

Análise da aplicação de projeto de experimentos nos parâmetros dos algoritmos genéticos

Alexandre Ferreira de Pinho^{1,2}, pinho@unifei.edu.br

José Arnaldo Barra Montevechi¹, montevechi@unifei.edu.br

Fernando Augusto Silva Marins², fmarins@feg.unesp.br

¹ Universidade Federal de Itajubá (UNIFEI), Itajubá, MG, Brasil

² Universidade Estadual Paulista (UNESP), Guratinguetá, SP, Brasil

*Recebido: Outubro, 2007 / Aceito: Dezembro, 2007

RESUMO

Este artigo apresenta uma análise da utilização de Projeto de Experimentos aplicado na definição dos parâmetros existentes nos Algoritmos Genéticos. O objetivo é demonstrar que através desta aplicação é possível realizar uma análise mais eficiente entre estes parâmetros e, principalmente, verificar o comportamento das suas interações, a fim de se obter um melhor planejamento do experimento. Inicialmente será apresentada uma revisão bibliográfica sobre Algoritmos Genéticos, com destaque aos seus parâmetros, seguida da revisão sobre Projeto de Experimentos. A condução deste trabalho se dá através da metodologia proposta por Montgomery, composta de sete passos: reconhecimento e definição do problema; escolha dos fatores, níveis e variações; seleção da variável de resposta; escolha do projeto experimental; execução do experimento; análise estatística dos dados; conclusões e recomendações. Finalmente as conclusões são apresentadas comprovando a eficácia da utilização de Projeto de Experimentos na definição dos parâmetros dos Algoritmos Genéticos.

Palavras-Chave: Algoritmos genéticos, Parâmetros dos Algoritmos Genéticos, Projeto de experimentos

1. INTRODUÇÃO

Há na literatura uma grande discussão sobre a determinação dos parâmetros dos algoritmos genéticos, ou seja, qual deverão ser os valores de configuração desses parâmetros para que se possa obter um melhor desempenho na utilização dos algoritmos genéticos.

Como exemplo, pode-se citar Cunha e Pinto (2001), que afirmam que não existem parâmetros ótimos para os operadores genéticos. Eles são especificados para cada problema, não podendo ser adotado de forma genérica, uma vez que comprometem o desempenho global do algoritmo.

Já Núñez-Letamendia (2005) afirma que não existe na literatura um consenso a respeito das definições dos parâmetros dos algoritmos genéticos. Rees e Koehler (2006) utilizaram em seu trabalho modelos de Markov para a estimativa destes parâmetros.

Siriwardene e Perera (2006) desenvolveram um modelo matemático para a seleção dos parâmetros dos Algoritmos Genéticos para o estudo de sistemas de escoamento urbano. Sadeghe (2006) estuda os efeitos destes parâmetros no resultado da função objetivo para problemas de programação da produção.

Projeto de Experimentos, segundo Montgomery (2001), refere-se ao processo de planejamento de experimentos de forma que dados apropriados possam ser analisados por métodos estatísticos, resultando em conclusões válidas e objetivas. Segundo Kelton (1999), uma das principais metas do projeto experimental é estimar como mudanças em fatores de entrada afetam os resultados, ou respostas do experimento.

Portanto, o objetivo deste artigo é analisar a utilização da técnica de Projeto de Experimentos aplicada aos parâmetros dos algoritmos genéticos, de tal forma a se obter um melhor resultado desses algoritmos. Além disso, espera-se obter uma análise mais eficiente entre estes parâmetros e suas interações, a fim de se obter um melhor planejamento do experimento.

Alguns trabalhos já foram publicados combinando essas duas técnicas. Como exemplo, pode-se citar Iyer e Saxena (2004) que utilizam o Projeto de Experimentos na escolha dos parâmetros de Algoritmo Genético. Tem-se também o trabalho de Cheng e Chang (2007) que utilizam Taguchi (uma técnica de projeto de experimentos) em problemas de programação da produção. Entretanto, nenhum desses trabalhos faz uma análise sobre o efeito das interações que pode existir entre os fatores dos algoritmos genéticos.

2. ALGORITMOS GENÉTICOS

Os algoritmos evolucionários se baseiam num processo de aprendizagem coletivo de uma população de indivíduos, cada um desses representando um ponto do espaço de busca de soluções potenciais de um determinado problema (BACK e SCHWEFEL, 1993). Esses indivíduos visam melhorar a sua adequação em relação ao meio ambiente, ou seja, o seu desempenho geral com respeito a um dado problema (GOLDBERG, 1989).

