

Comparação Estatística Entre Duas Variantes do Método de Otimização de Colisão de Partículas

Statistical Test to Compare Two Variants of the Particle Collision Algorithm

Ane Éilda Nogueira Frauches Almoia¹, Wagner Figueiredo Sacco², Antônio J. Silva Neto³.

RESUMO

Algumas variantes do algoritmo de colisão de partículas (Particle Collision Algorithm - PCA) vêm sendo propostas, e entre elas, o algoritmo de colisão de partículas com Hooke-Jeeves (HJPCA) e o algoritmo de colisão de partículas com seção de choque (CSPCA). O objetivo central deste artigo é utilizar-se de ferramentas estatísticas não-paramétricas a fim de comparar as taxas de sucesso quando estas duas variantes são expostas a problemas de natureza e complexibilidade distintas. Para isso, foi utilizado o teste dos postos sinalizados de Wilcoxon para determinar se estatisticamente um algoritmo apresenta um desempenho melhor do que o outro. Foram utilizadas funções teste desenvolvidas para o estudo do desempenho de métodos de otimização, bem como problemas modelados por sistemas de equações algébricas não-lineares. Ambos os algoritmos foram capazes de produzir resultados satisfatórios para as funções teste. Mas quando submetidos à segunda classe de problemas, seus desempenhos foram aparentemente distintos.

Palavras-chave: Algoritmo de colisão de partículas. Comparação estatística. Estatística não-paramétrica.

ABSTRACT

Several variants of the Particle Collision Algorithm (PCA) are being proposed, among them the Hooke-Jeeves Particle Collision Algorithm (HJPCA) and the Cross-Section Particle Collision Algorithm (CSPCA). The main objective of this paper is to use a non-parametric statistics tool, the Wilcoxon signed-rank test, in order to compare these two variants applied to benchmark problems, including complex test functions, non-linear systems and parameter-estimation problems. Both HJPCA and CSPCA were able to find the minimum global for the test functions. However, a different performance was noticed for the other set of problems.

Keywords: Particle collision algorithm. Statistical comparison. Nonparametric statistics.

¹ Mestre em Modelagem Computacional (IPRJ/UERJ). Engenheira de Controle e Automação (IFF).

E-mail: anefrauches@gmail.com

² D.Sc. em Engenharia Nuclear (COPPE/UFRRJ). Professor Associado na Universidade Federal do Oeste do Pará.

³ Ph.D. em Engenharia Mecânica (North Carolina State University). Professor Titular do Instituto Politécnico - UERJ.

1. INTRODUÇÃO

Um problema de otimização consiste na minimização ou maximização de uma determinada função objetivo, sendo esta dependente de um número finito de variáveis. Algoritmos de otimização metaheurísticos são considerados promissores em problemas que apresentam elevada complexidade, devido a sua capacidade de evitar os mínimos locais graças à presença de uma componente aleatória. O algoritmo de colisão de partículas (PCA) (SACCO; OLIVEIRA; PEREIRA, 2006; SACCO; OLIVEIRA, 2005) foi inspirado na física de colisão de partículas nucleares (DUDERSTADT; HAMILTON, 1976), com ênfase nas iterações de espalhamento e absorção. Algumas variantes deste método vêm sendo propostas, entre elas o método híbrido que combina o PCA com o algoritmo de busca Hooke-Jeeves (HJPCA) (RIOS-COELHO; SACCO; HENDERSON., 2010) e o algoritmo de colisão de partículas com seção de choque (CSPCA) (SACCO; RIOS-COELHO, 2016), que, conforme indica o nome, incorpora o conceito de seção de choque (DUDERSTADT; HAMILTON, 1976) ao algoritmo PCA.

Já foi provado por meio do “*No free lunch theorem*” (WOLPERT; MACREADY, 1997), que não existe um único algoritmo universalmente eficiente. Com isso, a comunidade de aprendizado de máquina tem utilizado os testes estatísticos como uma forma de comparar e validar os algoritmos desenvolvidos (DEMSAR, 2006). Como não se conhece exatamente a distribuição dos dados em questão, o melhor caminho é a utilização dos testes não-paramétricos (WACKERLY; MENDENHALL; SCHEAFFER, 2014).

Neste trabalho, os algoritmos HJPCA e CSPCA são aplicados a problemas *benchmark* e é empregado o teste não-paramétrico dos postos sinalizados de Wilcoxon, para determinar se existe alguma diferença significativa de desempenho entre os dois algoritmos em questão.

2. MATERIAIS E MÉTODOS

A seguir serão descritos os materiais e métodos considerados na realização desse trabalho.

2.1. Problemas de Otimização Considerados

Neste trabalho utilizou-se um conjunto de funções algébricas que apresentam comportamentos diferenciados, que as tornam capazes de cobrir diferentes tipos de dificuldades encontradas em problemas de otimização global (HEDAR; FUKUSHIMA,

2006). Além disso, utilizou-se também um conjunto de problemas que são modelados por meio de sistemas de equações algébricas não-lineares, entre eles estão incluídos alguns problemas práticos de estimação de parâmetros, que se utilizam de métodos de otimização para a obtenção da solução desejada. Tais funções e problemas são apresentados a seguir.

