INVESTIGAÇÃO SOBRE MUTAÇÕES q-GAUSSIANAS NA ESTRATÉGIA EVOLUTIVA COM ADAPTAÇÃO DA MATRIZ DE COVARIÂNCIA

Michel M. Santos, Renato Tinós

Departamento de Computação e Matemática - FFCLRP-USP

michelmozinho@gmail.com, rtinos@ffclrp.usp.br

Resumo – A Estratégia Evolutiva com Adaptação da Matriz de Covariância (CMA-ES) tem se mostrado como uma ótima opção em problemas de otimização contínua, sendo uma alternativa aos métodos baseados em gradiente quando estes apresentam dificuldades, como na otimização de problemas altamente multimodais. O algoritmo CMA-ES típico utiliza mutações com distribuição Gaussiana (G-CMA-ES), sendo que tal distribuição apresenta característica local determinada pelo desvio-padrão (ou intensidade de mutação, no contexto da computação evolutiva). No presente trabalho, é investigado o uso de mutações q-Gaussianas no CMA-ES (Q-CMA-ES). A distribuição q-Gaussiana permite controlar o caráter local ou global das mutações mudando a forma da distribuição através do parâmetro q. Consideramos aqui dois tipos de distribuições isotrópicas (ou esféricas) que diferem entre si pelo modo como o raio da hiperesfera é obtido: na primeira, o raio é uma variável aleatória com distribuição q-Gaussiana (respectivo ao algoritmo Q1-CMA-ES); na segunda, o raio é obtido através da norma de um vetor aleatório cujos elementos correspondem a n variáveis q-Gaussianas independentes (respectivo ao algoritmo Qn-CMA-ES). Ao nosso conhecimento, esta é a primeira vez que a mutação q-Gaussiana neste último formato é empregada em algoritmos evolutivos. A fim de avaliar o efeito da mutação q-Gaussiana no desempenho dos algoritmos, foi conduzida uma série de testes para um conjunto de funções *benchmark*. Os resultados do algoritmo G-CMA-ES foram comparados às versões do Q-CMA-ES para diferentes valores de q fixados. O algoritmo Qn-CMA-ES foi superior ao algoritmo original em alguns problemas para $q = -1$. Já no caso $q = 1$, o algoritmo Qn-CMA-ES é equivalente ao G-CMA-ES e por isso apresenta desempenho semelhante. O outro algoritmo modificado, a versão Q1-CMA-ES, em geral, teve o pior desempenho.

Palavras-chave – CMA-ES, Computação Evolutiva, Distribuição q-Gaussiana, Otimização Contínua.

Abstract – The Evolution Strategy with Covariance Matrix Adaptation (CMA-ES) has been considered a great option for continuous optimization problems, being an alternative to gradient-based methods when these are not appropriated, as in highly multimodal optimization problems. The typical CMA-ES algorithm uses mutations with Gaussian distribution (G-CMA-ES), such distribution has a local characteristic determined by the standard deviation (or mutation strength, in evolutionary computation context). In this paper, we investigated the use of q-Gaussian mutation in CMA-ES (Q-CMA-ES). The q-Gaussian distribution allows to control the local or global character of the mutations changing the shape of the distribution by means of q parameter. We consider here two types of isotropic (or spherical) distributions which differ each other by how the radius of the hypersphere is obtained: in first one, the radius is a random variable with q-Gaussian distribution (respective to Q1-CMA-ES algorithm); in second one, the radius is obtained by the norm of a random vector whose elements correspond to n independent q-Gaussians (respective to Qn-CMA-ES algorithm). As far we know, this is the first time that the q-Gaussian mutation in the latter format is used in evolutionary algorithms. In order to evaluate the effect of q-Gaussian mutation on the performance of the algorithms, we conducted several tests of optimization for a set of benchmark functions. The results of G-CMA-ES were compared to the results of Q-CMA-ES versions for different values of q fixed. The Qn-CMA-ES was superior to the original algorithm in some problems for $q = -1$. In the case $q = 1$, the Qn-CMA-ES is equivalent to the G-CMA-ES and therefore has similar performance. The other modified algorithm, the version Q1-CMA-ES, in general, had the worst performance.

Keywords – CMA-ES, Evolutionary Computation, q-Gaussian Distribution, Continuous Optimization.

1. INTRODUÇÃO

Nas últimas décadas tem havido um grande interesse em utilizar Algoritmos Evolutivos para a otimização contínua. Os principais atrativos de tais algoritmos em relação às técnicas tradicionais são o uso de populações de soluções e o uso de operadores de transformação estocásticos, ao invés de uma única solução otimizada em série e de modo determinístico. Entre os operadores de transformação, destaca-se o uso do operador de mutação que introduz variações aleatórias nas soluções ocasionando saltos no espaço de busca.

Em problemas de otimização contínuos, ao explorar o espaço de busca, eventualmente, pode ser favorável visitar pontos distantes para escapar de ótimos locais ou para tentar avançar mais rapidamente em direção ao ótimo global. Por outro lado, a prospecção de soluções locais pode ser útil para obter soluções cada vez mais precisas. Boa parte dos algoritmos evolutivos empregam mutações com distribuição Gaussiana, sendo que tal distribuição apresenta caráter local: mais de 99% da densidade de probabilidade concentra-se a uma distância de até 3 desvios-padrão do centro da distribuição. Empregando a distribuição

Gaussiana, a questão da busca local ou global pode ser abordada através do controle de intensidade (ou força) de mutação, que é o mecanismo utilizado para redimensionar a escala da distribuição. Quanto maior a força de mutação, maiores são, em média, os saltos no espaço de busca.

Outro modo de determinar a extensão do espaço visitado é usar outra distribuição de probabilidade para as mutações. Por exemplo em [1], a distribuição de Cauchy é utilizada com o objetivo de aumentar a possibilidade de fuga de ótimos locais em problemas multimodais. Já em [2, 3] é usada a distribuição q-Gaussiana, que possibilita alterar, além da escala, a forma da curva de densidade de probabilidade em função do parâmetro real $q \in (-\infty, 3)$. Assim, a distribuição q-Gaussiana pode apresentar característica local ou global dependendo do valor de q: quando $q \to -\infty$ a distribuição tende a ficar uniforme entre [-1,1]; no outro extremo, tomando $q \to 3$ a distribuição tende a ficar completamente espalhada (*flat*) [4].

Dentre os algoritmos evolutivos utilizados em otimização contínua, a estratégia evolutiva com adaptação da matriz de covariância (CMA-ES) tem se destacado, apresentando desempenho superior em uma série de problemas [5]. O CMA-ES adapta a matriz de covariância das mutações de forma a direcionar a busca para regiões promissoras do espaço de *fitness*. Na mutação, o CMA-ES tipicamente utiliza desvios aleatórios produzidos por distribuições Gaussianas (G-CMA-ES), sendo que, de acordo com o conhecimento dos autores, não foi ainda investigado se outras distribuições podem vir a ser úteis. No presente trabalho, o CMA-ES é alterado para utilizar mutações q-Gaussianas (Q-CMA-ES). Duas versões do algoritmo modificado usando distribuições diferentes são consideradas, sendo os algoritmos propostos comparados ao algoritmo original quanto ao erro da melhor solução em relação ao ótimo global.