Segundo Linden (2005), os Algoritmos Genéticos (AG) são um ramo dos algoritmos evolucionários e como tal podem ser definidos como uma técnica de busca baseada numa metáfora do processo biológico de evolução natural.

Definindo de outra maneira, pode-se dizer que algoritmos genéticos são algoritmos de busca baseados nos mecanismos de seleção natural e genética. Eles combinam a sobrevivência entre os melhores indivíduos com uma forma estruturada de troca de informação genética entre dois indivíduos para formar uma estrutura heurística de busca (MITCHELL, 1996).

Segundo Goldberg (1989), o AG difere dos métodos tradicionais de busca e otimização, principalmente em quatro aspectos:

- Opera numa população de pontos e não a partir de um ponto isolado;
- Opera num espaço de soluções codificadas e não no espaço de busca diretamente;
- Necessita somente de informação sobre o valor de uma função objetivo para cada membro da população e não requerem derivadas ou qualquer outro tipo de conhecimento;
- Usa transições probabilísticas e não regras determinísticas.

Nos algoritmos genéticos, populações de indivíduos são criadas e submetidas aos operadores genéticos, que usualmente consistem em seleção, cruzamento e mutação.

Segundo Bäck, Fogel e Michalewicz (2000), os operadores de seleção são usados para direcionar o processo para as melhores regiões do espaço de busca. Sua função é selecionar indivíduos da população para a reprodução, dando preferência aos indivíduos mais adaptados ao ambiente (MITCHELL, 1996).

Barbosa (2005) afirma que os operadores de cruzamento são responsáveis pela propagação das características dos indivíduos mais aptos da população (pais) por meio da troca de segmentos de informações entre estes, o que dará origem a novos indivíduos.

Os Operadores de mutação, segundo Goldberg (1989) executam um papel secundário, porém necessário, pois possibilita restaurar a diversidade genética eventualmente perdida durante o processo evolutivo.

3. PARÂMETROS DOS ALGORITMOS GENÉTICOS

Além dos critérios de seleção, cruzamento e mutação, existem alguns parâmetros dos Algoritmos Genéticos que podem ser escolhidos para melhorar seu desempenho (BARBOSA, 2005). Neste artigo optou-se trabalhar com quatro parâmetros, denominados: Tamanho da população, Número de gerações, Taxa de cruzamento e Taxa de mutação.3.1. TAMANHO DA POPULAÇÃO

Segundo Reeves (1995), pequenas populações podem provocar um sério risco de não obter cobertura do espaço de busca, enquanto que grandes populações podem exigir um esforço computacional excessivo para a resolução do problema.

Já Tanomaru (1995) afirma que, na prática, valores entre 50 e 200 cromossomos resolvem a maior parte dos problemas, mas populações maiores podem ser necessárias para problemas complexos.

Tem-se também o trabalho de Mitchell (1996), que diz que o melhor tamanho para a população é entre 50 e 100 indivíduos.

Laura Núñez-Letamendia (2007) utilizou em seus experimentos um tamanho de população igual 100. Já Cheng e Chang (2007) utilizaram valores entre 20 e 40 para o tamanho da população.

3.2. NÚMERO DE GERAÇÕES

O número de gerações varia de acordo com a complexidade do problema em questão e deve ser determinado experimentalmente. Com o AG resolve problemas de otimização, o ideal seria que o algoritmo terminasse assim que o ponto ótimo fosse encontrado.

Na prática, na maioria das aplicações, não se pode afirmar com certeza, que um dado ponto ótimo encontrado seja considerado um ótimo global. Desta forma algum critério deve ser adotado, para o término do processamento do algoritmo. Normalmente, usa-se o critério do número máximo de gerações ou tempo limite de processamento para o encerramento do processo.

3.3. TAXA DE CRUZAMENTO

É a taxa que determina se será feito o cruzamento entre dois cromossomos. Para tal, gera-se um número aleatório no intervalo [0,1] e compara-se à taxa. Se o número for menor que a taxa o cruzamento é efetuado.

Tanomaru (1995) relata que bons resultados geralmente são obtidos com alto valor da taxa de recombinação (maior que 0,7). Já Mitchell (1996) sugere uma taxa de 0,6, enquanto que Núñez-Letamendia (2007) definiram esse valor em 0,95.