2.1.1. Branin

A função conhecida como *Branin* (DIXON; SZEGO, 1978) é definida por:

$$f(x) = a(x_2 - b x_1^2 + c x_1 - r)^2 + s(1 - t)\cos(x_1) + s$$

onde $a = 1$, $b = \frac{5,1}{4\pi^2}$, $c = \frac{5}{\pi}$, $r = 6$, $s = 10$ e $t = \frac{1}{8\pi}$. Sua representação gráfica é mostrada na Figura 1.

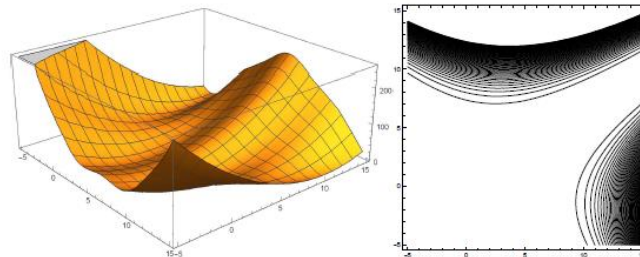


Figura 1. Função Branin: representação em 3D e curvas de nível

Dimensão: 2.

Domínio: $[-5, 15]$.

Mínimo global: $f(x^*) = 0,397887$, quando $x^* = (-\pi; 12,275)$, $(\pi; 2,275)$ e $(9,42478; 2,475)$.

2.1.2. Easom

A função conhecida como *Easom* (MICHALEWICZ, 1996) é definida pela equação abaixo. Sua representação gráfica é mostrada na Figura 2.

$$f(x) = -\cos(x_1)\cos(x_2)\exp(-(x_1 - \pi)^2 - (x_2 - \pi)^2)$$

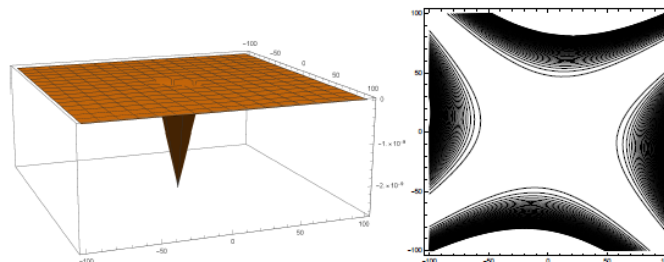


Figura 2. Função Easom: representação em 3D e curvas de nível

Dimensão: 2.

Domínio: $[-100, 100]$.

Mínimo global: $f(x^*) = -1$, quando $x^* = (\pi; \pi)$.

2.1.3. Goldstein-Price

A função conhecida como *Goldstein-Price* (DIXON; SZEGO, 1978) é definida pela equação abaixo. Sua representação gráfica é mostrada na Figura 3. Essa função apresenta vários mínimos locais e um único mínimo global.

$$f(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] \\ \times [30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$$

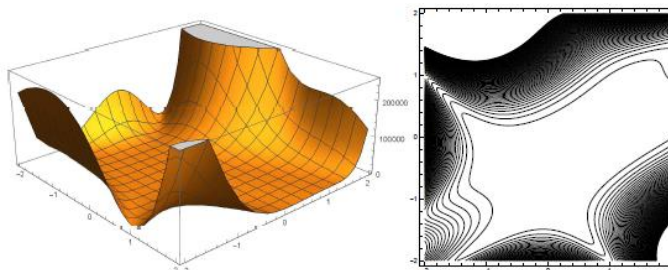


Figura 3. Função Goldstein-Price: representação em 3D e curvas de nível

Dimensão: 2.

Domínio: $[-2, 2]$.

Mínimo global: $f(x^*) = 3$, quando $x^* = (0; -1)$.

2.1.4. Rosenbrock

As funções conhecidas como *Rosenbrock* (MORE; GARBOW; HILLSTROM, 1943) são definidas pela equação abaixo. Sua representação gráfica, na configuração de 2 dimensões, é mostrada na Figura 4. Neste trabalho foram utilizadas as funções *Rosenbrock-02*, *Rosenbrock-05* e *Rosenbrock-10*, onde o número após o nome representa a dimensão d .

$$f(x) = \sum_{i=1}^{d-1} 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2$$

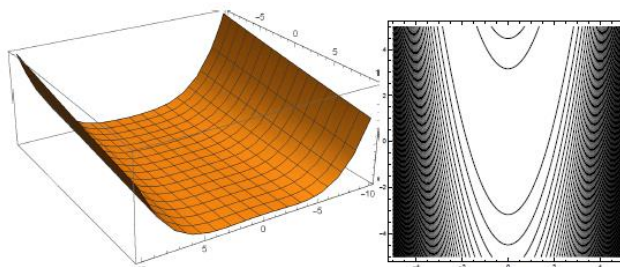


Figura 4. Função Rosenbrock-02: representação em 3D e curvas de nível

Dimensão: d .

Domínio: $[-5, 5]$.

Mínimo global: $f(x^*) = 0$, quando $x^* = (1; \dots; 1)$.

