



**UNIVERSIDADE PRESBITERIANA MACKENZIE
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**



**Estratégias Evolutivas com Mutações
Governadas por Distribuições Estáveis**

Agostinho Benigno Monteiro Gutierrez

Orientador: Prof. Dr. Pedro Paulo Balbi de Oliveira

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, como requisito das exigências do exame para obtenção do título de Mestre em Engenharia Elétrica.

São Paulo
2007

Agostinho Benigno Monteiro Gutierrez

**Estratégias evolutivas com mutações
governadas por distribuições estáveis**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, como requisito das exigências do exame para obtenção do título de Mestre em Engenharia Elétrica.

Aprovada em Junho de 2007.

Banca Examinadora

Dr. José Demísio Simões da Silva
Lab. Associado de Comp. e Matemática Aplicada
Instituto Nacional de Pesquisas Espaciais. LAC-INPE

Dr. Nizam Omar
Universidade Presbiteriana Mackenzie

Dr. Pedro Paulo Balbi de Oliveira
Universidade Presbiteriana Mackenzie

São Paulo
2007

Dedicatória

Dedico este trabalho a toda a minha família que sempre acreditou em mim, e em especial à minha querida esposa Marilise Souza de Carvalho Gutierrez, que com muito amor, ternura, companheirismo, amizade e apoio nos momentos mais duros e difíceis de minha vida, me estimula e incentiva a manter a garra necessária para o término desta realização.

*A palavra de Deus ilumina o caminho do homem.
Tu és meu benfeitor:
ensina-me os teus estatutos.
(Salmos 119:68)*

Agradecimentos

Ao amigo Hugo Santana Lima pela paciência em transferir seus conhecimentos adquiridos durante a sua brilhante carreira que muito me serviram para impulsionar neste trabalho.

Às funcionárias Priscilla Imperial Amadco Gomes de Farias, Carla Pegorelli e Sheila Carla de Souza por sua atenção e presteza ao atender aos meus pequenos pedidos.

Aos amigos Francisco Lopes da Silva, Renato Victor Mejias e Simone Erbs da Costa que com extra dedicação, e apoio tornaram concreta as minhas idéias, possibilitando realizar este trabalho com a qualidade esperada.

A minha filha Carolina de Carvalho Gutierrez, que com suas massagens relaxantes me aliviavam da tensão desse trabalho e meu filho Felipe de Andrade Rosa Gutierrez, que sempre quando possível compreendia a minha ausência.

A minha querida sobrinha Érika Ferreira de Carvalho, que com todo amor e paciência típica da sua personalidade colaborou com a sua dedicação e apoio nos momentos importantes.

A todos os professores deste excelente curso, por transmitirem seus conhecimentos de forma clara, objetiva e sem limitações, criando em mim uma nova estrada do saber.

Este trabalho conta com o apoio financeiro da FAPESP, através do projeto de pesquisa Processo 2005/04696-3, e da Wolfram Research, através do Mathematica Academic Grant 1149, e a fase inicial do trabalho se desenvolveu no contexto de projeto de pesquisa financiado pelo MACKPESQUISA, Edital 2004, todos concedidos a meu orientador. A essas instituições, nossos sinceros agradecimentos..

Em especial ao meu orientador professor Dr. Pedro Paulo Balbi de Oliveira que soube compreender as minhas dificuldades, para alcançar em tempo e momento os objetivos necessários. Por acreditar no meu potencial e julgá-lo merecedor da sua dedicação e incentivo. Pela paciência e aconselhamento vitais para a conclusão desta etapa em minha vida.

Sumário

LISTA DE FIGURAS	04
LISTA DE TABELAS	06
LISTA DE EQUAÇÕES	09
RESUMO	10
ABSTRACT	11
CAPÍTULO I	12
INTRODUÇÃO	12
CAPÍTULO II	16
COMPUTAÇÃO EVOLUTIVA	16
2.1. Introdução: Uma visão da inteligência computacional	16
2.2. Elementos básicos da biologia genética	17
2.3. Modalidades de computação evolutiva	18
2.4. Algoritmos evolutivos	20
2.4.1. Codificação de indivíduos: Representação genética	21
2.4.2. Definição da população inicial	22
2.4.3. Soluções e representações	23
2.4.4. Operadores genéticos	23
2.4.4.1. O operador de recombinação (<i>crossover</i>)	24
2.4.4.2. O operador de mutação	25
2.4.4.3. Seleção de indivíduos	29
2.5. Aplicações de computação evolutiva	30
2.6. Conclusões	31
CAPÍTULO III	32
ESTRATÉGIAS EVOLUTIVAS	32
3.1. Introdução	32
3.2. O algoritmo básico das EE – A notação $(\mu/\rho+\lambda)$	33
3.3. Ingredientes para o projeto dos algoritmos das estratégias evolutivas	35
3.3.1. Representação de indivíduos	35
3.3.2. Mutação	36
3.3.3. Recombinação	39

CAPÍTULO IV	43
DISTRIBUIÇÕES ESTÁVEIS E FUNÇÕES DE TESTE	43
4.1. Introdução às distribuições estáveis	43
4.2. Propriedades básicas e de parâmetros	46
4.3. Funções de teste	50
4.3.1. Função de Rastrigin	50
4.3.2. Função vale de Rosenbrock (Função 2 De Jong)	51
4.3.3. Função de Schwefel	52
4.3.4. Função de Griewangk	53
CAPÍTULO V	54
EXPERIMENTOS COM AS FUNÇÕES DE TESTE	54
5.1. Experimentos preliminares realizados	54
5.1.1. Definição dos experimentos	54
5.1.2. Resultado e discussões	57
5.2. Experimentos na função de teste de Rastrigin nas dimensões: 2, 5, 10 e 30	59
5.2.1. Experimentos com função de Rastrigin: 2 dimensões	61
5.2.2. Experimentos com função de Rastrigin: 5 dimensões	62
5.2.3. Experimentos com função de Rastrigin: 10 dimensões	64
5.2.4. Experimentos com função de Rastrigin: 30 dimensões	65
5.2.5. Dados dos experimentos para a distribuição Estável 3	66
5.2.6. Auto-adaptação dos parâmetros estratégicos α e γ em 30 dimensões	67
5.2.7. Auto-adaptação dos parâmetros estratégicos α e γ único em 30 dimensões	69
5.2.8. Conclusões sobre os experimentos com a função de Rastrigin	70
5.3. Experimentos na função de teste de Rosenbrock nas dimensões: 2, 5, 10 e 30	71
5.3.1. Experimentos com função de Rosenbrock: 2 dimensões	71
5.3.2. Experimentos com função de Rosenbrock: 5 dimensões	72
5.3.3. Experimentos com função de Rosenbrock: 10 dimensões	73
5.3.4. Experimentos com função de Rosenbrock: 30 dimensões	75
5.3.5. Dados dos experimentos para a distribuição S-Lévy	76
5.3.6. Auto-adaptação dos parâmetros estratégicos α e γ em 30 dimensões	77
5.3.7. Auto-adaptação com parâmetro estratégico α e γ único em 30 dimensões	78
5.3.8. Conclusões sobre os experimentos com a função de Rosenbrock	79

5.4.	Experimentos na função de teste de Schwefel nas dimensões: 2, 5, 10 e 30	81
5.4.1.	Experimentos com função de Schwefel: 2 dimensões	81
5.4.2.	Experimentos com função de Schwefel: 5 dimensões	82
5.4.3.	Experimentos com função de Schwefel: 10 dimensões	83
5.4.4.	Experimentos com função de Schwefel: 10 dimensões	84
5.4.5.	Dados dos experimentos para a distribuição S-Lévy.....	85
5.4.6.	Auto-adaptação dos parâmetros estratégicos α e γ em 30 dimensões	86
5.4.7.	Auto-adaptação com parâmetro estratégico α e γ único em 30 dimensões	87
5.4.8.	Conclusões sobre os experimentos com a função de Schwefel	88
CAPÍTULO VI		
	CONCLUSÕES	90
	REFERÊNCIAS BIBLIOGRÁFICAS	92

LISTA DE FIGURAS

Figura 2.1:	Metodologia hierárquica na computação natural	19
Figura 2.2:	Ilustração de um cromossomo salientando o locus, os genes, e os alelos	20
Figura 2.3:	Estrutura de um algoritmo evolutivo	21
Figura 2.4:	Distribuição Gaussiana	26
Figura 2.5:	Distribuição de Cauchy	26
Figura 2.6:	Distribuição de Lévy	26
Figura 2.7:	Distribuições estáveis (Gauss: $\alpha=2$; Cauchy: $\alpha=1$; Lévy: $\alpha=0.5$, e demais) [Vicent,05]	28
Figura 3.1:	Procedimento $(\mu/\rho+\lambda)$ - EE	33
Figura 3.2:	Cromossomo EE com parâmetros estratégicos	35
Figura 3.3:	Distribuições normais para uma média constante $\mu=0$	37
Figura 3.4:	Mutação de cromossomos EE [Jacob, 01]	38
Figura 3.5:	Recombinação discreta de dois cromossomos EE com $\rho=\rho_p = \rho_s$	40
Figura 3.6:	Recombinação intermediária de dois cromossomos EE por $\rho = \rho_p = \rho_s$ [Jacob,01]	40
Figura 3.7:	Esquema de recombinação múltipla local e discreta	42
Figura 4.1	Família de distribuições estáveis	44
Figura 4.2:	Família de distribuições de Lévy	47
Figura 4.3:	Família de distribuição distribuições de Cauchy	47
Figura 4.4:	Família de distribuições de Gauss	48
Figura 4.5:	Densidades estáveis dentro de $S(\alpha, 0.5, 1,0; 0)$	49
Figura 4.6:	Função de Rastrigin	50
Figura 4.7:	Função vale de Rosenbrock (Função 2 De Jong)	51
Figura 4.8:	Função de Schwefel	52
Figura 4.9:	Função de Griewang	53

Figura 5.1: Auto adaptividade na função de Rastrigin: distribuição Estável 3, com 70
variação livre e fixa dos parâmetros estratégicos em 30 dimensões

LISTA DE TABELAS

Tabela 2.1:	Modalidades de implementação da inteligência computacional	17
Tabela 2.2:	Correlação da natureza com os elementos dos algoritmos genéticos e seus componentes	18
Tabela 2.3:	As quatro subáreas da computação evolutiva	19
Tabela 2.4:	Comparação simplificada entre EE, AG, PE e PG	20
Tabela 2.5:	Tipos básicos de mutação	25
Tabela 5.1:	Experimentos preliminares com famílias de distribuições estáveis	54
Tabela 5.2:	Resultados dos experimentos preliminares	57
Tabela 5.3:	Correlação entre tipo de distribuição e parâmetro de estabilidade α	59
Tabela 5.4:	Visualização dos dados nos experimentos nas diversas funções	61
Tabela 5.5:	Resultados com a função de Rastrigin: Gauss e Cauchy em 2 dimensões	62
Tabela 5.6:	Resultados com a função de Rastrigin: S-Lévy e Lévy em 2 dimensões ..	62
Tabela 5.7:	Resultados com a função de Rastrigin: Gauss e Cauchy em 5 dimensões	63
Tabela 5.8:	Resultados com a função de Rastrigin: S-Lévy, Lévy e Estável 1 a Estável 4 em 5 dimensões	63
Tabela 5.9:	Resultados com a função de Rastrigin: Gauss e Cauchy em 10 dimensões	64
Tabela 5.10:	Resultados com a função de Rastrigin: S-Lévy, Lévy e Estável 1 a Estável 4 em 10 dimensões	64
Tabela 5.11:	Resultados com a função de Rastrigin: Gauss e Cauchy em 30 dimensões	65
Tabela 5.12:	Resultados com a função de Rastrigin: S-Lévy, Lévy e Estável 1 a Estável 4 em 30 dimensões	65
Tabela 5.13:	Análise da distribuição Estável 3 para 500 a 10000 gerações	66
Tabela 5.14:	Resultados com a função de Rastrigin: Auto-adaptação livre em 30 dimensões	68
Tabela 5.15:	Resultados com a função de Rastrigin: Auto-adaptação fixa nas 30 dimensões de 500 a 1500 gerações	69

Tabela 5.16:	Auto adaptatividade na função de Rastrigin: distribuição Estável 3, com variação livre e fixa dos parâmetros estratégicos com 30 dimensões	66
Tabela 5.17:	Resultados com a função de Rosenbrock: Gauss e Cauchy em 2 dimensões	71
Tabela 5.18:	Resultados com a função de Rosenbrock: S-Lévy e Lévy em 2 dimensões	71
Tabela 5.19:	Resultados com a função de Rosenbrock: Gauss e Cauchy em 5 dimensões	72
Tabela 5.20:	Resultados com a função de Rosenbrock: S-Lévy, Lévy e Estável 1 a Estável 4 em 5 dimensões	69
Tabela 5.21:	Resultados com a função de Rosenbrock: Gauss e Cauchy em 10 dimensões	70
Tabela 5.22:	Resultados com a função de Rosenbrock: S-Lévy, Lévy e Estável 1 a Estável 4 em 10 dimensões	70
Tabela 5.23:	Resultados com a função de Rosenbrock: Gauss e Cauchy em 30 dimensões	71
Tabela 5.24:	Resultados com a função de Rosenbrock: S-Lévy, Lévy e Estável 1 a Estável 4 em 30 dimensões	71
Tabela 5.25:	Resultados com a função de Rosenbrock: S-Lévy de 500 e 1500 gerações	72
Tabela 5.26:	Análise na função de Rosenbrock: Auto-adapt. livre nas n-dim de 500 a 10000 gerações	74
Tabela 5.27:	Análise na função de Rosenbrock: Auto-adapt. fixa nas n-dim. de 500 a 10000 gerações	75
Tabela 5.28:	Comparação na função de Rosenbrock: S-Lévy, Auto-adaptação livre e fixa nas n-dim	77
Tabela 5.29:	Resultados com a função de Schwefel: Gauss e Cauchy em 2 dimensões	Xx
Tabela 5.30:	Resultados com a função de Schwefel: S-Lévy e Lévy em 2 dimensões ..	Xx
Tabela 5.31:	Resultados com a função de Schwefel: Gauss e Cauchy em 5 dimensões	Xx
Tabela 5.32:	Resultados com a função de Schwefel: S-Lévy, Lévy e Estável 1 a Estável 4 em 5 dimensões	Xx
Tabela 5.33:	Resultados com a função de Schwefel: Gauss e Cauchy em 10 dimensões	Xx

Tabela 5.34:	Resultados com a função de Schwefel: S-Lévy, Lévy e Estável 1 a Estável 4 em 10 dim	Xx
Tabela 5.35:	Resultados com a função de Schwefel: Gauss e Cauchy em 30 dimensões	Xx
Tabela 5.36:	Resultados com a função de Schwefel: S-Lévy, Lévy e Estável 1 a Estável 4 em 30 dim	Xx
Tabela 5.37a:	Resultados com a função de Schwefel: S-Lévy de 500 e 1500 gerações ..	Xx
Tabela 5.38:	Resultados com a função de Schwefel: Auto-adapt. livre nas n-dim de 500 a 10000 gerações	Xx
Tabela 5.39:	Resultados com a função de Schwefel: Auto-adapt. fixa nas n-dim. de 500 a 10000 gerações	Xx
Tabela 5.40	Comparação na função de Schwefel: S-Lévy, Auto-adaptação livre e fixa nas n-dim	xx
Tabela 5.41:	Resultados com a função de Griewangk: Gauss e Cauchy em 2 dimensões	Xx
Tabela 5.42:	Resultados com a função de Griewangk: S-Lévy e Lévy em 2 dimensões	Xx
Tabela 5.43:	Resultados com a função de Griewangk: Gauss e Cauchy em 5 dimensões	Xx
Tabela 5.44:	Resultados com a função de Griewangk: S-Lévy, Lévy e Estável 1 a Estável 4 em 5 dimensões	Xx
Tabela 5.45:	Resultados com a função de Griewangk: Gauss e Cauchy em 10 dimensões	Xx
Tabela 5.46:	Resultados com a função de Griewangk: S-Lévy, Lévy e Estável 1 a Estável 4 em 10 dim	Xx
Tabela 5.47:	Resultados com a função de Griewangk: Gauss e Cauchy em 30 dimensões	Xx
Tabela 5.48:	Resultados com a função de Griewangk: S-Lévy, Lévy e Estável 1 a Estável 4 em 30 dim	Xx
Tabela 5.49:	Resultados com a função de Griewangk: S-Lévy de 500 e 1500 gerações	Xx
Tabela 5.50:	Análise na função de Griewangk: Auto-adapt. livre nas n-dim de 500 a 10000 gerações	Xx

LISTA DE EQUAÇÕES

(2.1)	Recombinação aritmética <i>para os indivíduos</i> \mathbf{x}_1 e \mathbf{x}_2 e filho \mathbf{x}'_1	25
(2.2)	Recombinação aritmética <i>para os indivíduos</i> \mathbf{x}_1 e \mathbf{x}_2 e filho \mathbf{x}'_2	25
(2.3)	Forma geral de mutação	27
(2.4)	Mutação que gera um vetor de descendente	27
(2.5)	Função de Densidade de Probabilidade para os vetores de valores reais	28
(3.1)	Vetor \mathbf{g} representa um cromossomo EE	35
(3.2)	Cromossomo EE representado como um vetor em par	35
(3.3)	Representação do cromossomo	37
(3.4)	Operador de mutação ϖ_{mut}	37
(3.5)	Vetor de mutação dos parâmetros objetivos \mathbf{P}_{mut}	37
(3.6)	O cromossomo EE que sofreu mutação no parâmetro objetivo	37
(3.7)	Representação da distribuição normal	38
(3.8)	Parâmetros objetivos e estratégicos com a componente de mutação	38
(3.9)	Operador de recombinação ϖ_{rec}	39
(3.10)	Operador de multicombinação ϖ_{rec}	41
(3.11)	Recombinação em pares r do cromossomo re combinado \mathbf{g}	41
(4.1)	Parâmetros de distribuições estáveis	46
(4.2)	Parâmetro (M) de Zolotarev	48
(4.3)	Função de Rastrigin	50
(4.4)	Função vale de Rosenbrock (Função 2 De Jong)	51
(4.5)	Função de Schwefel	52
(4.6)	Função de Griewangk	53

RESUMO

Usualmente, as estratégias evolutivas utilizam as distribuições Gaussianas para governar as mutações sobre valores reais. Já que na natureza e na matemática existem outros tipos de distribuições, tais como de Cauchy, de Lévy e de S-Lévy, além de uma infinidade de distribuições estáveis, é razoável se pensar em expandir a abordagem tradicional, utilizando-se um algoritmo baseado em outras distribuições existentes, ou mesmo que possibilite a escolha de uma distribuição estável, de forma auto-adaptativa. Esta idéia é aqui ilustrada, no contexto de populações de indivíduos que evoluem em busca do mínimo de uma função de teste (no caso, a função de Rastrigin, vale de Rosenberg, Griewangk e Schwefel em n-dimensões) através de estratégias evolutivas cujas mutações são guiadas por oito tipos específicos de distribuições e de um esquema auto-adaptativo em um subconjunto das distribuições estáveis.

Durante a evolução dos experimentos observa-se uma forte influência da escolha adequada da família de distribuição na correlação da busca do mínimo global na função de teste. Este fato se deve a diversidade utilizada na forma da distribuição: assimétrica e cauda longa (Lévy) e simétrica com vários tipos de cauda nas demais. A escolha do tipo de distribuição ocorre determinando-se adequadamente quatro parâmetros: Índice de estabilidade (α), assimétrico (β), escala (γ) e posição (δ).

A escolha do tipo de distribuição ocorre determinando-se os quatro parâmetros acima que fazem parte do cromossomo que também contém as possíveis coordenadas do ponto de mínimo global que serão mudadas com base distribuição escolhida. Com aplicação desta mutação diferenciada no processo evolutivo chegasse ao mínimo global da função de teste escolhida.

Os resultados indicaram que a utilização conjunta de distribuições estáveis governando as mutações das coordenadas podem acarretar uma melhora de desempenho com respeito à convergência e conseqüente determinação da solução, quando aplicadas sobre funções de teste delimitadas espacialmente.

ABSTRACT

Evolutionary strategies normally use the Gaussian distributions in order to control the mutations over real values. Since there are other kinds of distributions in nature and in mathematics, such as those of Cauchy, Lévy and S-Lévy, in addition to several stable distributions, it seems a natural step to extend the standard approach, by using an algorithm that would be based upon other existing distributions, or that would even allow the choice of a stable distribution in a self-adaptive way. Such an idea is briefly sketched herein, in the context of populations of individuals that evolve towards the minimum of a test function (namely, the n-dimensional Rastrigin, Rosenberg, Griewangk and Schwefel functions) by means of evolutionary strategies, whose mutations are guided by eight types of specific types of distributions and by a self-adaptive scheme over a subset of the possible stable distributions.

During the evolution of the experiment a remarkable influence on the right choice of the distribution family can be noted related to the search for the global minimum of a test function. This is due to the diversity used in the form of distribution: asymmetric and long tale (Lévy) and symmetric with various type of tale on the others. The choice of the type of distribution occurs determining four parameters properly: stability rate, asymmetric, scale and position.

The choice of the type of distribution occurs determining if the four parameters above mentioned are part of the chromosome that also contains the possible coordinates of the global minimum that will be mutated according to the chosen distribution. Having applied this different mutation in the evolutionary process will lead to the global minimum of the chosen test function.

The results indicate that the combined use of stable distribution controlling the mutations of the coordinates can result in a performance improvement regarding the convergence and consequent determination of the solution, when applied to spatially constrained benchmark functions.

CAPÍTULO I

INTRODUÇÃO

A computação evolutiva é formada por algoritmos inspirados na evolução biológica natural [Carrapiço, 01; Castro, 02], os quais podem ser considerados como técnicas computacionais de otimização. Nesses algoritmos, inicialmente uma população de soluções candidatas é gerada por meios aleatórios e, então avaliada de forma a se medir o quão próximo elas estão da solução buscada. Com base nessa avaliação, seleciona-se um subconjunto da população servindo de base para a geração de uma nova, esperando-se que esta população seja formada por um conjunto de soluções melhores do que as que serviram de base para a sua formação. Ao longo das gerações, a população evolui até chegar a soluções satisfatórias. Este processo de busca pode ser aplicado em praticamente qualquer área em que se procure a solução ótima de um problema, uma vez que ele possui menos restrições de aplicabilidade em relação a métodos de busca exatos, que necessitem de informações tais como a derivada da função objetivo.

Na computação evolutiva, pode-se identificar quatro famílias principais de algoritmos: a programação evolutiva, a programação genética, os algoritmos genéticos e as estratégias evolutivas [Eiben e Smith, 03a]. As estratégias evolutivas partiram da idéia de tentar modificações aleatórias nas soluções candidatas de um problema de busca, usualmente representados como vetores de números reais, implementando-se os operadores genéticos de mutação e seleção também utilizados nos outros algoritmos evolutivos. Esse esquema de mutação, tradicionalmente implementado através da mutação Gaussiana (ou Normal) pode ser expandido, como será visto no presente trabalho, para outros tipos de distribuições, como as de Cauchy [Rudolph, 98] e Lévy [Lee e Yao, 04; Rimmer e Nolan, 05], entre outras não nomeadas, todas de natureza *estável*, isto é, em que a resultante global de várias mutações em seqüência possui o mesmo tipo de distribuição das mutações individuais. A distribuição Gaussiana é a mais famosa e tradicional das distribuições estáveis.

Mais formalmente, considerando-se X_k uma variável aleatória com determinada distribuição de probabilidade, e definindo-se a somatória de um conjunto dessas variáveis, se esta nova distribuição resultante também possuir o mesmo tipo de distribuição que as componentes X_k , então, diz-se que a distribuição envolvida é *estável* [Tsallis, 00]. A motivação pelo uso de distribuições estáveis em algoritmos evolutivos é sua ampla ocorrência na natureza e na matemática, inclusive em processos que, até muito recentemente, acreditava-se serem regidos por distribuições Gaussianas [Tsallis, 00].

As estratégias evolutivas tiveram sua origem no mecanismo denominado (1+1), um esquema simples de seleção-mutação que opera em um único indivíduo e gera um único descendente por geração, através de mutação Gaussiana. Esta abordagem evoluiu para o chamado mecanismo ($\mu+1$), no qual uma população de μ indivíduos se recombina para formar um descendente, o qual, após sofrer mutação, substitui o pior elemento da população. Ainda que este mecanismo nunca tenha sido largamente usado, ele permitiu a transição para os mecanismos denominados ($\mu+\lambda$) e (μ,λ), em que, no primeiro caso, os μ pais e os λ filhos convivem, enquanto no segundo, os μ pais “morrem”, deixando apenas os λ filhos “vivos” [Bittencourt, 06]. Os seis passos a seguir ilustram os procedimentos utilizados nas estratégias evolutivas de forma simplificada:

1. Iniciar a população com μ_0 pais;
2. Realizar a operação de recombinação sobre os μ pais para gerar λ filhos;
3. Realizar a operação de mutação sobre os λ filhos;
4. Avaliar os indivíduos da população;
5. Selecionar μ indivíduos para a nova população, incluindo-se os pais – estratégia ($\mu+\lambda$) – ou excluindo-os – estratégia (μ,λ);
6. Caso o critério de término não seja satisfeito, voltar ao passo 2.

O objetivo deste trabalho é avaliar a influência de diferentes distribuições estáveis na mutação dos valores que estão sendo buscados pelo processo evolutivo (os parâmetros *objetivos*), os quais são aqui as coordenadas que representam os mínimos globais de algumas funções de teste n -dimensionais (para n 2, 5, 10 e 30); para tanto, observa-se a evolução de uma população de 100 indivíduos (soluções candidatas do problema de minimização), ao longo de 500 gerações, em 30 execuções distintas. Como o tipo de distribuição estável utilizada é determinada pelos chamados parâmetros *estratégicos*, estuda-se aqui a influência deles sobre o processo, em termos da busca pela obtenção do menor erro entre o valor mínimo da função e o valor representado por uma solução candidata do problema. Foram utilizados vários tipos de distribuições estáveis, e também abordou-se o problema de auto-adaptação na distribuição, dentro do espaço de todas as possíveis.

A seguir descreve-se o conteúdo dos capítulos seguintes neste trabalho. O Capítulo 2 trata de computação evolutiva, fazendo sua ligação com elementos básicos da genética, seus paralelos com a família específica dos algoritmos genéticos, e mencionando aplicações em diversas áreas da tecnologia.

Já o Capítulo 3, aprofundando o anterior, foca especificamente nas estratégias evolutivas, apresentando exemplos de representação de indivíduos para números reais, esclarecendo a noção de parâmetros objetivos e estratégicos, e tratando de aspectos envolvidos nos operadores de mutação, recombinação e seleção.

As distribuições estáveis e suas aplicações são abordadas no Capítulo 4, passando-se por uma breve caracterização sobre seus conceitos e importância. Ainda neste capítulo apresentam-se as funções de teste utilizadas no trabalho: Rastrigin, Rosenbrock, Schwefel e Griewangk.

Os experimentos realizados de computação evolutiva com distribuições estáveis são apresentados e discutidos no Capítulo 5, no contexto da busca do mínimo global das funções de teste mencionadas. Apresentam-se resultados decorrentes de experimentos com 2, 5, 10 e 30 dimensões, verificando-se o efeito do aumento da população e do número de gerações sobre a qualidade dos resultados; para tanto, várias distribuições estáveis são utilizadas, notadamente Gauss, Cauchy e Lévy,

mas também algumas distribuições estáveis não denominadas, e outras resultantes de um mecanismo de auto-adaptação no espaço das distribuições estáveis possíveis.

O Capítulo 6 encerra o texto, apresentando as conclusões do trabalho, e estabelecendo os passos a serem seguidos em sua continuação.

CAPÍTULO II

COMPUTAÇÃO EVOLUTIVA

2.1. Introdução: Uma visão da inteligência computacional

A solução de problemas práticos de grande complexidade requer a construção de sistemas capazes de integrar coerentemente as técnicas e metodologias originárias de diferentes áreas de pesquisa. A inteligência computacional compreende paradigmas computacionais que procuram desenvolver sistemas que apresentem alguma forma de inteligência similar à exibida por determinados sistemas biológicos [Cercone e McCalla, 94; Fogel, 00]. Esta área da ciência objetiva o desenvolvimento de sistemas inteligentes que imitam também aspectos do comportamento humano, tais como: aprendizado, percepção, raciocínio, evolução e adaptação [Anderson, 00].

Um sistema é *computacionalmente inteligente* [Bäck, Hammel e Schwefel, 97] quando exhibe (ou começa a exhibir): adaptabilidade computacional; tolerância computacional a falhas; velocidade de processamento comparável a de processos cognitivos humanos; e taxas de erro que se aproximam do desempenho humano. A inteligência computacional procura implementar essas idéias, englobando quatro diferentes técnicas, como resume a Tabela 2.1.

A computação evolutiva pode realizar a simulação de mecanismos de evolução através do uso de computadores, possibilitando o surgimento de uma nova filosofia de máquinas inteligentes. Considera-se inteligência como a capacidade de um sistema adaptar seu comportamento para atingir seus objetivos numa variedade de situações. Sob este ponto de vista, a inteligência e a evolução são dois conceitos intimamente relacionados. A simulação de processos evolutivos biológicos permite o entendimento de como a evolução guia os seres vivos na direção de um nível maior de inteligência [Atmar, 94; Fogel, 95].

Técnica	Descrição	Inspiração Natural
<i>Computação evolutiva</i>	Formada por algoritmos inspirados na teoria da evolução natural de Darwin [Carrapiço, 01; Castro, 02], tem sido muito aplicada nos problemas de otimização, em especial naqueles em que as técnicas tradicionais não são aplicáveis (como aqueles com funções descontínuas) ou que apresentam desempenho insatisfatório [Klug e Cummings, 00].	<i>Evolução biológica</i>
<i>Lógica nebulosa</i>	Modela o modo aproximado de raciocínio humano, visando desenvolver sistemas capazes de tomar decisões racionais em um ambiente de incerteza e imprecisão, tais como os conceitos de muito, pouco, pequeno, alto, quente, frio, entre outros, fornecendo uma resposta aproximada para uma questão, baseada em um conhecimento que é inexato, incompleto ou não totalmente confiável [Laprade e Torres, 93; Rao e Rao, 93; Almeida e Evsukoff, 03].	<i>Processo lingüístico</i>
<i>Redes neurais artificiais</i>	São modelos computacionais não-lineares, inspiradas na estrutura e operação do cérebro humano, que procuram reproduzir características humanas, tais como: aprendizado, associação, generalização e abstração [Fogel e Porto, 90]. A arquitetura de multicamadas é a mais popular. As propriedades são a sua capacidade de aproximação universal e de aprendizado [Azevedo, Brasil e Oliveira, 00; Haykin, 01; Damazio, Seixas e Soares, 03].	<i>Neurônios biológicos</i>
<i>Sistemas especialistas</i>	São programas computacionais destinados a solucionar problemas em um campo especializado do conhecimento humano. Usam técnicas de inteligência artificial [Russell e Norvig, 04], base de conhecimento e raciocínio inferencial [Cunha e Ribeiro, 87; Chorafas, 88; Levine, Drang e Edelson, 88; Weiss e Kulikowski, 88; Paraíso <i>et al.</i> , 93].	<i>Inferência humana</i>

Tabela 2.1: Modalidades de implementação da inteligência computacional.

2.2. Elementos básicos da biologia genética

Os principais termos empregados na computação evolutiva representam uma analogia entre as entidades biológicas reais, e as entidades computacionais de forma mais simplificada. Os princípios da natureza nos quais os algoritmos evolutivos se inspiram são simples. De acordo com a teoria de Charles Darwin, o princípio da seleção natural privilegia os indivíduos mais aptos, com maior longevidade e, portanto, com maior probabilidade de reprodução. Indivíduos com mais descendentes têm mais chance de perpetuarem seus códigos genéticos nas próximas gerações [Castro, 02]. Tais códigos genéticos constituem a identidade de cada indivíduo e estão representados nos cromossomos. Estes princípios são imitados na construção de algoritmos computacionais que buscam a melhor solução para um determinado problema, por meio da evolução de populações de soluções codificadas e também por cromossomos artificiais.

Nos algoritmos evolutivos um cromossomo é uma estrutura de dados que representa uma das possíveis soluções do espaço de busca do problema. Os cromossomos são então submetidos a um processo evolutivo, que envolve a avaliação, a seleção, a recombinação sexual e a mutação. Após vários ciclos de evolução a população deverá conter indivíduos mais aptos. A analogia entre os algoritmos genéticos e o sistema natural é representada na Tabela 2.2 abaixo:

Natureza	Algoritmos Evolutivos	Componentes dos Alg. Evolutivos
Cromossomo	Palavra binária, vetor, etc.	Representação (definição do indivíduo).
Gene	Característica do problema.	Função de evolução (função de aptidão).
Alelo	Valor da característica.	População.
Locus	Posição na palavra, vetor.	Mecanismo de seleção dos pais.
Genótipo	Estrutura (codificação).	Operadores genéticos (mutação, recombinação, etc).
Fenótipo	Possível solução do problema.	Mecanismo de seleção da próxima população.
Indivíduo	Solução.	Inicialização.
Geração	Ciclo.	Condição de término.

Tabela 2.2: Correlação da natureza com os elementos dos algoritmos genéticos e seus componentes.

As técnicas de computação evolutiva fornecem então um mecanismo de busca adaptativa que se baseia no princípio Darwiniano de reprodução e sobrevivência dos mais aptos [Michalewicz, 99; Fernandes, 03]. Observa-se, no entanto, que nos sistemas biológicos reais, não existem superfícies de adaptação estáticas. O ambiente está em constante mudança, fazendo com que populações estejam em constante evolução em direção aos novos pontos de ótimo.