3.4. TAXA DE MUTAÇÃO

Essa taxa determina se os genes dos cromossomos da população selecionada sofrerão mutação ou não. Para tal, gera-se um número aleatório no intervalo [0,1] e compara-se à taxa de mutação. Se o número for menor que a taxa o gene será modificado de acordo com o operador escolhido.

Tanomaru (1995) sugere um baixo valor para a taxa de mutação (menor que 0,01). Já Mitchell (1996) sugere uma taxa de 0,001, enquanto que Núñez-Letamendia (2007) definiram esse valor em 0,01.

4. PROJETO DE EXPERIMENTOS

A palavra experimento é utilizada de forma bastante precisa para indicar uma investigação onde o sistema sob estudo está sob o controle do investigador. Ao contrário, para um estudo observacional, algumas características estarão fora do controle do investigador (Cox e Reid, 2000).

De acordo com Montgomery (2001), o experimento pode ser visto como um teste, ou uma série de testes, nos quais mudanças propostas são aplicadas nas variáveis de entrada de um processo ou sistema, para então se observar e identificar as mudanças ocorridas nas variáveis de saída.

Uma grande variedade de abordagens, métodos e técnicas de análise, conhecidas coletivamente como projeto experimental, foi documentada em trabalhos como de Box *et al.* (1978) e Montgomery (2001).

Alguns termos são comumente utilizados em projeto de experimentos. Mason *et al.* (2003) definem fator como uma variável experimental controlável, cuja variação influencia a variável de resposta. Cada fator deve assumir alguns valores, definidos como níveis. As mudanças ocasionadas na média dos valores da variável de resposta correspondem aos efeitos.

Além dos efeitos causados pelos fatores, podem-se determinar os efeitos gerados pelas interações dos fatores. Estas interações correspondem a efeitos combinados, onde o efeito de cada fator depende dos níveis dos outros fatores.

A seqüência dos experimentos deve ser mantida em ordem aleatória, com o objetivo de proteger o experimento contra desconhecidas ou imensuráveis fontes indesejáveis de influência. A aleatorização também auxiliará, segundo Mason *et al.* (2003), a validação de determinadas suposições necessárias para a aplicação de certas técnicas estatísticas.

A aleatorização também se justifica devido a possíveis perturbações nos equipamentos utilizados no experimento. Estas perturbações podem fazer com que os experimentos realizados a posteriori sejam influenciados em relação aos primeiros experimentos realizados.

O fato é que esta aleatorização não impede a ocorrência destas perturbações, mas sim auxilia na garantia de que todos os níveis dos fatores tenham iguais chances de serem afetados por estas perturbações no equipamento.

A escolha do número de fatores e níveis irá implicar diretamente na escolha do arranjo experimental. Nessa decisão há um *trade-off* entre desempenho *versus* custo, onde a necessidade do conhecimento mais acurado de um número maior de fatores e interações acarreta um número maior de experimentos e replicações.

Um tipo de projeto experimental é o fatorial. Nos projetos fatoriais, cada fator testado é combinado com todos os níveis de todos os outros fatores. É comum no estudo de um número grande de fatores a escolha de dois níveis, chamados de nível alto e nível baixo. Isto se dá pelo intuito de se manter o tamanho do experimento em padrões satisfatórios.

Segundo Cox e Reid (2000), as vantagens do uso de um fatorial completo se dão na maior eficiência em estimar os efeitos principais dos fatores sob a variável em análise, e, principalmente, na definição da interação entre todos os fatores.

De acordo com Sanchez *et al.* (2006), muitos estudos ligados à administração de operações utilizam o projeto experimental fatorial completo devido à sua simplicidade e ao fato de que este projeto permite ao analista identificar interações entre os fatores, assim como os efeitos principais. Como exemplos destes estudos, pode-se citar trabalhos como Enns (1995), que utilizou o projeto fatorial para avaliar a taxa de utilização e técnicas de sequenciamento em um processo.

Já Malhotra e Ritzman (1994) consideraram um projeto fatorial 2^4 para avaliar o impacto da variabilidade da demanda, capacidade de utilização e flexibilidade de rotas em estações de serviço postal.

Por outro lado, problemas relativos à administração de operações podem se tornar complexos, com muitos fatores e importantes interações, além de relações não-lineares entre os fatores e a resposta, dificultando o uso de projetos fatoriais completos. O próprio trabalho de Sanchez *et al.* (2006), analisando um sistema JIT, destaca 34 fatores, definidos durante o experimento. Um fatorial completo exigiria 17,2 bilhões de combinações.