2.1.5. Zakharov

As funções conhecidas como *Zakharov* (HEDAR; FUKUSHIMA, 2006) são definidas pela equação abaixo. Sua representação gráfica, na configuração de duas dimensões, é mostrada na Figura 5. Neste trabalho foram utilizadas as funções *Zakharov-05* e *Zakharov-10*, onde o número após o nome representa a dimensão d .

$$f(x) = \sum_{i=1}^d x_i^2 + \left(\sum_{i=1}^d 0,5 i x_i \right)^2 + \left(\sum_{i=1}^d 0,5 i x_i \right)^4$$

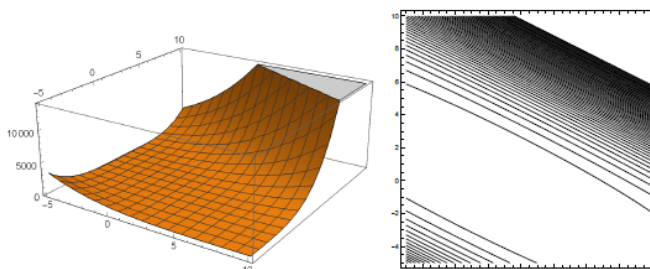


Figura 5. Função Zakharov-02: representação em 3D e curvas de nível

Dimensão: d .

Domínio: $[-5, 10]$.

Mínimo global: $f(x^*) = 0$, quando $x^* = (0; \dots; 0)$.

2.1.6. Shubert

A função conhecida como *Shubert* (LEVY; MONTALVO, 1985) é definida pela equação abaixo. Sua representação gráfica é mostrada na Figura 6. Essa função apresenta vários mínimos globais.

$$f(x) = 1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)$$

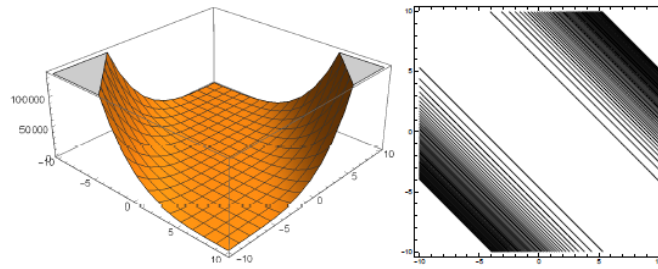


Figura 6. Função Shubert: representação em 3D e curvas de nível

Dimensão: 2.

Domínio: $[-10, 10]$.

Mínimo global: $f(\mathbf{x}^*) \approx -186,7309$, quando $\mathbf{x}^* \approx (-7,0835; 4,8580)$,

$(-7,0835; -7,7083)$, $(-1,4251; -7,0835)$, $(5,4828; 4,8580)$, $(-1,4251; -0,8003)$,
 $(4,8580; 5,4858)$, $(-7,7083; -7,0835)$, $(-7,0835; -1,4251)$, $(-7,7083; -0,8003)$,
 $(-7,7083; 5,4828)$, $(-0,8003; -7,7083)$, $(-0,8003; -1,4251)$, $(-0,8003; 4,8580)$,
 $(-1,4251; 5,4828)$, $(5,4828; -7,7083)$, $(4,8580; -7,0835)$, $(5,4828; -1,4251)$.

2.1.7. Hartmann As funções conhecidas como *Hartmann* (DIXON; SZEGO, 1978)

são definidas pela equação abaixo. Neste trabalho foram utilizadas as funções *Hartmann-03* e *Hartmann-06*, onde o número após o nome representa a dimensão n .

$$f_{n,m}(\mathbf{x}) = - \sum_{i=1}^m \alpha_i \exp \left(- \sum_{j=1}^n A_{ij}^{(n)} (x_j - P_{ij}^{(n)})^2 \right)$$

onde para $n = 3$ e $m = 4$:

$$\alpha = (1,0; 1,2; 3,0; 3,2)$$

$$A^{(3)} = \begin{pmatrix} 3,0 & 10 & 30 \\ 0,1 & 10 & 35 \\ 0,1 & 10 & 35 \\ 0,1 & 10 & 35 \end{pmatrix}$$

$$P^{(3)} = 10^{-4} \begin{pmatrix} 3689 & 1170 & 2673 \\ 4699 & 4387 & 7470 \\ 1091 & 8732 & 5547 \\ 381 & 5743 & 8828 \end{pmatrix}$$

e para $n = 6$ e $m = 4$:

$$A^{(6)} = \begin{pmatrix} 10 & 3 & 17 & 3,50 & 1,7 & 8 \\ 0,05 & 10 & 17 & 0,1 & 8 & 14 \\ 3 & 3,5 & 1,7 & 10 & 17 & 8 \\ 17 & 8 & 0,05 & 10 & 0,1 & 14 \end{pmatrix}$$

$$P^{(6)} = 10^{-4} \begin{pmatrix} 1312 & 1696 & 5569 & 124 & 8283 & 5886 \\ 2329 & 4134 & 8307 & 3736 & 1004 & 9991 \\ 2348 & 1451 & 3522 & 2883 & 3047 & 6650 \\ 4047 & 8828 & 8732 & 5743 & 1091 & 381 \end{pmatrix}$$

Dimensão: n .

Domínio: $[0,1]$.

Mínimo global:

$n = 3, m = 4 : f_{3,4}(x^*) \approx -3,86278$ quando $x^* \approx (0,114614; 0,555649; 0,852547)$.

$n = 6, m = 4 : f_{6,4}(x^*) \approx -3,32237$ quando $x^* \approx (0,20169; 0,150011; 0,476874; 0,275332; 0,311652; 0,6573)$.