2.3. Modalidades de computação evolutiva

A computação evolutiva define uma área comum que engloba várias pesquisas que vinham seguindo por caminhos distintos, mas conectadas pelo fato de simularem os vários aspectos da evolução. As técnicas de programação evolutiva, programação genética, algoritmos genéticos e estratégias evolutivas representadas na Figura 2.1 possuem algo fundamental em comum: cada uma delas trata da reprodução, variação aleatória, competição e seleção dos indivíduos de uma população. Estes quatro elementos formam a essência da evolução [Atmar, 94]. Se existir algum método

específico que resolva um dado problema em estudo, as técnicas de computação evolutiva podem ser usadas em associação a eles, podendo surpreender na qualidade dos resultados obtidos [Lewis, 00].

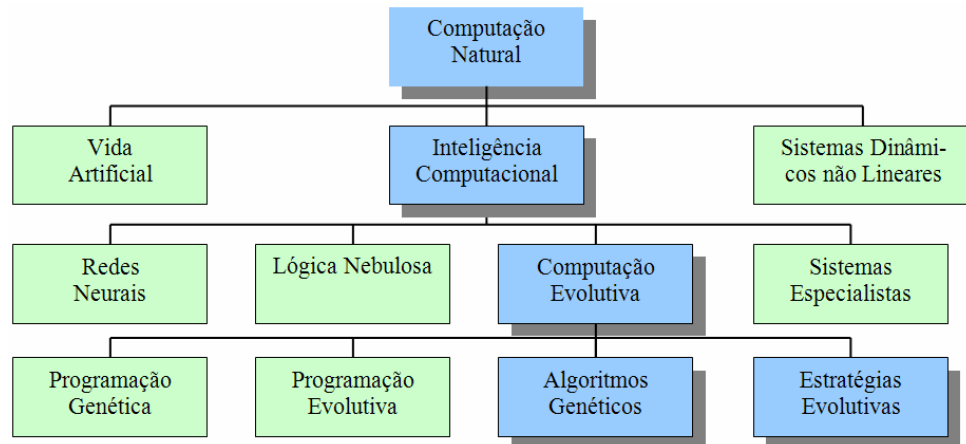


Figura 2.1: Metodologia hierárquica na computação natural.

Dentro da computação evolutiva têm-se quatro subáreas como na Tabela 2.3 a seguir:

Subáreas	Descrição
<i>Programação evolutiva (PE)</i>	Desenvolvida por Lawrence J. Fogel, em 1960, tinha a motivação de constituir um modelo de aprendizado, ou seja, seu objetivo estava em desenvolver modelos comportamentais [Porto, 00; Eiben e Smith, 03a; Castro, 06]. Recentemente expandiu seus objetivos, tornando-se um método genérico de otimização estocástica, como os outros algoritmos evolutivos. Representa as soluções candidatas em alto-nível, o mais próximo possível do problema em questão, e baseia-se no uso apenas de mutação (sem recombinação) e seleção.
<i>Estratégias evolutivas (EE)</i>	Utilizam mutações com alguma distribuição estável para modificar vetores reais e recombinação como operadores essenciais. O operador de seleção é determinístico e o tamanho da população de pais e de filhos pode ser distinto.
<i>Algoritmos genéticos (AG)</i>	Enfatizam a recombinação como o principal operador de busca e aplicam mutação com baixas probabilidades (operador secundário). Operam com representação binária de indivíduos e possui seleção probabilística (usualmente proporcional ao valor da função de adaptação dos indivíduos).
<i>Programação genética (PG)</i>	As estruturas de dados são representadas por árvores (usualmente programas de computador), e utilizam-se os operadores de recombinação e mutação. O processo de seleção tem forma similar à dos algoritmos genéticos, ou seja, probabilística e proporcional à função de adaptação [Kinnear, 00; Banzhaf <i>et al</i> , 02; Carvalho, Braga e Ludermir, 03].

Tabela 2.3: As quatro subáreas da computação evolutiva.

O interesse na busca por sistemas computacionais que explorem possíveis combinações entre os paradigmas acima cresceu de forma expressiva nos últimos tempos. Assim, muitas pesquisas foram

realizadas no sentido de investigar possíveis formas de cooperação entre estes métodos. O principal objetivo é desenvolver sistemas que sejam eficientes, robustos, fáceis de operar e capazes de fornecer soluções de qualidade para problemas complexos. Analisando de forma comparativa cada uma das abordagens mencionadas acima tem-se a Tabela 2.4:

	EE	AG	PE	PG
<i>Representação</i>	Valores reais	Cadeias binárias	Vetores reais	Árvores
<i>Auto-adaptação</i>	Desvio padrão e covariância	Nenhuma	Desvio padrão e coeficiente de correlação	Nenhuma
<i>Função de avaliação</i>	Valor da função objetivo	Valor escalonado da função objetivo	Valor escalonado da função objetivo	Valor escalonado da função objetivo
<i>Mutação</i>	Principal operador	Operador secundário	Único operador	Um dos operadores
<i>Recombinação</i>	Diferentes variações, uso com auto-adaptação	Principal operador	Nenhuma	Um dos operadores
<i>Seleção</i>	Determinística	Probabilística	Probabilística	Probabilística

Tabela 2.4: Comparação simplificada entre EE, AG, PE e PG.

2.4. Algoritmos evolutivos

Os algoritmos evolutivos empregam uma terminologia originada na teoria da evolução natural e da genética. Um indivíduo da população é representado por um único *chromossomo*, o qual contém a *codificação* de uma possível solução do problema. Os cromossomos são implementados como uma sequência de características (possivelmente, formando um vetor), onde cada uma é conhecida como *gene*. Os possíveis valores que um determinado gene pode assumir são denominados *alelos*, conforme a Figura 2.2.

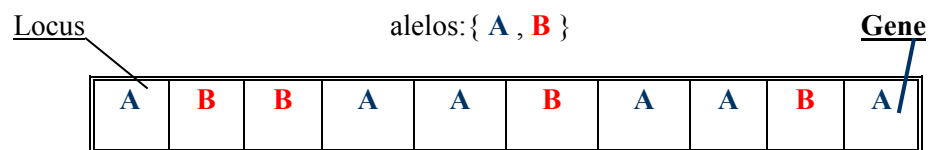


Figura 2.2: Ilustração de um cromossomo salientando o locus, os genes e os alelos.

Pode-se descrever um algoritmo evolutivo como ilustra a Figura 2.3. Durante a iteração t , o algoritmo mantém uma população de soluções potenciais (cromossomos, vetores), $P(t) = \{x_1^t, x_2^t, \dots, x_n^t\}$. Cada solução x_i^t é avaliada e produz uma medida de sua adaptação ou aptidão (*fitness*). Uma nova população (iteração $t + 1$) é então formada, privilegiando a participação dos indivíduos mais adaptados. Alguns membros da nova população passam por alterações através de operadores – usualmente, recombinação e mutação – que permitem formar novas soluções potenciais. Este processo se repete até que um número pré-determinado de iterações seja atingido, ou até que o nível de adaptação esperado seja alcançado.

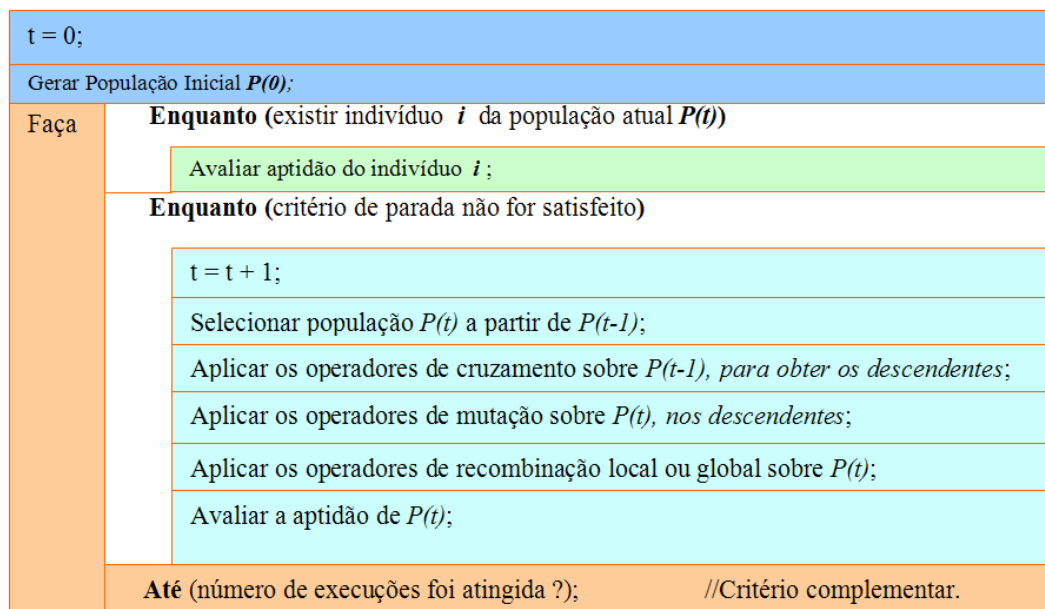


Figura 2.3: Estrutura de um algoritmo evolutivo.

2.4.1. Codificação de indivíduos: Representação genética

Cada indivíduo de uma população representa um candidato em potencial, que pode evoluir para a solução do problema em questão. No algoritmo genético clássico, as soluções candidatas são codificadas em arranjos binários de tamanho fixo, cuja motivação vem da teoria dos esquemas (*schemata theory*), onde mostra-se que essa representação maximiza o paralelismo implícito inerente ao algoritmo genético [Holland, 92].

Contudo, nas diversas aplicações práticas a utilização de codificação binária leva a um desempenho insatisfatório. Nos problemas de otimização numérica com parâmetros reais, os algoritmos evolutivos com representação em ponto flutuante freqüentemente apresentam desempenho superior à codificação binária.

Michalewicz [99] argumenta que o desempenho de um algoritmo evolutivo com codificação binária é pobre quando o espaço de busca tem dimensão elevada. No entanto, o espaço de busca por si só (sem levar em conta a escolha da representação) não determina a eficiência do algoritmo genético. Espaços de busca de dimensão elevada podem às vezes ser explorados eficientemente, enquanto que os espaços de busca de dimensão reduzida podem apresentar dificuldades significativas [Fogel, 94a].

Fica claro, portanto, que a codificação é uma das etapas mais críticas na definição de um algoritmo evolutivo. A definição inadequada da codificação pode levar a problemas de convergência prematura do algoritmo. A estrutura que um cromossomo deve representar numa solução completa deve ser a mais simples possível. Nos problemas de otimização restrita (por exemplo, minimizar $f(x, y)$ sujeita à restrição de que $x^2+y^2=5$) a codificação adotada pode fazer com que os indivíduos modificados por recombinação e/ou mutação sejam inválidos. Nesses casos, cuidados especiais devem ser tomados na definição da codificação e/ou dos operadores.

2.4.2. Definição da população inicial

O método mais comum utilizado na criação da população é a inicialização aleatória dos indivíduos. Se algum conhecimento inicial a respeito do problema estiver disponível, este pode ser utilizado na inicialização da população. Por exemplo, se sabe que a solução final (assumindo codificação binária) vai apresentar mais 0s do que 1s, então esta informação pode ser utilizada, mesmo que não se saiba exatamente a proporção. Já em problemas com restrição, deve-se tomar cuidado para não se gerar indivíduos inválidos na etapa de inicialização [Mitchell, 04].

2.4.3. Soluções e representações

Todo algoritmo de otimização e busca lida com soluções, cada uma representando uma instanciamento do problema em questão. Assim, a representação deve ser tal que pode ser completamente realizada na prática. Na maioria dos problemas de engenharia, a solução é um vetor de valores reais que especifica dimensões aos parâmetros-chaves do problema. Por exemplo, nos problemas de sistema de controle, uma solução é a variação de frequência e tempo funcional dependente dos parâmetros-chave de controle; já em um jogo, a solução é uma estratégia ou um algoritmo para resolver uma tarefa específica. Assim, fica claro que o significado de uma solução é uma característica natural do problema em questão e da sua representação [Deb, 00; Reisinger e Mäikkiläinen, 06].

Como a estrutura de um problema pode variar de um problema para outro, a solução de um problema específico pode ser representada de várias formas. Geralmente, o método de busca é mais eficiente ao lidar com uma representação em específico do que com várias representações. Assim, a escolha de um esquema de representação eficiente não depende apenas do problema em questão, mas também do método de busca. A eficiência e complexidade de um algoritmo de busca dependem muito de como as soluções foram representadas e o quão adequada é a representação para os operadores de busca em questão. Em alguns casos, um problema difícil pode ser simplificado pela escolha adequada da representação, com eficiência e utilizando um algoritmo específico.

2.4.4. Operadores genéticos

Todos os algoritmos evolutivos trabalham por meio de um processo de seleção e um mecanismo que produz variações. O mecanismo mais conhecido para a produção de variações é a mutação, onde um alelo de um gene é aleatoriamente substituído por outro. Em outras palavras, novas tentativas de soluções são criadas fazendo pequenas alterações nas soluções anteriores. Um padrão de mutação comumente usado é de 1 mutação por tamanho da representação; por exemplo, se um cromossomo tiver 100 bits de tamanho, a mutação seria constituída de forma que cada bit tenha a probabilidade de 0.01 de ser escolhido.

Apesar da maioria dos algoritmos evolutivos usarem mutação juntamente com a recombinação genética entre cromossomos, a mutação é algumas vezes tratada como se fosse um operador menor, de fundo, que apenas garante que a população consistirá em um agrupamento de informações diversas de alelos que podem ser aproveitadas por meio da recombinação. Em muitos problemas de otimização, contudo, um algoritmo evolutivo que usa mutação sem recombinação pode ser muito eficaz [Eshelman, 00].

2.4.4.1. O operador de recombinação (*crossover*)

O operador de recombinação ou *crossover* cria novos indivíduos através da combinação de dois ou mais indivíduos. A idéia intuitiva por trás do operador de recombinação é a troca de informação entre diferentes soluções candidatas. Nos algoritmos evolutivos atribui-se uma probabilidade de recombinação fixa aos indivíduos da população. Os operadores de recombinação mais comumente empregados são:

1) *Um e dois pontos*. Para a aplicação destes operadores, são selecionados dois indivíduos pais e, a partir de seus cromossomos, são gerados dois novos indivíduos filhos. Para gerar os filhos, seleciona-se um ou dois pontos de corte de forma aleatória nos cromossomos dos pais, e os segmentos de cromossomo criados a partir destes pontos de corte são trocados.

2) *Recombinação uniforme* [Michalewicz, 99]. Para cada gene é atribuída uma probabilidade fixa p (usualmente 0,5), de qual pai vai contribuir com seu valor para aquela posição; ou seja, com probabilidade p , cada gene do filho receberá o alelo do pai₁ e com probabilidade $1-p$ receberá o alelo do pai₂. Como a recombinação uniforme troca alelos ao invés da seqüência de alelos, pode-se combinar características independentemente da sua posição relativa no cromossomo.

3) *Aritmético* [Eiben e Smith, 03b]. Este operador é definido como uma combinação linear de dois vetores (cromossomos) x_1 e x_2 .

$$\mathbf{x}'_1 = a \mathbf{x}_1 + (1-a)\mathbf{x}_2 \quad (2.1)$$

$$\mathbf{x}'_2 = (1-a) \mathbf{x}_1 + a \mathbf{x}_2 \quad (2.2)$$

onde a é um número aleatório pertencente ao intervalo $[0, 1]$.

Outros exemplos de recombinação desenvolvidos para utilização em problemas de otimização numérica restrita e codificação em ponto flutuante são o *crossover geométrico*: $Z = \{ \sqrt{x_1 y_1}, \sqrt{x_2 y_2}, \dots, \sqrt{x_n y_n} \}$ e o *crossover esférico* [Schoenauer e Michalewicz, 97; Booker *et al*, 00; Medeiros, 06].

2.4.4.2. O operador de mutação

O operador de mutação modifica aleatoriamente um ou mais genes de um cromossomo. A probabilidade de ocorrência de mutação em um gene é denominada *taxa de mutação*. O valor da taxa de mutação pode variar desde valores pequenos, como ocorre nos algoritmos genéticos, até valores altos, como é o caso da programação evolutiva. A idéia intuitiva por trás do operador de mutação é criar uma variabilidade extra na população, mas sem destruir o progresso já obtido com a busca. Os tipos básicos de mutação são mostrados na Tabela 2.5:

Tipos de codificação	Descrição
<i>Binária</i>	Simplesmente troca-se o valor de um gene em um cromossomo [Holland, 92]. Assim, se um gene selecionado para mutação tem valor 1, o seu valor passará a ser 0 após a aplicação da mutação, e vice-versa.
<i>Ponto flutuante</i>	O mais comum neste caso é a <i>mutação Gaussiana</i> , que modifica todos os componentes de um cromossomo $\mathbf{x} = [x_1 \dots x_n]$ na forma [Bäck <i>et al</i> , 00]: $\mathbf{x}' = \mathbf{x} + \mathcal{N}(0, \sigma)$ onde $\mathcal{N}(0, \sigma)$ é um vetor de variáveis aleatórias Gaussianas independentes, com média zero e desvio padrão σ .

Tabela 2.5: Tipos básicos de mutação.

No presente trabalho, além da distribuição Gaussiana, utilizam-se outras distribuições (estáveis) como a de Cauchy, a de Lévy e outras não nomeadas, gerando valores aleatórios de mutação em valores reais.

Com base no exposto, pode-se melhor apreciar a relevância e motivação da pesquisa aqui relatada. Quando se analisam as distribuições Gaussiana, de Cauchy e de Lévy nas Figuras 2.4, 2.5 e 2.6, respectivamente, pode-se questionar qual seria o comportamento de uma mutação em valores reais baseada em cada uma delas.

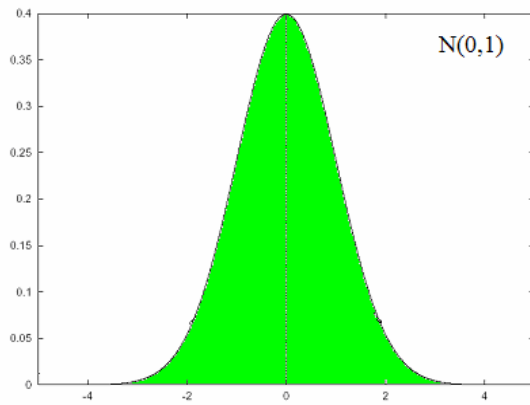


Figura 2.4: Distribuição Gaussiana.

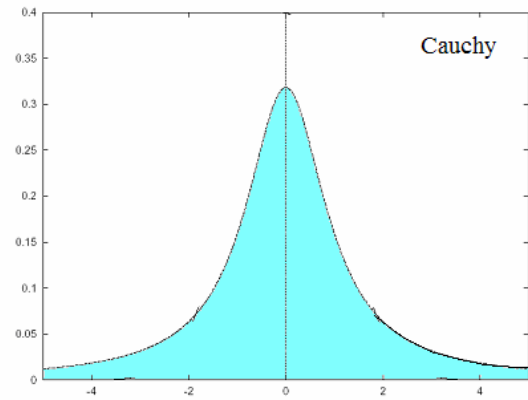


Figura 2.5: Distribuição de Cauchy.

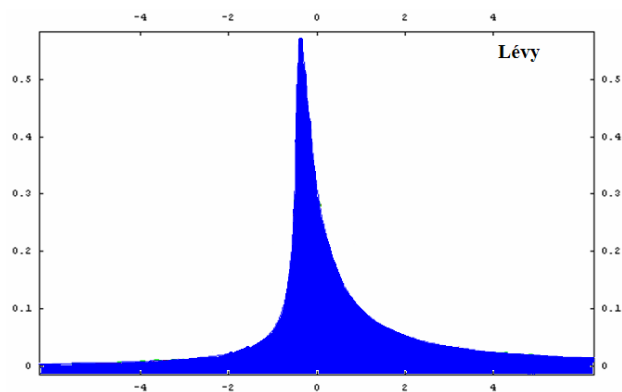


Figura 2.6: Distribuição de Lévy.

Uma distribuição de cauda pesada possui esse nome porque a “cauda” da função não tende rapidamente a zero, ou seja, a probabilidade de variáveis aleatórias com valores grandes é significativamente maior que zero [Silva, Campos e Cunha, 01; Soler, 07]. Pois bem, observando-se as distribuições de Lévy e de Cauchy, constata-se que estas apresentam caudas leves (não nulas) que permitem “saltos” na mutação mais longos, evitando-se a busca local em funções de teste padrão ou falso mínimo [Rudolph, 98; Lee e Yao, 04; Nolan, 01 e 05].

Dada uma representação de valor real onde cada elemento em uma população é um vetor n -dimensional $x \in \mathbb{R}^n$, há muitos métodos para criar novos elementos (descendentes) usando mutação [Bäck *et al*, 00]. A forma geral de mutação pode ser escrita como:

$$x' = m(x) \quad (2.3)$$

onde x é o vetor de pais, m é a função de mutação, e x' é o vetor de descendentes resultantes.

A aplicação do operador de mutação é simples, baseia-se na utilização de uma função de densidade de probabilidade (PDF) [Bäck *et al*, 00] que, aplicada ao vetor de pais gera o vetor x' de descendentes. A forma mais comum de mutação que gera um vetor de descendente é

$$x' = x + M \quad (2.4)$$

onde a mutação M é uma variável aleatória. M é frequentemente de média zero.

As mutações M podem assumir formas diferentes. Por exemplo, M poderia ser a variável aleatória uniforme $U(a,b)^n$, onde a e b são os limites inferior e superior, respectivamente. Nesse caso, a é geralmente igual a “ $-b$ ”, resultando da Equação 2.4 um descendente $x+U(-b,b)^n$. De forma semelhante, o operador de mutação progressiva (*creep mutation*) [Boyd,94; Nehab,04] assume probabilidade fixa de alterar cada componente de x para cima e para baixo por meio de um pequeno valor aleatório.

A mutação para os vetores de valores reais tem sido a Gaussiana [Fogel e Atmar, 90; Michalewicz, 99; De Jong *et al*, 00], cuja função de densidade de probabilidade é definida como

$$g(x) = [\sigma (2\pi)^{1/2}]^{-1} \exp[-0.5(x-\mu)^2/\sigma^2] \quad (2.5)$$

onde μ é a média da distribuição e σ é o desvio padrão. No uso de mutações Gaussianas de média zero geram-se descendentes que cuja média é a de seus pais, e que têm menor probabilidade de serem diferentes dos pais.

Outras funções de densidade também foram implementadas. Yao e Liu [96] propuseram o uso das distribuições de Cauchy para auxílio na fuga de mínimos locais (a distribuição de Cauchy possui uma cauda mais longa que a Gaussiana) e demonstraram que as mutações de Cauchy podem ser vantajosas.

Estudos posteriores levaram ao uso da distribuição de Lévy que se mostrou útil no uso conjunto com as mutações para valores reais devido ao fato da sua cauda ser mais longa e ter uma maior probabilidade para mutações pequenas e permitir mutações mais amplas. Nos extremos das caudas das distribuições de Lévy e de Cauchy, a primeira distribuição apresenta neste ponto uma probabilidade maior [Lee e Yao, 04].

Outro tipo possível é a distribuição de Laplace [François, 99], mas não há evidências de que ela seja mais adequada do que a Gaussiana ou a de Cauchy para os problemas de valores reais [Bäck *et al*, 00]. A Figura 2.7 mostra diversos tipos de distribuição estável em função do índice de estabilidade α , cujo valor define famílias estáveis, conforme a seguir:

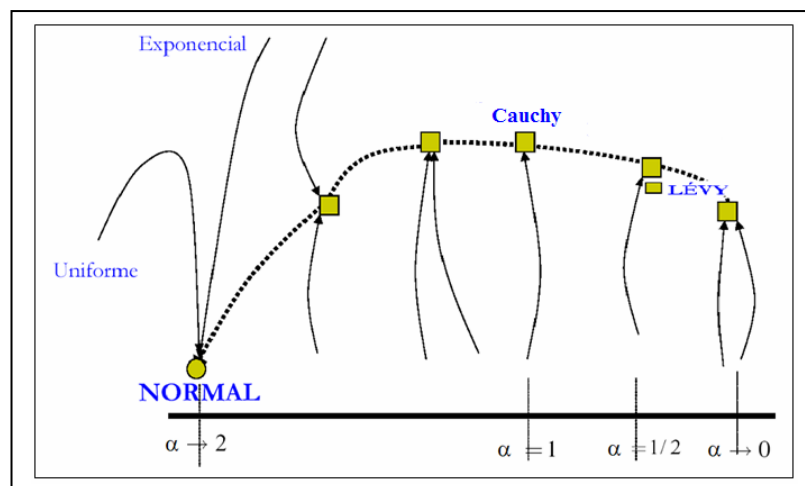


Figura 2.7: Distribuições estáveis (Gauss: $\alpha=2$; Cauchy: $\alpha=1$; Lévy: $\alpha=0.5$, e demais) [Vicente, 05].

No próximo capítulo, a mutação de números reais será tratada com base na aplicabilidade de distribuições estáveis, no contexto das estratégias evolutivas, visando analisar a influência da escolha adequada da distribuição em relação à função de teste escolhida.

2.4.4.3. Seleção de indivíduos

O algoritmo genético clássico utiliza um esquema de seleção de indivíduos para a próxima geração chamada método da roleta [Michalewicz, 99]. Este atribui a cada indivíduo de uma população uma probabilidade de passar para a próxima geração proporcional a sua aptidão medida, em relação à somatória das aptidões de todos os indivíduos da população. Assim, quanto maior aptidão um indivíduo tiver, maior a probabilidade de passar para a próxima geração. Na seleção de indivíduos pelo método da roleta pode fazer com que o melhor indivíduo da população seja perdido, ou seja, não passe para a próxima geração. Uma alternativa é escolher como solução o melhor indivíduo encontrado em todas as gerações do algoritmo. Outra opção é simplesmente sempre manter o melhor indivíduo da geração atual na geração seguinte, estratégia conhecida como *seleção elitista* [Fogel, 94a; Michalewicz, 99].

Outro exemplo de mecanismo de seleção é a *seleção baseada em ordenação (rank)* [Eiben e Smith, 03a]. Esta estratégia utiliza as posições dos indivíduos quando ordenados de acordo com sua aptidão para determinar a probabilidade de seleção. Podem ser usados mapeamentos lineares ou não-lineares para determinar a probabilidade de seleção [Michalewicz, 99]. Uma variação deste mecanismo é simplesmente passar os melhores indivíduos N para a próxima geração. A seguir, alguns mecanismos de seleção:

1. *Seleção por diversidade*: são selecionados os indivíduos mais diversos da população.
2. *Seleção bi-classista*: são selecionados os $P\%$ melhores indivíduos e os $(100 - P)\%$ piores indivíduos.

3. *Seleção aleatória*: são selecionados aleatoriamente N indivíduos da população. Pode-se subdividir este mecanismo de seleção em salvacionista (o melhor indivíduo e os outros casualmente) ou não-salvacionista.

2.5. Aplicações da computação evolutiva

As aplicações da computação evolutiva podem ser encontradas em diversas áreas, por exemplo: planejamento, projeto, simulação e identificação, e classificação [Freitas *et al*, 93; Beasley, 00; Tinos e Carvalho, 03].

No planejamento tem-se o roteamento com as seguintes áreas de aplicação: caixeiro viajante, roteamento de veículos, problema de transporte e robótica. Já na aplicação de seqüenciamento de tarefas desenvolve-se um plano para executar uma determinada quantidade de tarefas, em um período de tempo onde os recursos são limitados, há restrições e pode haver mais de um objetivo a ser otimizado, como: programação de trabalho, tabelas de horários, processamento computacional e empacotamento.

Nas aplicações em projeto tem-se: filtros, processamento de sinais e sistemas inteligentes.

A simulação envolve a determinação de como o sistema irá se comportar baseando-se em um modelo ou projeto, como: a determinação de equilíbrio de sistemas químicos e estruturas de proteínas. A identificação envolve a determinação do projeto de um sistema, dado seu comportamento, por exemplo, a determinação de pólos e zeros de um sistema.

Os sistemas classificadores são utilizados como partes de outros sistemas de controle em: economia, biologia, processamento de imagens e mineração de dados.

2.6. Conclusões

Neste capítulo, são apresentados os principais conceitos relacionados à computação evolutiva, mais especificamente os algoritmos evolutivos. Apesar dessa área crescer de forma vertiginosa nos últimos anos, ainda há várias questões possíveis de serem tratadas com esta ferramenta de modo alternativo.

CAPÍTULO III

ESTRATÉGIAS EVOLUTIVAS

3.1. Introdução

Esse capítulo fornece uma visão geral nas estratégias evolutivas sendo na Seção 3.2 as idéias básicas e os seus princípios. A Seção 3.3 são trabalhados os ingredientes para o projeto dos algoritmos das estratégias evolutivas tais como: operadores de mutação, recombinação, reprodução e seleção com exemplos curtos de implementação.

As estratégias evolutivas (EEs) foram desenvolvidas para a resolução de problemas técnicos de otimização em engenharia. Atualmente, constituem-se de importantes algoritmos computacionais em problemas de otimização de parâmetros.

A primeira EE desenvolvida foi a EE-(1+1), proposta por Rechenberg e Schwefel [Beyer,01], em experimentos no contexto de um túnel de vento [Bäck, Hammel e Schwefel, 97]. A EE-(1+1) foi então generalizada, por maiores valores do número μ de ancestrais (pais) e do número λ de descendentes (filhos) por geração. As estratégias evolutivas com multimembros são divididas de acordo com o mecanismo de seleção em:

- (i) Estratégia *soma* ($\mu+\lambda$) que sugerem a μ pais produzir λ filhos, após os μ pais e os λ filhos competirem pela sobrevivência.
- (ii) Estratégia *vírgula* (μ, λ), onde os λ filhos competem para sobreviver e os pais são completamente substituídos a cada geração.

Os μ pais podem ter incorporados um fator ρ , que especifica quantos deles serão utilizados, por recombinação, para que seus filhos sejam criados, que origina às estratégias evolutivas ($\mu/\rho+\lambda$) e ($\mu/\rho, \lambda$).

3.2. O algoritmo básico das EE - A notação $(\mu/\rho\ddagger\lambda)$

A meta comum de uma estratégia evolutiva é otimizar alguma função com respeito a um conjunto de variáveis de decisão ou parâmetros de controle $\mathbf{y} = (y_1, y_2, \dots, y_n)$. Para tanto, define-se a função $F(\mathbf{y})$, chamada de função de *aptidão* ou *adaptação*, que, atuando sobre os dados – no presente caso, o espaço de busca n-dimensional de valores reais, isto é, $\mathbf{y} \in \mathbb{R}^n$ – permite avaliar sua proximidade do valor buscado que otimiza F .

As estratégias evolutivas atuam em uma população de indivíduos, onde cada um possui um valor da função de aptidão, cada um representando uma coordenada no espaço de busca (os parâmetros *objetivos*) e os valores que determinam a característica da distribuição utilizada para a mutação de cada indivíduo (os parâmetros *estratégicos*). A Figura 3.1 mostra o pseudocódigo da EE $(\mu/\rho\ddagger\lambda)$ e a seguir descreve-se brevemente o algoritmo.

```

Início
g = 0;
Inicializar população ( $\beta_{\mu}^{(0)} = \{(y_m^{(0)}, s_m^{(0)}, F(y_m^{(0)}))\}$ )
Repetir
  Para l = 1 até  $\lambda$ 
     $\xi_l = \text{casamento}(\beta_{\mu}^{(g)}, \rho);$  (1)
     $s_l = \text{s\_recombinação}(\xi_l, \rho);$  (2)
     $y_l = \text{y\_recombinação}(\xi_l, \rho);$  (3)
     $s'_l = \text{s\_mutação}(s_l);$  (4)
     $y'_l = \text{y\_mutação}(y_l, s'_l);$  (5)
     $F'_l = F(y'_l);$ 
  Fim;
   $\beta_{\lambda}^{(g)} = \{(y'_l, s'_l, F'_l)\};$ 
  SE (tipo de seleção ==  $(\mu, \lambda)$ )
     $(\mu, \lambda) : \beta_{\mu}^{(g+1)} = \text{seleção } F_{\mu}(\beta_{\lambda}^{(g)});$ 
  SE (tipo de seleção ==  $(\mu + \lambda)$ )
     $(\mu + \lambda) : \beta_{\mu}^{(g+1)} = \text{seleção } F_{\mu}(\beta_{\lambda}^{(g)}, \beta_{\lambda}^{(g)});$ 
  g = g + 1;
Até condição de término;
Fim;

```

Figura 3.1: Procedimento $(\mu/\rho\ddagger\lambda)$ - EE [Beyer, 01].

Na geração $g=0$, a população $\beta_{\mu}^{(0)} = (\beta_1, \dots, \beta_{\mu})$, é inicializada. Após a inicialização o procedimento de laço “repetiraté (condição de término)” é executado. A partir da população $\beta_{\mu}^{(g)}$ na geração g , uma nova população de descendentes $\beta_{\lambda}^{(g)}$ é produzida pela execução de λ vezes no laço interno “**Para** $l=1$ **até** λ ”. Cada laço gera um descendente. Os principais pontos do algoritmo podem ser acompanhados na Figura 3.1 pelos comentários abaixo, cuja numeração encontra-se especificada no algoritmo:

1. Neste passo (casamento), uma família de pais ξ de tamanho ρ é escolhida aleatoriamente a partir de um subconjunto de μ pais. Esse processo de seleção de casamento é aleatório e independente dos valores objetivos de pais F . No caso especial de $\rho = \mu$, todos os pais tornam-se membros de uma família de pais ξ em contraste às técnicas de seleção usuais nos algoritmos genéticos [Michalewicz, 99].
2. Recombinação dos parâmetros estratégicos.
3. Recombinação dos parâmetros objetivos. Observe-se que, se $\rho=1$, o recombinante é simplesmente uma cópia de um pai.
4. Mutação dos parâmetros estratégicos: segue uma distribuição estável.
5. Mutação dos parâmetros objetivos.