Uma desvantagem do uso do fatorial completo se dá na quantidade de tempo e de experimentos a serem realizados. Segundo Kelton (1999), quando o número de fatores torna-se moderadamente grande, o número de experimentos (ou rodadas na simulação) explode. Um possível caminho para esta situação é o uso do fatorial fracionado, no qual somente uma fração de todas as possíveis combinações é experimentada. Este caminho é indicado para situações onde existe um grande número de fatores a serem analisados, e somente os efeitos principais dos fatores e interações de baixa ordem são considerados importantes.

5. METODOLOGIA

Segundo Silva e Menezes (2005), uma pesquisa é experimental quando se determina um objeto de estudo, selecionam-se as variáveis que seriam capazes de influenciá-lo, e definem-se as formas de controle e de observação dos efeitos que a variável produz no objeto.

De maneira geral, a condução de um projeto experimental pode-se dar através do fluxograma proposto por Montgomery (2001) e apresentado na Figura 1.

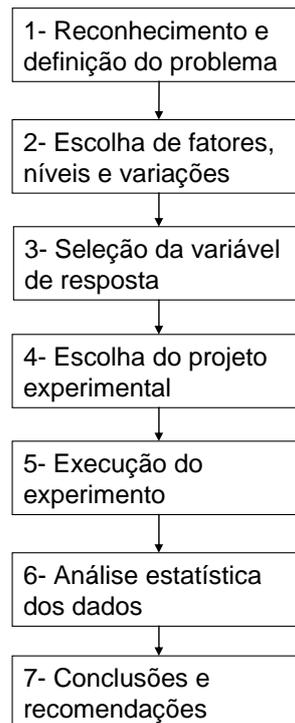


Figura 1: Passos para condução de um projeto de experimentos
Adaptado de Montgomery (2001)

Na prática, os passos 2 e 3 podem ser realizados de forma simultânea ou em ordem contrária.

6. APLICAÇÃO DA METODOLOGIA PROPOSTA

A seguir descrevem-se como os sete passos da metodologia proposta anteriormente por Montgomery (2001) foram conduzidos durante esta pesquisa.

6.1. RECONHECIMENTO E DEFINIÇÃO DO PROBLEMA

Suponha que se deseja otimizar a função apresentada pela equação (1). O objetivo desse projeto de experimentos será qual a escolha dos parâmetros dos algoritmos genéticos dará melhor precisão à resposta.

$$F(x, y) = 0,5 - \frac{(\sin \sqrt{x^2 + y^2})^2 - 0,5}{(1,0 + 0,001(x^2 + y^2))^2} \quad (1)$$

A Figura 2 apresenta o gráfico da função. Nota-se que esta função é difícil de otimizar uma vez que ela possui vários pontos de máximos locais. Entretanto esta função possui um ponto de ótimo conhecido e igual a 1, para x e y assumindo o valor de 0. Ou seja, $F(0,0) = 1$.