2.1.8. Shekel

As funções conhecidas como *Shekel* (DIXON; SZEGO, 1978) são definidas pela equação abaixo. Neste trabalho foram utilizadas as funções *Shekel-05*, *Shekel-07* e *Shekel-10*, onde o número após o nome representa a ordem m .

$$f_{4,m}(x) = \sum_{i=1}^m \left(\sum_{j=1}^4 [(x_j - \alpha_{ij})^T (x - \alpha_i) + c_i] \right)^{-1}$$

onde para $m = 10$:

$$\alpha = \begin{pmatrix} 4,0 & 4,0 & 4,0 & 4,0 \\ 1,0 & 1,0 & 1,0 & 1,0 \\ 8,0 & 8,0 & 8,0 & 8,0 \\ 6,0 & 6,0 & 6,0 & 6,0 \\ 7,0 & 3,0 & 7,0 & 3,0 \\ 2,0 & 9,0 & 2,0 & 9,0 \\ 5,0 & 5,0 & 3,0 & 3,0 \\ 8,0 & 1,0 & 8,0 & 1,0 \\ 6,0 & 2,0 & 6,0 & 2,0 \\ 7,0 & 3,6 & 7,0 & 3,6 \end{pmatrix}$$

$$c = (0,1; 0,2; 0,2; 0,4; 0,4; 0,6; 0,3; 0,7; 0,5; 0,5)$$

Dimensão: 4.

Domínio: $[0, 10]$.

Mínimo global:

$$m = 5 : f_{4,5}(x^*) \approx -10,1532 \text{ quando } x^* \approx (4; 4; 4; 4).$$

$$m = 7 : f_{4,7}(x^*) \approx -10,4029 \text{ quando } x^* \approx (4; 4; 4; 4).$$

$$m = 10 : f_{4,10}(x^*) \approx -10,5364 \text{ quando } x^* \approx (4; 4; 4; 4).$$

2.1.9. Equilíbrio Químico Com 10 Variáveis (EQ10)

Este problema se refere à combustão de propano (C_3H_8) no ar (O_2 e N_2) a fim de formar dez produtos. A reação química gera um sistema de dez equações e dez variáveis apresentado a seguir (MEINTJES; MORGAN, 1990):

$$\left\{ \begin{array}{l} f_1 = x_1 + x_4 - 3 \\ f_2 = 2x_1 + x_2 + x_4 + x_7 + x_8 + x_9 + 2x_{10} - R \\ f_3 = 2x_2 + 2x_5 + x_6 + x_7 - 8 \\ f_4 = 2x_3 + x_9 - 4R \\ f_5 = K_5 x_2 x_4 - x_1 x_5 \\ f_6 = K_6 x_2^{(\frac{1}{2})} x_4^{(\frac{1}{2})} - x_1^{(\frac{1}{2})} x_6 \left(\frac{p}{n_T}\right)^{(\frac{1}{2})} \\ f_7 = K_7 x_2^{(\frac{1}{2})} x_4^{(\frac{1}{2})} - x_4^{(\frac{1}{2})} x_7 \left(\frac{p}{n_T}\right)^{(\frac{1}{2})} \\ f_8 = K_8 x_1 - x_4 x_8 \left(\frac{p}{n_T}\right) \\ f_9 = K_9 x_1 x_3^{(\frac{1}{2})} - x_4 x_9 \left(\frac{p}{n_T}\right)^{(\frac{1}{2})} \\ f_{10} = K_{10} x_1^2 - x_4^2 x_{10} \left(\frac{p}{n_T}\right) \end{array} \right.$$

onde:

$$\left\{ \begin{array}{l} n_T = \sum_{i=1}^{10} x_i \\ p = 40 \text{ atm} \\ R = 10 \\ K_5 = 0,193 \\ K_6 = 0,002597 \\ K_7 = 0,003448 \\ K_8 = 0,00001799 \\ K_9 = 0,0002155 \\ K_{10} = 0,00003846 \end{array} \right.$$

Dimensão: 10.

Domínio: $[0, 100]$.

Mínimo global: $f(x^*) = 0$, quando $x^* = (2,910; 3,960; 19,987; 0,0898; 0,0236; 0,0007; 0,0324; 0,0004; 0,0260; 0,0356)$.

2.1.10. Equilíbrio Químico (EQ)

O problema apresentado anteriormente pode ser reduzido para um sistema de cinco equações e cinco variáveis conforme apresentado a seguir:

$$\left\{ \begin{array}{l} f_1 = x_1x_2 + x_1 - 3x_5 \\ f_2 = 2x_1x_2 + x_1 + x_2x_3^2 + R_8x_2 - Rx_5 + 2R_{10}x_2^2 + R_7x_2x_3 + R_9x_2x_4 \\ f_3 = 2x_2x_3^2 + 2R_5x_3^2 - 8x_5 + R_6x_3 + R_7x_2x_3 \\ f_4 = R_9x_2x_4 + 2x_4^2 - 4Rx_5 \\ f_5 = x_1(x_2 + 1) + R_{10}x_2^2 + x_2x_3^2 + R_8x_2 + R_5x_3^2 + x_4^2 + 1 + R_6x_3 - R_7x_2x_3 + R_9x_2x_4 \end{array} \right.$$

onde:

$$\left\{ \begin{array}{l} R = 10 \\ K_5 = 0,193 \\ K_6 = \frac{0,002597}{\sqrt{40}} \\ K_7 = \frac{0,003448}{\sqrt{40}} \\ K_8 = \frac{0,00001799}{\sqrt{40}} \\ K_9 = \frac{0,0002155}{\sqrt{40}} \\ K_{10} = \frac{0,00003846}{\sqrt{40}} \end{array} \right.$$

Dimensão: 5.