Após ter completado uma população de descendentes $\beta_{\lambda}^{(g)}$, a seleção é executada com o resultado de uma nova população de pais $\beta_{\mu}^{(g+1)}$.

Por fim, a condição de término é checada. Esta condição pode ser baseada no número máximo de gerações ou no tempo máximo de CPU. Alternativamente, pode-se ainda optar pela convergência dos valores da aptidão, dos parâmetros objetivos (proximidade do valor desejado) ou dos próprios parâmetros estratégicos (no caso de ser desejada uma determinada distribuição).

3.3. Ingredientes para o projeto dos algoritmos das estratégias evolutivas

3.3.1. Representação de indivíduos

Nas implementações de computação com EE os *vetores de números reais* são utilizados como estruturas de dados básicas a serem otimizadas. Para os parâmetros n a serem adaptados, as estratégias evolutivas podem trabalhar com um vetor \mathbf{g} da seguinte maneira:

$$\text{cromossomo EE:} \quad \mathbf{g} = (p_1, p_2, \dots, p_n) \text{ com } p_i \in \mathfrak{R}. \quad (3.1)$$

A letra \mathbf{g} para o símbolo do vetor representa *genótipo* em linha. Para versões EE simples são considerados apenas os parâmetros a serem ajustados no cromossomo, nos trabalhos de otimização mais complexos, no entanto, vem a ser vantajoso o processo evolutivo estender o conjunto de decisão aos parâmetros objetivos, (p_1, p_2, \dots, p_n) , bem como no vetor de *parâmetros estratégicos*, (s_1, \dots, s_n) . Esses parâmetros adicionais servem como variações ou desvios para o controle do tamanho de passo da mutação para cada um dos parâmetros objetivos. Daí um cromossomo EE pode ser representado como um vetor em par:

$$\mathbf{g} = (\mathbf{p}, \mathbf{s}) = ((p_1, \dots, p_n), (s_1, \dots, s_n)) \quad \text{com } p_i, s_i \in \mathfrak{R}. \quad (3.2)$$

Os parâmetros de controle \mathbf{s} são considerados modelos internos de condições ambientais, sob o processo evolutivo. A adaptação desses parâmetros de controle é uma parte integrante da busca absoluta pelos parâmetros objetivos. Um exemplo de cromossomo EE é mostrado na Figura 3.2.



Figura 3.2: Cromossomo EE com parâmetros estratégicos.

3.3.2. Mutação

A mutação, da mesma forma que a seleção, é uma das forças que impulsionam a evolução. A atual definição da mutação varia com o problema específico do contexto, variando desde uma mera mudança de letra em um nome, até as mutações representações mais complexas, envolvendo inúmeras características, todas elas integrantes do cromossomo.

Partindo de uma população de indivíduos (pais) iniciais, cujas coordenadas são armazenadas nos parâmetros objetivos de seus cromossomos, pode-se escolher o melhor indivíduo ou parte da melhor população para prosseguir na próxima geração. A busca pelo ponto de ótimo buscado é feita a partir de mutações de pais, dando origem a filhos, cujas coordenadas são levemente diferentes de cada pai individual. Os parâmetros objetivos dos cromossomos mutados (filhos) geralmente mostram diferenças pequenas e por vezes afastadas do cromossomo original.

Com isso, tem-se então uma nova população de indivíduos, que pode ser avaliada com respeito à tarefa de otimização apresentada. Certamente é vantajoso escolher alguns dos melhores indivíduos como novo ponto de início, para a geração seguinte. Nas estratégias evolutivas os operadores de mutação são preferidos, pois resultam em alterações que busquem outros pontos do espaço de busca igualmente válidos e afastados. Sendo assim, as soluções não devem se fixar (no caso de um processo de minimização) em mínimos locais e sim em um mínimo global.

Para a mobilidade de uma população, no sentido de sua habilidade de explorar os arredores mais próximos e mais distantes, é particularmente importante que o sobrevivente da próxima geração não apenas herde os melhores parâmetros objetivos, mas também seus parâmetros estratégicos. Estes determinam o modo de uma exploração mais adequada que acontecerá na próxima população de filhos. Para um cromossomo EE, os filhos mutantes encontram-se dentro de um raio maior em torno do pai, sendo que as distâncias entre pais e filhos tendem a ser menores.

Os parâmetros objetivos são modificados pelo uso dos parâmetros estratégicos. O operador de mutação ω_{mut} é definido como uma adição relacionada a um componente que normalmente origina-se

de uma distribuição estável, neste caso de Gauss. O vetor de mutação dos parâmetros objetivos p_{mut} é calculado como:

$$\omega_{mut}(p, s) = p_{mut} = p + N_0(s) \quad (3.3)$$

$$p_{mut} = p_1 + N_0(s_1), \dots, p_n + N_0(s_n) \quad (3.4)$$

onde $N_0(s) = (N_0(s_1), \dots, N_0(s_n))$ é um vetor de números aleatórios Gaussianos s_i . O cromossomo que sofreu mutação aparece da seguinte forma:

$$g_{mut} = (p_{mut}, s) \quad (3.5)$$

Observa-se assim, o papel central das distribuições Gaussianas nas mutações em EE, a qual se caracteriza pela forma

$$\mathcal{N}_m \left[\frac{p - m}{2ps^2} \right] \quad (3.6)$$

onde os parâmetros μ e σ denotam a média e o desvio padrão, respectivamente. Frequentemente a variância σ^2 é dada ao invés do desvio padrão, e a distribuição Gaussiana é referida como uma distribuição $N_\mu(\sigma^2)$. A Figura 3.3 mostra a distribuição Gaussiana $N_0(\sigma)$ para três desvios padrões diferentes, $\sigma=0.5, 1.0, e 2.0$.

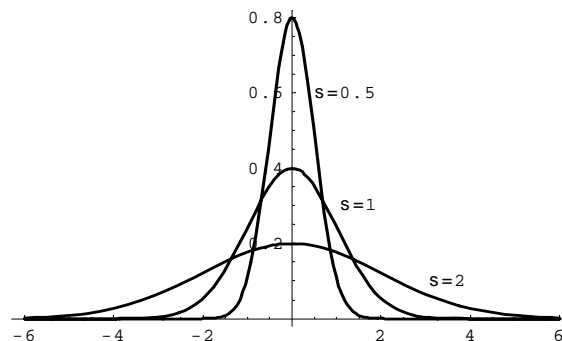


Figura 3.3: Distribuições normais para uma média constante $\mu=0$.

Note-se que a Gaussiana retorna valores pequenos (mais próximos da média zero) com mais freqüência que valores grandes; este é o aspecto central envolvido neste trabalho, em que, para outras distribuições, isto não acontece [Nolan, 06b].

Os valores dos parâmetros de um cromossomo base \mathbf{p} (pais) e um cromossomo mutante \mathbf{p}_{mut} (filhos) são mais parecidos, com certo intervalo de variação, após a aplicação dos parâmetros estratégicos conforme a Figura 3.4:

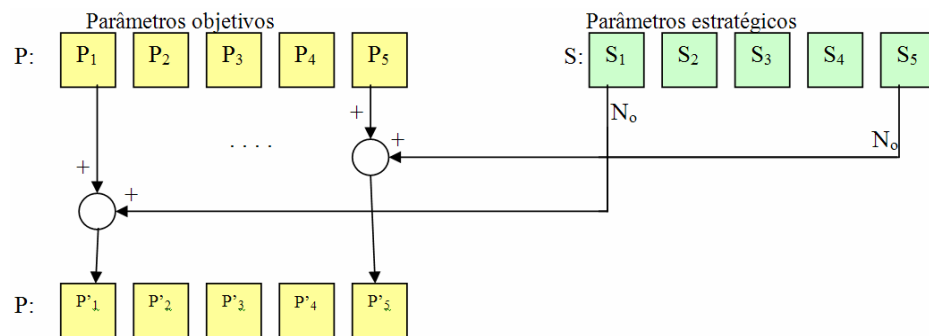


Figura 3.4: Mutação de cromossomos EE [Jacob, 01].

O intervalo de variação para os parâmetros objetivos pode ser controlado por um grupo de distribuições estáveis \mathbf{s} , que é o motivador para a definição de uma mutação EE como introduzido anteriormente e mostrado na Figura 3.4.

Utilizando a distribuição Gaussiana, geram-se alterações nos pares dos parâmetros objetivo e estratégico. A estrutura de dados \mathbf{p} representa o parâmetro objetivo a ser mutado, enquanto os parâmetros estratégicos do cromossomo \mathbf{s} sofrem uma alteração momentânea pela adição ou subtração do componente da distribuição, que gera novos cromossomos.

A mutação não é só restrita aos parâmetros objetivos, mas é também estendida aos parâmetros estratégicos conforme abaixo,

$$\mathbf{g}_{mut} = (\mathbf{p}_{mut}, \mathbf{s}_{mut}) \quad (3.7)$$

$$\mathbf{p}_{mut} = \mathbf{p} + N_0(\mathbf{s}), \mathbf{s}_{mut} = \phi(\mathbf{s}) \quad (3.8)$$

onde ϕ é uma função conhecida como de 'adaptação do passo de mutação' (*mutative step size adaptation*, ou MSA) [Jacob, 01; Nieberg e Beyer, 07]. Conseqüentemente, o esquema de mutação da Figura 3.4 deve ser estendido pela operação respectiva para a adaptação do tamanho de passo. Agora o algoritmo evolutivo deve lidar com uma tarefa de otimização mais difícil. As combinações mais favoráveis não apenas devem ser encontradas para os parâmetros objetivos, como também para os devidos parâmetros estratégicos.

3.3.3. Recombinação

As estratégias evolutivas utilizam um operador que recombina os parâmetros objetivos e estratégicos, respectivamente. Para dois cromossomos $\mathbf{a} = (p_a, s_a)$ e $\mathbf{b} = (p_b, s_b)$ pode-se definir um operador de recombinação ω_{rec} como segue:

$$\omega_{\text{rec}}(\mathbf{a}, \mathbf{b}) = (\mathbf{p}', \mathbf{s}') = ((p'_{1,}, \dots, p'_{n,}), (s'_{1,}, \dots, s'_{n,})) \quad (3.9)$$

$$\text{onde,} \quad \mathbf{p}'_i = \rho_p(p'_{a,i}, p'_{b,i}) \quad \forall i \in \{1, \dots, n\}$$

$$\mathbf{s}'_i = \rho_s(s'_{a,i}, s'_{b,i})$$

As novas entradas p'_i e s'_i do cromossomo recombinado são calculadas pelo mapeamento das funções de recombinação ρ_p e ρ_s nos correspondentes parâmetros objetivos e estratégicos.

No caso da chamada recombinação *discreta*, cada um dos elementos internos do vetor que contém parâmetros objetivos e estratégicos tem uma probabilidade definida através de uma função que gera números aleatórios no intervalo de $[0,1]$, definindo a participação ou não de cada elemento no novo vetor gerado. Os elementos são escolhidos com probabilidades iguais a 50% na Figura 3.5. A recombinação discreta nas EEs é semelhante ao operador de cruzamento uniforme nos algoritmos genéticos.

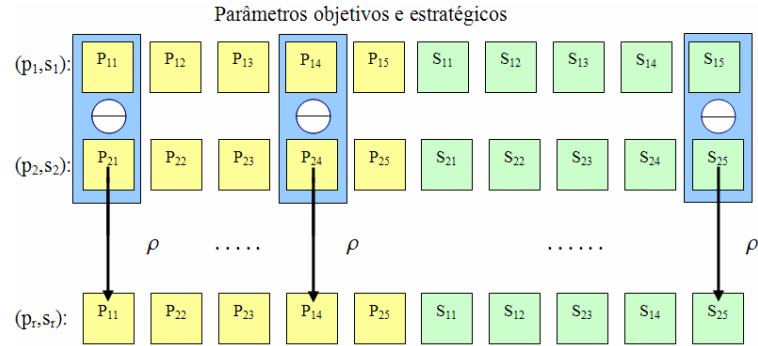


Figura 3.5: Recombinação discreta de dois cromossomos em EEs com $\rho = \rho_p = \rho_s$ [Jacob, 01].

Existe ainda o caso da recombinação *intermediária*, em que os valores médios dos respectivos componentes são calculados, como ilustra a Figura 3.6:

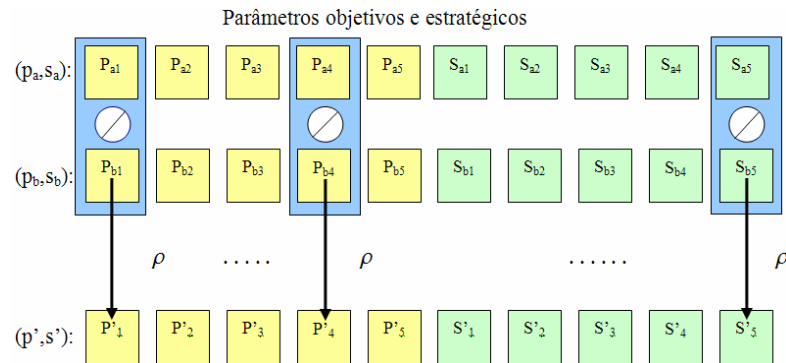


Figura 3.6: Recombinação intermediária de dois cromossomos em EEs por $\rho = \rho_p = \rho_s$ [Jacob, 01].

Este operador é usado para fazer recombinações, em especial entre parâmetros estratégicos. A obtenção da média eventualmente resulta, em uma diversidade reduzida de “intervalo”. Se a recombinação intermediária é aplicada com muita frequência, os novos filhos tornam-se mais e mais semelhantes. Isso pode levar a uma convergência prematura da combinação total de genes da população. A recombinação de operadores descrita pode com certeza ser estendida para o caso geral onde tem-se um número maior que dois de pais participando da recombinação. Não se deve esquecer que o custo computacional aumenta diretamente com o aumento do número de pais envolvidos na recombinação de cromossomos.

Pode-se ainda definir um operador de recombinação que trabalhe em populações de cromossomos, como foi feito para os operadores de mutação na Seção 3.3.2. Com esta definição estendida de recombinação também podem-se considerar algumas variantes de esquemas de recombinação [Booker *et al*, 00], como se utilizar quatro operadores de recombinação (recombinação local: discreta / intermediária e recombinação global: discreta / intermediária) para trabalhar apenas com uma subpopulação ou escolher seus pares de recombinação.

Assumindo uma população $P = (g_1, g_2, \dots, g_N)$ de cromossomos N da forma $g_i = (g_{i1}, \dots, g_{iN})$ com $1 \leq i \leq N$, um operador de multicombinação ω_{rec} , que combina r cromossomos, pode ser definido como

$$\omega_{\text{rec}}(g_{i_1}, \dots, g_{i_r}) = g' = (g'_1, \dots, g'_N) \quad (3.10)$$

onde todos os índices $i_1, \dots, i_r \in \{1, \dots, N\}$ com $i_j \neq i_k$, isto é, nenhum cromossomo aparece mais de uma vez, como recombinação em pares ($1 \leq r \leq N$).

Ambas variantes, recombinação local e global, podem ser usadas tanto na função de recombinação discreta como na intermediária. Para as entradas de recombinação em pares r do cromossomo recombinado g' é calculado o novo vetor filho a seguinte maneira:

$$\omega_{\text{rec}}^{\text{local}}(g_{i_1}, \dots, g_{i_r}) = (g'_{i_1}, \dots, g'_{i_r}) \quad (3.11)$$

Esse tipo de recombinação é denominada *local* porque a seleção da recombinação em pares ocorre antes da recombinação real. A recombinação é restrita a um grupo local da população inteira. Os parâmetros foram definidos em analogia à Figura 3.7 como vetores com componentes $P[i, k]$.

Os parâmetros objetivos $P[i, k]$ não são explicitamente levados em consideração, desde que sua recombinação seja realizada da mesma maneira, dentro do mesmo tipo de parâmetro. A função de recombinação aqui apresentada permite recombinar os parâmetros objetivos e estratégicos diferentemente, ou seja, discreta ou intermediária para cada um dos tipos de parâmetros, selecionando-se as recombinações em par r de acordo com o esquema na Figura 3.7.

Parâmetros objetivos e estratégicos

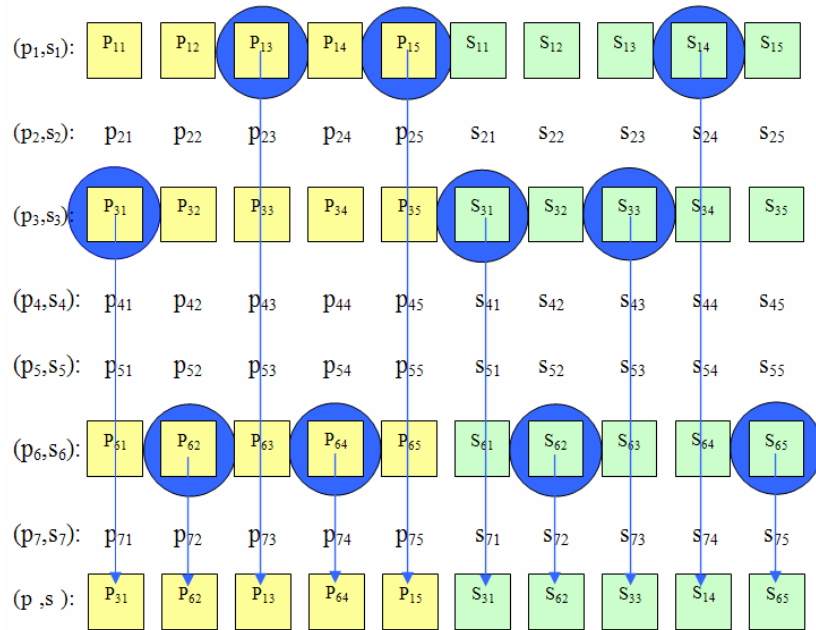


Figura 3.7: Esquema de recombinação múltipla local e discreta EE [Jacob, 01] .

CAPÍTULO 4

4. DISTRIBUIÇÕES ESTÁVEIS E FUNÇÕES DE TESTE

As distribuições estáveis são o motivador desse trabalho. Como as distribuições de Cauchy, de Lévy apresentam nas extremidades uma cauda pesada e longa, que representam uma probabilidade significativa para valores distantes da média, este fato justifica o desenvolvimento de experimentos que exploram essa características dessas distribuições, expandindo a tradição de uso das distribuições Gaussianas [Yao e Liu 96; Yao, Liu e Lin 99; Lee e Yao, 04].

4.1. Introdução às distribuições estáveis

A distribuição de probabilidade associada à variável aleatória contínua X é dita distribuição contínua de probabilidade. A distribuição Normal ou Gaussiana é a uma distribuição contínua de probabilidade de uma variável aleatória X [Larson e Faber, 04; Milone, 04], bem como as distribuições de Cauchy, de Lévy.

As distribuições de Gauss e Cauchy são distribuições simétricas em relação à média. A função de Cauchy tem média sempre zero, sua forma é sino ou cauda neutra, já a função de Gauss também tem média sempre zero, mas sua forma é sino ou cauda leve [Soler, 07]. Já a distribuição de Lévy é assimétrica [Bertoin, 02; Borak, 05], que também apresenta características de uma distribuição cauda longa só à direita ou à esquerda. No entanto, define-se no contexto deste trabalho uma versão simétrica da distribuição Lévy, aqui denominada de S-Lévy, com forma cauda longa, lembrando sua forma aguda em torno da média.

Todas essas distribuições consideradas até aqui são ditas distribuições estáveis. Considerando-se X_k uma variável aleatória com determinada distribuição de probabilidade, e definindo-se a somatória de um conjunto dessas variáveis, se a nova distribuição resultante também possuir o mesmo tipo de distribuição que as componentes X_k , então, diz-se que a distribuição envolvida é *estável*. De forma simplificada pode-se visualizar os grupos de distribuições conforme a Figura 4.1.

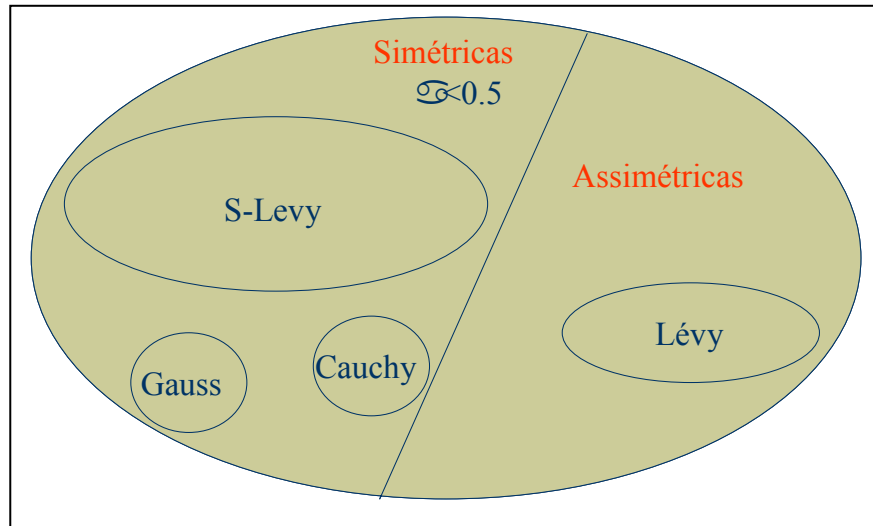


Figura 4.1: Famílias de distribuições estáveis.

A distribuição estável geral é descrita por quatro parâmetros:

- Índice de estabilidade $\alpha \in (0, 2]$,
- Parâmetro de assimetria $\beta \in [-1, 1]$,
- Parâmetro de escala $\gamma > 0$,
- Parâmetro de posição $\delta \in \mathbb{R}$.

Pode-se mostrar que é viável ajustar os modelos estáveis aos dados e usar diagnósticos para avaliar a boa qualidade de ajuste. As distribuições estáveis foram propostas como um modelo para diversos tipos de sistemas físicos e econômicos [Tsallis, 00; Nolan, 05]. Na física, o reflexo que sai de um espelho rotativo produz uma distribuição de Cauchy; períodos de movimentos brownianos produzem uma distribuição de Lévy; o campo gravitacional das estrelas produz uma distribuição estável pouco conhecida, chamada de distribuição de Holtsmark [Nolan, 01].

Outra motivação para o estudo dessas distribuições é o *teorema de limite central generalizado* estabelecendo que o único limite não-trivial possível de termos de somas normalizadas é estável. Observa-se que existem exemplos onde a somatória de muitos termos pequenos – o preço de um cabo,

o ruído presente em um sistema de comunicação – gera um modelo estável que deve ser utilizado para descrever tais sistemas, no caso, financeiro e de comunicação.

Um terceiro argumento para a modelagem com distribuições estáveis é empírica: muitos bancos de dados extensos exibem bordas longas e assimétricas. A forte evidência empírica para estes aspectos combinados com o teorema limite central generalizado é utilizado por muitos para justificar o uso de modelos estáveis.

Céticos dos modelos estáveis recuam da suposição implícita da variação infinita do modelo estável não-Gaussiano e propõem outros modelos para os bancos de dados assimétricos e de bordas longas, por exemplo, os modelos mistos, as variantes de tempo variável, dentre outros. As mesmas pessoas que afirmam que a população é inerentemente limitada e por este motivo deve haver uma variação finita, regularmente usam uma distribuição normal – com apoio ilimitado – como modelo para essa mesma população. A variação acontece, mas existe uma medida de extensão para uma distribuição, e não é apropriada para todos os problemas, o que geralmente pode causar preocupação é em relação à captura do formato de uma distribuição.

Agora é possível usar diagnósticos para avaliar quando um modelo estável descreve precisamente os dados ou não. Em alguns casos há razões solidamente teóricas para se acreditar que um modelo estável é apropriado. Se uma distribuição estável descreve os dados com precisão e parsimoniosamente com quatro parâmetros, então se aceita como um modelo para os dados observados.

Na próxima seção descrevem-se algumas formas de parâmetros para as distribuições estáveis e algumas propriedades básicas. Utiliza-se o programa *STABLE* desenvolvido por [Nolan, 06a e 06b], que implementa os modelos de distribuições estáveis e parametrizados.

4.2. Propriedades básicas e de parâmetros

Existem diferentes modelos matemáticos utilizando parâmetros de distribuições estáveis. Todos envolvem diferentes especificações da função característica e são úteis por várias razões técnicas [Nolan, 01]. A parametrização mais usada é a seguinte: $X \sim S(\alpha, \beta, \gamma, \delta; \mathbf{1})$, com a função característica dada por:

$$E \exp \{iXt\} = \exp \left\{ -|t|^\alpha \left[\frac{1-\beta}{2} + i \beta \operatorname{sign}(t) \right] \right\}, \quad (4.1)$$

onde, X é variável aleatória.

defini-se uma função sinal dada por: $\operatorname{sign}(t)$ igual a:

-1 para $t < 0$

0 para $t = 0$

+1 para $t > 0$

A melhor maneira de descrever todas as possíveis distribuições estáveis é através da sua função característica de X ou transformada de Fourier. Esta função pode ser expressa de varias formas neste estudo, $M=1$, processo padrão de Samorodnitsky e Taqu classificado como padrão [Nolan, 01] e $M=0$, processo de Zolotarev. Nas Figuras 4.2, 4.3 e 4.4 mostra-se a influência da escolha do parâmetro M na curva de representação das distribuições de Lévy, de Cauchy e de Gauss respectivamente, na parametrização $\beta=0$ (preto), $\beta=0.25$ (verm), $\beta=0.5$ (verde), $\beta=0.75$ (amarelo), $\beta=1$ (azul) com $\gamma=1$ e $\delta=0$. Em particular para $\alpha=0.5$ e $\beta=0$ tem-se S-Lévy na Figura 4.2.

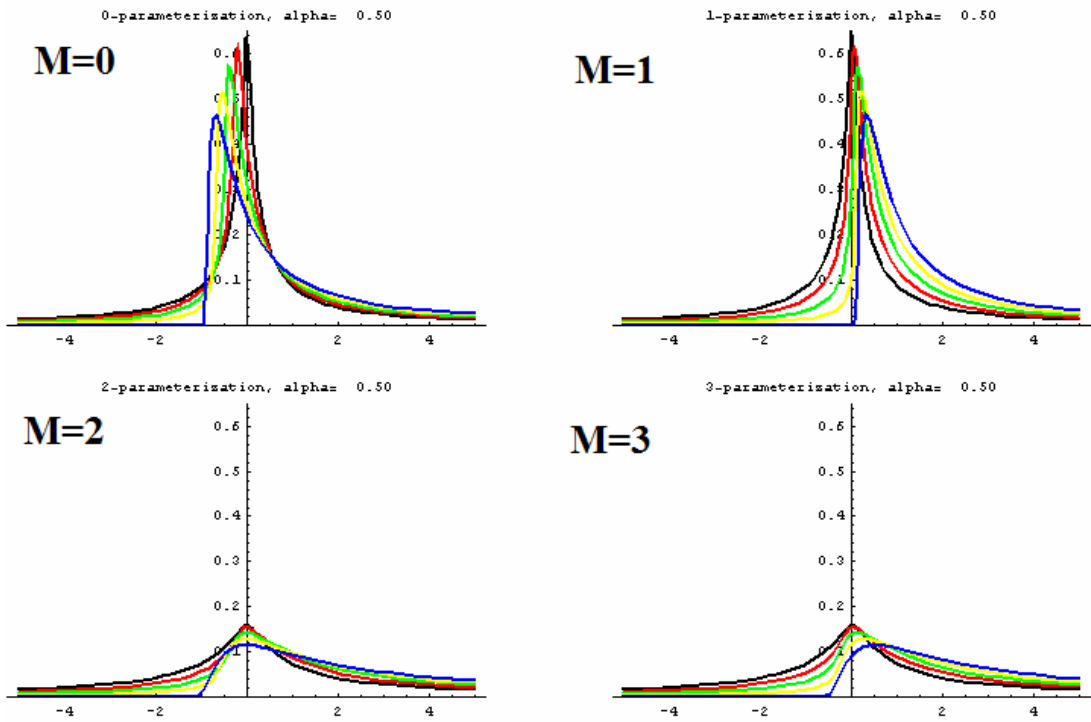


Figura 4.2: Família de distribuições de Lévy.

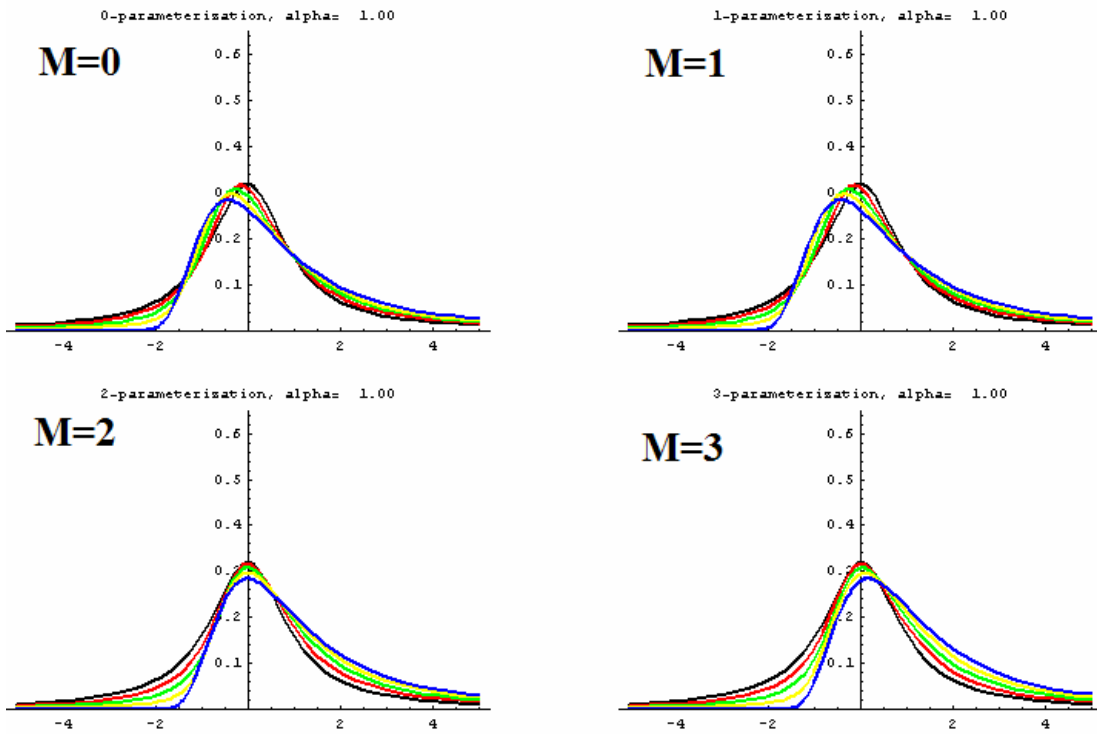


Figura 4.3: Família de distribuições de Cauchy.

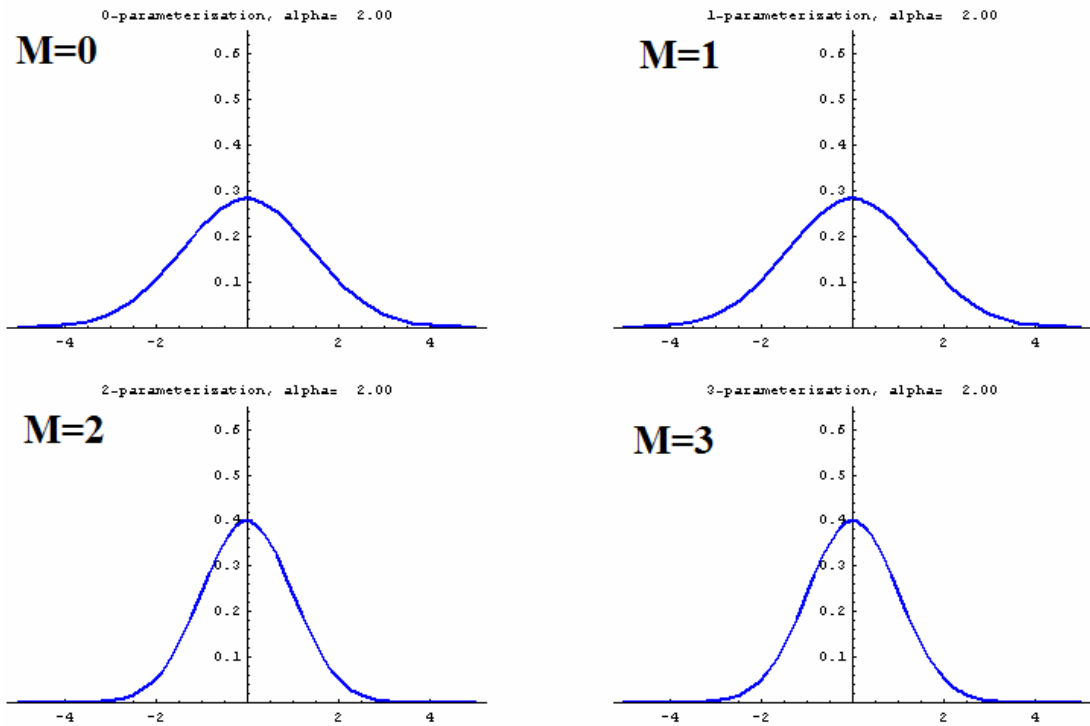


Figura 4.4: Família de distritribuições de Gauss.