As próximas seções estão organizadas conforme se segue: as seções 2 e 3, respectivamente, revisam alguns aspectos das distribuições empregadas e do algoritmo G-CMA-ES; a seção 4 apresenta o algoritmo modificado; experimentos e resultados são expostos e analisados na seção 5; por fim, as conclusões estão na seção 6.

2 DISTRIBUIÇÃO Q-GAUSSIANA

A distribuição q-Gaussiana pode ser obtida maximizando a entropia generalizada por Tsallis [4, 6] dada por $S_q = \frac{1 - \int_{-\infty}^{+\infty} p(x)^q dx}{1-q}$ em que $q \in (-\infty, 3)$ é o índice entrópico. A função densidade de probabilidade resultante é

$$p(x; \bar{\mu}_q, \bar{\sigma}_q) = \frac{\sqrt{B_q}}{A_q} e_q^{-B_q(x-\bar{\mu}_q)^2}, \tag{1}$$

com parâmetros q-média e q-variância definidos, respectivamente, por $\bar{\mu}_q = E_q[X]$ e $\bar{\sigma}_q = E_q[(X-\bar{\mu}_q)^2]$ através da q-esperança

$$E_q[f(X)] = \frac{\int f(x)[p(x)]^q dx}{\int [p(x)]^q dx}. \tag{2}$$

Na Equação 1, o fator de normalização é calculado por $A_q = \int_{-\infty}^{\infty} e_q^{-x^2} dx$, a escala da distribuição é caracterizada por $B_q = [(3-q)\bar{\sigma}_q]^{-1}$ e a função q-exponencial é definida por

$$e_q^y = \begin{cases} [(1+(1-q)y)^{-\frac{1}{q-1}}] & \text{se } 1+(1-q)y \geq 0 \\ 0 & \text{caso contrário} \end{cases}. \tag{3}$$

Ao tomar $\bar{\mu}_q = 0$ e $\bar{\sigma}_q = 1$, obtêm-se a q-Gaussiana padrão cuja curva de densidade de probabilidade $p(x)$ é apresentada, para alguns valores do parâmetros q, na Figura 1.

Muitos problemas de otimização envolvem múltiplas dimensões, neste caso é necessário estender o caso univariado da q-Gaussiana. A alternativa mais simples é usar uma variável aleatória para cada dimensão. No entanto, a distribuição resultante, dependendo do valor de q, pode ficar anisotrópica, isto é, algumas direções são favorecidas com maior probabilidade em relação ao restante do espaço (Figura 2).

Baseado em [7], um vetor aleatório $\boldsymbol{z}$ com distribuição isotrópica (ou esférica) é gerado por meio de um vetor aleatório $\boldsymbol{u}$ com distribuição uniforme na superfície da esfera multidimensional, indicando uma direção, que é multiplicado por uma variável aleatória escalar r, representando um raio aleatório:

$$\boldsymbol{z} \sim r\boldsymbol{u}. \tag{4}$$

Em [3], o raio r da equação 4 foi gerado através de uma distribuição q-Gaussiana, com o intuito de criar uma distribuição isotrópica. Denotamos a distribuição do vetor aleatório gerado com este procedimento como $\boldsymbol{z} \sim \mathcal{Q}$.

Uma outra maneira de gerar r é através da norma euclidiana do vetor composto por n-variáveis aleatórias [8]. No presente trabalho, tal procedimento é proposto para gerar a distribuição q-Gaussiana isotrópica usada nas mutações, sendo que a distribuição do vetor aleatório obtido é denotada como $\boldsymbol{z} \sim \mathcal{Q}^n$. Ressalta-se que neste caso, para $q = 1$, a distribuição resultante é a Gaussiana multivariada usual [8].

As variáveis unidimensionais com distribuição q-Gaussiana podem ser geradas conforme o método em [6]. Já o vetor aleatório $\boldsymbol{u}$ com distribuição uniforme na superfície da esfera multidimensional pode ser gerado tal que $\boldsymbol{u} \sim \frac{\boldsymbol{v}}{||\boldsymbol{v}||}$, em que $\boldsymbol{v}$ tem distribuição Gaussiana n-variada e $||.||$ é a norma euclidiana. A distribuição de $\boldsymbol{z}$ é chamada q-Gaussiana esférica n-dimensional com raio dependente de n ou não conforme o caso.
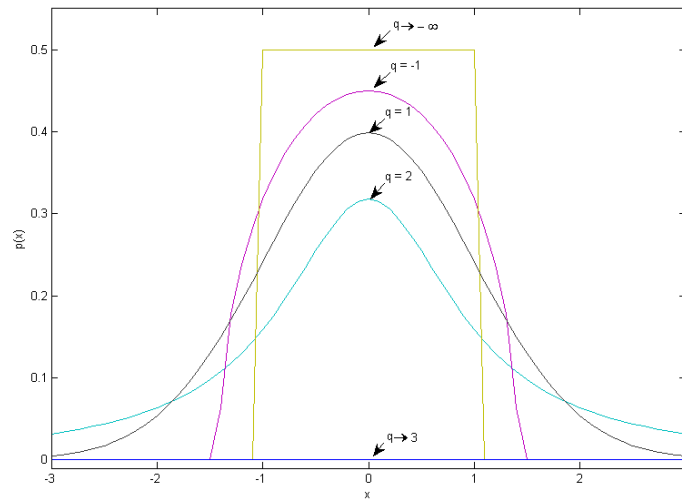
Figura 1: Densidades de probabilidade da distribuição q-Gaussiana para diferentes valores de q e as respectivas equivalências: para $q \to -\infty$ a distribuição q-Gaussiana tende à distribuição uniforme em $[-1, 1]$; para $q = -1$, recupera a distribuição semicírculo de Wigner; a distribuição Gaussiana é reproduzida em $q = 1$; o caso $q = 2$ é a distribuição de Cauchy; se $q \to 3$, a q-Gaussiana tende a uma distribuição completamente espalhada
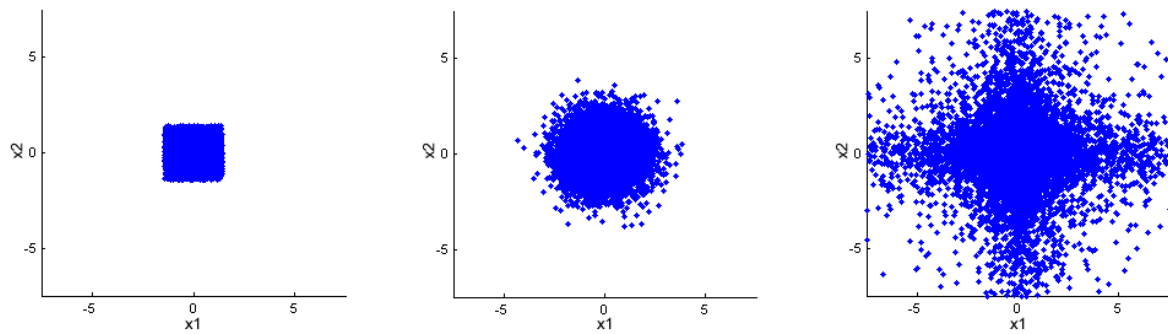

Figura 2: Os pontos correspondem a amostras de duas q-Gaussianas independentes. Note que algumas direções são favorecidas nos casos anisotrópicos: à esquerda, para $q = -1$ as direções diagonais aos eixos coordenados; à direita, quando $q = 2$ as direções ao longo dos eixos coordenados. Ao centro, o caso isotrópico em $q = 1$

3 O ALGORITMO G-CMA-ES

A principal característica do CMA-ES está em adaptar a matriz de covariância usada para gerar as mutações, o que tem o efeito de aproximar a curvatura da função de *fitness*, similar aos métodos Quase-Newton na área de otimização baseada em derivadas [9].