Domínio: [0, 100].

Mínimo global: $f(x^*) = 0$, quando $x^* = (0,0031; 34,59; 0,0650; 0,8594; 0,0369)$.

2.1.11. Potencial de Lennard-Jones (LJ)

O problema de otimização dos aglomerados de Lennard-Jones (LJ) (FLOUDAS; PARDALOS, 1999) é um problema de dois corpos usado para simular o aglomerado de gases raros com átomos pesados, como por exemplo o argônio, o xenônio e o criptônio. A energia potencial LJ entre dois átomos com uma distância r é dada pela equação: $LJ(r) = \frac{1}{r^{12}} - \frac{2}{r^6}$.

A energia potencial total V de um grupo de N átomos é dado por: $V = \sum_{i < j} LJ(r_{ij})$ onde r_{ij} é a distância Euclideana entre os átomos i e j .

A região de busca Ω é dada por: $\Omega = \cup_i \Omega_i$, onde:

$$\Omega_i = \left\{ \begin{array}{l} [0,4], \text{ se } i = 1 \text{ ou } 2 \\ [0, \pi], \text{ se } i = 3 \\ \left[-4 - \frac{1}{4} \left[\frac{i-4}{3} \right], 4 + \frac{1}{4} \left[\frac{i-4}{3} \right] \right], \text{ se } i = 4, \dots, 3N - 6. \end{array} \right.$$

Esse problema apresenta um mínimo global dependente do número de átomos. A seguir são mostrados os mínimos globais para os casos testados neste trabalho:

$N = 3$: Energia = - 3,000000

$N = 4$: Energia = - 6,000000

$N = 5$: Energia = - 9,103852

$N = 6$: Energia = - 12,7121

2.1.12. Problema do Fígado de Porco (PFP)

O problema do fígado de porco (ALI; STOREY; TORN, 1996) é um problema de estimação de parâmetros formulado sob uma ótica de otimização. A função logarítmica de verossimilhança a ser maximizada está representada pela equação:

$$f(k_m, \epsilon^2, \sigma_1, \dots, \sigma_5, V_{\max_1}, \dots, V_{\max_5}) = - \sum_{i=1}^5 \left(n_i \ln \sigma_i + \frac{R_i^2}{2\sigma_i^2} \right)$$

Os dados são analisados utilizando a técnica estatística Bayesiana, sendo o erro de $\ln V_{ij}$ normalmente distribuído. Com isso tem-se a equação:

$$R_i^2(V_{\max_1}, k_m, \epsilon^2) = \sum_{j=1}^{n_{hi}} (\ln V_{ij} - \ln \widehat{V}_{ij})^2$$

Os valores de \widehat{V}_{ij} são encontrados ao resolver o sistema de equações não-lineares mostrado na equação:

$$\frac{V_{\max_i}}{\widehat{V}_{ij}} \left(1 - \frac{\epsilon^2 \frac{V_{\max_i}}{2F_i k_m}}{\left(1 + \frac{\widehat{V}_{ij}/F_i k_m}{\exp(V_{\max_i} - \widehat{V}_{ij})/F_i k_m} - 1 \right)^2} \right) = \frac{k_m}{\widehat{c}_{ij}} + 1$$

$$\widehat{c}_{ij} = \frac{c_{ij} - \overline{c}_{ij}}{\ln(c_{ij}/\overline{c}_{ij})}$$

onde: V_{ij} são as medidas experimentais das taxas de eliminação para o j-ésimo experimento no i-ésimo fígado de porco (pig liver), k_m é a constante de Michaelis para a interação da enzima-substrato, ϵ^2 é o coeficiente de variação para as propriedades dos capilares, σ_i é o desvio padrão para cada fígado de porco i , V_{\max_i} é a taxa máxima de eliminação do fígado inteiro, n_i é o número de experimento no i-ésimo fígado de porco, F_i , c_{ij} e \overline{c}_{ij} são as constantes conhecidas associadas com o j-ésimo experimento para o i-ésimo fígado de porco.

Este problema apresenta $f^* = 59,8403$ como seu mínimo global quando $k_m = 0,2247$ e $\epsilon^2 = 0,1689$ e a estimação das demais variáveis são: $(\sigma_1, \dots, \sigma_5, V_{\max_1}, \dots, V_{\max_5}) = (0,035; 0,091; 0,051; 0,060; 0,031; 0,398; 0,428; 0,604; 0,457; 0,738)$.

2.1.13. Estimação de Parâmetros de um Experimento com Fertilizante (ExpF)

A ideia da estimação de parâmetros de um experimento com fertilizante (HARTLEY, 1961) é ajustar os dados experimentais em uma lei exponencial, descrita através da equação $y = k_1 + k_2 \exp k_3 x$, onde a variável y representa um campo de trigo com seis taxas de aplicação de fertilizante x , em escala codificada. Os parâmetros a serem estimados empregando um problema de mínimos quadrados são representados pelas variáveis k_1 , k_2 e k_3 .