A amplitude dos parâmetros é $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\gamma > 0$, e $\delta \in \mathbb{R}$. É preferível não usar σ para o parâmetro de escala, uma vez que o desvio padrão só tem sentido para $\alpha = 2$. Da mesma forma, é preferível usar δ ao invés de μ para o parâmetro de posição, porque nem sempre existem médias, e mesmo existindo, o parâmetro de posição e a média diferem.

Um parâmetro útil nas aplicações é a variação ($M=0$) de Zolotarev, que é definido após o ponto e vírgula na expressão geral: $X \sim S(\alpha, \beta, \gamma, \delta_0; \mathbf{0})$, com a função característica de X dada por:

$$E \exp \{ i X t \} = \exp \left\{ - \frac{\gamma}{2} \left[\frac{1 - \cos \left(\frac{2 \alpha t}{\gamma} \right)}{\alpha} + \frac{1 - \cos \left(\frac{2 \beta t}{\gamma} \right)}{\beta} \right] \right\} \quad \text{val } 1, \quad a = 1, \quad (4.2)$$

O valor desta representação é que as funções características (e daí as densidades correspondentes e definições) são conjuntamente contínuas em todos os quatro parâmetros. Cálculos numéricos precisos das densidades correspondentes mostram que nesta representação α e β têm uma significância mais clara quanto às medidas de densidade dos parâmetros de assimetria e borda na Figura 4.5. Em contraste, o parâmetro padrão $M=1$, $X \sim S(\alpha, \beta, \gamma, \delta_1; \mathbf{1})$ com $\beta \neq 0$ tende a $(\text{sign } \beta) \infty$ com $\alpha \uparrow 1$, é próximo δ_1 quando $\alpha = 1$, e tende a $-(\text{sign } \beta) \infty$ com $\alpha \downarrow 1$.

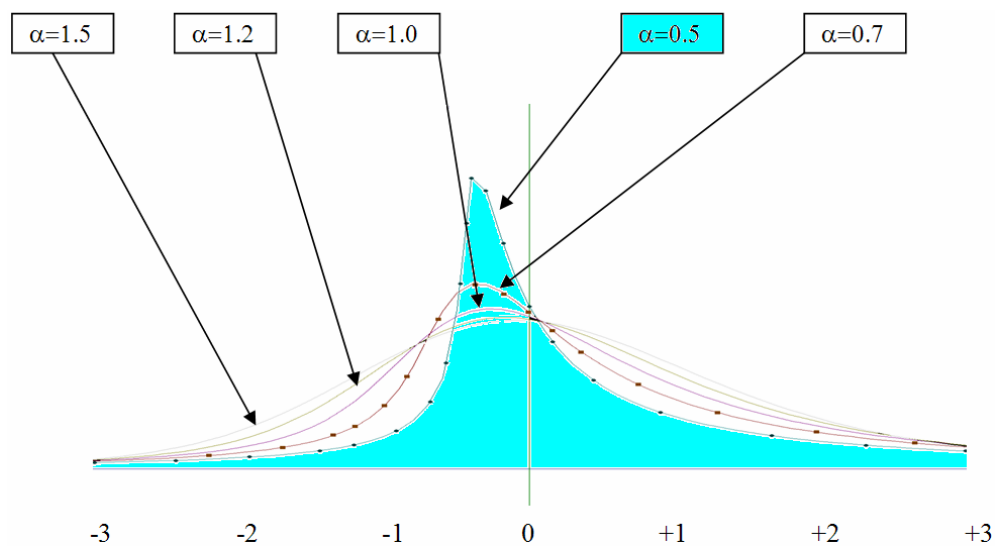


Figura 4.5: Densidades estáveis dentro de $S(\alpha, 0.5, 1, 0; 0)$.

Os parâmetros α , β e γ têm a mesma significância, enquanto o parâmetro δ_1 é diferente de δ_0 nas duas representações para $M=0$ ou 1 , e estão relacionados por:

$$\delta_1 = \delta_0 - \beta \tan(\pi\alpha/2)\gamma \quad \alpha \neq 1$$

$$\delta_1 = \delta_0 - \beta(2/\pi)\gamma \ln \gamma \quad \alpha = 1.$$

Observa-se que os parâmetros $(\alpha, \beta, \gamma, \delta)$ podem ser alterados com facilidade [Nolan, 01 e 05]. A seguir, o Capítulo 5 trata de relacionar as estratégias evolutivas com mutações governadas por distribuições estáveis.

4.3. FUNÇÕES DE TESTE

As funções padrão de teste apresentam diferentes pontos de mínimo global e, representam diferentes desafios com graus de complexidade passíveis de gerar uma falsa convergência ou convergência prematura. Utilizam-se no trabalho quatro funções de teste, analisando-se o comportamento da evolução da população ao longo da busca pelo mínimo global das funções. Utilizou-se o programa *Mathematica*, versão 5.2. [Wolfram, 99].

4.3.1. Função de Rastrigin

Essa função constitui um típico exemplo de função multimodal e não-linear, sendo dada pela expressão analítica abaixo, onde n é o número de componentes x_i , ou seja, a dimensão do espaço de busca na qual a função é definida. A função apresenta um grande número de pontos de mínimo, porém somente um mínimo global, em $f(x) = 0$, para $x_i = 0$.

$$f(x) = A * n + \sum_i^n [x_i^2 - 10 \cos(2\pi x_i)] \quad (4.3)$$

$$-5.12 < x_i < 5.12$$

A superfície desta função é parametrizada por intermédio de duas variáveis externas, A e ω , sendo que esses parâmetros controlam respectivamente a amplitude e a frequência de modulação da função; neste trabalho adotam-se os valores típicos $A=10$ e $\omega = 2\pi$. Na Figura 4.6 mostra-se a função de Rastrigin para $n=2$:

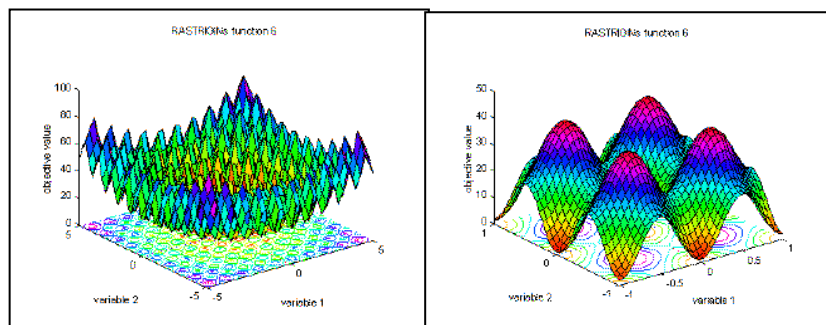


Figura 4.6: Função de Rastrigin [GEATbx, 07].

4.3.2. Função vale de Rosenbrock (Função 2 De Jong)

A função de Rosenbrock, também identificada como segunda função de De Jong, é uma função linear e unimodal, ou seja, contém apenas um ponto de mínimo, sendo que funções multimodais possuem vários pontos de mínimo local, mas apenas um mínimo global. A minimização da função de Rosenbrock é considerada um processo difícil pelo fato de ela possuir um pico muito reduzido e estreito. O topo do pico é muito pouco acentuado e sua superfície se apresenta na forma de uma parábola [Lima, 04]. A função de Rosenbrock é dada pela função abaixo e ilustrada na Figura 4.7:

$$f(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2] \quad (4.4)$$

$$-2.048 < x_i < 2.048$$

A função tem o ponto de mínimo global em $f(x) = 0$, para $x_i = 1$.

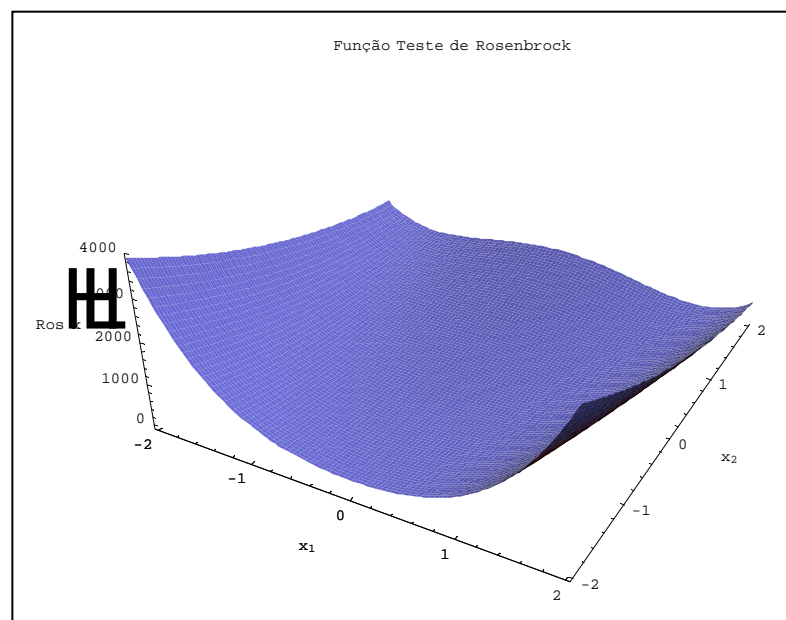


Figura 4.7: Função vale de Rosenbrock (Função 2 De Jong).

4.3.3. Função de Schwefel

A função de Schwefel é uma função não-linear e multimodal, caracterizada por possuir um segundo melhor ponto de mínimo que se encontra bem distante do ponto mínimo global, conforme ilustra a Figura 4.8.

A função de Schwefel é dada pela expressão

$$f(x) = nV + \sum_{i=1}^n -x_i \sin\left(\sqrt{|x_i|}\right) ; \text{ para } -500 < x_i < 500, \quad (4.5)$$

onde, n corresponde às dimensões e o V corresponde a um valor de ajuste que permite mover o ponto de mínimo global convenientemente.

Neste trabalho mantivesse o padrão utilizado em outros estudos, onde o valor de $V = 418,9829$ impõe um deslocamento do ponto de mínimo global $f(x) = n \cdot 418,9829$; $x_i = 420,9687$ [Gómez, 04].

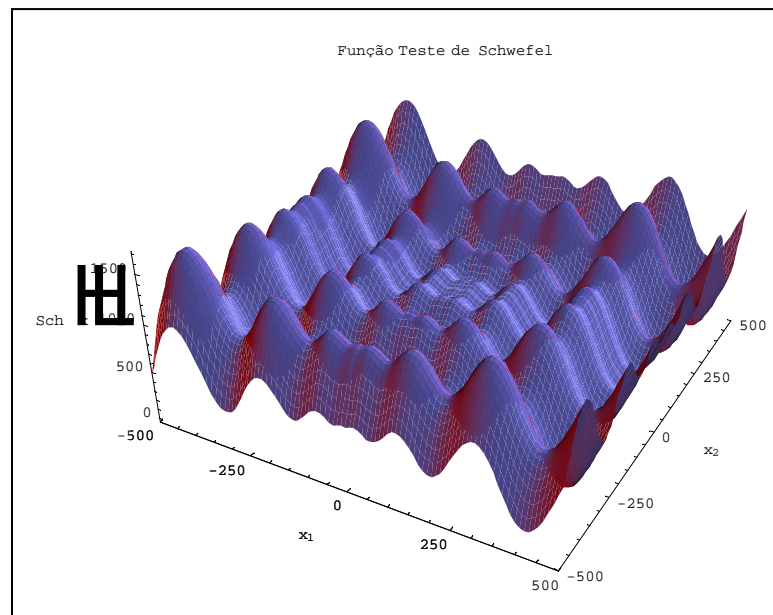


Figura 4.8: Função de Schwefel.

4.3.4. Função de Griewangk

A função de Griewangk é uma função não-linear e multimodal, com complexidade $O(n \cdot \ln(n))$, onde n é o número de parâmetros da função [Gómez, 04], e é dada pela expressão

$$f(x) = 1 + \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1, \quad (4.6)$$

$$-600 < x_i < 600$$

O termo em somatória produz uma parábola, enquanto o termo em produto impõe um aumento dos pontos de mínimos locais, os quais aumentam com a ampliação do espaço, sendo o mínimo global $f(x) = 0$, para $x_i = 0$, como ilustrado na Figura 4.9.

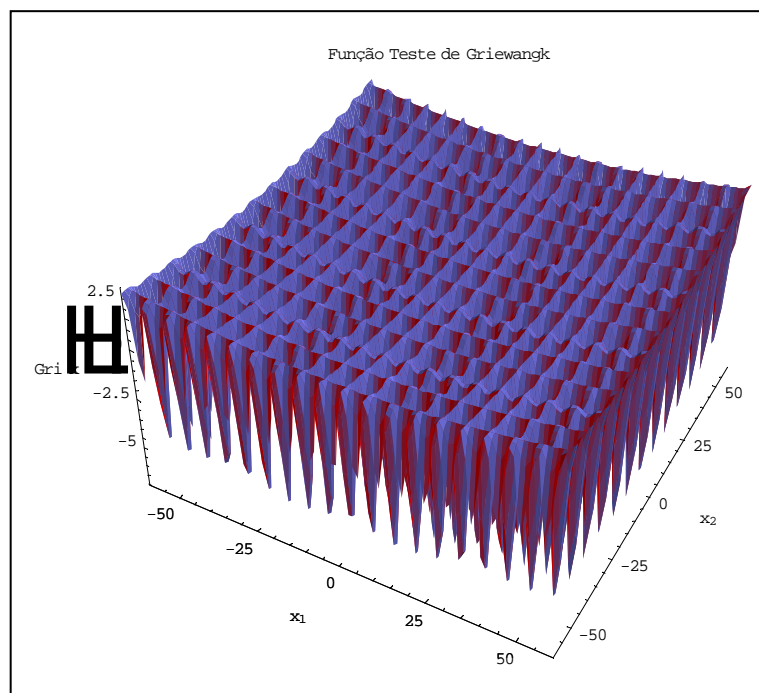


Figura 4.9: Função de Griewangk.

CAPÍTULO V

EXPERIMENTOS COM AS FUNÇÕES DE TESTE

5.1. Experimentos preliminares realizados

5.1.1. Definição dos experimentos

Os experimentos aqui relatados [de Oliveira e Gutierrez, 06] se deram com a função de teste de Rastrigin, exemplo típico de função multimodal e não-linear, com somente o mínimo global $f(x) = 0$, para $x_i = 0$; maiores detalhes foram descritos na Seção 4.3.1.

O objetivo era encontrar o ponto global da função em 2 dimensões, através de estratégias evolutivas em que as mutações são regidas por distribuições estáveis. Foram realizadas: 277 execuções, divididas em 7 parametrizações distintas, organizadas em 3 grupos, cada parametrização definida por um esquema de mutação próprio, governado por uma família de distribuições estáveis (FD). As famílias utilizadas são representadas na Tabela 5.1, e constituem quatro conjuntos genéricos de distribuições estáveis (prefixados pela letra E), e duas famílias específicas, distribuições de Gauss e de Lévy.

Quant. Execs.	FD	Parâmetros das Distribuições				$(\mu/\rho+\lambda)$
		α	β	γ	δ	
Grupo A						
80	E-III	(0, 2]	[0, 1]	1	0	(100/50+90)
40	E-I	(0, 2]	0	[0.01, 1/3]	0	(100/50+90)
80	E-I	(0, 2]	0	[0.01, 1/3]	0	(100/20+90)
40	E-I	(0, 2]	0	[0.01, 1/3]	0	(100/2+90)
Grupo B						
40	E-IV	(0.5, 2]	[0, 1]	1	0	(100/2+90)
40	E-II	(0.5, 2]	0	[0.01, 1/3]	0	(100/2+90)
Grupo C						
20	Gauss	2	0	[0.01, 1/3]	[-1, 1]	(100/2+90)
20	Lévy	0.5	1	[0.01, 1/3]	[-1, 1]	(100/2+90)

Tabela 5.1: Experimentos preliminares com famílias de distribuições estáveis.

A representação do indivíduo ou cromossomo tem a forma $[x_1, x_2, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \{\alpha_{21}, \beta_{22}, \gamma_{23}, \delta_{24}\}]$, onde: x_1, x_2 são os parâmetros objetivos, i.e., as coordenadas da função de teste cujo ponto de mínimo pode se localizar; e $\{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}$ e $\{\alpha_{21}, \beta_{22}, \gamma_{23}, \delta_{24}\}$ são os conjuntos de parâmetros estratégicos, que atuam no operador genético de mutação, respectivamente, das componentes x_1 e x_2 .

Os outros parâmetros de controle da evolução são listados a seguir:

- Intervalo de variação das coordenadas x_1 e x_2 : $\{-2.0, 2.0\}$
- Número de gerações: 500
- Método de seleção: aleatória em toda a população (para os cruzamentos) + elitismo puro de 10%
- Tamanho da população: 100
- Tipos de estratégia ($\mu/\rho+\lambda$):

Grupo A: (100/50+90) ou (100/20+90)

Grupos B e C: (100/2+90)

A diferença básica entre os grupos é que, enquanto o primeiro utiliza cruzamentos múltiplos (com 20 ou até 50 pais), nos outros dois grupos o cruzamento usa apenas um par de pais.

- Mutação dos parâmetros objetivos e estratégicos: Em cada indivíduo a mutação é governada pela distribuição de probabilidade nele definida em seus parâmetros estratégicos, observando-se, naturalmente, que o novo valor gerado esteja dentro do intervalo de variação permitido para o parâmetro em questão. Os parâmetros estratégicos são mutados através da função utilizada no parâmetro de tamanho de passo.

- **Recombinação:** Em todos os casos utilizou-se recombinação *discreta* com recombinação *local* para os parâmetros objetivos, e recombinação *intermediária* com abrangência *local* para os parâmetros estratégicos. As formas *discreta* ou *intermediária* de recombinação dizem respeito, respectivamente, ao fato dos valores que definem a estrutura de um filho ser obtido por um dos valores correspondentes (homólogos) em cada conjunto de pais envolvido, ou pela média dos valores homólogos. A abrangência *local* da recombinação refere-se ao conjunto de pais a serem envolvidos em cada etapa da recombinação ser fixo e pré-definido.
- **Erro:** É o módulo da diferença entre o valor ideal desejado e o valor encontrado na função de aptidão.
- **Tamanho de passo:** A evolução dos parâmetros estratégicos pode incluir a adaptação do tamanho de passo, facilmente incorporada na função e previamente definida como *mutação*. Sendo *tamanho de passo* uma função de um argumento, que modifica os respectivos valores dos parâmetros estratégicos, propiciando a mudança do tipo de distribuição estável, que evolui no decorrer das n-gerações do algoritmo evolutivo. Diferentes funções de adaptação podem agora ser testadas com facilidade [Eiben, Hinterding e Michalewicz, 99; Bäck *et al*, 00; Chen, Liu e Chen, 06; Whitacre, Pham e Sarker, 06].

Este parâmetro de controle possibilita nos experimentos desta seção algo bastante simples, ou seja, se o valor do parâmetro estratégico for menor que 0.5, multiplica-se este por 1.3, caso contrário, dividi-se por 1.3, respeitando-se os intervalos dos parâmetros estratégicos adotados. A escolha desses números “0.5”, se baseia no foco ao redor da distribuição de Lévy ou S-Lévy e o número “1.3” de forma experimental, realizando testes e observando o comportamento dos erros obtidos. Como já dito existem inúmeros métodos e funções que podem governar a mutação dos parâmetros estratégicos com características de mutação. Nesta e nas demais seções foi mantida esta função de adaptação de tamanho de passo. Esta abordagem pode ser modificada para cada um dos parâmetros estratégicos, variando o valor de tomada de decisão e velocidade de incremento ou decremento.

5.1.2. Resultados e discussões

A Tabela 5.2 apresenta um resumo dos resultados obtidos.

Famílias de Distributions	Erros Referentes ao Melhor Indivíduo de Cada Execução			
	Médio	Mínimo	Máximo	Desvio Padrão
Grupo A				
E-III	1.06E-01	1.72E-03	2.93E-01	7.24E-02
E-I	2.08E-04	6.06E-07	9.92E-04	2.32E-04
E-I	1.42E-04	2.33E-06	4.87E-04	1.18E-04
E-I	1.01E-04	3.04E-06	3.99E-04	1.06E-04
Grupo B				
E-IV	5.75E-03	2.38E-05	3.04E-02	5.56E-03
E-II	1.42E-04	5.57E-06	5.87E-04	1.31E-04
Grupo C				
Gauss	4.47E-03	2.12E-04	2.54E-02	6.37E-03
Lévy	3.58E-03	1.24E-05	1.62E-02	3.43E-03

Tabela 5.2: Resultados dos experimentos preliminares.

Para as análises foi considerada a última geração de cada evolução resultante, tomando-se seu melhor indivíduo. Do conjunto de melhores indivíduos tomaram-se então os erros mínimo e máximo associados, e também a média e o desvio padrão desses erros. Essas são as grandezas utilizadas a seguir para a apresentação e discussão dos resultados obtidos.

A primeira constatação a ser feita diz respeito à família E-I do grupo A, onde se observa que a redução da quantidade de pais envolvidos na recombinação (de 50 para 20) produziu maior robustez no processo evolutivo, refletida (conforme mostra a terceira linha do grupo A, na Tabela 5.2) no menor desvio padrão e nos menores valores de erro médio e máximo; entende-se aqui por robustez, o fato de se ter execuções com resultados mais homogêneos, em termos de seu patamar de qualidade. Apesar disso, o melhor resultado do conjunto foi obtido no contexto do maior número de pais (estratégia “100/50+90”), uma ocorrência naturalmente possível, dada a natureza estocástica da busca evolutiva.

A tentativa de diminuir ainda mais a quantidade de pais na estratégia evolutiva, de 20 para 2 (segunda linha do grupo B), não provocou melhora nem de robustez, nem na qualidade em si do melhor resultado, dado que nenhum dos valores das grandezas observadas chegou a diminuir. Ressalte-se, entretanto, que a família de distribuições utilizadas nesse contexto (E-II) não é exatamente a mesma que a utilizada com 20 pais (a família E-I). De fato, observa-se que a família E-II é um subconjunto da família E-I, obtido pela redução do intervalo de variação do parâmetro α . Ficou a dúvida, portanto, se as quedas de robustez e desempenho observadas na estratégia evolutiva “100/2+90” deveram-se à menor variabilidade desta estratégia ou à menor flexibilidade permitida para as mutações neste caso, decorrente da família E-II ser mais restrita em termos de suas distribuições associadas.

Procurando-se responder essa questão, por um lado procedeu-se à diminuição adicional da quantidade de pais na estratégia evolutiva, de 20 para 2, mantendo-se a mesma distribuição original E-I (conforme a quarta linha do grupo A, na Tabela 5.1). Observa-se com isso um novo aumento (ainda que pequeno) na robustez do processo evolutivo, como se pode constatar com os novos valores de desvio padrão e erro médio e máximo resultantes, todos menores ainda que os anteriores (conforme se observa na quarta linha do grupo A, na Tabela 5.2).

Adicionalmente, definiram-se as famílias E-III e E-IV, que diferem exatamente no mesmo intervalo de variação de α acima mencionado, e utilizam um valor fixo do parâmetro γ . Como por construção a família E-IV é um subconjunto da família E-III (segundo a mesma relação que envolve E-I e E-II), associou-se a estratégia com a maior variabilidade aqui utilizada (a “100/50+90”) à família mais ampla em questão (E-III), e com menor variabilidade (a “100/2+90”) à família E-IV. Com isso, os resultados obtidos – apresentados na primeira linha dos grupos A e B, respectivamente, da Tabela 5.2 – permitem constatar claríssima melhoria de robustez e desempenho na situação de menor variabilidade e mutações mais restritas, evidenciada por uma queda acentuada nas quatro figuras de mérito consideradas. Tudo indica, portanto, que a estratégia de menor variabilidade foi a maior responsável pelos ganhos observados, independentemente da família de distribuições consideradas.

Com base nessa conclusão e, agora utilizando esta última estratégia evolutiva (“100/2+90”), procedeu-se a uma comparação entre as famílias específicas de distribuições de Gauss e de Lévy, experimentos que definem o grupo C das Tabelas 5.1 e 5.2. Ressalte-se que ambas as famílias individualmente englobam todas as distribuições de probabilidade possíveis do tipo considerado, cada família parametrizada por quaisquer pares de valores (γ, δ) , respeitados seus domínios de variação. Enquanto a família das Gaussianas é bastante difundida na literatura de computação evolutiva, distribuições de Lévy são pouco utilizadas em conjunto com as mutações em estratégias evolutivas. Observando-se os resultados obtidos, conclui-se que todas as medidas resultantes de erro foram menores no contexto das distribuições de Lévy do que das distribuições de Gauss, uma indicação de que as primeiras possam, de fato, ter permitido uma melhor exploração do espaço de busca considerado.

5.2. Experimentos na função de teste de Rastrigin nas dimensões: 2, 5, 10 e 30

Nesta seção e nas subsequentes varia-se o tipo de distribuição definida pelo parâmetro de estabilidade (α), conforme a Tabela 5.3:

TIPO DE DISTRIBUIÇÃO	PARÂMETRO DE ESTABILIDADE (α)
Gauss	2
Cauchy	1
S-Lévy	0.5 ($\beta=0$)
Lévy	0.5 ($\beta=1$)
Estável 1	0.45
Estável 2	0.375
Estável 3	0.30
Estável 4	0.25

Tabela 5.3: Correlação entre tipo de distribuição e parâmetro de estabilidade (α).

Adicionalmente à escolha do parâmetro de estabilidade, em alguns experimentos varia-se ainda o parâmetro de escala, dentro do intervalo definido pelos valores mínimo $\gamma_{\min}=7.07E-09$ e

máximo $\gamma_{\text{máx}}=7.07\text{E-}01$. Além desses dois parâmetros de definição da família de distribuição devem ser definidos os seguintes parâmetros de controle da evolução:

- Intervalo de variação das coordenadas x_1, \dots, x_n : $\{-5.12, \dots, 5.12\}$
- Número de gerações: 500, 1500, 5000 ou 10000
- Número de execuções para cada experimento: 30
- Método de seleção: aleatória em toda a população (para os cruzamentos) + elitismo puro de 10%
- Tamanho da população: 100
- Tipo de estratégia ($\mu/\rho+\lambda$): (100/2+90)
- Mutação dos parâmetros objetivos e estratégicos: Em cada indivíduo a mutação é governada pela distribuição de probabilidade nele definida em seus parâmetros estratégicos.
- Recombinação: Em todos os casos utilizou-se recombinação *discreta* com abrangência *local* para os parâmetros objetivos, e recombinação *intermediária* com abrangência *local* para os parâmetros estratégicos.
- Erro: É o módulo da diferença entre o valor ideal desejado e o valor encontrado na função de aptidão.
- Auto-adaptação: Apresenta inicialmente a possibilidade de se realizar os seguintes ajustes [Ostermeier, 92; Angeline et al, 96; Weicker, 99]

Intervalo do tamanho de passo: Apresenta a possibilidade de se definir os valores limites de variação dos 4 parâmetros estratégicos de forma individualizada.

onde $0.25 \leq \alpha \leq 2$ e $7.07\text{E-}09 \leq \gamma_{\text{máx}} \leq 7.07\text{E-}01$.

Os resultados de cada experimento são decorrentes de 30 execuções, e apresentados em termos dos valores médios (sobre as 30 execuções) do erro mínimo, máximo, e médio obtidos, bem como do desvio padrão associado aos erros máximos; a Tabela 5.4 exemplifica a apresentação dos resultados:

PARÂMETROS (γ)	PARÂMETROS (α) PARA AS DIVERSAS DIMENSÕES (2, 5, 10 E 30)							
	Distribuições estáveis nomeadas				Distribuições estáveis "não" nomeadas			
	Gauss	Cauchy	S-Lévy	Lévy	Estável 1	Estável 2	Estável 3	Estável 4
	2.0	1.0	0.5($\beta=0$)	0.5($\beta=1$)	0.45	0.375	0.30	0.25
7.07E-01*	Os quatro dados básicos colocados em todas as tabelas são: erro mínimo, erro máximo, erro médio e o desvio padrão dos erros considerando 30 execuções para uma população de 100 indivíduos por 500 gerações.							
7.07E-02								
7.07E-03								
7.07E-04								
7.07E-05								
	Mínimo	Máximo	Média	Desvio Padrão				
7.07E-06	Com o desenvolvimento dos experimentos torna-se notória a dependência dos resultados da função de distribuição estável (α) e de seu respectivo parâmetro de escala (γ).							
7.07E-07								
7.07E-08								
7.07E-09								
7.07E-10								
7.07E-11**	* Especificamente realizados na função de teste de Griewangk. ** Especificamente realizados na função de teste de Schwefel							
7.07E-12**								
7.07E-13**								

Tabela 5.4: Esquema geral utilizado no texto para apresentação dos resultados.

5.2.1. Experimentos com função de Rastrigin: 2 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, x_2, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \{\alpha_{21}, \beta_{22}, \gamma_{23}, \delta_{24}\}]$, onde: x_1, x_2 são os parâmetros objetivos, i.e., as coordenadas da função de teste cujo ponto de mínimo pode se localizar; e $\{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}$ e $\{\alpha_{21}, \beta_{22}, \gamma_{23}, \delta_{24}\}$ são os conjuntos de parâmetros estratégicos, que atuam no operador genético de mutação, respectivamente, das componentes x_1 e x_2 .

Nas Tabelas 5.5 e 5.6 pode-se observar que, à medida o parâmetro de escala (γ) é diminuído, os resultados obtidos dos erros mínimos vão melhorando na distribuição de Gauss, Cauchy, S-Lévy e Lévy, respectivamente, até os experimentos 6, 7, 15 e 16, voltando a aumentar a partir destes, sinalizando que não convém reduzir mais o parâmetro de escala. Para simplificar os nomes das tabelas e seus respectivos cabeçalhos, as distribuições são mencionadas pelo seu nome característico, e, no caso da distribuição Gaussiana, esta é denominada simplesmente Gauss.

Exper.	Par. Escala	Gauss (0.1, 0.1, 0.1) -> Erro					Cauchy (0.1, 0.1, 0.1) -> Erro			
		γ	σ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média
1	7.07E-02	1.00E-01	9.21E-07	3.93E-04	1.15E-04	1.09E-04	7.57E-07	2.67E-04	7.97E-05	8.25E-05
2	7.07E-03	1.00E-02	6.91E-08	3.02E-06	7.96E-07	7.65E-07	7.80E-08	2.01E-06	5.96E-07	5.25E-07
3	7.07E-04	1.00E-03	1.15E-11	3.54E-08	9.05E-09	9.26E-09	3.97E-10	4.21E-08	9.81E-09	1.11E-08
4	7.07E-05	1.00E-04	2.00E-12	2.99E-10	7.67E-11	7.10E-11	4.71E-12	3.32E-10	7.91E-11	7.55E-11
5	7.07E-06	1.00E-05	2.13E-14	5.09E-03	1.70E-04	9.29E-04	1.42E-14	3.21E-12	8.12E-13	6.61E-13
6	7.07E-07	1.00E-06	3.55E-15	6.75E-14	1.09E-14	1.33E-14	3.55E-15	1.42E-14	5.68E-15	5.49E-15
7	7.07E-08	1.00E-07	2.58E-11	9.59E-03	3.29E-04	1.75E-03	0	0	0	0
8	7.07E-09	1.00E-08	7.60E-06	2.37E-05	1.04E-06	4.51E-06				

Tabela 5.5: Resultados com a função de Rastrigin: Gauss e Cauchy em 2 dimensões.

Exper.	Par. Escala	S-Lévy (0.1, 0.1, 0.1) -> Erro				Lévy (0.1, 0.1, 0.1) -> Erro			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
9	7.07E-02	4.57E-07	7.77E-05	1.86E-05	1.81E-05	1.24E-03	1.05E+00	1.52E-01	2.45E-01
10	7.07E-03	6.73E-11	7.45E-07	1.81E-07	2.11E-07	1.64E-05	2.02E-02	2.77E-03	5.35E-03
11	7.07E-04	6.52E-11	9.03E-09	2.57E-09	2.42E-09	6.34E-08	1.13E-04	1.95E-05	2.94E-05
12	7.07E-05	5.83E-13	8.56E-11	1.85E-11	1.77E-11	9.25E-10	4.77E-06	4.13E-07	9.69E-07
13	7.07E-06	2.13E-14	8.03E-13	2.19E-13	2.24E-13	6.54E-11	9.73E-05	3.25E-06	1.78E-05
14	7.07E-07	3.55E-15	7.11E-15	1.89E-15	2.59E-15	1.10E-13	8.20E-10	3.57E-11	1.49E-10
15	7.07E-08	0	0	0	0	3.55E-15	7.92E-09	3.00E-10	1.45E-09
16	7.07E-09					0	0	0	0

Tabela 5.6: Resultados com a função de Rastrigin: S-Lévy e Lévy em 2 dimensões.

O melhor resultado obtido foi na família de distribuição com cauda pesada ou seja, nas distribuições de Cauchy, S-Lévy e Lévy. Note-se que a função envolvida não oferece um grande desafio em 2 dimensões.

5.2.2. Experimentos com função de Rastrigin: 5 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, \dots, x_5, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \dots, \{\alpha_{51}, \beta_{52}, \gamma_{53}, \delta_{54}\}]$. Nas Tabelas 5.7 e 5.8 pode-se observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 20, 22, 32, 30, 40 (esquerda), 40 (direita), 48 (esquerda) e 48 (direita).