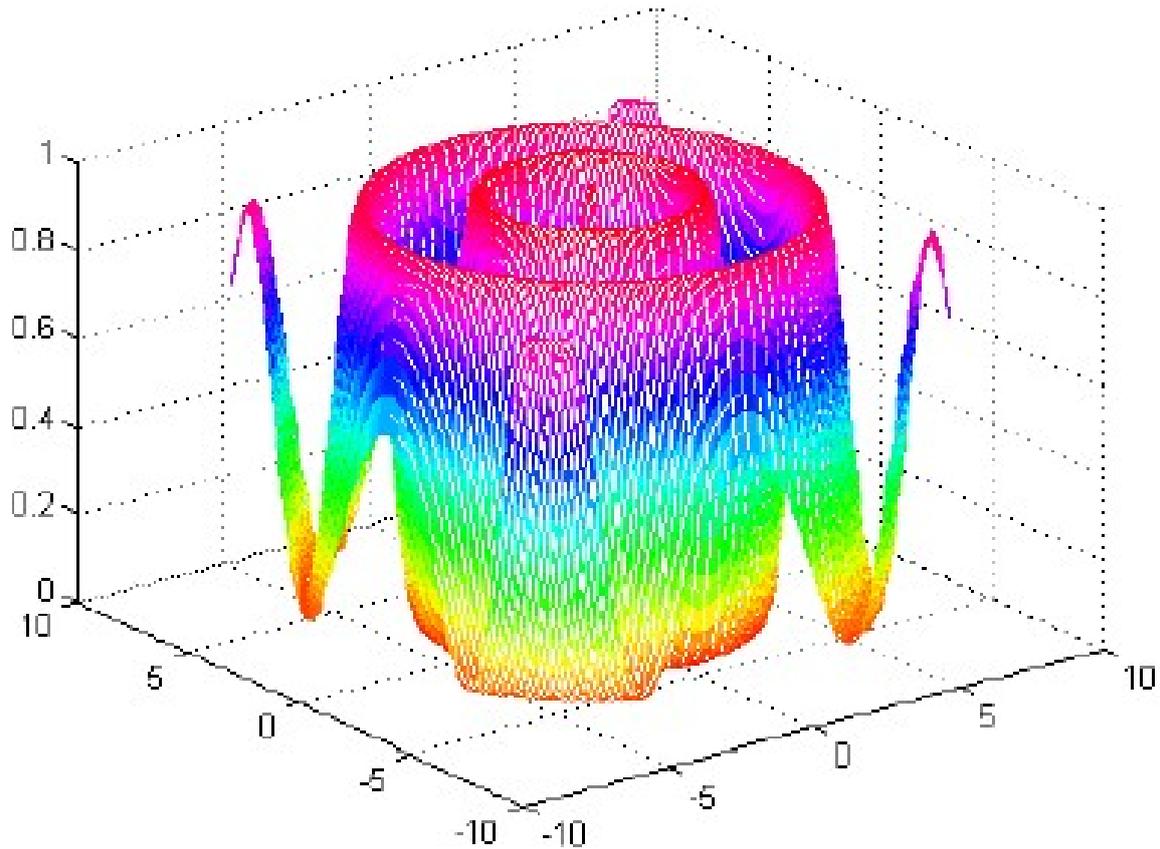


Figura 2: Gráfico da função

6.2. ESCOLHA DE FATORES, NÍVEIS E VARIAÇÕES

Os fatores selecionados para o projeto de experimentos são os parâmetros dos algoritmos genéticos (tamanho da população, taxa de cruzamento e taxa de mutação) discutidos no item 3 deste artigo. Optou-se em trabalhar com dois níveis para cada fator. Esses níveis e suas variações são apresentados na Tabela 1 a seguir.

Tabela 1: Fatores, níveis e variações do projeto de experimentos

Fatores		Níveis	
		-	+
A	Tamanho da População	50	200
B	Número de Gerações	50	100
C	Taxa de Cruzamento	0,1	0,8
D	Taxa de Mutação	0,01	0,3

6.3. SELEÇÃO DA VARIÁVEL DE RESPOSTA

A variável de resposta para esse projeto de experimentos será o valor maximizado da função apresentada anteriormente.

6.4. ESCOLHA DO PROJETO EXPERIMENTAL

Como observado anteriormente, este projeto possui três fatores com dois níveis cada. Desta forma, optou-se pelo uso do fatorial completo, pois conforme já mencionado na seção 4, as vantagens do uso de um fatorial completo se dá na maior eficiência em estimar os efeitos principais dos fatores sob a variável em análise, e, principalmente, na definição da interação entre todos os fatores.

A matriz com o planejamento dos experimentos é mostrada na Tabela 2. Nota-se que para um experimento de fatorial completo com três fatores e com dois níveis cada, serão necessários oito experimentos ($2^4 = 16$). As colunas A, B e C presentes na Tabela 2 são referentes aos fatores descritos no item 6.2.

Tabela 2: Matriz com o planejamento dos experimentos

Experimentos	A	B	C	D
1	-	-	-	-
2	+	-	-	-
3	-	+	-	-
4	+	+	-	-
5	-	-	+	-
6	+	-	+	-
7	-	+	+	-
8	+	+	+	-
9	-	-	-	+
10	+	-	-	+
11	-	+	-	+
12	+	+	-	+
13	-	-	+	+
14	+	-	+	+
15	-	+	+	+
16	+	+	+	+

6.5. EXECUÇÃO DO EXPERIMENTO

A execução do experimento foi realizada em um software de algoritmo genético, desenvolvido pelo primeiro autor desse artigo, que trabalha na busca dos pontos de máximo ou dos pontos de mínimo de funções matemáticas.

Para a execução do experimento os operados genéticos foram mantidos constantes. Apenas os parâmetros do algoritmo genético foram se alterando conforme a matriz de planejamento de experimentos apresentada na Tabela 2.

