnicas é mais estável.

Por fim, das Tabelas 4 e 5, a técnica PSOBL* obteve melhor desempenho no melhor caso e a técnica PSWARM* obteve o melhor desempenho no pior caso.

6 CONCLUSÕES

Visando melhorar o desempenho do PSO e alcançar bons resultados com menos cálculos de função, foram propostos a combinação do PSO com a Busca Padrão e uma técnica de redução de avaliações de função na Busca Padrão.

A partir dos resultados dos experimentos computacionais, a combinação dos métodos é bastante efetiva, inclusive seus resultados tendem ser tão estáveis quanto os resultados da implementação original do PSO. Também, o uso da Busca Padrão modificada garante melhor desempenho em relação ao uso Busca Padrão original de Kolda et al. (2003), indicando que pode-se esperar boas respostas evitando a realização de muitos cálculos de função acumulados num estágio da busca.

Nota-se que a Busca Padrão com o método de redução de cálculos de função aplicado ao PSwarm gera resultados bastante competitivos em relação à implementação original do PSwarm (Vaz and Vicente, 2007). Também é possível perceber que, em alguns casos, a téc-

Tabela 2: Médias dos Resultados

	PSO	PSOBL	PSWARM	PSWARM*	PSOBL*
f_{01}	8644.25292	3891.51593	1155.79831	160.19485	3711.94884
f_{02}	6371238.37903	2324435.89481	253586.02300	71034.64680	2240448.00341
f_{03}	231.19114	87.58554	168.00159	154.81582	10.89912
f_{04}	15.61644	5.13454	8.66166	4.44875	3.30421
f_{05}	24.59102	9.08691	11.40218	2.44176	8.16494
f_{06}	68.06103	9.19099	15.32471	7.43402	7.09755
f_{07}	15878.56136	9567.70045	12645.81543	11818.40706	9790.91133
f_{08}	-901.86667	-2644.90000	-2970.00000	-2970.00000	-2967.03333
f_{09}	3.15592	0.53428	0.47091	0.41832	0.53662
f_{10}	-3369.74740	-9244.95408	-8743.35000	-8841.74618	-9562.61156

Tabela 3: Desvio Padrão dos Resultados

	PSO	PSOBL	PSWARM	PSWARM*	PSOBL*
f_{01}	2452.48568	3723.02114	442.88020	212.64681	3732.65602
f_{02}	2945088.21614	2334068.17434	115080.95536	64122.77940	2308692.61380
f_{03}	24.88357	35.97504	19.25034	17.44717	4.44271
f_{04}	0.70836	1.02365	1.22728	1.33961	0.99764
f_{05}	7.36044	8.56875	3.98592	1.91382	8.73584
f_{06}	103.50350	7.65807	3.10133	4.34735	9.35036
f_{07}	5411.42633	2959.78598	3949.85257	3461.10893	3025.88379
f_{08}	118.62277	84.06211	0.00000	0.00000	1.37674
f_{09}	1.01535	0.19646	0.16956	0.13348	0.22516
f_{10}	469.14558	953.13296	533.70000	702.73236	1063.48698

Tabela 4: Melhores Resultados

	PSO	PSOBL	PSWARM	PSWARM*	PSOBL*
f_{01}	4630.62011	48.53892	501.77934	2.72006	32.94124
f_{02}	2663103.89600	4712.96010	38918.41930	1239.94001	1895.81945
f_{03}	182.98717	38.42866	108.55652	129.20280	0.41764
f_{04}	13.98402	3.63179	6.38570	2.13627	1.11701
f_{05}	11.13279	1.46886	5.51601	1.02433	1.01120
f_{06}	31.42554	3.22150	10.96314	1.82931	0.58921
f_{07}	8240.77108	6131,58038	5960.24597	6198,86834	5602.78425
f_{08}	-1200.00000	-2785.00000	-2970.00000	-2970.00000	-2969.00000
f_{09}	0.93468	0.21982	0.22415	0.15377	0.19597
f_{10}	-4447,94557	-10538.50309	-10008.83000	-10885,33141	-10920,40041

Tabela 5: Piores Resultados

	PSO	PSOBL	PSWARM	PSWARM*	PSOBL*
f_{01}	15631.49297	15631.49297	2945.31797	1103.92212	15631.49297
f_{02}	16034277.82000	10570159.72000	513353.97784	255497.27060	9792137.23400
f_{03}	291.34420	173.61043	204.00034	189.01985	16.94574
f_{04}	16.67526	8.04577	11.29123	7.90511	4.60577
f_{05}	39.73515	28.60586	27.50786	10.93530	26.86595
f_{06}	612.72747	39.99005	22.94679	16.57734	39.99005
f_{07}	26285.12587	17099.35558	19659.55740	19659.55741	17099.35558
f_{08}	-696.00000	-2442.00000	-2970.00000	-2970.00000	-2965.00000
f_{09}	5.15905	1.04200	0.96517	0.78285	0.93739
f_{10}	-2568.64444	-6503.63777	-7909.67000	-7675,94230	-7666,89647

nica híbrida proposta aqui supera o desempenho do PSwarm bem como a técnica híbrida com a Busca Padrão modificada.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro da FAPEMIG (PCE-01337-15) e do CNPq (proc. 310778/2013-1).

REFERÊNCIAS

- H. S. Bernardino, H. J.C. Barbosa, and L. G. Fonseca. Surrogate-assisted clonal selection algorithms for expensive optimization problems. *Evol. Intelligence*, 4(2):81–97, 2011.
- A. R. Conn, K. Scheinberg, and L. N. Vicente. *Introduction to Derivative-Free Optimization*. Society for Industrial and Applied Mathematics, 2009.
- A. L. Custódio. *Aplicações de Derivadas Simpléticas em Métodos de Procura Directa*. PhD thesis, New University of Lisbon, 2007.
- J. E. Dennis, Jr., and V. Torczon. Direct search methods on parallel machines. *SIAM Journal on Optimization*, 1:448–474, 1991.
- R. Hooke and T. A. Jeeves. “direct search” solution of numerical and statistical problems. *Journal of the Assoc. Comput. Mach*, 8(2):212–229, 1961.
- J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proc. of the IEEE International Conference on Neural Networks*, volume 4, pages 1942–1948. IEEE, 1995.
- T. G. Kolda, R. M. Lewis, and V. Torczon. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Review*, 45(3):385–482, 2003.
- R. M. Lewis and V. Torczon. Pattern search algorithms for bound constrained optimization. *SIAM Journal on Optimization*, 9(4):1082–1099, 1999.
- R. Polli, J. Kennedy, and T. Blackwell. Particle swarm optimization. *Swarm Intelligence*, 1(1): 33–57, 2007.
- A. I. F. Vaz and L. N. Vicente. A particle swarm pattern search method for bound constrained global optimization. *Journal of Global Optimization*, 39(2):197–219, 2007.