Exper.	Par. Escala	Gauss ()->Erro					Cauchy ()->Erro			
		σ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
17	7.07E-02	1.00E-01	4.28E-02	2.10E+00	5.52E-01	5.68E-01	2.85E-02	1.42E+00	2.36E-01	2.92E-01
18	7.07E-03	1.00E-02	3.21E-04	1.99E+00	6.64E-01	6.58E-01	1.70E-04	3.84E-03	1.06E-03	7.94E-04
19	7.07E-04	1.00E-03	1.70E-06	1.99E+00	6.63E-01	6.58E-01	3.89E-06	9.95E-01	1.36E-01	3.43E-01
20	7.07E-05	1.00E-04	4.88E-08	1.72E+00	4.92E-01	5.02E-01	3.03E-08	9.95E-01	4.32E-01	5.01E-01
21	7.07E-06	1.00E-05	1.66E-02	3.39E+00	9.60E-01	7.07E-01	3.90E-10	1.99E+00	6.96E-01	6.48E-01
22	7.07E-07	1.00E-06	3.14E-03	2.67E+00	7.39E-01	8.46E-01	4.86E-12	3.03E+00	7.31E-01	7.63E-01
23	7.07E-08	1.00E-07	3.77E-03	5.16E+00	1.66E+00	1.57E+00	1.58E-04	2.59E+00	9.19E-01	6.87E-01
24	7.07E-09	1.00E-08	3.49E-04	6.66E+00	1.09E+00	1.43E+00	3.03E-03	4.24E+00	1.15E+00	1.07E+00

Tabela 5.7: Resultados com a função de Rastrigin: Gauss e Cauchy em 5 dimensões.

Exper.	Par. Escala	S-Lévy ()->Erro				Lévy ()->Erro			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
25	7.07E-02	2.69E-02	1.05E+00	1.79E-01	2.27E-01	4.23E+00	9.76E+00	7.02E+00	1.43E+00
26	7.07E-03	4.24E-05	1.34E-03	5.70E-04	3.07E-04	5.33E-01	4.94E+00	2.35E+00	9.86E-01
27	7.07E-04	1.29E-06	9.64E-06	5.14E-06	2.04E-06	4.89E-02	2.89E+00	9.02E-01	8.77E-01
28	7.07E-05	8.74E-09	3.23E-07	6.91E-08	5.74E-08	2.82E-03	2.34E+00	6.34E-01	6.65E-01
29	7.07E-06	1.51E-10	9.95E-01	4.60E-02	1.88E-01	4.92E-06	2.01E+00	6.37E-01	6.42E-01
30	7.07E-07	1.90E-12	9.95E-01	1.70E-01	3.50E-01	2.73E-07	2.92E+00	8.26E-01	8.85E-01
31	7.07E-08	4.26E-14	1.99E+00	5.31E-01	6.25E-01	9.91E-06	2.15E+00	5.76E-01	5.77E-01
32	7.07E-09	0.00E+00	1.99E+00	5.25E-01	7.69E-01	2.93E-04	3.15E+00	6.56E-01	8.09E-01
Exper.	Par. Escala	Estável 1 ()->Erro				Estável 2 ()->Erro			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
33	7.07E-02	1.34E-02	1.32E+00	1.44E-01	2.47E-01	1.05E-02	1.11E+00	1.64E-01	2.67E-01
34	7.07E-03	7.37E-05	2.56E-03	6.05E-04	4.88E-04	7.16E-05	1.04E-03	4.95E-04	2.55E-04
35	7.07E-04	4.80E-07	1.41E-05	4.93E-06	2.83E-06	1.65E-07	1.53E-05	4.79E-06	3.04E-06
36	7.07E-05	5.05E-09	1.58E-07	5.54E-08	3.39E-08	9.01E-09	8.30E-08	4.37E-08	1.96E-08
37	7.07E-06	1.57E-10	9.95E-01	4.90E-02	1.97E-01	1.36E-10	1.65E-09	5.68E-10	3.53E-10
38	7.07E-07	1.09E-12	9.95E-01	4.62E-02	1.88E-01	1.14E-12	1.90E-01	6.33E-03	3.47E-02
39	7.07E-08	2.13E-14	1.99E+00	3.83E-01	6.08E-01	2.84E-14	1.04E+00	2.29E-01	4.15E-01
40	7.07E-09	0.00E+00	1.99E+00	5.75E-01	6.68E-01	0.00E+00	1.99E+00	3.26E-01	5.44E-01
Exper.	Par. Escala	Estável 3 ()->Erro				Estável 4 ()->Erro			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
41	7.07E-02	3.96E-02	1.20E+00	2.93E-01	2.78E-01	5.20E-02	1.63E+00	4.31E-01	3.73E-01
42	7.07E-03	9.68E-05	1.14E+00	4.07E-02	2.08E-01	5.41E-05	2.47E-01	1.15E-02	4.64E-02
43	7.07E-04	1.61E-06	2.15E-05	5.02E-06	3.81E-06	3.14E-07	1.45E-05	4.40E-06	2.56E-06
44	7.07E-05	9.35E-09	1.28E-07	4.87E-08	2.61E-08	9.85E-09	1.24E-07	5.55E-08	3.11E-08
45	7.07E-06	2.37E-11	3.08E-09	5.62E-10	5.43E-10	5.97E-11	4.14E-09	9.47E-10	9.29E-10
46	7.07E-07	2.27E-12	3.45E-08	1.16E-09	6.29E-09	2.81E-12	1.72E-08	5.82E-10	3.13E-09
47	7.07E-08	2.84E-14	4.08E-09	1.37E-10	7.45E-10	4.26E-14	3.64E-12	5.02E-13	9.07E-13
48	7.07E-09	0.00E+00	9.66E-01	4.60E-02	1.90E-01	0.00E+00	8.36E-06	2.78E-07	1.53E-06

Tabela 5.8: Resultados com a função de Rastrigin: S-Lévy, Lévy, Estável 1 a Estável 4 em 5 dimensões.

Obteve-se o melhor resultado na família de distribuição com cauda simétrica, ou seja, nas distribuições de Cauchy, S-Lévy e Estável 1 a Estável 4.

5.2.3. Experimentos com função de Rastrigin: 10 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, \dots, x_{10}, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \dots, \{\alpha_{101}, \beta_{102}, \gamma_{103}, \delta_{104}\}]$. Nas Tabelas 5.9 e 5.10 pode-se observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 51, 52, 62, 63, 70, 71, 81 (esquerda) e 81 (direita).

Exper.	Par. Escala	Gauss (Gauss) --> Erro					Cauchy (Cauchy) --> Erro			
		γ	σ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média
49	7.07E-02	1.00E-01	1.17E+00	6.78E+00	2.83E+00	1.25E+00	7.91E-01	5.46E+00	2.95E+00	1.35E+00
50	7.07E-03	1.00E-02	1.42E-02	3.00E+00	1.61E+00	8.51E-01	9.49E-03	2.40E+00	5.18E-01	7.04E-01
51	7.07E-04	1.00E-03	1.32E-04	5.73E+00	2.49E+00	1.55E+00	1.46E-04	3.31E+00	9.08E-01	8.66E-01
52	7.07E-05	1.00E-04	3.84E+00	3.15E+01	1.55E+01	6.93E+00	1.60E-06	4.97E+00	1.68E+00	1.05E+00
53	7.07E-06	1.00E-05	6.49E+00	2.93E+01	1.69E+01	6.20E+00	1.50E-02	6.15E+00	2.33E+00	1.59E+00
54	7.07E-07	1.00E-06	3.10E+00	3.36E+01	1.70E+01	8.13E+00	6.11E+00	2.95E+01	1.45E+01	5.39E+00
55	7.07E-08	1.00E-07	6.13E+00	3.35E+01	1.81E+01	7.64E+00	2.89E+00	2.80E+01	1.32E+01	7.12E+00
56	7.07E-09	1.00E-08	5.14E+00	3.01E+01	1.88E+01	7.13E+00	1.63E+00	3.11E+01	1.84E+01	6.28E+00

Tabela 5.9: Resultados com a função de Rastrigin: Gauss, Cauchy em 10 dimensões.

Exper.	Par. Escala	S-Lévy (S-Lévy) --> Erro				Lévy (Lévy) --> Erro			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média
57	7.07E-02	8.13E+00	2.53E+01	1.56E+01	3.93E+00	3.09E+01	5.21E+01	3.99E+01	4.98E+00
58	7.07E-03	1.40E-02	2.32E+00	8.00E-01	7.44E-01	9.48E+00	2.08E+01	1.50E+01	2.86E+00
59	7.07E-04	2.90E-04	1.11E+00	1.45E-01	3.12E-01	3.47E+00	1.20E+01	7.80E+00	2.23E+00
60	7.07E-05	3.16E-06	2.00E+00	8.00E-02	3.66E-01	1.43E+00	6.90E+00	4.16E+00	1.55E+00
61	7.07E-06	5.95E-08	1.99E+00	4.58E-01	6.01E-01	1.92E+00	8.09E+00	4.02E+00	1.63E+00
62	7.07E-07	1.96E-09	3.98E+00	1.20E+00	1.14E+00	1.24E+00	7.11E+00	4.17E+00	1.63E+00
63	7.07E-08	3.41E-05	3.98E+00	1.29E+00	1.23E+00	1.10E+00	1.61E+01	5.89E+00	3.53E+00
64	7.07E-09	1.63E-03	5.17E+00	1.71E+00	1.29E+00	4.61E+00	1.72E+01	7.61E+00	3.62E+00
Exper.	Par. Escala	Estável 1 (Estável 1) --> Erro				Estável 2 (Estável 2) --> Erro			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média
65	7.07E-02	7.84E+00	2.61E+01	1.71E+01	4.15E+00	1.51E+01	3.06E+01	2.26E+01	3.94E+00
66	7.07E-03	3.78E-02	4.27E+00	1.26E+00	1.16E+00	5.16E-02	4.25E+00	1.35E+00	1.12E+00
67	7.07E-04	3.26E-04	2.01E+00	1.38E-01	4.00E-01	6.89E-04	1.56E+00	2.20E-01	4.32E-01
68	7.07E-05	3.28E-06	9.97E-01	4.13E-02	1.82E-01	3.05E-06	6.42E-03	4.36E-04	1.28E-03
69	7.07E-06	4.30E-08	1.30E+00	1.46E-01	3.77E-01	1.27E-07	1.01E+00	1.82E-01	3.56E-01
70	7.07E-07	8.94E-10	1.99E+00	4.07E-01	6.69E-01	3.62E-09	1.44E+00	3.38E-01	4.80E-01
71	7.07E-08	2.09E-08	3.98E+00	1.33E+00	1.01E+00	1.33E-10	2.45E+00	6.40E-01	8.43E-01
72	7.07E-09	4.63E-03	5.00E+00	2.14E+00	1.37E+00	3.97E-09	3.98E+00	1.37E+00	9.95E-01
73	7.07E-10	1.82E-02	4.25E+00	2.24E+00	1.23E+00	3.44E-04	3.98E+00	1.44E+00	1.12E+00
Exper.	Par. Escala	Estável 3 (Estável 3) --> Erro				Estável 4 (Estável 4) --> Erro			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média
74	7.07E-02	1.86E+01	4.18E+01	2.99E+01	6.14E+00	2.07E+01	4.78E+01	3.37E+01	7.78E+00
75	7.07E-03	2.85E-01	5.54E+00	2.50E+00	1.64E+00	2.07E+00	2.00E+01	1.01E+01	3.66E+00
76	7.07E-04	1.92E-03	1.42E+00	3.43E-01	4.78E-01	9.40E-03	3.48E+00	8.52E-01	9.23E-01
77	7.07E-05	1.18E-05	1.13E+00	6.20E-02	2.31E-01	1.85E-04	3.47E+00	3.40E-01	7.90E-01
78	7.07E-06	5.72E-07	1.11E+00	7.11E-02	2.68E-01	2.03E-06	1.01E+00	1.14E-01	2.81E-01
79	7.07E-07	8.02E-09	1.07E+00	1.11E-01	3.17E-01	6.84E-08	1.00E+00	3.40E-02	1.83E-01
80	7.07E-08	4.78E-10	3.98E+00	2.00E-01	7.45E-01	3.12E-09	1.99E+00	1.44E-01	4.32E-01
81	7.07E-09	2.91E-10	3.99E+00	4.93E-01	8.95E-01	1.03E-09	1.09E+00	2.50E-01	3.93E-01
82	7.07E-10	8.02E-08	2.58E+00	7.87E-01	7.60E-01	6.54E-09	2.99E+00	4.82E-01	7.48E-01

Tabela 5.10: Resultados com a função de Rastrigin: S-Lévy, Lévy e Estável 1 a Estável 4 em 10 dimensões.

O melhor resultado foi na família de distribuição com cauda de simétrica Estável 2.

5.2.4. Experimentos com função de Rastrigin: 30 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, x_2, \dots, x_{29}, x_{30}, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \{\alpha_{21}, \beta_{22}, \gamma_{23}, \delta_{24}\}, \dots, \{\alpha_{29\ 1}, \beta_{29\ 2}, \gamma_{29\ 3}, \delta_{29\ 4}\}, \{\alpha_{30\ 1}, \beta_{30\ 2}, \gamma_{30\ 3}, \delta_{30\ 4}\}]$. As Tabelas 5.11 e 5.12 permitem observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 84, 85, 90, 89, 95, 96, 97, 100 (esquerda) e 100 (direita).

Exper.	Par. Escala	Gauss ()-->Erro					Cauchy ()-->Erro			
		γ	σ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média
83	7.07E-02	1.00E-01	1.73E+01	3.23E+01	2.63E+01	3.69E+00				
84	7.07E-03	1.00E-02	3.20E+00	2.01E+01	8.04E+00	3.62E+00	2.48E+00	1.23E+01	7.03E+00	2.76E+00
85	7.07E-04	1.00E-03	5.57E+01	1.31E+02	9.39E+01	1.68E+01	2.37E+00	1.07E+01	6.11E+00	1.84E+00
86	7.07E-05	1.00E-04	1.09E+02	1.78E+02	1.49E+02	1.72E+01	4.08E+00	1.23E+01	7.93E+00	2.34E+00
87	7.07E-06	1.00E-05					3.44E+01	9.45E+01	6.38E+01	1.33E+01

Tabela 5.11: Resultados com a função de Rastrigin: Gauss e de Cauchy em 30 dimensões.

Exper.	Par. Escala	S-Lévy ()-->Erro				Lévy ()-->Erro			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
88	7.07E-05	3.93E+00	1.81E+01	1.02E+01	3.36E+00	2.23E+01	4.19E+01	3.18E+01	4.93E+00
89	7.07E-06	2.09E+00	1.14E+01	5.79E+01	2.33E+01	9.54E+00	3.73E+01	2.29E+01	6.42E+00
90	7.07E-07	1.05E+00	1.06E+01	6.41E+00	2.47E+00	1.14E+01	3.47E+01	2.32E+01	6.65E+00
91	7.07E-08	5.10E+00	1.62E+01	9.47E+00	2.76E+00	1.72E+01	5.08E+01	3.32E+01	9.64E+00
Exper.	Par. Escala	Estável 1 ()-->Erro				Estável 2 ()-->Erro			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
92	7.07E-05	1.23E+00	1.01E+01	5.66E+00	2.09E+00	3.27E+00	9.29E+00	5.84E+00	1.69E+00
93	7.07E-06	1.03E+00	1.20E+01	5.50E+00	2.39E+00				
94	7.07E-07	5.17E-01	1.01E+01	5.36E+00	2.00E+00	1.04E+00	1.30E+01	5.01E+00	2.70E+00
95	7.07E-08	3.03E+00	1.25E+01	7.21E+00	2.23E+00	1.06E+00	1.05E+01	6.30E+00	2.07E+00
96	7.07E-09					9.13E-01	1.38E+01	7.16E+00	2.63E+00
97	7.07E-10					3.77E+00	1.44E+01	9.90E+00	2.45E+00
Exper.	Par. Escala	Estável 3 ()-->Erro				Estável 4 ()-->Erro			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
98	7.07E-07					3.20E+00	1.34E+01	6.92E+00	2.53E+00
99	7.07E-08	1.16E+00	1.51E+01	5.59E+00	2.95E+00	1.40E+00	1.20E+01	5.59E+00	2.01E+00
100	7.07E-09	3.23E-01	1.31E+01	6.44E+00	2.22E+00	4.54E-01	1.05E+01	5.58E+00	2.29E+00
101	7.07E-10	2.80E+00	1.53E+01	6.94E+00	2.98E+00	1.96E+00	1.03E+01	5.85E+00	2.24E+00

Tabela 5.12: Resultados com a função de Rastrigin: S-Lévy, de Lévy e Estável 1 a Estável 4 em 30 dimensões.

O melhor resultado dos experimentos de 1 a 101 foi na família de distribuição com cauda simétrica Estável 3 (na cor azul), com erros de até 10^{-1} .

Em conclusão, o conjunto de experimentos aqui relatados, com progressivo aumento da dimensão da função de Rastrigin indica que 2 dimensões mostraram-se não relevantes na condução dopara o resultado final, podendo-se, portanto, iniciar-se os experimentos posteriores em 5 dimensões.

5.2.5. Dados dos experimentos para a distribuição Estável 3

Como verificado na seção anterior, o melhor resultado foi obtido com a distribuição Estável 3, o que torna esta distribuição merecedora de maior atenção. Nesse sentido, verifica-se agora a evolução dos dados com quantidade variável de execuções, de 3 a 30, independentes entre si, 30-dimensão na função de teste de Rastrigin, e variando-se o número de gerações de 500, 1500, 5000 e 10000.

Com isso, tem-se o propósito de verificar em que condições é mais interessante executar experimentos no futuro, ou seja, com um número maior de gerações associado a um número menor de execuções, ou o contrário. Vale lembrar que reduzindo significativamente o número de execuções, pode-se perder a relevância da amostra (repetibilidade).

Na Tabela 5.13 observa-se que os resultados obtidos para os erros mínimos vão melhorando na distribuição Estável 3, quando aumenta-se o esforço computacional até 10000 gerações, mesmo com a redução do número de execuções.

Exper.	Execuções	Estável 3 (2011-04-19 09:32:00-09:41)					Estável 3 (2011-04-19 09:32:00-09:41)			
		Erro para 500 gerações					Erro para 1500 gerações			
		Mínimo	Máximo	Média	Desv. Padrão		Mínimo	Máximo	Média	Desv. Padrão
102	3.00E+00	2.13E+00	1.23E+01	6.14E+00	5.39E+00	1.05E+00	3.98E+00	2.67E+00	1.49E+00	
103	6.00E+00	1.40E+00	1.03E+01	5.19E+00	3.29E+00	3.61E-06	1.57E+00	4.31E-01	6.86E-01	
104	9.00E+00	3.06E+00	9.12E+00	6.48E+00	2.08E+00	1.73E-04	4.01E+00	1.28E+00	1.23E+00	
105	1.20E+01	3.11E+00	8.15E+00	5.12E+00	1.73E+00	6.25E-06	3.06E+00	7.86E-01	9.60E-01	
106	1.50E+01	2.47E+00	9.13E+00	5.63E+00	1.61E+00	2.79E-06	4.15E+00	1.16E+00	1.31E+00	
107	1.80E+01	2.24E+00	1.03E+01	5.97E+00	2.13E+00					
108	2.10E+01	3.06E+00	1.23E+01	7.13E+00	2.39E+00					
109	2.40E+01	4.04E+00	1.38E+01	6.80E+00	2.26E+00					
110	2.70E+01	1.16E+00	9.34E+00	5.62E+00	2.21E+00					
111	3.00E+01	2.45E+00	1.19E+01	5.98E+00	2.25E+00					
Exper.	Execuções	Erro para 5000 gerações					Erro para 10000 gerações			
		Mínimo	Máximo	Média	Desv. Padrão		Mínimo	Máximo	Média	Desv. Padrão
		112	3.00E+00	2.87E-10	6.21E-10	4.02E-10	1.90E-10	7.73E-11	1.34E-10	1.07E-10
113	6.00E+00	2.38E-10	6.48E-10	5.37E-10	1.54E-10	8.23E-11	1.35E-10	1.11E-10	2.25E-11	
114	9.00E+00	2.89E-10	8.78E-10	5.33E-10	1.77E-10					

Tabela 5.13: Resultados com a função de Rastrigin: Estável 3 para 500, 1500, 5000 e 10000 gerações.

Os melhores resultados para 500, 1500, 5000 e 10000 gerações foram respectivamente nos experimentos 110, 106, 113 e 112, e, no conjunto de experimentos, verifica-se que sempre é vantajoso aumentar o número de gerações, independente do número de execuções.

5.2.6. Auto-adaptação dos parâmetros estratégicos α e γ em 30 dimensões

A motivação para executar os experimentos desta seção reside no fato de analisar se é possível deixar o processo evolutivo descobrir uma família de distribuições estáveis que acarrete melhor nível de qualidade do que os experimentos com famílias pré-definidas. Ressalte-se que este novo conjunto de experimentos se encontra dentro das mesmas parametrizações que as existentes na seção anterior mas distintas das usadas na Seção 5.1, que também tratou de auto-adaptação.

Sendo assim, inicialmente define-se um intervalo de valores válidos para os parâmetros estratégicos: índice de estabilidade $\alpha \in [0.25, 2]$; índice de assimetria $\beta = 0$; de escala $\Upsilon \in [7.07E-9, 7.07E-2]$, e de posição $\delta = 0$. Os parâmetros estratégicos podem variar livremente para cada dimensão de forma independente. Foi observado o valor dos índices de estabilidade e escala para a melhor execução, no conjunto de execuções de cada experimento, conforme a Tabela 5.14.

Exper.	Execuç.	Distrib. Estáveis (Distribuição de Cauchy)						Distrib. Estáveis (Distribuição de Cauchy)						
		Erro para 500 gerações						Erro para 1500 gerações						
		Mínimo	Máximo	Média	Desv. Padr.	Estabil. (σ)	Escala (γ)	Mínimo	Máximo	Média	Desv. Padr.	Estabil. (σ)	Escala (γ)	
115	3.00E+00	3.16E+01	5.13E+01	4.20E+01	9.90E+00	1.14E+00	1.47E-01	4.05E+01	4.61E+01	4.30E+01	2.86E+00	1.17E+00	4.51E-02	
116	6.00E+00	4.67E+01	5.75E+01	5.40E+01	3.98E+00	1.37E+00	3.51E-02	2.97E+01	4.31E+01	3.76E+01	4.95E+00	1.09E+00	3.33E-02	
117	9.00E+00	4.14E+01	6.37E+01	5.33E+01	7.61E+00	1.18E+00	1.88E-02	3.11E+01	4.72E+01	3.83E+01	5.86E+00	1.15E+00	5.04E-02	
118	1.20E+01	3.96E+01	5.57E+01	4.71E+01	5.89E+00	1.11E+00	1.02E-01	2.97E+01	5.22E+01	3.79E+01	6.80E+00	1.17E+00	8.33E-02	
119	1.50E+01	3.11E+01	7.85E+01	4.79E+01	1.09E+01	1.13E+00	2.95E-02	3.37E+01	5.70E+01	4.29E+01	7.00E+00	9.64E-01	4.12E-02	
120	1.80E+01	2.56E+01	6.40E+01	4.74E+01	1.09E+01	1.13E+00	8.72E-02	2.73E+01	5.15E+01	3.99E+01	6.66E+00	1.23E+00	5.80E-02	
121	2.10E+01	3.08E+01	6.90E+01	4.91E+01	8.15E+00	1.02E+00	5.59E-02	2.68E+01	4.79E+01	3.87E+01	5.17E+00	1.10E+00	2.52E-02	
122	2.40E+01	3.18E+01	6.18E+01	4.86E+01	7.64E+00	1.16E+00	4.42E-02	3.17E+01	5.36E+01	4.08E+01	6.05E+00	1.01E+00	7.48E-02	
123	2.70E+01	3.15E+01	7.03E+01	4.82E+01	1.01E+01	1.07E+00	5.47E-02	3.10E+01	5.84E+01	4.11E+01	5.89E+00	1.16E+00	3.90E-02	
124	3.00E+01	3.19E+01	6.93E+01	4.88E+01	8.00E+00	1.18E+00	5.45E-02	2.80E+01	5.11E+01	4.00E+01	6.95E+00	1.25E+00	2.75E-02	
Exper.	Execuç.	Erro para 5000 gerações						Erro para 10000 gerações						
		Mínimo	Máximo	Média	Desv. Padr.	Estabil. (σ)	Escala (γ)	Mínimo	Máximo	Média	Desv. Padr.	Estabil. (σ)	Escala (γ)	
		125	3.00E+00	2.00E+01	3.93E+01	2.77E+01	1.02E+01	1.26E+00	8.28E-02	3.00E+01	3.26E+01	3.16E+01	1.35E+00	1.22E+00
126	6.00E+00	2.77E+01	3.28E+01	3.00E+01	1.89E+00	1.34E+00	1.65E-02	2.44E+01	3.15E+01	2.81E+01	2.59E+00	1.11E+00	4.50E-02	
127	9.00E+00	2.90E+01	5.23E+01	4.05E+01	8.09E+00	1.07E+00	1.88E-02	2.34E+01	4.02E+01	3.14E+01	5.47E+00	1.11E+00	6.01E-02	
128	1.20E+01	3.32E+01	5.49E+01	4.19E+01	7.33E+00	1.18E+00	4.78E-02	2.00E+01	3.44E+01	2.82E+01	4.27E+00	1.20E+00	2.05E-02	
129	1.50E+01	2.34E+01	4.27E+01	3.19E+01	4.98E+00	1.01E+00	6.18E-02	1.46E+01	3.99E+01	3.06E+01	5.95E+00	1.15E+00	6.05E-02	
130	1.80E+01	3.01E+01	5.24E+01	3.89E+01	5.08E+00	1.14E+00	8.16E-02	2.10E+01	4.07E+01	3.07E+01	5.97E+00	9.98E-01	3.58E-02	
131	2.10E+01	2.62E+01	5.00E+01	3.91E+01	6.97E+00	1.13E+00	3.46E-02	2.23E+01	3.75E+01	2.97E+01	4.42E+00	1.09E+00	8.98E-02	
132	2.40E+01	2.50E+01	5.57E+01	3.87E+01	8.61E+00	1.09E+00	1.09E+00	7.46E-02	2.33E+01	4.11E+01	3.03E+01	4.17E+00	1.06E+00	8.18E-02
133	2.70E+01	2.85E+01	5.67E+01	4.10E+01	6.66E+00	1.15E+00	1.17E-01	2.05E+01	3.75E+01	2.98E+01	4.77E+00	1.06E+00	3.32E-02	
134	3.00E+01	2.61E+01	5.60E+01	4.16E+01	6.92E+00	1.04E+00	1.01E-01	2.08E+01	4.07E+01	2.93E+01	5.57E+00	1.20E+00	6.91E-02	

Tabela 5.14: Resultados com a função de Rastrigin: Auto-adaptação livre nas 30 dimensões..

Os melhores resultados para 500, 1500, 5000 e 10000 gerações ocorreram nos experimentos 120, 121, 125 e 129, respectivamente. Observando-se os valores dos parâmetros de estabilidade e de escala, e comparando-se com os obtidos na seção anterior, constata-se que os valores encontrados no presente experimento ficaram muito aquém do obtido com 10000 gerações da distribuição Estável 3 (experimento 112), cujo erro mínimo foi de $7.73E-11$.

Na Seção 5.2.8 discute-se melhor a interrelação dos dados experimentais, após a ampliação dos experimentos da seção seguinte, em se que fixam os parâmetros estratégicos para todas as 30 dimensões.

5.2.7. Auto-adaptação dos parâmetros estratégicos α e γ único nas 30 dimensões

Nesta seção, fixa-se a aplicação do parâmetro estratégico nas 30 dimensões, com o propósito de diminuir a eventual turbulência gerada pela liberdade total de se aplicar 30 parâmetros estratégicos independentes nas 30 dimensões, que pode ter sido a causa do fraco resultado do experimento anterior. Tal idéia se alinha com uma das conclusões obtidas em [Lima e de Oliveira, 06], onde a estratégia mais vantajosa de controlar automaticamente a pressão seletiva em um processo evolutivo ocorreu quando sua variação foi feita em blocos de gerações, e não a cada geração individual..

A auto-adaptação dos parâmetros estratégicos possibilita uma evolução dentre as famílias de distribuições estáveis onde, inicialmente define-se um intervalo de valores válidos para os parâmetros: índice de estabilidade $\alpha \in [0.25, 2]$; assimétrico $\beta = 0$; escala $\gamma \in [7.07E-9, 7.07E-2]$ e posição $\delta = 0$. Como dito, o parâmetro estratégico é fixo para todas 30 dimensões, sendo observado o valor do índice de estabilidade e escala para a melhor execução dentre as execuções de cada experimento, conforme as Tabelas 5.15

Exper.	Execuç.	Distrib. Estáveis							Distrib. Estáveis						
		Erro para 500 gerações							Erro para 1500 gerações						
		Mínimo	Máximo	Média	Dev. Padr.	Estabil. (%)	Escala (%)	Mínimo	Máximo	Média	Dev. Padr.	Estabil. (%)	Escala (%)		
135	3.00E+00	4.07E+01	5.49E+01	4.66E+01	7.38E+00	1.27E+00	8.63E-02	3.38E+01	4.74E+01	3.91E+01	7.29E+00	1.14E+00	3.73E-02		
136	6.00E+00	4.51E+01	5.76E+01	5.09E+01	4.80E+00	1.23E+00	2.49E-02	3.68E+01	5.18E+01	4.44E+01	5.22E+00	1.14E+00	7.69E-02		
137	9.00E+00	2.92E+01	6.08E+01	4.87E+01	1.07E+01	1.12E+00	3.24E-02	2.75E+01	5.31E+01	3.80E+01	8.23E+00	1.03E+00	4.32E-02		
138	1.20E+01	3.75E+01	6.28E+01	4.69E+01	6.88E+00	1.14E+00	4.97E-02	3.06E+01	4.57E+01	3.83E+01	4.32E+00	1.22E+00	1.73E-02		
139	1.50E+01	3.85E+01	6.39E+01	4.73E+01	6.46E+00	1.20E+00	3.22E-02	2.92E+01	6.05E+01	4.13E+01	8.72E+00	1.15E+00	8.34E-02		
140	1.80E+01	3.30E+01	7.28E+01	5.13E+01	9.23E+00	1.25E+00	3.01E-02	3.08E+01	5.18E+01	4.05E+01	6.42E+00	1.15E+00	7.36E-02		
141	2.10E+01	3.34E+01	5.84E+01	4.65E+01	7.42E+00	1.11E+00	2.71E-02	3.11E+01	4.85E+01	3.88E+01	5.07E+00	1.03E+00	9.43E-02		
142	2.40E+01	3.23E+01	6.65E+01	4.68E+01	7.74E+00	1.13E+00	5.77E-02	2.38E+01	4.86E+01	3.70E+01	6.23E+00	1.25E+00	1.89E-01		
143	2.70E+01	2.98E+01	6.43E+01	4.70E+01	8.70E+00	1.11E+00	4.47E-02	2.97E+01	5.00E+01	4.03E+01	6.03E+00	1.15E+00	3.23E-02		
144	3.00E+01	3.71E+01	6.53E+01	4.99E+01	6.64E+00	1.23E+00	2.99E-02	2.33E+01	5.18E+01	4.00E+01	7.20E+00	1.14E+00	4.16E-02		
Exper.	Execuç.	Erro para 5000 gerações							Erro para 10000 gerações						
		Mínimo	Máximo	Média	Dev. Padr.	Estabil. (%)	Escala (%)	Mínimo	Máximo	Média	Dev. Padr.	Estabil. (%)	Escala (%)		
		145	3.00E+00	3.25E+01	3.84E+01	3.47E+01	3.25E+00	1.29E+00	1.86E-01	2.81E+01	3.52E+01	3.06E+01	4.53E+00	1.09E+00	1.64E-01
146	6.00E+00	2.91E+01	5.31E+01	3.55E+01	8.89E+00	1.08E+00	8.88E-02	1.90E+01	4.02E+01	3.02E+01	7.75E+00	1.18E+00	1.38E-01		
147	9.00E+00	2.95E+01	4.29E+01	3.55E+01	4.75E+00	1.18E+00	3.42E-02	2.56E+01	4.00E+01	3.14E+01	3.99E+00	1.10E+00	3.15E-02		
148	1.20E+01	2.69E+01	4.49E+01	3.41E+01	6.52E+00	1.00E+00	6.99E-02	2.01E+01	4.30E+01	3.14E+01	6.52E+00	1.10E+00	7.52E-02		
149	1.50E+01	2.73E+01	4.29E+01	3.37E+01	4.84E+00	1.15E+00	9.52E-02	2.38E+01	3.58E+01	2.91E+01	3.77E+00	1.08E+00	1.44E-01		
150	1.80E+01	2.28E+01	4.19E+01	3.45E+01	5.38E+00	1.16E+00	2.10E-02	1.68E+01	4.73E+01	2.94E+01	7.67E+00	1.20E+00	3.38E-02		
151	2.10E+01	2.47E+01	4.21E+01	3.31E+01	4.65E+00	1.16E+00	1.09E-01	2.07E+01	4.45E+01	2.97E+01	6.08E+00	1.21E+00	7.98E-02		
152	2.40E+01	2.41E+01	4.27E+01	3.33E+01	5.12E+00	1.23E+00	5.85E-02	2.25E+01	4.94E+01	3.31E+01	6.90E+00	1.10E+00	1.01E-01		
153	2.70E+01	2.13E+01	4.07E+01	3.11E+01	5.13E+00	1.19E+00	6.84E-02	2.43E+01	4.53E+01	3.09E+01	5.07E+00	1.16E+00	1.08E-01		
154	3.00E+01	2.29E+01	4.87E+01	3.27E+01	5.60E+00	1.22E+00	3.90E-02	1.94E+01	4.02E+01	3.00E+01	5.28E+00	1.23E+00	7.76E-02		

Tabela 5.15: Resultados com a função de Rastrigin: Auto-adapt. fixa nas 30 dimensões de 500 a 10000 gerações.