Os principais passos da versão do Q-CMA-ES utilizada aqui são apresentados a seguir:

- **Passo 1 (inicialização)**: uma população é gerada aleatoriamente com probabilidade uniforme no espaço inicial de busca $[A, B]^n$. Baseado nesta população um centroide $\boldsymbol{m}^{(0)}$ é calculado. A intensidade de mutação inicial é $\sigma^{(0)} = (A - B)/2$, sendo os demais parâmetros iniciados com valores fixados e, em especial, ao contador de gerações g é atribuído zero.

- **Passo 2 (mutação)**: são gerados $i = 1, \ldots, \lambda$ novos indivíduos $\boldsymbol{x}_i$ por meio de mutações com distribuição Gaussiana n-variada denotada por $\mathcal{N}(\boldsymbol{m}, \sigma^2 \boldsymbol{C})$ com $\boldsymbol{m}^{(g)}$ sendo o centroide da população anterior:

$$\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}). \tag{5}$$

- **Passo 3 (seleção)**: os indivíduos gerados são avaliados e ordenados segundo o *fitness* obtido. Assim, denota-se $\boldsymbol{x}_{i:\lambda}$ para o i-ésimo indivíduo de menor *fitness* segundo a ordem $f(\boldsymbol{x}_{1:\lambda}) \leq f(\boldsymbol{x}_{2:\lambda}) \leq \ldots \leq f(\boldsymbol{x}_{i:\lambda}) \leq \ldots \leq f(\boldsymbol{x}_{\lambda:\lambda})$. Os μ melhores indivíduos são selecionados para a recombinação.

- **Passo 4 (recombinação)**: na recombinação é calculado um novo centroide através da média ponderada dos μ indivíduos selecionados, atribuindo maior peso aos indivíduos de melhor *fitness*:

$$\boldsymbol{m}^{(g+1)} = \sum_{i=1}^{\mu} w_i \boldsymbol{x}_{i:\lambda}, \tag{6}$$

com $w_1 > \ldots > w_\mu$ sendo os pesos reais positivos e $\sum_{i=1}^{\mu} w_i = 1$.

- **Passo 5 (controle de σ)**: o controle da intensidade global de mutação σ depende do caminho de evolução $\boldsymbol{p}_\sigma$ que representa o vetor resultante do caminho percorrido pelo algoritmo no espaço de busca através dos centroides obtidos nas gerações mais recentes. Para adaptar σ, o tamanho de $\boldsymbol{p}_\sigma$ é comparado ao tamanho esperado sob a hipótese de seleção ao acaso. Este tamanho esperado é obtido calculando-se $E[\|\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})\|]$. Quando a norma de $\boldsymbol{p}_\sigma$ é menor que o valor esperado significa que indivíduos mais próximos do centroide atual foram selecionados e σ deve diminuir para concentrar as mutações. Similarmente, se a norma de $\boldsymbol{p}_\sigma$ é maior que o valor esperado, então σ deve aumentar para expandir as mutações, ou seja,

$$\sigma^{(g+1)} = \sigma^{(g)} \times \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|\boldsymbol{p}_\sigma\|}{E\|\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})\|} - 1\right)\right), \tag{7}$$

em que c_σ e d_σ são constantes reais.

- **Passo 6 (adaptação da matriz de covariância)**: de maneira análoga ao Passo 5, outro caminho de evolução $\boldsymbol{p}_c$ é usado para adaptar a matriz de covariância, além disso o cálculo da covariância amostral é feito de modo ponderado, similar ao cálculo do centroide (Equação 6), atribuindo maior peso aos indivíduos de melhor *fitness*:

$$\boldsymbol{C}^{(g+1)} = (1 - c_1 - c_\mu)\boldsymbol{C}^{(g)} + c_1 \boldsymbol{p}_c \boldsymbol{p}_c^T + c_\mu \sum_{i=1}^{\mu} w_i \left(\frac{\boldsymbol{x}_{1:\lambda}^{(g+1)} - \boldsymbol{m}^{(g)}}{\sigma^{(g)}}\right)\left(\frac{\boldsymbol{x}_{1:\lambda}^{(g+1)} - \boldsymbol{m}^{(g)}}{\sigma^{(g)}}\right)^T, \tag{8}$$

em que c_1, c_μ são constantes reais.

- **Passo 7 (término)**: se alguma condição de término é satisfeita encerre a execução, senão retorne ao Passo 2.

4. O ALGORITMO PROPOSTO: Q-CMA-ES

No algoritmo Q-CMA-ES investigado aqui, a primeira modificação em relação ao algoritmo original consiste na substituição da distribuição normal multivariada (Equação 5), pela q-Gaussiana isotrópica conforme a Equação 4 com raio dependendo de n

$$\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma \mathcal{Q}^n(\boldsymbol{0}, \boldsymbol{C}) \tag{9}$$

ou não

$$\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma \mathcal{Q}(\boldsymbol{0}, \boldsymbol{C}) \tag{10}$$

que dizem respeito às versões específicas do algoritmo, respectivamente, Qn-CMA-ES e Q1-CMA-ES.

A outra modificação visa tornar o controle de intensidade global de mutação robusto às amostras com grande desvio absoluto (*outliers*) o que é característico quando a distribuição q-Gaussiana passa a ter segundo momento infinito, isto é, para $q \geq 5/3$. Neste caso, o principal efeito sobre o controle da intensidade global ocorre na magnitude dos vetores, por exemplo para $q = 2$ os tamanhos esperados de $\|\boldsymbol{p}_\sigma\|$, sob a hipótese de seleção ao acaso e usando as distribuições alternativas, $E[\|\mathcal{Q}(\boldsymbol{0}, \boldsymbol{I})\|]$ e $E[\|\mathcal{Q}^n(\boldsymbol{0}, \boldsymbol{I})\|]$ ambos divergem, já que o primeiro momento da Cauchy é infinito. Assim, a magnitude de $\|\boldsymbol{p}_\sigma\|$ pode assumir valores muito altos, o que levaria σ a rapidamente divergir. Logo é importante limitar o impacto da magnitude dos vetores usados no controle de intensidade de mutação. Um requisito adicional é que as modificações sejam válidas para diversos valores de q. Considerando tais fatores, o tamanho médio $E[\|\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})\|]$ é substituído pelo tamanho mediano $\tilde{p} = \text{Mediana}[\|\mathcal{Q}(\boldsymbol{0}, \boldsymbol{I})\|]$ (ou $\tilde{p} = \text{Mediana}[\|\mathcal{Q}^n(\boldsymbol{0}, \boldsymbol{I})\|]$) e o impacto de $\|\boldsymbol{p}_\sigma\|$ é limitado por $\min(\|\boldsymbol{p}_\sigma\|, a\tilde{p})$, em que $a > 1$ é o parâmetro determinando a magnitude em relação ao tamanho mediano. Assim, o Passo 5 do algoritmo original é substituído por