A região de busca deste problema é $k_1 = [-10^5, 10^5]$, $k_2 = [-10^5, 10^5]$ e $k_3 = [-10^2, 10^2]$ e apresenta $f^* = 11390,23$ como mínimo global quando $k_1 = 523,3$, $k_2 = -159,9$ e $k_3 = -0,1997$.

2.1.14. Ajuste Exponencial com Quatro Parâmetros (AExp)

O problema do ajuste exponencial com quatro parâmetros (NIELSEN, 2000) representado pela equação abaixo foi formulado para simular um problema exponencial de ajuste de dados a partir de dados experimentais da variável y_i .

$$f_i(x) = y_i - (x_3 \exp x_1 t_i + x_4 \exp x_2 t_i)$$
$$t_i = 0,02i$$

A região de busca deste problema é $[-100,100]$ e apresenta $f(x^*) = 0,005$ como mínimo global quando $x^* = (-4; -5; 4; -4)$.

2.2. Descrição Das Variantes Utilizadas Do Método De Colisão De Partículas

2.2.1. PCA Canônico

Este algoritmo foi inspirado nas interações de colisão nuclear, especialmente o espalhamento e a absorção. A interação espalhamento ocorre quando um nêutron é espalhado através da colisão com um núcleo alvo. Já a absorção ocorre quando o nêutron é absorvido pelo núcleo alvo.

Primeiramente, deve-se escolher aleatoriamente uma configuração inicial e então ocorrerá a sua modificação a fim de gerar uma configuração nova. A função objetivo das

duas configurações é comparada de forma a determinar se a nova configuração apresenta melhor resultado. Em caso positivo, esta configuração terá sua vizinhança avaliada aleatoriamente por meio da função “*Exploitation*”. Caso contrário, o algoritmo seguirá realizando uma busca aleatória de uma nova configuração. Este passo é chamado de “*Scattering*”. O PCA pode ser considerado um algoritmo de Metropolis (METROPOLIS ET AL., 1953), já que uma nova solução de pior qualidade que a melhor corrente pode ser aceita com uma certa probabilidade, com o intuito de evitar a convergência para mínimos locais. A seguir é apresentado o pseudocódigo do PCA em sua versão canônica.

Algoritmo 1 Pseudocódigo do algoritmo de colisão de partículas canônico.

```
1: OldConfig ← random(range)
2: enquanto critério de parada não for satisfeito faça
3:   NewConfig ← random(range)
4:   se  $f(\textit{NewConfig}) > f(\textit{OldConfig})$  então
5:     OldConfig ← NewConfig
6:     Exploitation() - busca em uma vizinhança da solução atual -
7:   senão
8:     Scattering() - gera um novo candidato de forma aleatória -
9:   fim se
10: fim enquanto
```

2.2.2. PCA com Hooke-Jeeves – Hooke-Jeeves PCA (HJPCA)

O método de Hooke-Jeeves (HOOKE; JEEVES, 1961) é um método determinístico de busca direta que realiza dois tipos de busca: uma busca exploratória e uma busca de padrões.

O algoritmo é iniciado em um ponto x_0 e as soluções nas direções positiva e negativa são comparadas para encontrar o melhor ponto x_1 . Caso a solução encontrada x_1 seja melhor do que x_0 , x_1 passará a ser utilizado como base para a próxima iteração. Caso não seja possível encontrar nenhum ponto melhor do que x_0 , é reduzido o passo, empregando um parâmetro α . Em seguida, a busca feita utilizando este passo reduzido. A cada iteração, o passo pode ser reduzido até atingir um valor mínimo de ϵ .

No HJPCA, a etapa de “*Exploitation*” é realizada com o método de Hooke-Jeeves. Ou seja, toda vez que a solução da nova configuração apresentar um resultado melhor do que a configuração antiga, este resultado será utilizado como ponto inicial para o método determinístico escolhido. Este algoritmo será responsável por realizar uma busca local por uma solução ainda melhor que a corrente.

2.2.3. PCA com Seção de Choque - Cross Section PCA (CSPCA)

O algoritmo é iniciado com a geração de NP pontos no espaço de busca, que irão caracterizar as suas regiões de maneira analógica à representação da seção de choque em um reator nuclear. Estes pontos são gerados a partir do gerador Sobol (SOBOL, 1967) e são ordenados de acordo com os seus valores da função objetivo.

A solução tentativa será gerada pela equação abaixo inspirada pelo mecanismo de mutação encontrado no método de evolução diferencial (STORN; PRICE, 1997), onde *atual* é a solução corrente e x_{r1} e x_{r2} são pontos do espaço de busca escolhidos aleatoriamente dentre aqueles gerados inicialmente pela sequência de Sobol. Se a solução *teste* obtida for melhor que *atual*, procede-se com a busca local pelo Hooke-Jeeves tal qual no HJPCA.

$$teste \leftarrow atual + F(x_{r1} - x_{r2})$$

No CSPCA, a etapa de “*Scattering*” se dá pela probabilidade de escolha da melhor solução pela técnica “*rank weighting*” (HAUPT; HAUPT, 2004).