Os melhores resultados para 500, 1500, 5000 e 10000 gerações foram, respectivamente, nos experimentos 137, 144, 153 e 150. Observando-se os valores dos parâmetros de estabilidade e de escala e, novamente comparando-se com os resultados da distribuição Estável 3 com 10000 gerações (experimento 112), os valores aqui encontrados não se aproximam do resultado daquele (com erro mínimo de 7.73E-11).

5.2.8. Conclusões sobre os experimentos com a função de Rastrigin

Para uma melhor visualização dos resultados experimentais realizados nas Seções 5.2.5 a 5.2.7 é criada a Tabela 5.16 e a Figura 5.1, onde se constata que nenhum dos processos empregados nas Seções 5.2.6 e 5.2.7 (respectivamente, parametrização livre e fixa dos parâmetros estratégicos para as 30 dimensões) conseguiu aproximar-se da parametrização pré-definida com a distribuição Estável 3.

Geração	Erros entre f(x) mínimo, máximo e média utilizando a função de Rastrigin											
	Estável 3 ($\sigma=0.3$ e $\mu=7.07E-09$)				Auto-adapt. variação livre: par. estr. n-dim.				Auto-adapt. par. estr. iguais para n-dim.			
	Mínimo	Máximo	Média	Desv.Padr.	Mínimo	Máximo	Média	Desv.Padr.	Mínimo	Máximo	Média	Desv.Padr.
500	1.16E+00	9.34E+00	5.62E+00	2.21E+00	2.56E+01	6.40E+01	4.74E+01	1.09E+01	2.92E+01	6.08E+01	4.87E+01	1.07E+01
1500	2.79E-06	4.15E+00	1.16E+00	1.31E+00	2.68E+01	4.79E+01	3.87E+01	5.17E+00	2.33E+01	5.18E+01	4.00E+01	7.20E+00
5000	2.38E-10	6.48E-10	5.37E-10	1.54E-10	2.00E+01	3.93E+01	2.77E+01	1.02E+01	2.13E+01	4.07E+01	3.11E+01	5.13E+00
10000	7.73E-11	1.34E-10	1.07E-10	2.85E-11	1.46E+01	3.99E+01	3.06E+01	5.95E+00	1.68E+01	4.73E+01	2.94E+01	7.67E+00

Tabela 5.16: Auto-adaptatividade na função de Rastrigin: distribuição Estável 3, com variação livre e fixa dos parâmetros estratégicos em 30 dimensões.

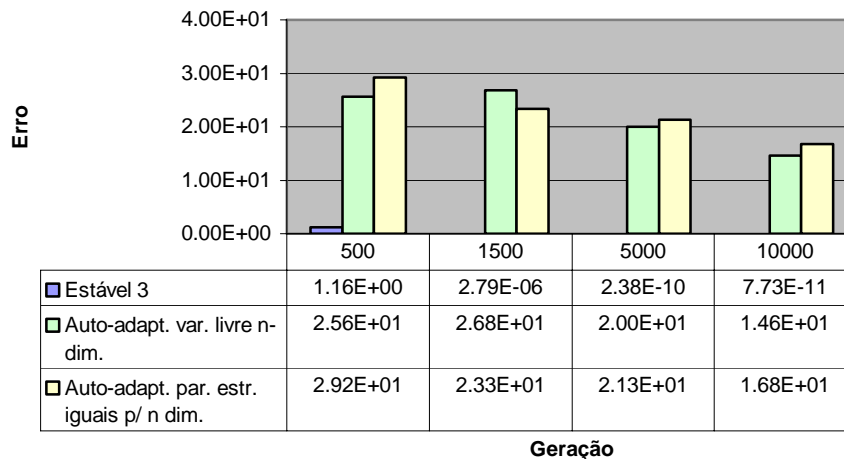


Figura 5.1: Auto-adaptatividade na função de Rastrigin: distribuição Estável 3, com variação livre e fixa dos parâmetros estratégicos em 30 dimensões.

Apesar dos resultados auto-adaptativos aqui obtidos terem sido fracos, naturalmente eles não invalidam resultados positivos com algoritmos evolutivos auto-adaptativos segundo outras abordagens [Saravanan *et al*, 95; Eiben, Hinterding e Michalewicz, 99; Beyer, 01; Gómez, 04; Nieberg e Beyer, 07].

5.3. Experimentos na função de teste de Rosenbrock nas dimensões: 2, 5, 10 e 30

Nesta seção utiliza-se a função de teste de Rosenbrock, que segue as mesmas parametrizações da Seção 5.2, executando-se o seguinte parâmetro de controle de evolução:

- Intervalo de variação das coordenadas x_1, \dots, x_n : $\{-2.048, \dots, 2.048\}$

A minimização desta função é considerada um processo difícil pelo fato dela possuir um pico muito reduzido e estreito. O topo do pico é muito pouco acentuado e sua superfície se apresenta na forma de uma parábola [Lima, 04].

5.3.1. Experimentos com função de Rosenbrock: 2 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, x_2, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \{\alpha_{21}, \beta_{22}, \gamma_{23}, \delta_{24}\}]$. As Tabelas 5.17 e 5.18 permitem observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 162(esquerda), 162(direita), 169(esquerda) e 169(direita).

Exper.	Par. escala	Gauss (Gauss)					Cauchy (Cauchy)				
		γ	σ	Erro				Erro			
				Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
155	7.07E-02	1.00E-01	1.79E-07	2.73E-05	7.32E-06	7.00E-06	9.69E-08	4.15E-05	5.82E-06	8.14E-06	
156	7.07E-03	1.00E-02	5.64E-09	3.99E-06	3.33E-07	7.56E-07	1.15E-09	8.15E-07	1.26E-07	1.53E-07	
157	7.07E-04	1.00E-03	3.47E-10	1.01E-01	1.43E-02	2.54E-02	7.79E-11	5.11E-02	9.85E-03	1.45E-02	
158	7.07E-05	1.00E-04	2.83E-06	2.64E-01	3.38E-02	5.28E-02	3.17E-12	2.50E-01	2.78E-02	4.76E-02	
159	7.07E-06	1.00E-05	2.09E-16	1.74E-01	4.12E-02	4.42E-02	5.87E-15	7.51E-02	1.28E-02	1.99E-02	
160	7.07E-07	1.00E-06	9.61E-07	1.39E-01	3.56E-02	3.61E-02	3.01E-08	2.25E-01	4.50E-02	5.25E-02	
161	7.07E-08	1.00E-07	2.56E-18	1.78E-01	3.62E-02	4.95E-02	1.27E-09	1.12E-01	2.84E-02	3.73E-02	
162	7.07E-09	1.00E-08	1.20E-19	1.57E-01	3.19E-02	3.63E-02	3.54E-21	1.41E-01	4.13E-02	4.42E-02	

Tabela 5.17: Resultados com a função de Rosenbrock: Gauss e Cauchy em 2 dimensões.

Exper.	Par. de escala	S Lévy (S Lévy)				Lévy (Lévy)				
		γ	Erro				Erro			
			Mínimo	Máximo	Média	Desv. padrão	Mínimo	Máximo	Média	Desv. Padrão
163	7.07E-02	3.05E-07	3.41E-05	6.56E-06	6.75E-06	1.69E-06	2.14E-04	3.82E-05	4.31E-05	
164	7.07E-03	4.86E-09	4.62E-06	4.29E-07	9.97E-07	1.91E-09	1.08E-06	2.49E-07	2.51E-07	
165	7.07E-04	2.23E-11	6.20E-04	3.64E-05	1.16E-04	7.03E-11	8.19E-09	3.24E-09	2.46E-09	
166	7.07E-05	1.43E-12	1.13E-01	1.55E-02	2.73E-02	1.82E-13	8.55E-02	1.99E-02	2.59E-02	
167	7.07E-06	2.54E-14	1.55E-01	2.58E-02	3.55E-02	5.18E-14	7.27E-02	1.13E-02	1.65E-02	
168	7.07E-07	4.81E-16	1.73E-01	3.40E-02	4.10E-02	5.18E-16	1.59E-01	3.03E-02	4.38E-02	
169	7.07E-08	8.04E-19	3.18E-01	3.98E-02	6.48E-02	9.22E-18	1.05E-01	2.76E-02	3.04E-02	
170	7.07E-09	3.34E-08	1.34E-01	2.72E-02	3.33E-02	8.73E-06	1.54E-01	2.54E-02	3.54E-02	

Tabela 5.18: Resultados com a função de Rosenbrock: S-Lévy e Lévy em 2 dimensões.

O melhor resultado foi na família de distribuição de, Cauchy. Note-se que a função envolvida não oferece um grande desafio em 2 dimensões.

5.3.2. Experimentos com função de Rosenbrock: 5 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, \dots, x_5, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \dots, \{\alpha_{51}, \beta_{52}, \gamma_{53}, \delta_{54}\}]$. As Tabelas 5.19 e 5.20 permitem observar que os resultados obtidos dos erros

Obteve-se o melhor resultado foi na família de distribuição com cauda longa, ou seja, na distribuição de Lévy

5.3.3. Experimentos com função de Rosenbrock: 10 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, \dots, x_{10}, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \dots, \{\alpha_{10\ 1}, \beta_{10\ 2}, \gamma_{10\ 3}, \delta_{10\ 4}\}]$. As Tabelas 5.21 e 5.22 permitem observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 204(esquerda), 204(direita), 212, 211, 221, 219, 228 e 227.

Análise das distribuições estáveis de Gauss e Cauchy em 10 dimensões para a Função de Rosenbrock										
Exper.	Par. escala	Gauss					Cauchy			
		σ	Erro				Erro			
			Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
203	7.07E-02	1.00E-01	7.64E+00	9.12E+00	8.32E+00	3.89E-01	8.17E+00	9.26E+00	8.77E+00	2.73E-01
204	7.07E-03	1.00E-02	5.35E+00	6.60E+00	5.84E+00	3.02E-01	6.70E+00	7.14E+00	6.97E+00	1.08E-01
205	7.07E-04	1.00E-03	8.37E+00	1.79E+01	1.11E+01	2.19E+00	6.84E+00	7.21E+00	7.08E+00	8.93E-02
206	7.07E-05	1.00E-04	8.08E+00	1.79E+01	1.05E+01	2.18E+00	7.59E+00	8.25E+00	7.91E+00	1.21E-01
207	7.07E-06	1.00E-05	8.83E+00	1.51E+01	1.09E+01	1.44E+00	7.85E+00	9.01E+00	8.51E+00	2.90E-01
208	7.07E-07	1.00E-06	9.20E+00	1.47E+01	1.14E+01	1.42E+00	8.16E+00	1.50E+01	1.04E+01	1.58E+00
209	7.07E-08	1.00E-07	8.82E+00	1.84E+01	1.18E+01	2.45E+00	8.37E+00	1.47E+01	1.11E+01	1.63E+00

Tabela 5.21: Resultados com a função de Rosenbrock: Gauss e Cauchy em 10 dimensões.

Exper.	Par. escala	S-Lévy (erro)				Lévy (erro)			
		Erro				Erro			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média
210	7.07E-02	8.87E+00	1.07E+01	9.77E+00	4.23E-01	4.17E+00	2.19E+01	1.35E+01	4.16E+00
211	7.07E-03	7.37E+00	8.10E+00	7.80E+00	1.90E-01	4.68E-01	3.26E+01	1.13E+01	1.03E+01
212	7.07E-04	7.01E+00	7.52E+00	7.30E+00	9.58E-02	7.99E+00	2.28E+01	1.24E+01	3.08E+00
213	7.07E-05	7.07E+00	7.47E+00	7.29E+00	9.64E-02	8.04E+00	1.87E+01	1.03E+01	1.95E+00
214	7.07E-06	7.31E+00	7.73E+00	7.55E+00	1.27E-01	7.03E+00	1.72E+01	9.37E+00	1.96E+00
215	7.07E-07	7.53E+00	8.00E+00	7.87E+00	1.12E-01	7.51E+00	1.45E+01	9.52E+00	1.80E+00
216	7.07E-08	7.33E+00	8.43E+00	8.09E+00	2.16E-01	7.93E+00	1.39E+01	9.64E+00	1.46E+00
Exper.	Par. escala	Estável 1 (erro)->Erro				Estável 2 (erro)->Erro			
		Erro				Erro			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média
217	7.07E-02	8.51E+00	1.12E+01	9.81E+00	5.60E-01	9.60E+00	1.30E+01	1.12E+01	8.39E-01
218	7.07E-03	7.13E+00	8.23E+00	7.88E+00	2.57E-01	7.28E+00	8.31E+00	8.04E+00	2.12E-01
219	7.07E-04	7.20E+00	7.54E+00	7.37E+00	8.89E-02	7.13E+00	7.77E+00	7.53E+00	1.51E-01
220	7.07E-05	7.06E+00	7.50E+00	7.26E+00	1.01E-01	7.19E+00	7.64E+00	7.38E+00	9.50E-02
221	7.07E-06	7.01E+00	7.63E+00	7.48E+00	1.21E-01	7.30E+00	7.59E+00	7.46E+00	8.56E-02
222	7.07E-07	7.53E+00	7.98E+00	7.75E+00	1.10E-01	7.30E+00	7.78E+00	7.64E+00	1.20E-01
223	7.07E-08	7.54E+00	8.18E+00	8.01E+00	1.20E-01	7.54E+00	8.03E+00	7.82E+00	9.80E-02
Exper.	Par. escala	Estável 3 (erro)->Erro				Estável 4 (erro)->Erro			
		Erro				Erro			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média
224	7.07E-03	7.81E+00	8.60E+00	8.29E+00	1.81E-01	7.37E+00	9.06E+00	8.59E+00	3.30E-01
225	7.07E-04	7.35E+00	7.96E+00	7.74E+00	1.63E-01	7.33E+00	8.23E+00	7.92E+00	2.33E-01
226	7.07E-05	7.24E+00	7.79E+00	7.52E+00	1.25E-01	7.42E+00	7.99E+00	7.69E+00	1.24E-01
227	7.07E-06	7.17E+00	7.63E+00	7.43E+00	1.06E-01	7.16E+00	7.73E+00	7.52E+00	1.38E-01
228	7.07E-07	7.09E+00	7.68E+00	7.49E+00	1.20E-01	7.24E+00	7.70E+00	7.51E+00	1.11E-01
229	7.07E-08	7.45E+00	7.80E+00	7.65E+00	1.02E-01	7.18E+00	7.73E+00	7.61E+00	1.23E-01
230	7.07E-09	7.51E+00	7.95E+00	7.80E+00	9.94E-02	7.40E+00	7.98E+00	7.80E+00	1.32E-01

Tabela 5.22: Resultados com a função de Rosenbrock: S-Lévy, Lévy e Estável 1 a Estável 4 em 10 dimensões.

O melhor resultado foi na família de distribuição com cauda longa, ou seja, na distribuição de Lévy .

5.3.4. Experimentos com função de Rosenbrock: 30 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, \dots, x_{30}, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \dots, \{\alpha_{30\ 1}, \beta_{30\ 2}, \gamma_{30\ 3}, \delta_{30\ 4}\}]$. As Tabelas 5.23 e 5.24 permitem observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 232, 233, 239, 237, 245, 244, 250 e 253.

Análise das distribuições estáveis de Gauss e Cauchy em 30 dimensões para a Função de Rosenbrock										
Exper.	Par. escala	Gauss					Cauchy			
		Erro				Erro				
	γ	σ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
231	7.07E-02	1.00E-01	3.65E+01	4.04E+01	3.82E+01	9.25E-01	2.78E+01	2.93E+01	2.83E+01	3.36E-01
232	7.07E-03	1.00E-02	2.75E+01	2.92E+01	2.82E+01	4.58E-01	2.74E+01	2.78E+01	2.77E+01	7.70E-02
233	7.07E-04	1.00E-03	3.07E+01	6.45E+01	4.10E+01	7.18E+00	2.73E+01	2.78E+01	2.76E+01	1.22E-01
234	7.07E-05	1.00E-04	4.52E+01	9.63E+01	6.42E+01	1.30E+01	2.73E+01	2.93E+01	2.82E+01	4.38E-01
235	7.07E-06	1.00E-05					3.07E+01	6.45E+01	4.10E+01	7.18E+00

Tabela 5.23: Resultados com a função de Rosenbrock: Gauss e Cauchy em 30 dimensões.

Análise das distribuições estáveis de S-Lévy e Lévy em 30 dimensões para a Função de Rosenbrock										
Exper.	Par. escala	S-Lévy				Lévy				
		Erro				Erro				
	γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão	
236	7.07E-02					8.29E+02	1.95E+03	1.39E+03	2.45E+02	
237	7.07E-03					2.42E+01	1.17E+02	4.29E+01	1.71E+01	
238	7.07E-04	2.74E+01	2.78E+01	2.77E+01	8.46E-02	6.68E+01	1.22E+02	9.80E+01	1.29E+01	
239	7.07E-05	2.69E+01	2.80E+01	2.78E+01	2.28E-01					
240	7.07E-06	2.74E+01	2.78E+01	2.76E+01	1.13E-01					
241	7.07E-07	2.75E+01	2.84E+01	2.81E+01	1.66E-01					
Exper.	Par. escala	Estável 1				Estável 2				
		Erro				Erro				
	γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão	
242	7.07E-04	2.79E+01	2.90E+01	2.86E+01	2.47E-01	2.96E+01	3.14E+01	3.03E+01	4.56E-01	
243	7.07E-05	2.77E+01	2.81E+01	2.80E+01	1.15E-01	2.77E+01	2.84E+01	2.82E+01	1.51E-01	
244	7.07E-06	2.75E+01	2.80E+01	2.78E+01	1.14E-01	2.70E+01	2.81E+01	2.78E+01	2.99E-01	
245	7.07E-07	2.72E+01	2.81E+01	2.79E+01	1.59E-01	2.77E+01	2.81E+01	2.79E+01	1.11E-01	
246	7.07E-08	2.76E+01	2.94E+01	2.84E+01	3.92E-01					
Exper.	Par. escala	Estável 3				Estável 4				
		Erro				Erro				
	γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão	
247	7.07E-04	4.00E+01	6.43E+01	4.68E+01	5.00E+00	4.86E+01	1.19E+02	8.36E+01	1.54E+01	
248	7.07E-05	2.86E+01	2.99E+01	2.93E+01	3.35E-01					
249	7.07E-06	2.77E+01	2.84E+01	2.81E+01	2.05E-01	2.83E+01	3.03E+01	2.91E+01	4.55E-01	
250	7.07E-07	2.75E+01	2.82E+01	2.79E+01	1.63E-01	2.78E+01	2.85E+01	2.82E+01	2.04E-01	
251	7.07E-08	2.76E+01	2.81E+01	2.79E+01	1.36E-01	2.77E+01	2.83E+01	2.80E+01	1.28E-01	
252	7.07E-09	2.76E+01	2.82E+01	2.80E+01	1.39E-01	2.74E+01	2.82E+01	2.80E+01	1.71E-01	
253	7.07E-10					2.72E+01	2.83E+01	2.80E+01	2.29E-01	

Tabela 5.24: Resultados com a função de Rosenbrock: S-Lévy, Lévy e Estável 1 a Estável 4 em 30 dimensões.

O melhor resultado dos experimentos de 155 a 253 foi na família de distribuição com cauda simétrica S-Lévy (na cor azul), com erros de até 10^{+1} .

Em conclusão, o conjunto de experimentos aqui relatados, com progressivo aumento da dimensão da função de Rosenbrock indica que 2 dimensões mostraram-se não relevantes para o resultado final, podendo-se, portanto, iniciar-se os experimentos posteriores em 5 dimensões.

5.3.5. Experimentos com a distribuição S-Lévy

Como verificado na seção anterior o melhor resultado foi obtido, a distribuição S-Lévy, o que torna esta distribuição merecedora de maior atenção. Nesse sentido, verifica-se agora a evolução dos dados com quantidades variável de de execuções de 3 até 30, independente entre si e variando-se o número de gerações de 500, 1500, 5000 e 10000.

Com isso, tem-se o propósito de verificar em que condições, mais interessante executar experimentos no futuro, ou seja, com um número maior de gerações associado a um número menor de execuções, ou o contrário. Vale lembrar que reduzindo significativamente o número de execuções, pode-se perder a relevância da amostra (repetibilidade).

Exper.	Execuções	S-Lévy				S-Lévy			
		Erro para 500 gerações				Erro para 1500 gerações			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
254	3.00E+00	2.76E+01	2.79E+01	2.78E+01	1.25E-01	2.71E+01	2.72E+01	2.72E+01	4.78E-02
255	6.00E+00	2.72E+01	2.79E+01	2.77E+01	2.65E-01	2.69E+01	2.73E+01	2.72E+01	1.28E-01
256	9.00E+00	2.76E+01	2.80E+01	2.79E+01	1.33E-01	2.68E+01	2.73E+01	2.71E+01	1.70E-01
257	1.20E+01	2.76E+01	2.80E+01	2.79E+01	1.27E-01	2.71E+01	2.73E+01	2.72E+01	7.96E-02
258	1.50E+01	2.74E+01	2.80E+01	2.78E+01	1.73E-01	2.70E+01	2.73E+01	2.71E+01	8.63E-02
259	1.80E+01	2.74E+01	2.80E+01	2.78E+01	1.44E-01	2.70E+01	2.73E+01	2.71E+01	8.49E-02
260	2.10E+01	2.75E+01	2.80E+01	2.78E+01	1.12E-01	2.68E+01	2.73E+01	2.71E+01	1.32E-01
261	2.40E+01	2.68E+01	2.80E+01	2.77E+01	2.50E-01	2.68E+01	2.73E+01	2.71E+01	1.27E-01
262	2.70E+01	2.75E+01	2.80E+01	2.78E+01	1.34E-01	2.69E+01	2.74E+01	2.72E+01	1.24E-01
263	3.00E+01	2.73E+01	2.80E+01	2.78E+01	1.53E-01	2.69E+01	2.73E+01	2.72E+01	1.04E-01
Exper.	Execuções	Erro para 5000 gerações				Erro para 10000 gerações			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
		264	3.00E+00	2.60E+01	2.60E+01	2.60E+01	1.61E-02	2.49E+01	2.49E+01
265	6.00E+00	2.59E+01	2.61E+01	2.60E+01	7.63E-02	2.47E+01	2.49E+01	2.48E+01	7.45E-02
266	9.00E+00	2.59E+01	2.61E+01	2.60E+01	6.66E-02	2.48E+01	2.49E+01	2.49E+01	5.04E-02
267	1.20E+01	2.59E+01	2.61E+01	2.60E+01	5.16E-02				
268	1.50E+01	2.58E+01	2.61E+01	2.60E+01	7.44E-02				
269	1.80E+01	2.59E+01	2.61E+01	2.60E+01	5.41E-02				
270	2.10E+01	2.59E+01	2.61E+01	2.60E+01	5.41E-02				

Tabela 5.25: Resultados com a função de Rosenbrock: S-Lévy de 500 e 10000 gerações.

Na Tabelas 5.25 pode-se observar-se que os resultados obtidos para os erros mínimos não melhoram na distribuição S-Lévy, quando aumenta-se o esforço computacional até 10000 gerações mesmo com a redução do número de execuções.

Os melhores resultados para 500, 1500, 5000 e 10000 gerações foram respectivamente nos experimentos 261(esquerda), 261(direita), 268 e 265 e no conjunto de experimentos, verifica-se que nem sempre é vantajoso aumentar o número de gerações independente do número de execuções.

5.3.6. Auto-adaptação dos parâmetros estratégicos α e γ em 30 dimensões

A motivação para executar os experimentos desta seção reside no fato de analisar se é possível deixar o processo evolutivo descobrir uma família de distribuições estáveis que acarrete melhor nível de qualidade do que os experimentos com famílias pré-definidas. Resalte-se que este novo conjunto de experimentos se encontra dentro das mesmas parametrizações que as existentes na seção anterior, mas distintas das usadas na Seção 5.1, que também tratou de auto-adaptação.

Sendo assim, inicialmente define-se um intervalo de valores válidos para os parâmetros estratégicos: índice de estabilidade $\alpha \in [0.25, 2]$; índice assimetria $\beta = 0$; de escala $\Upsilon \in [7.07E-9, 7.07E-2]$ e de posição $\delta = 0$. Os parâmetros estratégicos podem variar livremente para cada dimensão de forma independente. Foi observado o valor dos índices de estabilidade e escala para a melhor execução no conjunto de execuções de cada experimento, conforme, a Tabela 5.26.

Os melhores resultados para 500, 1500, 5000 e 10000 gerações ocorreram nos experimentos 274, 280, 287 e 283. Observando-se os valores dos parâmetros de estabilidade e de escala, e comparando-se com os, obtidos na seção anterior, os valores encontrados nos experimentos 271 a 288 se aproximam do resultado desejado, ou seja, erro mínimo $2.47E+01$ obtido no experimento 265 em relação ao erro mínimo $3.18E+01$ no experimento 283. Na Seção 5.3.8 conjuga-se melhor a interlação dos dados experimentais com a ampliação dos experimentos da seção seguinte, que fixa os parâmetros estratégicos para todas as 30 dimensões.

Exper.	Execuç.	Distribuições Estáveis							Distribuições Estáveis						
		Erro para 500 gerações							Erro para 1500 gerações						
		Mínimo	Máximo	Média	Dev. Padr.	Estabil. (σ)	Escala (γ)	Mínimo	Máximo	Média	Dev. Padr.	Estabil. (σ)	Escala (γ)		
271	3.00E+00	3.83E+01	4.22E+01	4.04E+01	1.95E+00	1.21E+00	9.77E-02	3.58E+01	3.79E+01	3.68E+01	1.04E+00	9.45E-01	4.70E-02		
272	6.00E+00	3.56E+01	4.15E+01	3.89E+01	2.06E+00	1.10E+00	7.33E-02	3.55E+01	3.88E+01	3.70E+01	1.22E+00	1.06E+00	5.29E-02		
273	9.00E+00	3.67E+01	4.03E+01	3.90E+01	1.17E+00	1.11E+00	5.53E-02	3.54E+01	3.92E+01	3.77E+01	1.21E+00	1.08E+00	2.33E-02		
274	1.20E+01	3.34E+01	4.22E+01	3.89E+01	2.30E+00	1.10E+00	2.16E-02	3.56E+01	3.93E+01	3.72E+01	1.16E+00	1.23E+00	2.69E-02		
275	1.50E+01	3.64E+01	4.29E+01	3.97E+01	1.86E+00	1.16E+00	3.36E-02	3.62E+01	3.94E+01	3.79E+01	9.74E-01	1.23E+00	3.18E-02		
276	1.80E+01	3.75E+01	4.19E+01	3.99E+01	1.27E+00	1.04E+00	8.72E-02	3.44E+01	3.94E+01	3.71E+01	1.32E+00	1.24E+00	3.54E-02		
277	2.10E+01	3.58E+01	4.22E+01	3.97E+01	1.78E+00	1.15E+00	6.56E-02	3.50E+01	4.07E+01	3.73E+01	1.62E+00	1.22E+00	5.97E-02		
278	2.40E+01	3.39E+01	4.22E+01	3.97E+01	2.08E+00	1.04E+00	3.70E-02	3.53E+01	4.14E+01	3.75E+01	1.37E+00	1.21E+00	5.42E-02		
279	2.70E+01	3.64E+01	4.29E+01	4.01E+01	1.66E+00	1.12E+00	9.67E-02	3.52E+01	3.95E+01	3.76E+01	1.22E+00	1.31E+00	3.16E-02		
280	3.00E+01	3.63E+01	4.29E+01	3.96E+01	1.77E+00	1.11E+00	4.65E-02	3.37E+01	3.98E+01	3.77E+01	1.26E+00	1.12E+00	4.07E-02		
Exper.	Execuç.	Distribuições Estáveis							Distribuições Estáveis						
		Erro para 5000 gerações							Erro para 10000 gerações						
		Mínimo	Máximo	Média	Dev. Padr.	Estabil. (σ)	Escala (γ)	Mínimo	Máximo	Média	Dev. Padr.	Estabil. (σ)	Escala (γ)		
281	3.00E+00	3.80E+01	4.26E+01	4.09E+01	2.48E+00	1.19E+00	4.65E-02	3.54E+01	3.59E+01	3.56E+01	2.71E-01	1.19E+00	3.24E-02		
282	6.00E+00	3.44E+01	3.63E+01	3.54E+01	7.62E-01	1.14E+00	4.38E-02	3.51E+01	3.62E+01	3.57E+01	4.09E-01	1.28E+00	4.87E-02		
283	9.00E+00	3.33E+01	3.78E+01	3.63E+01	1.36E+00	1.24E+00	3.71E-02	3.18E+01	3.64E+01	3.49E+01	1.37E+00	1.19E+00	1.13E-01		
284	1.20E+01	3.46E+01	3.85E+01	3.62E+01	1.20E+00	1.12E+00	2.87E-02	3.36E+01	3.66E+01	3.54E+01	8.15E-01	1.12E+00	4.84E-02		
285	1.80E+01	3.40E+01	3.78E+01	3.57E+01	1.11E+00	1.08E+00	3.48E-02	3.34E+01	3.88E+01	3.57E+01	1.18E+00	1.21E+00	4.01E-02		
286	2.10E+01	3.42E+01	3.77E+01	3.59E+01	9.42E-01	1.23E+00	3.40E-02	3.33E+01	3.78E+01	3.54E+01	9.29E-01	1.14E+00	3.33E-02		
287	2.40E+01	3.27E+01	3.83E+01	3.55E+01	1.42E+00	1.03E+00	7.81E-02	3.27E+01	3.66E+01	3.48E+01	1.09E+00	1.35E+00	8.22E-02		
288	3.00E+01	3.37E+01	3.85E+01	3.58E+01	1.25E+00	1.16E+00	3.13E-02	3.39E+01	3.91E+01	3.59E+01	1.27E+00	1.09E+00	4.39E-02		

Tabela 5.26: Resultados com a função de Rosenbrock: Auto-adapt. livre nas n-dim de 500 a 10000 gerações.

5.3.7. Auto-adaptação com parâmetro estratégico α e γ único nas 30 dimensões

Nesta seção, fixa-se a aplicação do parâmetro estratégico as 30 dimensões com o propósito de diminuir a turbulência gerada pela liberdade total de se aplicar 30 parâmetros estratégicos independentes nas 30 dimensões [Lima e de Oliveira, 06].

A auto-adaptação dos parâmetros estratégicos possibilita uma evolução dentre as famílias de distribuições estáveis onde, inicialmente define-se um intervalo de valores válidos para os parâmetros: índice de estabilidade $\alpha \in [0.25, 2]$; assimétrico $\beta = 0$; escala $\gamma \in [7.07E-9, 7.07E-1]$ e posição $\delta = 0$. Como dito, o parâmetro estratégico é fixo para todas 30 dimensões, sendo observado o valor do índice de estabilidade e escala para a melhor execução dentre as execuções de cada experimento, conforme a Tabela 5.27.

Exper.	Execuç.	Distribuições Estáveis						Distribuições Estáveis					
		Erro para 500 gerações						Erro para 1500 gerações					
		Mínimo	Máximo	Média	Dev. Padr.	Estabil. (E)	Escala (Y)	Mínimo	Máximo	Média	Dev. Padr.	Estabil. (E)	Escala (Y)
289	3.00E+00	3.63E+01	3.88E+01	3.79E+01	1.40E+00	1.05E+00	6.19E-02	3.56E+01	3.97E+01	3.75E+01	2.07E+00	1.09E+00	6.24E-02
290	6.00E+00	3.92E+01	4.17E+01	4.03E+01	1.03E+00	1.28E+00	5.07E-02	3.62E+01	3.97E+01	3.73E+01	1.32E+00	1.07E+00	5.00E-02
291	9.00E+00	3.73E+01	4.09E+01	3.89E+01	1.35E+00	1.03E+00	3.50E-02	3.45E+01	4.04E+01	3.73E+01	1.89E+00	1.20E+00	6.48E-02
292	1.20E+01	3.61E+01	4.18E+01	3.90E+01	1.76E+00	1.22E+00	5.78E-02	3.47E+01	3.92E+01	3.73E+01	1.49E+00	1.23E+00	5.15E-02
293	1.50E+01	3.81E+01	4.22E+01	3.99E+01	1.21E+00	1.17E+00	6.11E-02	3.53E+01	4.05E+01	3.75E+01	1.39E+00	1.14E+00	5.38E-02
294	1.80E+01	3.65E+01	4.30E+01	3.99E+01	1.84E+00	1.24E+00	3.57E-02	3.44E+01	3.98E+01	3.72E+01	1.63E+00	1.14E+00	6.05E-02
295	2.10E+01	3.60E+01	4.18E+01	3.87E+01	1.66E+00	1.04E+00	1.13E-01	3.60E+01	3.94E+01	3.77E+01	1.09E+00	1.22E+00	3.73E-02
296	2.40E+01	3.69E+01	4.27E+01	3.92E+01	1.46E+00	1.06E+00	3.77E-02	3.53E+01	3.97E+01	3.75E+01	1.27E+00	1.10E+00	1.06E-01
297	2.70E+01	3.66E+01	4.26E+01	3.96E+01	1.53E+00	9.95E-01	6.97E-02	3.40E+01	3.98E+01	3.72E+01	1.28E+00	1.02E+00	5.58E-02
298	3.00E+01	3.60E+01	4.44E+01	3.96E+01	1.94E+00	1.09E+00	1.01E-01	3.53E+01	4.01E+01	3.76E+01	1.29E+00	1.15E+00	3.07E-02
Exper.	Execuç.	Erro para 5000 gerações						Erro para 10000 gerações					
		Mínimo	Máximo	Média	Dev. Padr.	Estabil. (E)	Escala (Y)	Mínimo	Máximo	Média	Dev. Padr.	Estabil. (E)	Escala (Y)
		299	3.00E+00	3.46E+01	3.56E+01	3.52E+01	5.01E-01	1.16E+00	6.84E-02				
300	6.00E+00	3.50E+01	3.69E+01	3.58E+01	6.80E-01	1.13E+00	5.11E-02						
301	9.00E+00	3.48E+01	3.77E+01	3.65E+01	8.80E-01	1.05E+00	6.84E-02						
302	1.20E+01	3.45E+01	3.65E+01	3.57E+01	6.06E-01	1.10E+00	4.73E-02						
303	1.50E+01	3.42E+01	3.72E+01	3.57E+01	1.07E+00	1.17E+00	6.84E-02						
304	1.80E+01	3.39E+01	3.89E+01	3.59E+01	1.45E+00	1.05E+00	3.28E-02						
305	2.10E+01	3.30E+01	3.74E+01	3.55E+01	1.18E+00	1.11E+00	2.96E-02						
306	2.40E+01	3.33E+01	3.85E+01	3.63E+01	1.14E+00	1.15E+00	9.78E-02						
307	2.70E+01	3.32E+01	3.88E+01	3.56E+01	1.37E+00	1.12E+00	1.09E-01						
308	3.00E+01												

Tabela 5.27: Resultados com a função de Rosenbrock: Auto-adapt. fixa nas n-dim. de 500 a 10000 gerações.