$$\sigma^{(g+1)} = \sigma^{(g)} \times \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\min(\|\boldsymbol{p}_\sigma\|, a\tilde{p})}{\tilde{p}} - 1\right)\right), \tag{11}$$

em que o valor de $\tilde{p}$ é obtido através de 10.000 amostras Monte Carlo antes do laço principal do algoritmo e o valor adotado para o parâmetro a foi 2.

Tabela 1: Características dos problemas de teste

Prob	Descrição	$f(\boldsymbol{x}^\star)$	Intervalo	Propriedades
f_1	Função esfera deslocada	-450	$[-100, 100]$	U, S
f_2	Problema 1.2 de Schwefel deslocado	-450	$[-100, 100]$	U, NS
f_3	Função elíptica deslocada, rotacionada e altamente condicionada	-450	$[-100, 100]$	U, NS, Ro
f_4	Problema 1.2 de Schwefel deslocado e com ruído	-450	$[-100, 100]$	U, NS, Ru
f_5	Problema 2.6 de Schwefel com ótimo na fronteira	-310	$[-100, 100]$	U, NS
f_6	Função de Rosenbrock deslocada	390	$[-100, 100]$	M, NS
f_7	Função de Griewank rotacionada e sem fronteiras	-180	$[-\infty, \infty]$	M, NS, Ro
f_8	Função de Ackley deslocada, rotacionada e com ótimo global na fronteira	-140	$[-32, 32]$	M, NS, Ro
f_9	Função de Rastrigin deslocada	-330	$[-5, 5]$	M, S
f_{10}	Função de Rastrigin deslocada e rotacionada	-330	$[-5, 5]$	M, NS, Ro
f_{11}	Função de Weierstrass deslocada e rotacionada	90	$[-0.5, 0.5]$	M, NS, Ro

O vetor $\boldsymbol{x}^\star$ corresponde ao mínimo global, o intervalo é referente a cada elemento de $\boldsymbol{x}$ e as abreviações significam: U unimodal, M multimodal, S separável, NS não-separável, Ro rotacionada e Ru com ruído

5 EXPERIMENTOS & RESULTADOS

O algoritmo G-CMA-ES original foi comparado as versões modificadas para usar mutações q-Gaussianas considerando diferentes valores de q. Os algoritmos foram avaliados em experimentos numéricos de otimização em um conjunto de 11 funções f_1-f_{11} para avaliação de desempenho [10]. Algumas das características destes problemas estão listadas na Tabela 1.

Nos experimentos, o valor inicial $\boldsymbol{x}^{(0)}$ foi escolhido uniformemente no intervalo de inicialização $[A, B]^n$ (a busca não foi restrita ao intervalo). A intensidade de mutação inicial foi $\sigma^{(0)} = (B - A)/2$. Já a condição de término foi: atingir o número máximo de interações dado por $10^3 \times n$; ou alcançar erro menor ou igual a 10^{-8}. Os valores para o parâmetro q foram escolhidos conforme a equivalência da q-Gaussiana com outras distribuições, sendo selecionados os casos em que $q = -1$, $q = 1$ e $q = 2$ com equivalência, respetivamente, para as distribuições semicírculo de Wigner, Gaussiana e Cauchy. Outro valor escolhido foi $q = 5/3$, pois para $q \geq 5/3$ a distribuição q-Gaussiana passa a ter segundo momento infinito. As funções de teste foram avaliadas em dimensão $n = 10$ e $n = 30$. Os experimentos empregaram a versão "pura" do G-CMA-ES disponível em http://www.lri.fr/ hansen/cmaesintro.html, que não possui extensões ao algoritmo original como reinício do algoritmo, adaptação ativa da matriz de covariância e controle do intervalo de busca. Cada algoritmo foi executado 25 vezes, sendo que ao término de cada execução do algoritmo, o menor valor de f na execução $f(\boldsymbol{x}_{melhor})$ foi usado para calcular o erro em relação ao mínimo global $f(\boldsymbol{x}^\star) - f(\boldsymbol{x}_{melhor})$. Estatísticas do erro foram computadas e o teste de postos com sinais de Wilcoxon pareado foi aplicado para testar, ao nível de significância $\alpha = 1\%$, a hipótese nula H_0 de que "a diferença entre os resultados pareados tem distribuição com mediana igual a zero" considerando o G-CMA-ES e cada uma das versões do Q-CMA-ES para os valores de q selecionados.

Nas Tabelas 2 e 3, respectivamente, em dimensão $n = 10$ e $n = 30$, são apresentadas estatísticas para o erro $f(\boldsymbol{x}^\star) - f(\boldsymbol{x}_{melhor})$ em 25 rodadas de otimização para cada um dos 11 problemas de teste. No campo "término" consta o número de vezes em que a execução foi encerrada por um dos seguintes motivos: o erro foi pequeno o suficiente para considerar que o algoritmo convergiu para o mínimo global, representado por "G"; ou o número máximo de interações foi atingido, contabilizado em "M". No campo "p-valor", consta o resultado da análise estatística comparando o erro $f(\boldsymbol{x}^\star) - f(\boldsymbol{x}_{melhor})$ entre G-CMA-ES e uma das respectivas versões modificadas. Considerando tal comparação estatística, logo após cada p-valor, entre parênteses, são apresentados os símbolos: "n", se o teste não resultou em diferença significativa; "+", se o algoritmo modificado tem melhor desempenho; "-", se o algoritmo modificado tem pior desempenho. Na última linha o campo "Escore" totaliza para cada algoritmo modificado os resultados dos testes estatísticos usando os valores 1 para "+", -1 para "-" e 0 para "n". Assim o escore fornece uma medida de comparação entre cada um dos algoritmos modificados contra o algoritmo original.

O algoritmo Qn-CMA-ES, para $q = -1$, foi estatisticamente superior ao algoritmo original nos problemas f_8 e f_{11} em dimensão $n = 10$ e no problema f_{11} em dimensão $n = 30$. Além disso, não houve diferença significativa em relação ao G-CMA-ES na maior parte dos outros problemas, o que seria esperado apenas para $q = 1$. Tal desempenho superior do algoritmo modificado não pode ser atribuído a capacidade de escapar de mínimos locais, ainda que as funções f_8 e f_{11} sejam multimodais, pois para $q = -1$ a distribuição resultante apresenta característica local. De outro modo, o resultado pode ser justificado por maior precisão alcançada ao prospectar uma região de mínimo local.