3. RESULTADOS

Os métodos de otimização utilizados neste trabalho foram implementados em linguagem C. Todos os experimentos foram realizados em um PC com processador Intel® Core™ i5-2410M CPU@2.30GHz, com 8,00 GB de RAM, rodando em um ambiente Windows 10 Home. Os códigos foram compilados com o Dev-C++. Para a parte estocástica relativo a determinação de novas configurações, foi utilizado o gerador pseudoaleatório denominado Mersenne Twister (MATSUMOTO; NISHIMURA, 1998).

Os métodos de otimização foram rodados cem vezes utilizando como critério de parada a equação abaixo, onde $f(x^*)$ é o mínimo global conhecido, $f(x)$ é o melhor valor até a iteração corrente, o coeficiente $\epsilon_1 = 10^{-4}$ corresponde ao erro relativo e $\epsilon_2 = 10^{-6}$ corresponde ao erro absoluto (SIARRY ET AL., 1997).

$$|f(x^*) - f(x)| \leq \epsilon_1 |f(x^*)| + \epsilon_2$$

No algoritmo de Hooke-Jeeves, presente em ambos os métodos, utilizou-se $\epsilon = 10^{-6}$ e $\alpha = 0,8$.

Os resultados relacionados à taxa percentual de sucesso dos algoritmos são apresentados na Tabela 1.

Tabela 1. Taxa de sucesso dos métodos de otimização.

	HJPC A	CSPCA		HJPCA	CSPCA
Branin	100%	100%	Shekel-07	100%	100%
Easom	100%	100%	Shekel-10	100%	100%
Goldstein- Price	100%	100%	EQ10	95%	8%
Rosenbrock- 02	100%	100%	EQ	77%	100%
Rosenbrock- 05	100%	100%	LJ-3	100%	100%
Rosenbrock- 10	100%	100%	LJ-4	100%	100%
Zakharov-05	100%	100%	LJ-5	100%	100%
Zakharov-10	100%	100%	LJ-6	100%	100%
Shubert	100%	100%	PFP	100%	100%
Hartmman- 03	100%	100%	ExpF	100%	100%
Hartmman- 06	100%	100%	AExp	99%	100%
Shekel-05	100%	100%			

4. DISCUSSÃO

Na suíte de testes de funções, que, na tabela, vai da função Branin até a Shekel-10, ambos os algoritmos apresentaram 100% de taxa de sucesso para todas as funções, a despeito desse conjunto ser bastante desafiador (HEDAR; FUKUSHIMA, 2006).

Os problemas “*Equilíbrio químico*” (EQ) e “*Equilíbrio químico com 10 variáveis*” (EQ10) são *benchmarks* para resolução de sistemas não-lineares. O HJPCA obteve o seu pior desempenho no primeiro problema, enquanto o CSPCA obteve o seu pior desempenho no segundo problema.

O problema “*Ajuste exponencial com quatro parâmetros*” (AExp) é um problema inverso de estimação de parâmetros que se utiliza da otimização como mecanismo de busca na sua solução. Para este problema, o HJPCA obteve um resultado inferior ao do CSPCA, porém bem próximo (99% contra 100%).

Ao analisar os resultados, o CSPCA se mostra aparentemente mais robusto. Entretanto, tal fato se torna duvidoso devido à grande margem apresentada para o problema “*Equilíbrio químico com 10 variáveis*”. Sendo assim, tornou-se necessária a utilização de métodos estatísticos não-paramétricos para determinar se existe um método mais eficiente do que o outro.

O teste dos postos sinalizados de Wilcoxon é a alternativa não-paramétrica ao teste t-pareado e serve para indicar estatisticamente se duas amostras representam duas diferentes populações (MCDONALD, 2014).

A hipótese nula (H_0) afirma que a diferença mediana entre os pares é zero, ou seja, que não existe diferenças significativas entre as amostras.

Primeiramente deve-se calcular a diferença de resultados entre as amostras e seu valor absoluto deve ser ordenado do menor para o maior. Com isso, assume-se posto um àquele com a menor diferença absoluta. Caso haja diferenças iguais, deve-se considerar o valor médio entre os postos. Caso a primeira amostra tenha sido melhor do que a segunda, o posto adquirirá o sinal positivo (R^+), caso contrário sinal negativo (R^-). Soma-se todos os postos com sinais positivos e os postos com sinais negativos. Caso a diferença entre as amostras seja zero, metade do valor de seu posto deve ser somado a R^+ e metade a R^- . Chama-se de T o menor valor entre R^+ e R^- . Caso T seja menor ou igual ao valor de distribuição de Wilcoxon para N conjuntos de dados, indicará que estatisticamente existem diferenças significativas entre os métodos de otimização empregados nos testes.

Ao realizar o teste dos postos sinalizados de Wilcoxon, foi obtido o valor $T = 128$. O valor crítico encontrado na tabela é igual a 73. Com isso pode-se concluir que existem evidências estatísticas suficientes para sugerir que não há diferenças significativas entre os métodos HJPCA e CSPCA.