Os melhores resultados para 500, 1500 e 5000 gerações foram respectivamente nos experimentos 298, 297 e 305. Observando-se os valores dos parâmetros de estabilidade e de escala e comparando-se com os dados obtidos na Seção 5.3.5, os valores encontrados nos experimentos 289 a 308 se aproximam do resultado desejado, ou seja, erro mínimo $2.58E+01$ obtido no experimento 265 em relação ao erro mínimo $3.30E+01$ no experimento 305.

5.3.8. Conclusões sobre os experimentos com a função de Rosenbrock

Para uma melhor visualização dos resultados experimentais realizados nas Seções 5.3.5 a 5.3.7 é criada a Tabela 5.28 e a Figura 5.2, onde se constata que ambos os processos empregados nas Seções 5.3.6 e 5.3.7, respectivamente parametrização livre dos parâmetros estratégicos e fixa para as 30 dimensões conseguiram aproximar-se da parametrização pré-definida

Para uma melhor visualização dos resultados experimentais realizados nesta seção criou-se a Tabela 5.28, onde constata-se que na busca de uma família de distribuição estável a S-Lévy continua apresentando melhores resultados.

Geração	Erros entre f(x) mínimo, máximo e média utilizando a função de Rastrigin											
	S-Lévy				Auto-adapt. variação livre: par. estr. n-dim.				Auto-adapt. par. estr. iguais para n-dim.			
	Mínimo	Máximo	Média	Desv.Padr.	Mínimo	Máximo	Média	Desv.Padr.	Mínimo	Máximo	Média	Desv.Padr.
500	2.68E+01	2.80E+01	2.77E+01	2.50E-01	3.34E+01	4.22E+01	3.89E+01	2.30E+00	3.60E+01	4.44E+01	3.96E+01	1.94E+00
1500	2.68E+01	2.73E+01	2.71E+01	1.27E-01	3.37E+01	3.98E+01	3.77E+01	1.26E+00	3.40E+01	3.98E+01	3.72E+01	1.28E+00
5000	2.58E+01	2.61E+01	2.60E+01	7.44E-02	3.27E+01	3.83E+01	3.55E+01	1.42E+00	3.30E+01	3.74E+01	3.55E+01	1.18E+00
10000	2.47E+01	2.49E+01	2.48E+01	7.45E-02	3.18E+01	3.64E+01	3.49E+01	1.37E+00				

Tabela 5.28: Comparação na função de Rosenbrock: S-Lévy, Auto-adaptação livre e fixa nas n-dim.

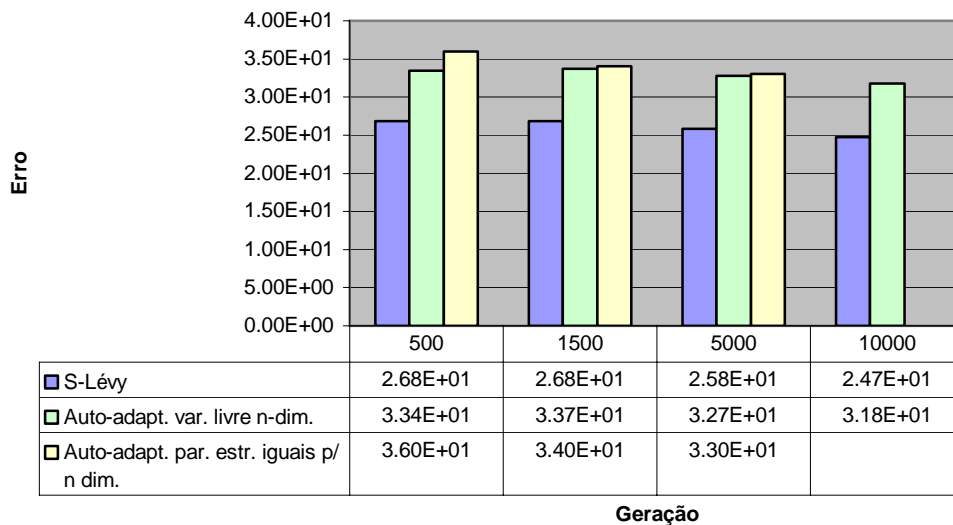


Figura 5.2: Gráfico na função de Rosenbrock: S-Lévy, Auto-adaptação livre e fixa nas n-dim.

Nesta função de teste observa-se que a utilização da distribuição S-Lévy apresenta melhores resultados independente do esforço computacional (número de gerações). Quando utiliza-se auto-adaptação em ambos processos de controle dos parâmetros estratégicos nas n-dimensões: variação livre ou fixa, obteve-se resultados próximos ($10E-1$), nos experimentos com 1500 e 5000 gerações, e apresentando uma ligeira vantagem para o processo de auto-adaptação com variação livre nas n-dimensões em 500 gerações.

5.4. Experimentos na função de teste de Griewangk nas dimensões: 2, 5, 10 e 30

Nesta seção utiliza-se a função de teste de Griewangk, que segue as mesmas parametrizações da Seção 5.2, exetuado-se o seguinte parâmetro de controle de evolução:

- Intervalo de variação das coordenadas x_1, \dots, x_n : $\{-600, \dots, 600\}$

5.4.1. Experimentos com função de Griewangk: 2 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, x_2, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \{\alpha_{21}, \beta_{22}, \gamma_{23}, \delta_{24}\}]$. Nas Tabelas 5.29 e 5.30 pode-se observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 316, 314, 321 e 322.

Análise das distribuições estáveis de Gauss e Cauchy em 2 dimensões na Função de Griewangk										
Exper.	Par. de escala	Gauss (Gauss)					Cauchy (Cauchy)			
		γ	σ	Erro				Erro		
			Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
308	7.07E-02	1.00E-01	1.34E-08	7.40E-03	2.47E-04	1.35E-03	2.65E-09	7.40E-03	4.93E-04	1.88E-03
309	7.07E-03	1.00E-02	4.19E-11	7.40E-03	4.93E-04	1.88E-03	2.66E-11	7.40E-03	4.93E-04	1.88E-03
310	7.07E-04	1.00E-03	5.03E-13	6.43E-11	2.00E-11	1.64E-11	2.08E-13	7.40E-03	2.47E-04	1.35E-03
311	7.07E-05	1.00E-04	1.20E-14	7.40E-03	3.79E-04	1.40E-03	1.33E-15	7.40E-03	2.47E-04	1.35E-03
312	7.07E-06	1.00E-05	2.22E-16	9.86E-03	8.98E-04	2.53E-03	0.00E+00	7.40E-03	4.95E-04	1.88E-03
313	7.07E-07	1.00E-06	0.00E+00	7.41E-03	8.14E-04	1.89E-03	0.00E+00	2.36E-03	9.99E-05	4.41E-04
314	7.07E-08	1.00E-07	0.00E+00	7.67E-03	1.41E-03	2.89E-03	0.00E+00	7.40E-03	8.08E-04	2.11E-03
315	7.07E-09	1.00E-08	0.00E+00	1.40E-02	5.90E-04	2.60E-03	0.00E+00	7.40E-03	3.05E-04	1.35E-03

Tabela 5.29: Resultados com a função de Griewangk: Gauss e Cauchy em 2 dimensões.

Análise das distribuições estáveis de S-Lévy e Lévy em 2 dimensões na Função de Griewangk									
Exper.	Par. de escala	S Lévy (S-Lévy)				Lévy (Lévy)			
		γ	Erro				Erro		
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
316	7.07E-02	2.87E-10	1.34E-07	3.43E-08	3.21E-08	2.84E-06	9.76E-03	1.94E-03	2.76E-03
317	7.07E-03	5.82E-12	7.40E-03	2.47E-04	1.35E-03	5.51E-08	9.55E-03	4.70E-04	1.76E-03
318	7.07E-04	3.21E-13	7.40E-03	4.93E-04	1.88E-03	2.60E-10	2.33E-03	8.41E-05	4.26E-04
319	7.07E-05	4.44E-16	9.86E-03	5.75E-04	2.21E-03	1.83E-12	7.40E-03	5.11E-04	1.87E-03
320	7.07E-06	0.00E+00	2.22E-15	4.07E-16	4.99E-16	6.73E-14	7.40E-03	2.47E-04	1.35E-03
321	7.07E-07	0.00E+00	7.40E-03	4.93E-04	1.88E-03	0.00E+00	8.21E-05	4.85E-06	1.65E-05
322	7.07E-08	0.00E+00	7.40E-03	2.47E-04	1.35E-03	0.00E+00	7.40E-03	6.03E-04	1.80E-03
323	7.07E-09	0.00E+00	7.40E-03	9.88E-04	2.56E-03	0.00E+00	7.40E-03	3.84E-04	1.42E-03

Tabela 5.30: Resultados com a função de Griewangk: S-Lévy e Lévy em 2 dimensões.

O melhor resultado foi na família de distribuição com cauda leve, S-Lévy com erros até 10^{-16} .

5.4.2. Experimentos com função de Griewangk: 5 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, \dots, x_5, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \dots, \{\alpha_{51}, \beta_{52}, \gamma_{53}, \delta_{54}\}]$. Nas Tabelas 5.31 e 5.32 pode-se observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 327, 326, 338a, 338b, 343, 347, 355 e 356.

Análise das distribuições estáveis de Gauss e Cauchy em 5 dimensões na Função Griewangk											
Exper.	Par. de escala	Gauss						Cauchy			
		Erro						Erro			
		γ	σ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
324	7.07E-02	1.00E-01	9.91E-03	5.05E-01	6.64E-02	9.13E-02	1.24E-04	5.43E-02	1.93E-02	1.40E-02	
325	7.07E-03	1.00E-02	9.86E-03	1.53E-01	5.05E-02	3.45E-02	5.46E-07	1.40E-01	3.51E-02	3.01E-02	
326	7.07E-04	1.00E-03	7.40E-03	2.24E-01	4.60E-02	4.21E-02	7.40E-03	1.85E-01	4.79E-02	4.79E-02	
327	7.07E-05	1.00E-04	1.70E-02	2.05E-01	5.45E-02	3.90E-02	7.40E-03	9.61E-02	4.22E-02	2.53E-02	
328	7.07E-06	1.00E-05	1.95E-02	1.86E-01	6.02E-02	3.57E-02	7.40E-03	1.95E-01	5.06E-02	4.29E-02	
329	7.07E-07	1.00E-06	1.08E-02	1.99E-01	6.87E-02	4.86E-02	1.94E-02	2.32E-01	6.87E-02	5.04E-02	
330	7.07E-08	1.00E-07	1.48E-02	2.15E-01	7.11E-02	4.49E-02	2.73E-02	2.03E-01	8.24E-02	4.88E-02	
331	7.07E-09	1.00E-08	1.98E-02	2.72E-01	8.42E-02	5.30E-02	8.29E-03	2.70E-01	1.03E-01	7.25E-02	

Tabela 5.31: Resultados com a função de Griewangk: Gauss e Cauchy em 5 dimensões.

Análise das distribuições estáveis de S-Lévy e Lévy em 5 dimensões na Função Griewangk											
Exper.	Par. de escala	S Lévy						Lévy			
		Erro						Erro			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão	
332	7.07E-02	2.95E-05	3.32E-02	1.75E-02	9.12E-03	2.79E-02	4.28E-01	1.13E-01	7.51E-02		
333	7.07E-03	2.40E-07	3.45E-02	1.84E-02	9.17E-03	8.25E-03	1.47E-01	6.01E-02	3.90E-02		
334	7.07E-04	6.53E-09	3.45E-02	1.77E-02	9.24E-03	5.67E-04	9.99E-02	4.02E-02	2.60E-02		
335	7.07E-05	7.35E-11	7.39E-02	1.99E-02	1.51E-02	1.22E-02	1.28E-01	4.59E-02	3.17E-02		
336	7.07E-06	4.94E-13	8.38E-02	3.33E-02	2.05E-02	1.73E-02	1.23E-01	5.18E-02	2.77E-02		
337	7.07E-07	9.77E-15	1.01E-01	3.70E-02	2.79E-02	5.86E-05	2.09E-01	6.68E-02	5.16E-02		
338	7.07E-08	7.40E-03	1.18E-01	4.54E-02	3.00E-02	1.23E-02	2.34E-01	6.32E-02	4.77E-02		
339	7.07E-09	9.86E-03	2.19E-01	4.91E-02	4.39E-02	1.14E-02	2.04E-01	6.42E-02	4.96E-02		
Exper.	Par. de escala	Estável 1						Estável 2			
		Erro						Erro			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão	
340	7.07E-02	1.11E-05	3.46E-02	1.32E-02	8.12E-03	3.68E-05	5.25E-02	2.19E-02	1.26E-02		
341	7.07E-03	1.37E-03	1.99E+00	4.06E-01	6.94E-01	4.68E-06	3.45E-02	2.13E-02	8.60E-03		
342	7.07E-04	5.41E-09	3.45E-02	1.86E-02	9.57E-03	5.23E-09	3.50E-02	2.21E-02	1.02E-02		
343	7.07E-05	7.40E-03	5.58E-02	2.02E-02	1.19E-02	4.86E-11	4.43E-02	2.04E-02	1.10E-02		
344	7.07E-06	7.40E-03	1.23E-01	3.10E-02	2.40E-02	7.44E-06	4.19E-02	1.90E-02	9.97E-03		
345	7.07E-07	7.98E-03	1.23E-01	5.05E-02	3.51E-02	4.71E-14	7.64E-02	2.89E-02	1.82E-02		
346	7.07E-08	7.40E-03	3.89E-01	6.09E-02	7.90E-02	0.00E+00	1.33E-01	3.87E-02	3.20E-02		
347	7.07E-09	7.40E-03	8.87E-02	3.64E-02	2.53E-02	9.86E-03	1.87E-01	6.16E-02	4.69E-02		
Exper.	Par. de escala	Estável 3						Estável 4			
		Erro						Erro			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão	
348	7.07E-02	2.54E-03	2.99E-02	1.59E-02	7.88E-03	4.45E-03	5.20E-02	2.46E-02	1.13E-02		
349	7.07E-03	3.34E-06	2.71E-02	1.53E-02	7.37E-03	6.54E-07	2.96E-02	1.42E-02	7.78E-03		
350	7.07E-04	1.18E-08	3.56E-02	2.10E-02	1.01E-02	7.40E-03	4.08E-02	2.02E-02	8.31E-03		
351	7.07E-05	2.06E-11	3.45E-02	2.00E-02	8.96E-03	4.67E-11	3.40E-02	1.76E-02	9.95E-03		
352	7.07E-06	1.66E-10	3.54E-02	1.67E-02	9.32E-03	1.46E-12	3.25E-02	1.92E-02	9.53E-03		
353	7.07E-07	7.11E-15	5.17E-02	2.19E-02	1.37E-02	2.35E-09	3.94E-02	1.84E-02	8.98E-03		
354	7.07E-08	4.44E-16	6.65E-02	2.56E-02	1.42E-02	4.42E-03	5.91E-02	2.08E-02	1.11E-02		
355	7.07E-09	7.40E-03	1.26E-01	3.61E-02	2.54E-02	1.33E-15	5.17E-02	1.95E-02	1.24E-02		

Tabela 5.32: Resultados com a função de Griewangk: S-Lévy, Lévy e Estável 1 a Estável 4 em 5 dimensões.

O melhor resultado foi na família de distribuição com cauda leve, Estável 4 com erros 10^{-15} .

5.4.3. Experimentos com função de Griewangk: 10 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, \dots, x_{10}, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \dots, \{\alpha_{10\ 1}, \beta_{10\ 2}, \gamma_{10\ 3}, \delta_{10\ 4}\}]$. Nas Tabelas 5.33 e 5.34 pode-se observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 358a, 358b, 366, 368, 376, 377, 385a e 385b.

Análise das distribuições estáveis de Gauss e Cauchy em 10 dimensões na Função Griewangk										
Exper.	Par. de escala	Gauss ($\mathcal{N}(\mu, \sigma^2)$)					Cauchy ($\mathcal{C}(x_0, \gamma)$)			
		γ	σ	Erro				Erro		
			Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
356	7.07E-02	1.00E-01	1.14E-01	1.76E+00	6.18E-01	4.41E-01	3.09E-02	2.26E-01	1.29E-01	5.05E-02
357	7.07E-03	1.00E-02	7.05E-02	1.24E+00	6.58E-01	3.10E-01	2.71E-02	1.17E+00	3.87E-01	2.86E-01
358	7.07E-04	1.00E-03	6.94E-01	2.14E+00	1.31E+00	3.94E-01	6.89E-02	1.01E+00	4.45E-01	2.69E-01
359	7.07E-05	1.00E-04	7.04E-01	2.92E+00	1.31E+00	4.64E-01	1.35E-01	1.70E+00	7.18E-01	3.60E-01
360	7.07E-06	1.00E-05	7.92E-01	2.51E+00	1.42E+00	4.55E-01	4.43E-01	2.96E+00	1.19E+00	5.67E-01
361	7.07E-07	1.00E-06	3.92E-01	2.77E+00	1.27E+00	4.99E-01	4.56E-01	2.88E+00	1.21E+00	5.23E-01
362	7.07E-08	1.00E-07	7.28E-01	2.10E+00	1.29E+00	4.03E-01	4.65E-01	2.56E+00	1.07E+00	4.44E-01
363	7.07E-09	1.00E-08	7.04E-01	3.71E+00	1.35E+00	6.04E-01	7.11E-01	2.80E+00	1.43E+00	5.49E-01

Tabela 5.33: Resultados com a função de Griewangk: Gauss e Cauchy em 10 dimensões.

Análise das distribuições estáveis de S-Lévy e Lévy em 10 dimensões na Função Griewangk										
Exper.	Par. de escala	S Lévy ($\mathcal{S}\text{-Lévy}(\alpha, \beta, \gamma, \delta)$)					Lévy ($\mathcal{L}(\alpha, \beta, \gamma, \delta)$)			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
364	7.07E-02	9.93E-04	1.39E-01	3.70E-02	3.47E-02	3.87E-01	1.74E+00	9.35E-01	3.54E-01	
365	7.07E-03	3.18E-05	1.49E-01	6.85E-02	3.83E-02	1.76E-01	1.90E+00	5.12E-01	3.34E-01	
366	7.07E-04	9.86E-03	2.89E-01	8.96E-02	5.38E-02	1.60E-01	2.05E+00	6.66E-01	4.55E-01	
367	7.07E-05	4.22E-02	3.42E-01	1.74E-01	8.09E-02	1.00E-01	3.19E+00	6.41E-01	7.37E-01	
368	7.07E-06	9.58E-03	1.17E+00	2.17E-01	2.20E-01	1.46E-01	1.32E+00	5.28E-01	3.22E-01	
369	7.07E-07	9.60E-02	1.43E+00	4.48E-01	2.81E-01	1.45E-01	1.52E+00	6.93E-01	3.45E-01	
370	7.07E-08	2.59E-01	2.31E+00	7.13E-01	5.21E-01	3.07E-01	1.93E+00	9.15E-01	4.49E-01	
371	7.07E-09	1.02E-01	2.84E+00	8.22E-01	5.22E-01	3.89E-01	1.77E+00	1.01E+00	3.81E-01	

Exper.	Par. de escala	Estável 1 ($\mathcal{E}_1(\alpha, \beta, \gamma, \delta)$)					Estável 2 ($\mathcal{E}_2(\alpha, \beta, \gamma, \delta)$)			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
372	7.07E-02	2.24E-03	1.03E-01	3.36E-02	3.21E-02	3.29E-03	1.08E-01	2.71E-02	2.49E-02	
373	7.07E-03	7.42E-03	1.43E-01	6.52E-02	3.93E-02	3.53E-05	1.43E-01	5.85E-02	3.98E-02	
374	7.07E-04	9.86E-03	1.55E-01	9.17E-02	3.79E-02	3.48E-05	1.27E-01	5.97E-02	4.18E-02	
375	7.07E-05	9.83E-09	2.35E-01	1.13E-01	5.98E-02	2.49E-02	1.57E-01	9.13E-02	3.62E-02	
376	7.07E-06	2.67E-05	3.82E-01	1.70E-01	1.02E-01	2.33E-09	3.96E-01	1.16E-01	8.67E-02	
377	7.07E-07	4.68E-02	1.07E+00	3.77E-01	2.57E-01	2.46E-02	5.79E-01	2.14E-01	1.39E-01	
378	7.07E-08	1.06E-01	1.43E+00	5.37E-01	3.55E-01	2.49E-02	9.93E-01	2.85E-01	1.97E-01	
379	7.07E-09	2.22E-01	2.06E+00	7.39E-01	4.17E-01	1.20E-01	1.04E+00	5.19E-01	2.60E-01	

Exper.	Par. de escala	Estável 3 ($\mathcal{E}_3(\alpha, \beta, \gamma, \delta)$)					Estável 4 ($\mathcal{E}_4(\alpha, \beta, \gamma, \delta)$)			
		γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
380	7.07E-02	2.47E-02	3.05E-01	1.33E-01	7.96E-02	7.01E-02	5.83E-01	2.77E-01	1.06E-01	
381	7.07E-03	2.07E-04	1.26E-01	4.85E-02	3.90E-02	1.31E-03	8.75E-02	3.00E-02	2.40E-02	
382	7.07E-04	7.17E-06	1.20E-01	6.71E-02	3.80E-02	3.48E-05	1.27E-01	5.97E-02	4.18E-02	
383	7.07E-05	7.40E-03	1.43E-01	7.38E-02	4.24E-02	4.24E-02	1.36E-01	7.95E-02	3.81E-02	
384	7.07E-06	3.60E-09	2.66E-01	9.76E-02	5.30E-02	3.02E-08	1.51E-01	9.41E-02	3.69E-02	
385	7.07E-07	2.68E-06	2.41E-01	9.27E-02	5.28E-02	5.54E-08	1.51E-01	8.82E-02	3.40E-02	
386	7.07E-08	4.70E-06	3.16E-01	1.38E-01	7.38E-02	7.40E-03	1.79E-01	9.34E-02	4.31E-02	
387	7.07E-09	7.82E-03	4.85E-01	2.06E-01	1.19E-01	7.61E-03	3.07E-01	1.11E-01	8.00E-02	

Tabela 5.34: Resultados com a função de Griewangk: S-Lévy, Lévy e Estável 1 a Estável 4 em 10 dim.

O melhor resultado foi na família de distribuição com cauda leve, Estável 2 com erros até 10^{-9} .

5.4.4. Experimentos com função de Griewangk: 30 dimensões

A representação do indivíduo ou cromossomo tem a forma $[x_1, \dots, x_{30}, \{\alpha_{11}, \beta_{12}, \gamma_{13}, \delta_{14}\}, \dots, \{\alpha_{30\ 1}, \beta_{30\ 2}, \gamma_{30\ 3}, \delta_{30\ 4}\}]$. Nas Tabelas 5.35 e 5.36 pode-se observar que os resultados obtidos dos erros mínimos vão melhorando nos experimentos 389, 390, 394, 395, 398a, 398b, 403a e 403b.

Análise das distribuições estáveis de Gauss e Cauchy em 30 dimensões para a Função Griewangk										
Exper.	Par. de escala	Gauss ($\alpha, \beta, \gamma, \delta$)					Cauchy ($\alpha, \beta, \gamma, \delta$)			
		Erro				Erro				
	γ	σ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
388	7.07E-01	1.00E+00	3.03E-01	5.30E-01	4.17E-01	5.69E-02	6.13E-01	8.68E-01	7.83E-01	5.69E-02
389	7.07E-02	1.00E-01	3.81E+00	9.77E+00	6.22E+00	1.50E+00	1.47E-02	3.67E-02	2.39E-02	5.74E-03
390	7.07E-03	1.00E-02	4.18E+00	2.10E+01	9.12E+00	3.19E+00	1.05E+00	2.13E+00	1.25E+00	2.22E-01
391	7.07E-04	1.00E-03	5.03E+00	1.40E+01	9.86E+00	2.67E+00	2.62E+00	9.81E+00	4.89E+00	1.73E+00

Tabela 5.35: Resultados com a função de Griewangk: Gauss e Cauchy em 30 dimensões.

Análise das distribuições estáveis de S-Lévy e Lévy em 30 dimensões na Função vale de Griewangk									
Exper.	Par. de escala	S Lévy ($\alpha, \beta, \gamma, \delta$)				Lévy ($\alpha, \beta, \gamma, \delta$)			
		Erro				Erro			
	γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
392	7.07E-03	1.58E-02	4.22E-02	2.58E-02	6.09E-03				
393	7.07E-04	9.75E-03	1.44E-01	4.36E-02	3.07E-02	3.60E+00	1.04E+01	5.96E+00	1.74E+00
394	7.07E-05	4.82E-01	1.07E+00	8.99E-01	1.62E-01	2.17E+00	1.21E+01	5.54E+00	2.19E+00
395	7.07E-06					2.28E+00	9.74E+00	5.45E+00	1.64E+00
Exper.	Par. de escala	Estável 1 ($\alpha, \beta, \gamma, \delta$)-->Erro				Estável 2 ($\alpha, \beta, \gamma, \delta$)-->Erro			
		Erro				Erro			
	γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
396	7.07E-03	3.40E-02	9.52E-02	5.97E-02	1.30E-02	3.05E-01	6.16E-01	4.17E-01	7.15E-02
397	7.07E-04	3.89E-03	3.92E-02	2.04E-02	9.94E-03	2.18E-02	6.52E-02	4.22E-02	1.00E-02
398	7.07E-05	6.28E-02	7.46E-01	4.17E-01	1.74E-01	2.63E-02	2.00E-01	7.11E-02	3.83E-02
399	7.07E-06	4.91E-02	3.38E-01	1.67E-01	5.89E-02	1.33E-01	1.02E+00	7.42E-01	2.19E-01
400	7.07E-07					9.89E-01	1.15E+00	1.06E+00	3.86E-02
Exper.	Par. de escala	Estável 3 ($\alpha, \beta, \gamma, \delta$)-->Erro				Estável 4 ($\alpha, \beta, \gamma, \delta$)-->Erro			
		Erro				Erro			
	γ	Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
401	7.07E-05	5.71E-02	1.80E-01	1.08E-01	3.52E-02	4.19E-01	8.01E-01	5.59E-01	9.50E-02
402	7.07E-06	4.91E-02	3.38E-01	1.67E-01	5.89E-02	1.13E-01	3.94E-01	2.40E-01	7.55E-02
403	7.07E-07	1.22E-01	9.45E-01	6.14E-01	2.22E-01	1.52E-01	4.77E-01	2.86E-01	8.35E-02

Tabela 5.36: Resultados com a função de Griewangk: S-Lévy, Lévy e Estável 1 a Estável 4 em 30 dim.

Observa-se que a distribuição estável considerada como a melhor para a função de Griewangk é a Estável 1 com erros mínimos da ordem de 10^{-03} . A distribuição S-Lévy também apresenta resultados semelhantes porém levemente maiores, isto se justifica devido ao fato do parâmetro de estabilidade, que define a família de distribuição ser muito próximo, como já visto na Tabela 5.3.

5.4.5. Dados dos experimentos para a distribuição Estável 1

Como verificado na seção anterior o melhor resultado foi utilizando-se a distribuição de Estável 1, que desperta para este novo conjunto de experimentos. Neste caso verifica-se a evolução dos dados com o número de execuções de 3 até 30, para execuções independentes, variando-se o número de gerações de 500, 1500, 5000 e 10000.

Tem-se o propósito de verificar em que condições será mais interessante executar experimentos no futuro, ou seja, com um número maior de gerações associado a um número menor de execuções, ou ao contrário. Vale lembrar que, reduzindo significativamente o número de execuções, pode-se perder a relevância da amostra (repetibilidade).

Na Tabelas 5.37 pode-se observar que os resultados obtidos dos erros mínimos vão melhorando lentamente na distribuição Estável 1, quando aumenta-se o esforço computacional até 10000 gerações mesmo com a redução do número de execuções, considerando-se um mínimo de 6 execuções.

Exper.	Execuções	Estável 1 (91 90 91 92 93 94 95 96 97 98 99 100)-->Erro				Estável 1 (91 90 91 92 93 94 95 96 97 98 99 100)-->Erro			
		Erro para 500 gerações				Erro para 1500 gerações			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
404	3.00E+00	1.042E-02	3.170E-02	2.343E-02	1.141E-02	6.177E-05	7.453E-03	4.985E-03	4.264E-03
405	6.00E+00	1.576E-02	2.608E-02	2.013E-02	3.820E-03	4.499E-05	9.905E-03	2.928E-03	4.521E-03
406	9.00E+00	3.850E-05	9.891E-03	4.435E-03	4.227E-03	3.578E-05	7.448E-03	1.693E-03	3.257E-03
407	1.20E+01	8.724E-03	3.235E-02	1.857E-02	6.491E-03	4.642E-05	1.238E-02	3.547E-03	4.514E-03
408	1.50E+01	7.803E-03	3.467E-02	1.922E-02	7.907E-03	3.642E-05	7.450E-03	1.037E-03	2.601E-03
409	1.80E+01	6.876E-03	3.685E-02	2.114E-02	9.406E-03	4.622E-05	9.914E-03	3.615E-03	4.158E-03
410	2.10E+01	3.365E-03	5.500E-02	1.799E-02	1.036E-02	4.259E-05	1.236E-02	3.927E-03	4.308E-03
411	2.40E+01	5.950E-03	6.597E-02	1.801E-02	1.164E-02	4.029E-05	9.898E-03	3.239E-03	3.880E-03
412	2.70E+01	6.077E-03	3.723E-02	1.859E-02	8.432E-03	3.762E-05	1.237E-02	3.981E-03	4.282E-03
413	3.00E+01	8.538E-03	3.456E-02	1.830E-02	7.313E-03	3.157E-05	7.459E-03	2.519E-03	3.545E-03
Exper.	Execuções	Erro para 5000 gerações				Erro para 10000 gerações			
		Mínimo	Máximo	Média	Desv. Padrão	Mínimo	Máximo	Média	Desv. Padrão
414	3.00E+00	8.145E-06	1.345E-05	1.131E-05	2.798E-06	7.403E-03	7.403E-03	9.045E-03	1.421E-03
415	6.00E+00	1.218E-05	7.407E-03	1.245E-03	3.019E-03	7.452E-06	9.865E-03	5.348E-03	4.245E-03
416	9.00E+00	1.036E-05	9.872E-03	4.394E-03	4.230E-03	6.062E-06	1.479E-02	2.472E-03	5.226E-03
417	1.20E+01	9.976E-06	9.870E-03	3.299E-03	4.116E-03	9.300E-06	7.410E-03	1.861E-03	3.346E-03
418	1.50E+01	1.086E-05	7.411E-03	1.986E-03	3.386E-03	5.540E-06	7.404E-03	3.458E-03	3.819E-03
419	1.80E+01	9.541E-06	9.869E-03	3.025E-03	3.926E-03	1.056E-05	9.873E-03	3.437E-03	3.979E-03
420	2.10E+01	7.086E-06	1.233E-02	2.125E-03	3.996E-03	4.808E-06	9.865E-03	4.233E-03	4.208E-03
421	2.40E+01	8.523E-06	9.867E-03	3.196E-03	3.879E-03	5.606E-06	1.233E-02	2.062E-03	3.753E-03
422	2.70E+01	1.030E-05	1.233E-02	3.664E-03	4.289E-03	5.763E-06	9.865E-03	3.477E-03	4.002E-03
423	3.00E+01	1.023E-05	1.233E-02	3.792E-03	4.231E-03	6.591E-06	1.478E-02	4.609E-03	4.328E-03

Tabela 5.37: Resultados com a função de Griewangk: Estável 1 de 500 a 10000 gerações.

Os melhores resultados para 500, 1500, 5000 e 10000 gerações foram respectivamente nos experimentos 407, 414, 421a e 421b. Para a função de Griewangk no conjunto de experimentos de 405 a 424, verifica-se que é vantajoso aumentar o número de gerações considerando um mínimo de 6 execuções.

5.4.6. Auto-adaptação dos parâmetros estratégicos α e γ em 30 dimensões

A motivação para executar os experimentos desta seção, reside no fato de analisar se é possível deixar o processo evolutivo descobrir uma família de distribuições estáveis que acarrete melhor nível de qualidade do que os experimentos com famílias pre-definidas. Este novo conjunto de experimentos se encontra dentro das mesmas parametrizações que as existentes na seção anterior, que não é a mesma da Seção 5.1, fato que justifica a execução deste experimentos.