Ainda sobre o Qn-CMA-ES, o caso $q = 1$ apresentou concordância nos resultados experimentais com o algoritmo original em todos os problemas, quer fosse em $n = 10$ ou em $n = 30$, o que é interessante já que as modificações foram além da simples substituição de distribuição. Já para os outros valores de q, as amostras são geradas com maior dispersão, o que causou degradação de desempenho para $q = 5/3$ e, principalmente, para $q = 2$.

Observando apenas os escores, tanto para $n = 10$ como para $n = 30$, fica claro que o algoritmo Q1-CMA-ES teve desempenho bastante inferior ao Qn-CMA-ES. O algoritmo Q1-CMA-ES, em geral, foi inferior ao G-CMA-ES, apresentando desempenho semelhante apenas em poucos problemas.

Tabela 2: Estatísticas do erro obtido em 25 rodadas de otimização em dimensão $n = 10$

Pro.	G-CMA-ES	Q1-CMA-ES				Qn-CMA-ES			
		$q = -1$	$q = 1$	$q = 5/3$	$q = 2$	$q = -1$	$q = 1$	$q = 5/3$	$q = 2$
f_1									
mediana	8.39E-009	8.43E-009	8.21E-009	8.63E-009	8.51E-009	7.90E-009	7.80E-009	8.41E-009	7.14E-003
média	8.27E-009	7.61E-009	8.11E-009	8.42E-009	8.23E-009	7.54E-009	7.70E-009	7.93E-009	1.07E-001
desv pad	± 1.60E-009	± 2.15E-009	± 1.17E-009	± 1.33E-009	± 1.33E-009	± 2.00E-009	± 1.77E-009	± 1.50E-009	± 2.33E-001
mínimo	4.73E-009	3.14E-009	5.87E-009	5.26E-009	5.54E-009	2.20E-009	4.33E-009	4.88E-009	3.13E-007
término	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=0, M=25
p-valor		3.13E-001 (n)	6.19E-001 (n)	8.82E-001 (n)	9.04E-001 (n)	2.31E-001 (n)	4.50E-002 (n)	4.27E-001 (n)	1.23E-005 (-)
f_2									
mediana	8.61E-009	9.44E-009	9.61E-009	9.62E-009	9.40E-009	7.48E-009	8.32E-009	8.73E-009	9.25E-001
média	7.71E-009	9.18E-009	9.43E-009	9.33E-009	9.24E-009	7.41E-009	7.66E-009	8.02E-009	1.39E+001
desv pad	± 2.01E-009	± 6.44E-010	± 5.91E-010	± 7.81E-010	± 5.46E-010	± 1.24E-009	± 1.94E-009	± 1.55E-009	± 4.65E+001
mínimo	2.05E-009	7.67E-009	7.65E-009	6.35E-009	8.24E-009	4.95E-009	2.15E-009	4.72E-009	1.34E-005
término	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=0, M=25
p-valor		2.95E-003 (-)	1.13E-004 (-)	4.46E-004 (-)	8.91E-004 (-)	2.64E-001 (n)	9.04E-001 (n)	6.19E-001 (n)	1.23E-005 (-)
f_3									
mediana	6.77E-009	4.95E+005	1.11E+005	1.06E+005	6.96E+004	7.60E-009	8.39E-009	8.55E-009	7.54E+005
média	7.05E-009	5.19E+005	2.00E+005	1.32E+005	9.57E+004	7.42E-009	8.10E-009	8.23E-009	9.99E+005
desv pad	± 1.83E-009	± 4.83E+005	± 2.23E+005	± 1.12E+005	± 7.20E+004	± 1.79E-009	± 1.42E-009	± 1.66E-009	± 1.10E+006
mínimo	4.38E-009	3.77E+004	3.19E+004	1.06E+004	4.47E+003	3.42E-009	4.38E-009	2.29E-009	7.90E+002
término	G=25, M=0	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=25, M=0	G=25, M=0	G=25, M=0	G=0, M=25
p-valor		1.23E-005 (-)	1.23E-005 (-)	1.23E-005 (-)	1.23E-005 (-)	6.19E-001 (n)	5.44E-002 (n)	4.80E-002 (n)	1.23E-005 (-)
f_4									
mediana	7.14E-009	3.90E-004	1.17E-008	9.37E-009	9.93E-009	7.60E-009	8.55E-009	9.11E-009	4.09E+003
média	7.45E-009	3.06E-002	7.93E-006	1.78E-007	1.40E-004	7.45E-009	7.95E-009	8.41E-009	9.06E+003
desv pad	± 1.96E-009	± 6.99E-002	± 2.78E-005	± 6.32E-007	± 6.99E-004	± 1.71E-009	± 1.82E-009	± 1.83E-009	± 1.26E+004
mínimo	2.60E-009	8.90E-008	4.49E-009	5.95E-009	7.29E-009	4.29E-009	3.27E-009	4.21E-009	2.26E+001
término	G=25, M=0	G=0, M=25	G=12, M=13	G=22, M=3	G=16, M=9	G=25, M=0	G=25, M=0	G=25, M=0	G=0, M=25
p-valor		1.23E-005 (-)	3.62E-005 (-)	1.94E-004 (-)	1.40E-004 (-)	9.89E-001 (n)	3.00E-001 (n)	5.44E-002 (n)	1.23E-005 (-)
f_5									
mediana	8.71E+001	1.06E+003	6.22E+002	5.99E+002	5.80E+002	5.39E+001	1.02E+002	7.00E+001	6.76E+002
média	1.07E+002	1.15E+003	7.70E+002	7.90E+002	6.87E+002	7.62E+001	1.13E+002	1.04E+002	1.24E+003
desv pad	± 9.38E+001	± 7.38E+002	± 6.41E+002	± 6.18E+002	± 6.11E+002	± 6.20E+001	± 1.04E+002	± 1.15E+002	± 1.69E+003
mínimo	1.82E+000	1.00E+002	6.91E+001	3.44E+001	3.38E+001	5.69E+000	2.22E+000	1.08E+000	6.04E+001
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		1.23E-005 (-)	6.45E-005 (-)	2.00E-005 (-)	9.04E-005 (-)	1.50E-001 (n)	9.68E-001 (n)	8.82E-001 (n)	4.07E-005 (-)