5. CONSIDERAÇÕES FINAIS

Este trabalho apresentou duas variantes do algoritmo de colisão de partículas: o algoritmo de colisão de partículas com Hooke-Jeeves (HJPCA) e o algoritmo de colisão de partículas com seção de choque (CSPCA). Ambos algoritmos foram utilizados para resolver uma série de problemas teste a fim de determinar suas taxas de sucesso em cada problema. Com os resultados obtidos, tornou-se possível utilizar da estatística não-paramétrica para determinar se uma variante é mais robusta estatisticamente do que a outra. A estatística não-paramétrica precisou ser usada para evitar um resultado erroneamente tendencioso a um dos métodos, visto que nos três problemas onde ambos não obtêm 100% de sucesso, o CSPCA vence em dois e o HJPCA em um; porém com grande margem (95% a 8%). Após os testes, constatou-se que ambos os métodos são estatisticamente iguais no que diz respeito à eficiência em resolução de problemas.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro da Fundação Carlos Chagas de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Código de Financiamento 001).

REFERÊNCIAS

- ALI, M. M., STOREY, C., TORN, A. **Application of some recent stochastic global optimization algorithms to practical problems**. TUCS Technical Report No 47, 1996.
- DEMSAR, J. **Statistical comparisons of classifiers over multiple data sets**. Journal of Machine Learning, v. 7, p. 1-30, 2006.
- DIXON, L., SZEGO, G. **Towards global optimization 2**. North-Holland Pub. Co, 1978.
- DUDERSTADT, J. J.; HAMILTON, L. J. **Nuclear reactor analysis**. Wiley New York, 1976.
- FLOUDAS, C., PARDALOS, P. **Handbook of test problems in local and global optimization**. Springer, 1999.
- HARTLEY, H. O. **The modified gauss-newton method for the fitting of non-linear regression functions by least squares**. Technometrics, v. 3, n. 2, p. 269–280, 1961.
- HAUPT, R. L., HAUPT, S. E. **Practical genetic algorithms**. John Wiley & Sons, Inc, 2004.
- HEDAR, A. R., FUKUSHIMA, M. **Tabu search directed by direct search methods for nonlinear global optimization**. European Journal of Operational Research, v. 170, p. 329-346, 2006.
- HOOKE, R., JEEVES, T. A. **Direct search solution of numerical and statistical problems**. Journal of the ACM, v. 8, n. 2, p. 212-229, 1961.
- LEVY, A. V., MONTALVO, A. **The tunneling algorithm for the global optimization of functions**. Journal Sci. Stat. Comput., v. 6, p. 15-29, 1985.
- MATSUMOTO, M., NISHIMURA, T. **Mersenne Twister: a 623-dimensionally equidistributed uniform pseudo-random number generator**. ACM Transactions on Modeling and Computer Simulation (TOMACS), v. 8, n. 1, p. 3–30, 1998.
- MCDONALD, J. H. **Handbook of Biological Statistics**. Sparky House Publishing, 2014.
- MEINTJES, K., MORGAN, A. P. **Chemical equilibrium systems as numerical test problems**. ACM Trans. Math. Softw., v. 16, n. 2, p. 143-151, 1990.
- METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H., TELLER, E. **Equations of state calculations by fast computing machines**. J. Chem. Phys., vol 21, 1087-1092.

MICHALEWICZ, Z. **Genetic algorithms + data structures = evolution programs.** Springer Verlag, 1996.

MORE, J., GARBOW, B., HILLSTROM, K. **Testing unconstrained optimization software.** ACM Trans. Math. Software, v. 7, p. 17-41, 1943.

NIELSEN, H. **UCTP - Test problems for unconstrained optimization.** Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 2000.

RIOS-COELHO, A. C., SACCO, W. F., HENDERSON, N. **A Metropolis algorithm combined with Hooke-Jeeves local search method applied to global optimization.** Applied Mathematics and Computation, v. 217, p. 843-853, 2010.

SACCO, W. F., OLIVEIRA, C. R. E. **New Stochastic Optimization Algorithm based on Particle Collision.** Transactions of the American Nuclear Society, 2005.

SACCO, W. F., OLIVEIRA, C. R. E., PEREIRA, C. M. N. A. **Two stochastic optimization algorithms applied to nuclear reactor core design.** Progress in Nuclear Energy, v. 58, n. 6, p. 525-539, 2006.

SACCO, W. F., RIOS-COELHO, A. C. **A new Metropolis optimisation method, the cross-section particle collision algorithm: Some preliminary results.** International Journal of Nuclear Energy, Science and Technology, v. 10, p. 59-71, 2016.

SIARRY, P., BERTHIAU, G., DURDIN, F., HAUSSY, J. **Enhanced simulated annealing for globally minimizing functions of many-continuous variables.** ACM Transactions on Mathematical Software (TOMS), v. 23, n. 2, p. 209–228, 1997.

SOBOL, I. **On the distribution of points in a cube and the approximate evaluation of integrals.** USSR Computational Mathematics and Mathematical Physics, v. 7, n. 4, p. 86-112, 1967.

STORN, R., PRICE, K. **Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces.** Journal of Global Optimization, v. 11, n. 4, p. 341-359, 1997.

WACKERLY, D., MENDENHALL, W., SCHEAFFER, R. **Mathematical Statistics with Applications.** Cengage Learning, 2014.

WOLPERT, D., MACREADY, W. **No free lunch theorems for optimization.** IEEE Transactions on Evolutionary Computation, v. 1, n. 1, p. 67-82, 1997.