Sendo assim, inicialmente define-se um intervalo de valores válidos para os parâmetros: índice de estabilidade $\alpha \in [0.25, 2]$; assimétrico $\beta = 0$; escala $Y \in [7.07E-09, 7.07E+00]$ e posição $\delta = 0$. Os parâmetros estratégicos podem variar livremente para cada dimensão de forma independente. Foi

observado o valor dos índices de estabilidade e escala para a melhor execução dentre as execuções de cada experimento, conforme a Tabela 5.38.

Exper.	Execuç	Distribuições Estáveis						Distribuições Estáveis					
		Erro para 500 gerações						Erro para 1500 gerações					
		Mínimo	Máximo	Média	Desv. Padrão	Estabil. (α)	Escala (γ)	Mínimo	Máximo	Média	Desv. Padrão	Estabil. (α)	Escala (γ)
424	3.00E+00	8.836E-01	1.019E+00	9.696E-01	7.471E-02	1.089E+00	1.690E+00	8.771E-01	8.927E-01	8.847E-01	7.814E-03	1.299E+00	2.652E+00
425	6.00E+00	9.662E-01	1.025E+00	9.969E-01	1.959E-02	1.024E+00	7.575E-01	8.439E-01	9.785E-01	9.092E-01	5.687E-02	1.100E+00	1.661E+00
426	9.00E+00	9.351E-01	1.042E+00	1.001E+00	3.643E-02	1.152E+00	2.039E-01	7.496E-01	1.010E+00	8.704E-01	8.487E-02	1.213E+00	8.847E-01
427	1.20E+01	8.127E-01	1.026E+00	9.637E-01	6.320E-02	1.059E+00	2.330E-01	6.853E-01	1.008E+00	8.949E-01	9.351E-02	1.019E+00	2.281E-01

Tabela 5.38: Resultados com a função de Griewangk: Auto-adapt. livre nas n-dim de 500 a 1500 gerações.

Os melhores resultados para 500 e 1500 gerações foram respectivamente nos experimentos 427a, 427b. Observando-se os valores dos parâmetros de estabilidade e de escala, na comparação com os dados obtidos na seção anterior os valores encontrados nos experimentos 404 a 423 estes não se aproximam do resultado desejado, ou seja, erro mínimo $3.16E-05$ obtido no experimento 413 em relação ao erro mínimo $6.784E-01$ no experimento 427b. Na Seção 5.3.8. conjuga-se melhor a interlação dos dados experimentais com a ampliação dos experimentos da seção seguinte, que fixa os parâmetros estratégicos para todas as 30 dimensões.

5.4.7. Auto-adaptação com parâmetro estratégico α e γ único nas 30 dimensões

Nesta seção, fixa-se a aplicação do parâmetro estratégico as 30 dimensões com o propósito de diminuir a turbulência gerada pela liberdade total de se aplicar 30 parâmetros estratégicos independentes nas 30 dimensões [Lima e de Oliveira, 06].

A auto-adaptação dos parâmetros estratégicos possibilita uma evolução dentre as famílias de distribuições estáveis onde, inicialmente define-se um intervalo de valores válidos para os parâmetros: índice de estabilidade $\alpha \in [0.25, 2]$; assimétrico $\beta = 0$; escala $\gamma \in [7.07E-9, 7.07E+0]$ e posição $\delta = 0$. Como dito, o parâmetro estratégico é fixo para todas 30 dimensões, sendo observado o valor do índice de estabilidade e escala para a melhor execução dentre as execuções de cada experimento, conforme a Tabela 5.39.

Exper.	Execuç.	Distribuições Estáveis						Distribuições Estáveis					
		Erro para 500 gerações						Erro para 1500 gerações					
		Mínimo	Máximo	Média	Desv. Padr.	Estabil. (☒)	Escala (☒)	Mínimo	Máximo	Média	Desv. Padr.	Estabil. (☒)	Escala (☒)
428	3.00E+00	9.90E-01	1.03E+00	1.01E+00	1.89E-02	1.13E+00	7.41E-01	7.92E-01	9.02E-01	8.57E-01	5.78E-02	1.04E+00	1.56E+00
429	6.00E+00	9.34E-01	1.02E+00	9.75E-01	3.84E-02	1.21E+00	1.03E+00	8.02E-01	9.74E-01	8.66E-01	5.74E-02	1.25E+00	8.58E-01
430	9.00E+00	8.21E-01	1.02E+00	9.53E-01	8.65E-02	1.13E+00	1.81E+00	8.36E-01	9.80E-01	9.21E-01	5.06E-02	1.08E+00	2.17E+00
431	1.20E+01	8.58E-01	1.05E+00	9.73E-01	5.51E-02	1.12E+00	6.55E-01						

Tabela 5.39: Resultados com a função de Griewangk: Auto-adapt. fixa nas n-dim. de 500 a 1500 gerações.

Os melhores resultados para 500 e 1500 gerações foram respectivamente nos experimentos 436 e 434. Observando-se os valores dos parâmetros de estabilidade, escala e comparando-se com os dados obtidos na Seção 5.3.5, os valores encontrados nos experimentos 434 a 437 não se aproximam do resultado desejado, ou seja, erro mínimo $3.16E-05$ obtido no experimento 413 na distribuição Estável 1 é muito melhor em relação ao erro mínimo $8.21E-01$ no experimento 436.

Esta conclusão se baseia no fato que o menor erro encontrado determina a melhor busca na função de teste de Griewangk. Como os valores obtidos nesta seção e na anterior não foram adequados, os experimentos foram reduzidos nesta seção por se apresentarem muito distantes dos desejados, não justificando o esforço computacional demandado.

5.4.8. Conclusões sobre os experimentos com a função de Griewangk

Para uma melhor visualização dos resultados experimentais realizados nas Seções 5.4.5 a 5.4.7 é criada a Tabela 5.40 e a Figura 5.3, onde se constata que ambos os processos empregados nas Seções 5.4.6 e 5.4.7, respectivamente parametrização livre dos parâmetros estratégicos e fixa para as 30 dimensões não conseguiram aproximar-se da parametrização pré-definida, constatando-se que na busca de uma família de distribuição a Estável 1 apresenta melhores resultados.

Geração	Erros entre f(x) mínimo, máximo e média utilizando a função de Rastrigin											
	S-Lévy				Auto-adapt. variação livre: par. estr. n-dim.				Auto-adapt. par. estr. iguais para n-dim.			
	Mínimo	Máximo	Média	Desv. Padr.	Mínimo	Máximo	Média	Desv. Padr.	Mínimo	Máximo	Média	Desv. Padr.
500	3.85E-05	9.89E-03	4.44E-03	4.23E-03	8.13E-01	1.03E+00	9.64E-01	6.32E-02	8.21E-01	1.02E+00	9.53E-01	8.65E-02
1500	3.16E-05	7.46E-03	2.52E-03	3.54E-03	6.85E-01	1.01E+00	8.95E-01	9.35E-02	7.92E-01	9.02E-01	8.57E-01	5.78E-02
5000	7.09E-06	1.23E-02	2.12E-03	4.00E-03								

Tabela 5.40: Comparação na função de Griewangk: Estável, Auto-adaptação livre e fixa nas n-dim.

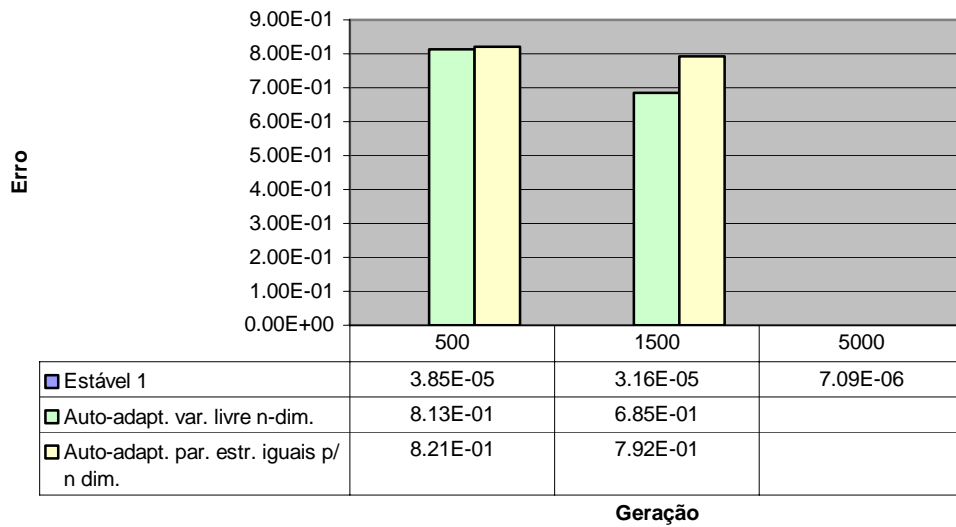


Figura 5.2: Gráfico na função de Rosenbrock: S-Lévy, Auto-adaptação livre e fixa nas n-dim.

Nesta função de teste observa-se que a utilização da distribuição S-Lévy apresenta melhores resultados independente do esforço computacional (número de gerações). Quando utiliza-se auto-adaptação em ambos processos de controle dos parâmetros estratégicos nas n-dimensões: variação livre ou fixa, obteve-se resultados próximos ($10E-1$), nos experimentos com 1500 e 5000 gerações, e apresentando uma ligeira vantagem para o processo de auto-adaptação com variação livre nas n-dimensões em 500 gerações.

CAPÍTULO VI

CONCLUSÕES

Existem inúmeros trabalhos que tratam do uso de distribuições Gaussianas em conjunto com estratégias evolutivas ou computação evolutiva em geral. Neste trabalho procuramos ampliar o conjunto de distribuições e analisar o desempenho na busca de mínimos globais, bem como a qualidade e o esforço computacional demandada pelo algoritmo evolutivo.

Com as análises dos dados obtidos nas dimensões 2, 5, 10 e 30 verificamos que a busca da solução encontrada na dimensão 2 não se aplicava diretamente na dimensão subsequente, embora indicasse um provável caminho na próxima dimensão pesquisada dentro da mesma função de teste.

Esta observação determina que o esforço despendido para descobrir passo-a-passo, a melhor distribuição, passando por um processo longo de parametrizações e testes, até se obter uma solução para o conjunto escolhido: número de elementos na população, quantidade de execuções, número de gerações, tipos de operadores, esquema de seleção e parametrização da própria distribuição.

Desta forma as opções de implementação bem como os conjuntos de valores utilizados com certeza determinam o sucesso ou não da obtenção dos resultados. Um fato importante que pode ser acompanhado em todos os resultados de forma inegável é que existe uma distribuição diferenciada para cada função de teste analisada em uma determinada dimensão.

Aparentemente, este trabalho mostra um caminho alternativo; no entanto, apesar de avançar no tema, sugere-se que um estudo mais aprofundado sobre as questões aqui levantadas seja conduzido, indo-se muito além dos experimentos reportados.

Na busca de distribuições não-nomeadas foram adicionadas mais quatro famílias, que exploram as distribuições com baixo índice de estabilidade ($0.25 \leq \alpha < 0.5$) e com parâmetros de escala extremamente pequenos ($0.707E-13 \leq \gamma < 0.707E-1$), valores estes decorrência do aprofundamento

passo-a-passo obtido no desenvolvimento dos experimentos. Certamente poderiam ser outros valores, gerando uma infinidade de experimentos, sempre na busca de melhores ajustes.

Os experimentos apresentados que utilizam auto-adaptação com variação em todas as dimensões ou mesmo os que fixam os parâmetros estratégicos nas n -dimensões, pouco se aproximaram dos valores obtidos na metodologia passo-a-passo. Embora a auto-adaptação não tenha neste trabalho apresentado um resultado adequado, inúmeros métodos alternativos são aplicados com sucesso em diversas funções de teste na literatura.

De fato, há muito que sistematizar em termos de experimentos, antes que conclusões mais fortes possam ser tiradas. Nesse sentido, deve-se avaliar o efeito de outros aspectos do algoritmo evolutivo empregado, como seu desempenho em outras funções de teste, bem como analisar com mais cautela o efeito dos próprios aspectos que foram discutidos.

Todas essas são direções necessárias para validar as observações quantitativas aqui realizadas. Indubitável, entretanto, é que este estudo envolvendo distribuições estáveis se alinha com inúmeros outros que têm aparecido na literatura em anos recentes, atestando a ubiquidade e importância das distribuições estáveis, especialmente a de Lévy, na natureza e na matemática, o que é forte sugestão de seu potencial também na computação evolutiva.

6.1. Sugestões para trabalhos futuros

REFERÊNCIAS BIBLIOGRÁFICAS

- [Almeida e Evsukoff, 03] Paulo Eduardo Maciel de Almeida e Alexandre Gonçalves Evsukoff. “Sistemas Fuzzy”. *Sistemas inteligentes: Fundamentos e aplicações*. Coordenadora Solange Oliveira Rezende. Editora Manole, 169-224, 2003.
- [Angeline *et al*, 96] P. J. Angeline, D.B. Fogel e L.J. Fogel. “A comparison of self-adaptation methods for finite state machines in a dynamic environment”. *Proc. 5th. Ann. Conf. on Evolutionary Programming*. Eds. L.J. Fogel, P. J. Angeline and T Bäck. Cambridge, MA. MIT Press, 441-9, 1996.
- [Anderson, 00] John R. Anderson. *Learning and memory an integrated approach*. Wiley. 2^a ed, 2000.
- [Atmar, 94] Wirt Atmar. “Notes on the simulation of evolution”. *IEEE Transactions on Neural Networks*, 5(1):130-148, 1994.
- [Azevedo, Brasil e Oliveira, 00] Fernando M. de Azevedo, Lourdes M. Brasil e Roberto Célio L. de Oliveira. *Redes neurais com aplicações em controle e em sistemas especialistas*. Visual Books. 1^a ed, 2000.
- [Bäck *et al*, 00] Thomas Bäck, David B. Fogel, Darrell Whitley e Peter J. Angeline. “Mutation operators”. In T. Bäck, D. B. Fogel e Z. Michalewicz eds. *Evolutionary computation 1- Basic algorithms and operators*. Bristol e Philadelphia: Institute of Physics Publishing, 237-243, 2000.
- [Bäck, Hammel e Schwefel, 97] Thomas Bäck, Ulrich Hammel e Hans-Paul Schwefel. “Evolutionary computational: comments on the history and current state”. *IEEE Transactions on Evolutionary Computation*, 1(1):3-17, 1997.
- [Banzhaf *et al*, 02] Wolfgang Banzhaf, Peter Nordin, Robert E. Keller e Frank D. Francone. *Genetic programming an introduction*. Morgan Kaufmann. 5^a ed, 2002.

- [Beasley, 00] David Beasley. "Possible applications of evolutionary computation". In T. Bäck, D. B. Fogel e Z. Michalewicz eds. *Evolutionary computation 1- Basic algorithms and operators*. Bristol e Philadelphia: Institute of Physics Publishing, 4-19, 2000.
- [Beyer, 01] Hans-Georg Beyer. *The Theory of evolution strategies*. Springer-Verlag. 1^a ed, 2001.
- [Bertoin, 02] Jean Bertoin. *Lévy Processes*. Cambridge University Press. 3^a ed, 2002.
- [Bittencourt, 06] Guilherme Bittencourt. *Inteligência artificial: Ferramentas e teorias*. Editora da UFSC. 3^a ed, 2006.
- [Body, 94] Ian D. Body, Patrick D. Surry e Nicholas J. Radcliffe. "Constrained gas network pipe sizing with genetic algorithms". *Edinburgh Parallel Computing Centre*. EPCC, TR94(11):1-7, 1994.
- [Borak, 05] Szymon Borak. Wolfgang Härdle e Rafal Weron. "Stable Distributions". *SFB 649 Economic risk*. Humboldt-Universität zu Berlin. Discussion Paper 2005(008):1-25, 2005.
- [Booker et al, 00] Lashon B. Booker, David B. Fogel, Darrell Whitley, Peter J. Angeline e A. E. Eiben. "Recombination". In T. Bäck, D. B. Fogel e Z. Michalewicz eds. *Evolutionary computation 1- Basic algorithms and operators*. Bristol e Philadelphia: Institute of Physics Publishing, 256-307, 2000.
- [Carrapiço, 01] Francisco J. Nascimento Carrapiço. "A origem da vida e a sua evolução-Uma questão central no âmbito da exobiologia". *Centro de Biologia Ambiental*. Departamento de Biologia Vegetal. Faculdade de Ciências da Universidade de Lisboa. p. 25-32, 2001. Disponível em <http://azolla.fc.ul.pt/astrobiologia/Exobiologia.pdf>.
- [Carvalho, Braga e Ludermir, 03] André Carlos P. de L. F. de Carvalho, Antônio de P. Braga e Teresa B. Ludermir. "Computação Evolutiva". *Sistemas inteligentes: Fundamentos e aplicações*. Coordenadora Solange Oliveira Rezende. Editora Manole. 225-248, 2003.

- [Cercone e McCalla, 94] Nicholas Cercone e Gordon McCalla. “Ten years of computational intelligence”. *Computational Intelligence*. Blackwell's Pub. 10(4):1-6, 1994.
- [Castro, 02] Leandro Nunes de Castro. Resumo do livro *The origin of species*. Charles R. Darwin. Wordsworth Editions Limited. 1998. 1-14, 2002.
- [Castro, 06] Leandro Nunes de Castro. *Fundamentals of Natural Computing: Basic Concepts, Algorithms, and Applications*. Chapman & Hall/CRC. 1^a ed, 2006.
- [Chen, Liu e Chen, 06] Chan-Hong Chen, Wei-Nan Liu e Ying-Ping Chen. “Adaptive discretization for probabilistic model building genetic algorithms”. *GECCO' 06. Session Genetic Algorithms-9: Adaptation*. 1103-10, 2006.
- [Chorafas, 88] Dimitris N. Chorafas. *Sistemas especialistas - Aplicações comerciais*. Tradução: Miriam Fonseca Diniz. Revisão Técnica: Nizam Omar. McGraw-Hill. 1^a ed, 1988.
- [Cunha e Ribeiro, 87] Horácio da Cunha e Sousa Ribeiro. *Introdução aos sistemas especialistas*. Livros Técnicos e Científicos Editora. 1^a ed, 1987.
- [Damazio, Seixas e Soares, 03] Denis O. Damazio. José M. Seixas e A. C. Soares. “Um classificador neural compacto e eficiente com capacidade de identificar contaminação em dados experimentais”. *Controle & Automação. Sociedade brasileira de automática*. 14(4):359-367, 2003.
- [de Oliveira e Gutierrez, 06] P.P.B. de Oliveira e A.B.M. Gutierrez. “Auto-adaptação em estratégias evolutivas com distribuições estáveis: resultados preliminares”. In: A.P.L.F. Carvalho e M.M.B.R. Vellasco, Eds. *Anais do I Workshop on Computational Intelligence - IX Simpósio Brasileiro de Redes Neurais*, CD-ROM, Sociedade Brasileira de Computação, Ribeirão Preto, SP, 23-26 Out 2006.

- [Deb, 00] Kalyanmoy Deb. "Introduction to representations". In T. Bäck, D. B. Fogel e Z. Michalewicz eds. *Evolutionary computation 1- Basic algorithms and operators*. Bristol e Philadelphia: Institute of Physics Publishing, 127-131, 2000.
- [Eiben, Hinterding e Michalewicz, 99] Ágoston E. Eiben, Robert Hinterding e Zbigniew Michalewicz. "Parameter control in evolutionary algorithms". *IEEE Transactions on Evolutionary Computation*, 3:124-141, 1999.
- [Eiben e Smith, 03a] Ágoston E. Eiben e J.E. Smith. *Introduction to evolutionary computing*. Springer-Verlag. 1^a ed, 2003.
- [Eiben e Smith, 03b] Ágoston E. Eiben e J.E. Smith. *Introduction to evolutionary computing*. Genetic algorithms. 3:28-30, 2003. Disponível em <http://www.cs.vu.nl/~gusz/ecbook/slides>.
- [Eshelman, 00] Larry J. Eshelman. "Genetic In T. Bäck, D. B. Fogel e Z. Michalewicz eds. *Evolutionary computation 1- Basic algorithms and operators*. Bristol e Philadelphia: Institute of Physics Publishing, 64-80, 2000.
- [Fernandes, 03] Anita Maria da Rocha Fernandes. *Inteligência artificial - Noções gerais*. Visual Books. 1^a ed, 2003.
- [Fogel e Atmar, 90] David B. Fogel e J. W. Atmar. "Comparing genetic operators with gaussian mutations in simulated evolutionary processes using linear systems". *Biological Cybernetics*, 63:111-114, 1990.
- [Fogel e Porto, 90] David B. Fogel. Lawrence J. Fogel. V. W. Porto. "Evolving neural networks". *Biological Cybernetics*, 63:487-493, 1990.
- [Fogel *et al*, 91] David B. Fogel, Lawrence J. Fogel e J. W. Atmar. "Meta-evolutionary programming". *Proc. 25th Asilomar conf. on signals systems and computers*. Pacific Grove. CA. Ed. R. R. Chen. 540-545, 1991.

- [Fogel, 94a] David B. Fogel. “An Introduction to Simulated Evolutionary Optimization”. *IEEE Trans. on Neural Networks*, 5(1):3-14, 1994.
- [Fogel, 94b] David B. Fogel. “Asymptotic convergence properties of genetic algorithms and evolutionary programming: Analysis and experiments”. *Cybernetics and Systems: An International Journal*. 389-407, 1994.
- [Fogel D.B., 95] David B. Fogel. “Review of computational intelligence: imitating life”. *IEEE Trans. Neural Networks*, 6(6):1562-1565, 1995.
- [Fogel, 00] David B. Fogel. “Principles of evolutionary processes”. In T. Bäck, D. B. Fogel e Z. Michalewicz eds. *Evolutionary computation 1- Basic algorithms and operators*. Bristol e Philadelphia: Institute of Physics Publishing, 23-26, 2000.
- [François, 99] Olivier François. “Controlling mutation/selection algorithms with stochastic approximation”. Proceedings of the 1999 Congress on Evolutionary computation-CEC99. Washington, USA. p. 1487-93, 1999.
- [Freitas *et al*, 93] A.A. Freitas. J.C. Anacleto. R. Morábito Neto e C. Kirner. “Algoritmos genéticos e sua aplicação ao problema do corte de barras”. In Gilberto S. Nakamiti, Hélio A. Navarro, Ivan R. Guilherme, José P. A. Prado e Ricardo L. Freitas eds. *I Simpósio brasileiro de automação inteligente*. Rio Claro, São Paulo. 38-47, 1993.
- [GEATbx, 07] “GEATbx – The genetic and evolutionary algorithm toolbox for Matlab”. GEATbx Support. 2007. Disponível em <http://www.geatbx.com/docu/fcnindex.html>
- [Gómez, 04] Jonatan Gómez. “Self adaptation of operator rates in evolutionary algorithms”. *GECCO 2004*. Eds. K. Deb *et al*. Springer-Verlag Berlin Heidelberg. 1162-1173, 2004.
- [Haykin, 01] Simon Haykin. *Redes neurais: princípios e prática*. Editora Bookman. 2^a ed, 2001.
- [Holland, 92] John H. Holland. *Adaptation in natural artificial systems*. MIT Press. 1^a ed, 1992.
- [Jacob, 01] Christian Jacob. *Illustrating evolutionary computation with Mathematica*. University of Calgary. Morgan Kaufmann Publishers. 1^a ed, 2001.

- [Kinnear, 00] Kenneth E. Kinnear Jr. “Derivative methods in genetic programming”. In T. Bäck, D. B. Fogel e Z. Michalewicz eds. *Evolutionary computation 1- Basic algorithms and operators*. Bristol e Philadelphia: Institute of Physics Publishing, 103-113, 2000.
- [Klug e Cummings, 00] William S. Klug e Michael R. Cummings. *Concepts of genetics*. Prentice Hall. 6^a ed, 2000.
- [Larson e Farber, 04] Ron Larson e Bestsy Farber. *Estatística aplicada*. Pearson Prentice Hall. 2^a Ed, 2004.
- [Laprade e Torres, 93] Alain Laprade e Germano Lambert-Torres. “Controle de um veículo utilizando a teoria dos conjuntos nebulosos”. In Gilberto S. Nakamiti, Hélio A. Navarro, Ivan R. Guilherme, José P. A. Prado e Ricardo L. Freitas eds. *I Simpósio brasileiro de automação inteligente*. Rio Claro, São Paulo. 165-174, 1993.
- [Lee e Yao, 04] Chang-Yong Lee e Xin Yao. “Evolutionary programming using mutations based on Lévy probability distribution”. *IEEE Transactions on evolutionary computation*, 8(1):1-13, 2004.
- [Levine, Drang e Edelson, 88] Robert I. Levine, Diane E. Drang e Barry Edelson. *Inteligência artificial e sistemas especialistas – aplicações e exemplos práticos*. McGraw-Hill. 1^a ed, 1988.
- [Lewis, 00] Harry R. Lewis e Christos H. Papadimitriou. *Elementos de teoria da computação*. Bookman. 2^a ed, 2000.
- [Lima, 04] Hugo Santana Lima. *Estratégias de variação automática da pressão seletiva em uma classe de algoritmos evolutivos*. Universidade Presbiteriana Mackenzie. 2004.
- [Lima e de Oliveira, 06] Hugo Santana Lima e Pedro Paulo Balbi de Oliveira. “Estratégias de variação automática da pressão seletiva em uma classe de algoritmos evolutivos”. Anais do XXVI Congresso da Sociedade Brasileira de Computação - SBC, XXXIII Seminário integrado de software e hardware - SEMISH. Campo Grande, MS, 151-165, 2006. Disponível em <http://natalnet.dca.ufrn.br/sbc2006/pdf/arq0113.pdf>.
- [Medeiros, 02] Felipe Leonardo Lobo Medeiros. *Algoritmo genético híbrido como um método de busca de estados estacionários de sistemas dinâmicos*. INPE. 2002.

- [Michalewicz, 99] Zbigniew Michalewicz. *Genetic algorithms + Data structures = Evolutionary programs*. Springer-Verlag. 3^a ed, 1999.
- [Milone, 04] Giuseppe Milone. *Estática geral aplicada*. Thomson. 1^a ed, 2004.
- [Mitchell, 04] Melanie Mitchell. *An introduction to genetic algorithms*. Prentice-Hall of India Private Limited. 1^a ed, 2004.
- [Nehab, 04] Diego F. Nehab e Marco Aurélio C. Pacheco. "Schemata theory for the real coding and arithmetical operators". *19th ACM Symposium on Applied Computing. Session: Evolutionary computation and optimization (ECO)*. ACM Press, 1-58113-812-1. Nicosia, Cyprus, 1006-12, 2004.
- [Nolan, 01] John P. Nolan. *Maximum likelihood estimation and diagnostics for stable distributions*. In O. E. Barndorff-Nielsen, T. Mikosch e S. I. Resnick eds. *Levy processes - Theory and applications*. Boston: Birkhäuser, 379-400, 2001.
- [Nolan, 02] John P. Nolan. *Stable Distributions*. Birkhauser. 1^a ed, 2002.
- [Nolan, 05] John P. Nolan. *Stable Distributions. Models for Heavy Tailed Data*. American University. 2005. Disponível em <http://academic2.american.edu/~jpnolan/stable/chap1.pdf>.
- [Nolan, 06a] John P. Nolan. *Information on stable distributions*. American University. 2006. Disponível em <http://academic2.american.edu/~jpnolan/stable/stable.html>.
- [Nolan, 06b] John P. Nolan. *Stable distributions pdf (probability density function). cdf (cumulative distribution function). quantiles (quantiles for stable distribution with parameters). random (stable random sample)*. American University. 2006. Disponível em <http://www.hostsrv.com/webmac/app1/MSPScripts/webm1016/Stable/SPDF.jsp>
- [Nieberg e Beyer, 07] Silja Meyer-Nieberg e Hans-Georg Beyer. "Self-adaptation in evolutionary algorithms". In F. Lobo, C. Lima e Z. Michalewicz, eds. *Parameter Setting in Evolutionary Algorithms*, Springer-Verlag, 2007
- [Ostermeier, 92] A. Ostermeier. "An evolution strategy with momentum adaptation of the random number distribution". In R. Männer e B. Manderick eds. *Proc. 2nd Int. Conf. on Parallel Problem Solving from Nature*. Brussels, Belgium, 199-208, 1992.

- [Paraíso *et al*, 93] E. C. Paraíso. L. M. Silveira. M. P. Ramos e C. A. A. Kaestner. “Sistema especialista em manutenção preditiva de equipamentos e plantas industriais”. In Gilberto S. Nakamiti, Hélio A. Navarro, Ivan R. Guilherme, José P. A. Prado e Ricardo L. Freitas eds. *I Simpósio brasileiro de automação inteligente*. Rio Claro, São Paulo, 96-104, 1993.
- [Porto, 00] V. William Porto. “Evolutionary Programming”. In T. Bäck, D. B. Fogel e Z. Michalewicz eds. *Evolutionary computation 1- Basic algorithms and operators*. Bristol e Philadelphia: Institute of Physics Publishing, 89-102, 2000.
- [Rao e Rao, 93] Valluru B. Rao e Hayagriva V. Rao. *C++ Neural networks and fuzzy logic*. MIS Press. 1^a ed, 1993.
- [Reisinger e Miikkulainen, 06] Joseph Reisinger e Risto Miikkulainen. “Selecting for Evolvable Representations”. *GECCO'06. Session Genetic Algorithms-7: Representations*. ACM 1-59593-186. Seattle, Washigton, USA, 2006.
- [Rimmer e Nolan, 05] Robert H Rimmer e John P. Nolan. “Stable distributions in Mathematica”. *The Mathematica Journal*, 9(4):776-789, 2005. Disponível em <http://www.mathematica-journal.com/issue/v9i4/>.
- [Rudolph, 98] Günter Rudolph. “Local convergence rates of simple evolutionary algorithms with Cauchy mutations”. Universität Dortmund. Fachbereich Informatik. LS(XI):1-20, 1998.
- [Russell e Norvig, 04] Stuart Russel e Peter Norvig. *Inteligência artificial*. Editora Campus. 3^a ed, 2004.
- [Saravanan *et al*, 95] N. Saravanan. David B. Fogel e Kevin M. Nelson. “A comparison of methods for self-adaptation in evolutionary algorithms”. *BioSystems*, 36(2):157-166, 1995.
- [Schoenauer e Michalewicz, 97] Marc Schoenauer e Zbigniew Michalewicz. “Boundary operators for constrained parameter optimization problems”. In Thomas Bäck ed. *7th International Conference on Genetic Algorithms*. Morgan Kaufmann. East Lansing, MI, USA, 320-329, 1997.
- [Silvia, Campos e Cunha, 01] Jorge Luiz de Castro e Silva, Marcilia Andrade Campos e Paulo Roberto Freire Cunha. “Modelagem estocástica de tráfego de redes de alta velocidade”. XXI

- Congresso da sociedade brasileira de computação SBC/JAI. Fortaleza, Ceará, 2001.
Disponível em www.lia.ufc.br/sbc2001/download/jai2/pdf
- [Soler, 07] Júlia M. Pavan Soler. “Técnicas computacionais em probabilidade e estatística I”. Instituto de matemática e estatística da Universidade São Paulo (IME/USP). 2007. Disponível em www.ime.usp.br/~pavan/pdf/MAE5704/_Material_Aula_3.pdf
- [Tinos e Carvalho, 03] Renato Tinos e André C. P. F. de Carvalho. “Alteração da probabilidade de mutação do gene em algoritmos genéticos aplicados a problemas não-estacionários”. *XXIII Congresso da Sociedade Brasileira de Computação. SBC2003*. Campinas, São Paulo, 277-286, 2003.
- [Tsallis, 00] Constantino Tsallis. “As distribuições de Lévy”. *Revista Brasileira de Ensino de Física*, 22(2):156-162, 2000.
- [Vicente, 05] Prof. Dr. Renato Vicente. “Análise de risco. Value-at-risk: Um overview. Parte 1”. Escola de Artes, Ciências e Humanidades. Universidade São Paulo. p. 16-19, 2005. Disponível em <http://www.uspleste.usp.br/rvicente/teach.html>.
- [Weicker, 99] Karsten Weicker e Nicole Weicker. ”On evolution strategy optimization in dynamic environments“. *Proc. of the 1999 Congress on evolutionary computation - CEC99*. Washington, D.C., USA. IEEE, 3:2039-2046, 1999.
- [Weiss e Kulikowski, 88] Sholom M. Weiss e Casimir A. Kulikowski. *Guia prático para projetar sistemas especialistas*. Livros Técnicos e Científicos Editora. 1^a ed, 1988.
- [Whitacre, Pham e Sarker, 06] James M. Whitacre, Tuan Q. Pham e Ruhul A. Sarker. “Use of statistical outlier detection method in adaptive evolutionary algorithms”. *GECCO’ 06. Session Genetic Algorithms-9: Adaptation*. ACM 1-59593-186-4. Seattle, Washigton, USA, 2006.
- [Wolfram, 99] Stephen Wolfram. *The Mathematica book*. Wolfram Media/Cambridge Universty Press. 1999.
- [Woyczynski, 01] Wojbor A. Woyczynski. “Levy processes in the physical sciences”. In O. E. Barndorff-Nielsen, T. Mikosch e S. I. Resnick eds. *Levy processes - Theory and applications*. Birkhäuser Boston, 241-266, 2001.

[Yao e Liu 96] Xin Yao e Yong Liu. "Fast evolutionary programming". In L. J. Fogel, P. J. Angeline e T Back eds. *Proc. 5th. Ann. Conf. on Evolutionary Programming*. Cambridge, MA. MIT Press. 1-10, 1996.

[Yao, Liu e Lin 99] X. Yao, Y. Liu and G. Lin. "Evolutionary programming made faster", *IEEE Transactions on Evolutionary Computation*, 3(2):82-102, 1999.