Pro.	G-CMA-ES	Q1-CMA-ES				Qn-CMA-ES			
		$q = -1$	$q = 1$	$q = 5/3$	$q = 2$	$q = -1$	$q = 1$	$q = 5/3$	$q = 2$
f_6									
mediana	8.75E-009	1.24E+002	7.41E+000	5.39E+000	3.74E+000	8.81E-009	9.68E-009	1.38E+000	1.49E+004
média	7.90E-001	1.32E+003	7.61E+002	1.07E+002	3.19E+001	1.84E+000	3.22E+000	4.78E+001	5.62E+005
desv pad	± 1.57E+000	± 3.83E+003	± 2.03E+003	± 2.86E+002	± 8.55E+001	± 3.84E+000	± 1.03E+001	± 1.52E+002	± 1.90E+006
mínimo	5.14E-009	1.86E+000	2.37E+000	4.15E-001	3.46E-002	3.62E-009	6.35E-009	9.32E-009	1.05E+001
término	G=17, M=8	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=16, M=9	G=16, M=9	G=2, M=23	G=0, M=25
p-valor		1.77E-005 (-)	3.22E-005 (-)	3.62E-005 (-)	8.09E-005 (-)	8.61E-001 (n)	2.31E-001 (n)	2.16E-004 (-)	1.23E-005 (-)
f_7									
mediana	9.86E-003	9.60E-002	6.64E-002	5.41E-002	7.87E-002	9.86E-003	9.86E-003	9.86E-003	3.74E-001
média	1.06E-002	1.28E-001	9.40E-002	1.03E-001	1.39E-001	1.33E-002	1.18E-002	1.10E-002	5.99E-001
desv pad	± 7.87E-003	± 1.12E-001	± 7.66E-002	± 1.05E-001	± 1.21E-001	± 1.11E-002	± 1.08E-002	± 9.15E-003	± 7.24E-001
mínimo	4.91E-009	2.46E-002	2.46E-002	8.95E-009	1.72E-002	2.88E-009	5.96E-009	6.54E-009	8.22E-004
término	G=6, M=19	G=0, M=25	G=0, M=25	G=1, M=24	G=0, M=25	G=6, M=19	G=7, M=18	G=6, M=19	G=0, M=25
p-valor		1.23E-005 (-)	1.23E-005 (-)	4.07E-005 (-)	1.23E-005 (-)	2.59E-001 (n)	8.51E-001 (n)	9.57E-001 (n)	1.57E-005 (-)
f_8									
mediana	6.05E+001	6.05E+001	6.05E+001	6.05E+001	6.05E+001	**6.02E+001**	6.05E+001	6.06E+001	6.05E+001
média	6.05E+001	6.05E+001	6.05E+001	6.05E+001	6.05E+001	**6.02E+001**	6.05E+001	6.06E+001	6.05E+001
desv pad	± 8.75E-002	± 1.36E-001	± 1.22E-001	± 9.59E-002	± 9.70E-002	**± 2.15E-001**	± 9.94E-002	± 1.07E-001	± 1.09E-001
mínimo	6.03E+001	6.02E+001	6.01E+001	6.03E+001	6.03E+001	**6.00E+001**	6.04E+001	6.02E+001	6.02E+001
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	**G=0, M=25**	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		4.50E-002 (n)	5.44E-002 (n)	2.01E-001 (n)	1.92E-001 (n)	**7.22E-005 (+)**	7.98E-001 (n)	2.88E-001 (n)	3.97E-001 (n)
f_9									
mediana	1.09E+001	1.49E+001	1.29E+001	1.39E+001	1.59E+001	1.39E+001	1.39E+001	1.59E+001	5.17E+001
média	1.22E+001	1.78E+001	1.56E+001	1.54E+001	1.79E+001	1.38E+001	1.36E+001	1.90E+001	5.14E+001
desv pad	± 6.58E+000	± 1.49E+001	± 7.21E+000	± 6.82E+000	± 6.50E+000	± 5.68E+000	± 5.52E+000	± 1.10E+001	± 1.46E+001
mínimo	3.98E+000	2.98E+000	5.97E+000	3.98E+000	7.96E+000	3.98E+000	4.97E+000	4.97E+000	3.10E+001
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		8.50E-002 (n)	2.38E-002 (n)	1.92E-001 (n)	1.80E-003 (-)	2.42E-001 (n)	4.59E-001 (n)	3.24E-002 (n)	1.23E-005 (-)
f_{10}									
mediana	9.95E+000	1.39E+001	1.59E+001	1.39E+001	1.99E+001	1.59E+001	1.29E+001	2.66E+001	5.29E+001
média	1.12E+001	1.48E+001	1.84E+001	1.91E+001	2.26E+001	1.55E+001	1.27E+001	2.44E+001	5.52E+001
desv pad	± 4.85E+000	± 6.87E+000	± 8.61E+000	± 1.38E+001	± 1.09E+001	± 4.95E+000	± 6.01E+000	± 1.09E+001	± 1.60E+001
mínimo	2.98E+000	5.97E+000	6.96E+000	5.97E+000	9.95E+000	6.96E+000	3.98E+000	6.96E+000	3.29E+001
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		2.78E-002 (n)	1.52E-003 (-)	1.36E-003 (-)	4.58E-005 (-)	1.66E-003 (-)	3.10E-001 (n)	1.94E-004 (-)	1.23E-005 (-)
f_{11}									
mediana	4.24E+000	6.41E+000	8.85E+000	1.02E+001	1.04E+001	**2.07E+000**	5.18E+000	1.09E+001	1.04E+001
média	4.98E+000	6.69E+000	8.32E+000	9.97E+000	1.01E+001	**2.18E+000**	5.68E+000	1.08E+001	1.02E+001
desv pad	± 3.02E+000	± 1.73E+000	± 2.37E+000	± 1.18E+000	± 1.24E+000	**± 1.63E+000**	± 4.51E+000	± 8.55E-001	± 1.15E+000
mínimo	9.31E-001	4.36E+000	3.23E+000	6.23E+000	5.05E+000	**4.66E-003**	3.77E+002	8.22E+000	6.05E+000
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	**G=0, M=25**	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		1.86E-002 (n)	6.65E-004 (-)	9.04E-005 (-)	4.57E-005 (-)	**4.46E-004 (+)**	5.81E-001 (n)	1.23E-005 (-)	1.23E-005 (-)
Escore		-6	-8	-8	-9	1	0	-3	-10

Em negrito estão os casos em que, considerando a comparação estatística, o Q-CMA-ES teve melhor desempenho em relação ao G-CMA-ES

Tabela 3: Estatísticas do erro obtido em 25 rodadas de otimização em dimensão $n = 30$

Pro.	G-CMA-ES	Q1-CMA-ES				Qn-CMA-ES			
		$q = -1$	$q = 1$	$q = 5/3$	$q = 2$	$q = -1$	$q = 1$	$q = 5/3$	$q = 2$
f_1									
mediana	9.12E-009	9.28E-009	9.46E-009	9.44E-009	9.35E-009	9.01E-009	8.83E-009	9.14E-009	6.92E+003
média	8.90E-009	9.17E-009	9.37E-009	9.39E-009	9.23E-009	8.69E-009	8.63E-009	9.00E-009	9.59E+003
desv pad	± 1.01E-009	± 7.13E-010	± 5.74E-010	± 5.24E-010	± 6.13E-010	± 1.27E-009	± 1.15E-009	± 6.14E-010	± 7.55E+003
mínimo	5.63E-009	6.85E-009	7.87E-009	8.13E-009	7.64E-009	4.88E-009	5.30E-009	7.32E-009	5.88E+002
término	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=25, M=0	G=0, M=25
p-valor		3.47E-002 (n)	5.78E-002 (n)	4.50E-002 (n)	2.64E-001 (n)	6.00E-001 (n)	7.78E-001 (n)	9.68E-001 (n)	1.23E-005 (-)
f_2									
mediana	8.80E-009	1.78E-001	1.35E-001	2.14E-001	1.97E-001	8.90E-009	9.09E-009	9.51E-009	1.97E+005
média	8.88E-009	2.37E-001	1.59E-001	2.32E-001	2.42E-001	8.80E-009	8.99E-009	9.43E-009	2.27E+005
desv pad	± 8.65E-010	± 2.04E-001	± 1.14E-001	± 1.37E-001	± 1.64E-001	± 1.03E-009	± 6.62E-010	± 3.77E-010	± 1.13E+005
mínimo	7.09E-009	8.09E-002	4.57E-002	6.67E-002	9.00E-002	6.04E-009	7.75E-009	8.52E-009	8.62E+004
término	G=25, M=0	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=25, M=0	G=25, M=0	G=25, M=0	G=0, M=25
p-valor		1.23E-005 (-)	1.23E-005 (-)	1.23E-005 (-)	1.23E-005 (-)	8.61E-001 (n)	7.37E-001 (n)	1.19E-002 (n)	1.23E-005 (-)
f_3									
mediana	1.71E+003	2.73E+006	3.17E+006	3.00E+006	3.47E+006	1.05E+005	2.51E+003	1.95E+005	2.58E+008
média	9.11E+003	3.02E+006	3.09E+006	3.19E+006	3.20E+006	1.44E+005	6.15E+003	2.91E+005	2.88E+008
desv pad	± 1.52E+004	± 1.57E+006	± 1.33E+006	± 1.31E+006	± 1.26E+006	± 1.33E+005	± 7.33E+003	± 2.11E+005	± 2.39E+008
mínimo	1.33E+001	8.89E+005	8.41E+005	7.15E+005	1.02E+006	3.66E+004	5.25E+002	2.85E+004	7.33E+007
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		1.23E-005 (-)	1.23E-005 (-)	1.23E-005 (-)	1.23E-005 (-)	1.23E-005 (-)	8.82E-001 (n)	1.23E-005 (-)	1.23E-005 (-)
f_4									
mediana	1.54E+003	5.09E+003	2.32E+003	2.50E+003	2.95E+003	2.13E+004	9.97E-009	1.06E+003	9.95E+005
média	8.94E+003	6.47E+003	2.86E+003	2.84E+003	3.03E+003	2.42E+004	3.59E+003	2.68E+003	1.05E+006
desv pad	± 1.48E+004	± 3.57E+003	± 1.73E+003	± 1.18E+003	± 1.04E+003	± 2.21E+004	± 7.57E+003	± 3.94E+003	± 5.76E+005
mínimo	8.18E-009	2.01E+003	6.15E+002	1.28E+003	1.26E+003	9.93E-009	7.92E-009	1.02E+001	3.87E+005
término	G=4, M=21	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=1, M=24	G=13, M=12	G=0, M=25	G=0, M=25
p-valor		8.40E-001 (n)	8.61E-001 (n)	4.43E-001 (n)	6.57E-001 (n)	2.14E-002 (n)	1.22E-001 (n)	2.42E-001 (n)	1.23E-005 (-)
f_5									
mediana	1.12E+003	1.68E+004	1.48E+004	1.11E+004	9.17E+003	1.25E+003	7.77E+002	3.02E+003	3.73E+004
média	1.14E+003	1.69E+004	1.46E+004	1.15E+004	9.39E+003	1.33E+003	9.13E+002	3.30E+003	3.51E+004
desv pad	± 6.47E+002	± 4.96E+003	± 4.00E+003	± 3.40E+003	± 2.53E+003	± 6.92E+002	± 5.22E+002	± 1.95E+003	± 9.29E+003
mínimo	2.77E+002	9.65E+003	8.53E+003	6.24E+003	5.52E+003	4.54E+002	3.10E+002	4.17E+002	1.54E+004
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		1.23E-005 (-)	1.23E-005 (-)	1.23E-005 (-)	1.23E-005 (-)	1.22E-001 (n)	2.76E-001 (n)	3.62E-005 (-)	1.23E-005 (-)

Pro.	G-CMA-ES	Q1-CMA-ES				Qn-CMA-ES			
		$q = -1$	$q = 1$	$q = 5/3$	$q = 2$	$q = -1$	$q = 1$	$q = 5/3$	$q = 2$
f_6									
mediana	1.21E+001	4.31E+002	3.44E+002	4.76E+002	3.79E+002	1.65E+001	1.77E+001	3.08E+001	1.11E+009
média	5.20E+001	1.88E+003	1.02E+003	1.93E+003	1.97E+003	1.11E+002	1.80E+002	3.52E+002	7.72E+009
desv pad	± 9.31E+001	± 3.44E+003	± 2.26E+003	± 4.36E+003	± 4.61E+003	± 1.57E+002	± 2.90E+002	± 6.75E+002	± 1.99E+010
mínimo	4.53E+000	2.30E+001	2.42E+001	2.25E+001	2.46E+001	6.73E+000	5.51E+000	1.82E+001	4.53E+007
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		1.26E-004 (-)	5.76E-005 (-)	1.40E-004 (-)	5.13E-005 (-)	1.28E-001 (n)	1.10E-002 (n)	5.35E-003 (-)	1.23E-005 (-)
f_7									
mediana	9.02E-009	9.86E-003	9.86E-003	2.46E-002	2.46E-002	9.24E-009	9.35E-009	9.86E-003	1.85E+002
média	1.38E-003	2.06E-002	1.38E-002	3.68E-002	2.29E-002	1.28E-003	1.28E-003	1.17E-002	2.57E+002
desv pad	± 3.96E-003	± 1.59E-002	± 8.60E-003	± 5.86E-002	± 1.63E-002	± 3.70E-003	± 3.90E-003	± 1.08E-002	± 1.95E+002
mínimo	6.98E-009	3.46E-008	9.86E-003	1.01E-007	4.49E-008	5.80E-009	6.83E-009	8.78E-009	4.99E+001
término	G=22, M=3	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=22, M=3	G=22, M=3	G=8, M=17	G=0, M=25
p-valor		1.77E-005 (-)	2.26E-005 (-)	1.77E-005 (-)	1.77E-005 (-)	7.75E-001 (n)	7.78E-001 (n)	2.40E-004 (-)	1.23E-005 (-)
f_8									
mediana	6.11E+001	6.10E+001	6.11E+001	6.10E+001	6.10E+001	6.10E+001	6.10E+001	6.10E+001	6.10E+001
média	6.11E+001	6.10E+001	6.11E+001	6.10E+001	6.10E+001	6.10E+001	6.10E+001	6.10E+001	6.11E+001
desv pad	± 4.53E-002	± 5.42E-002	± 5.22E-002	± 4.55E-002	± 5.75E-002	± 2.66E-001	± 6.19E-002	± 5.74E-002	± 4.94E-002
mínimo	6.10E+001	6.09E+001	6.09E+001	6.09E+001	6.09E+001	6.01E+001	6.09E+001	6.09E+001	6.09E+001
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		4.22E-002 (n)	5.27E-001 (n)	9.80E-002 (n)	6.53E-002 (n)	7.36E-002 (n)	8.75E-002 (n)	2.53E-001 (n)	2.76E-001 (n)
f_9									
mediana	4.88E+001	8.56E+001	7.56E+001	9.95E+001	1.13E+002	5.37E+001	4.78E+001	8.30E+001	2.90E+002
média	4.97E+001	1.37E+002	1.12E+002	9.76E+001	1.36E+002	5.64E+001	4.68E+001	1.30E+002	2.82E+002
desv pad	± 1.47E+001	± 1.31E+002	± 1.02E+002	± 3.44E+001	± 7.28E+001	± 1.54E+001	± 1.31E+001	± 8.16E+001	± 7.34E+001
mínimo	3.08E+001	2.98E+001	4.18E+001	3.38E+001	4.28E+001	3.28E+001	2.79E+001	3.58E+001	1.20E+002
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		1.26E-004 (-)	1.40E-004 (-)	5.13E-005 (-)	1.23E-005 (-)	2.64E-001 (n)	6.00E-001 (n)	3.22E-003 (-)	1.23E-005 (-)
f_{10}									
mediana	4.88E+001	7.06E+001	6.96E+001	1.17E+002	1.35E+002	4.58E+001	5.17E+001	6.67E+001	4.04E+002
média	5.15E+001	7.64E+001	7.61E+001	1.19E+002	1.42E+002	4.89E+001	4.97E+001	1.15E+002	4.19E+002
desv pad	± 1.48E+001	± 2.33E+001	± 3.78E+001	± 4.77E+001	± 4.95E+001	± 1.26E+001	± 1.08E+001	± 7.35E+001	± 1.34E+002
mínimo	2.79E+001	4.18E+001	3.78E+001	5.17E+001	4.88E+001	2.89E+001	3.08E+001	3.68E+001	2.50E+002
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		8.08E-004 (-)	5.45E-004 (-)	1.77E-005 (-)	1.57E-005 (-)	5.27E-001 (n)	7.16E-001 (n)	4.03E-004 (-)	1.23E-005 (-)
f_{11}									
mediana	4.01E+001	2.59E+001	2.71E+001	4.17E+001	4.21E+001	**6.60E+000**	3.96E+001	4.22E+001	4.16E+001
média	2.68E+001	2.58E+001	2.76E+001	3.83E+001	4.12E+001	**7.80E+000**	2.74E+001	4.19E+001	4.13E+001
desv pad	± 1.60E+001	± 2.92E+000	± 3.76E+000	± 6.39E+000	± 3.69E+000	**± 7.22E+000**	± 1.66E+001	± 1.17E+000	± 1.98E+000
mínimo	3.19E+001	1.99E+001	2.00E+001	2.41E+001	2.42E+001	**2.43E+000**	4.97E+000	3.84E+001	3.73E+001
término	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	G=0, M=25	**G=0, M=25**	G=0, M=25	G=0, M=25	G=0, M=25
p-valor		7.37E-001 (n)	5.27E-001 (n)	9.42E-003 (-)	4.16E-003 (-)	**1.77E-005 (+)**	9.89E-001 (n)	1.57E-003 (-)	5.82E-003 (-)
Escore		-7	-7	-8	-8	0	0	-7	-10

Em negrito está o caso em que, considerando a comparação estatística, o Q-CMA-ES teve melhor desempenho em relação ao G-CMA-ES

6   CONCLUSÃO

Foram apresentadas versões modificadas do G-CMA-ES para usar mutações com distribuição q-Gaussiana. A distribuição $\mathcal{Q}^n$ equivale a distribuição Gaussiana n-variada usual quando $q = 1$ [8], o que justifica, neste caso, o desempenho semelhante entre G-CMA-ES e Qn-CMA-ES. Para outros valores de q, a distribuição resultante é diferente da distribuição usada no G-CMA-ES e, mesmo assim, em vários casos o desempenho não foi muito diferente do algoritmo original. Ao que sabemos, é a primeira vez que a distribuição q-Gaussiana na forma $\mathcal{Q}^n$ é empregada em algoritmos evolutivos, exceto, obviamente, por alguns casos particulares da q-Gaussiana de uso bastante consolidado.

Em ambos os casos do algoritmo modificado, o valor de q foi mantido fixo. O desempenho do Q-CMA-ES poderia ser melhorado se o valor de q fosse adaptado ao longo da execução do algoritmo e os resultados para $q = -1$ apontam nesta direção. Futuros passos deste trabalho visam criar uma versão do Q-CMA-ES que faça auto-adaptação da forma da distribuição ao longo do processo de otimização e para isto o uso da distribuição $\mathcal{Q}^n$ nas mutações é mais adequado. Adaptar a forma da distribuição ao longo da otimização pode ser útil tanto para prospectar uma região de ótimo, como para escapar de ótimos locais, sem depender apenas da intensidade global de mutação σ.

REFERÊNCIAS

[1] X. Yao, Y. Liu and G. Lin. "Evolutionary programming made faster". *IEEE Transactions on Evolutionary Computation*, vol. 2, no. 3, pp. 82–102, 1999.

[2] M. Moret, P. Pascutti, P. Bisch, M. Mundim and K. Mundim. "Classical and quantum conformational analysis using Generalized Genetic Algorithm". *Physica A: Statistical Mechanics and its Applications*, vol. 363, no. 2, pp. 260 – 268, 2006.

[3] R. Tinós and S. Yang. "Use of the q-Gaussian mutation in evolutionary algorithms". *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, pp. 1–27, 2010. 10.1007/s00500-010-0686-8.

[4] D. Prato and C. Tsallis. "Nonextensive foundation of Lévy distributions". *Physical Review E*, vol. 60, no. 2, pp. 2398–2401, August 1999.

[5] S. García, D. Molina, M. Lozano and F. Herrera. "A study on the use of non-parametric tests for analyzing the evolutionary algorithms' behaviour: a case study on the CEC'2005 Special Session on Real Parameter Optimization". *Journal of Heuristics*, vol. 15, no. 6, pp. 617–644, December 2009.

[6] W. Thistleton, J. Marsh, K. Nelson and C. Tsallis. "Generalized Box Muller Method for Generating q-Gaussian Random Deviates". *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4805 –4810, dec. 2007.

[7] A. Obuchowicz. "Multidimensional mutations in evolutionary algorithms based on real-valued representation". *International Journal of Systems Science*, vol. 34, no. 7, pp. 469 – 483, 2003.

[8] N. Hansen, F. Gemperle, A. Auger and P. Koumoutsakos. "When Do Heavy-Tail Distributions Help?" In *Parallel Problem Solving from Nature*, pp. 62–71, 2006.

[9] N. Hansen. "The CMA Evolution Strategy: A Tutorial". Technical report, 2011.

[10] P. N. Suganthan, N. Hansen, J. J. Liang, K. Deb, Y. P. Chen, A. Auger and S. Tiwari. "Problem definitions and evaluation criteria for the CEC 2005 Special Session on Real Parameter Optimization". Technical report, Nanyang Technological University, 2005.