O resultado encontrado para cada experimento é mostrado na Tabela 3. Nota-se que cada experimento foi executado três vezes. Os resultados de cada uma das execuções são apresentados nas colunas Y1, Y2 e Y3 respectivamente. Esta tabela também apresenta a média dos três resultados encontrados.

Tabela 3: Execução do experimento

Experimento	A	B	C	D	Y1	Y2	Y3
1	-	-	-	-	0,961766	0,956667	0,958816
2	+	-	-	-	0,95999	0,990021	0,987279
3	-	+	-	-	0,92836	0,98703	0,961502
4	+	+	-	-	0,984563	0,988724	0,958782
5	-	-	+	-	0,87034	0,958545	0,921811
6	+	-	+	-	0,961706	0,990175	0,962776
7	-	+	+	-	0,99005	0,990266	0,988241
8	+	+	+	-	0,990284	0,990257	0,987996
9	-	-	-	+	0,92106	0,989962	0,961961
10	+	-	-	+	0,990176	0,990041	0,962571
11	-	+	-	+	0,990243	0,960061	0,990283
12	+	+	-	+	0,990247	0,988004	0,990274
13	-	-	+	+	0,962523	0,989841	0,988051
14	+	-	+	+	0,990284	0,962776	0,96277
15	-	+	+	+	0,990275	0,984984	0,911148
16	+	+	+	+	0,978208	0,988905	0,990283

6.6. ANÁLISE ESTATÍSTICA DOS DADOS

A partir dos resultados encontrados na Tabela 3, calculou-se os efeitos principais de cada fator e o efeito de suas interações. Para os cálculos desenvolvidos utilizou-se o software MiniTab® versão 14.1.

O peso dos efeitos pode ser visualizado na Figura 3. Nessa figura verifica-se que os fatores principais A (Tamanho da População) e B (Número de Gerações), assim como a interação entre os fatores BCD (número de gerações * taxa de cruzamento * taxa de mutação) são significativa para um nível de significância α de 5%. Cabe ressaltar que para efetuar o teste de significância, através do teste t, o MiniTab® calcula utilizando-se do erro experimental e a variância global, comparando-o com o valor crítico.

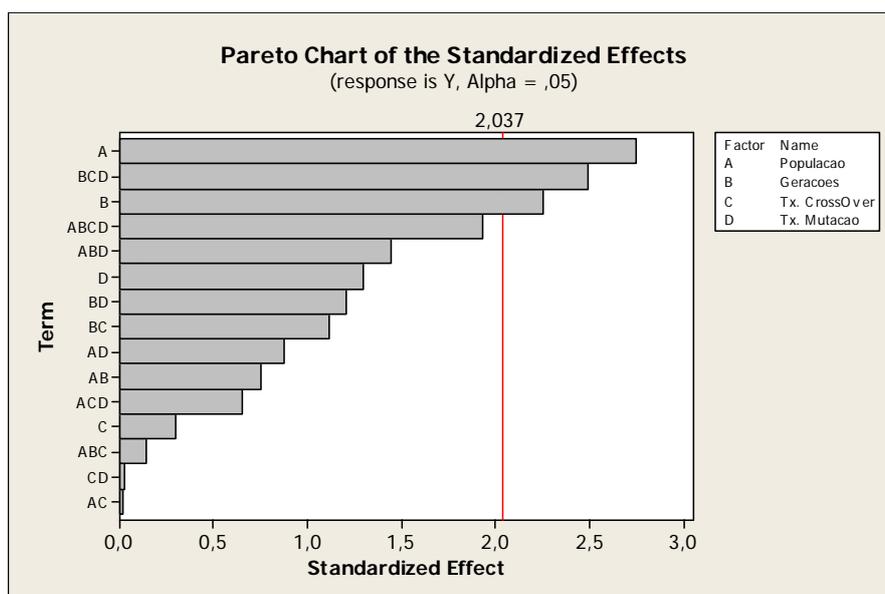


Figura 3: Gráfico de Pareto para os efeitos padronizados

A análise dos efeitos principais de cada fator, apresentada na Figura 4, mostra que os fatores “Tamanho da População”, “Número de Gerações” e “Taxa de Mutação” possuem forte efeito positivo sobre a resposta final, ou seja, o valor máximo da função. Isso significa que a alteração do nível (-) para o nível (+) melhora o resultado final. Já o fator “Taxa de Cruzamento” não possui efeitos significativos sobre o resultado final.

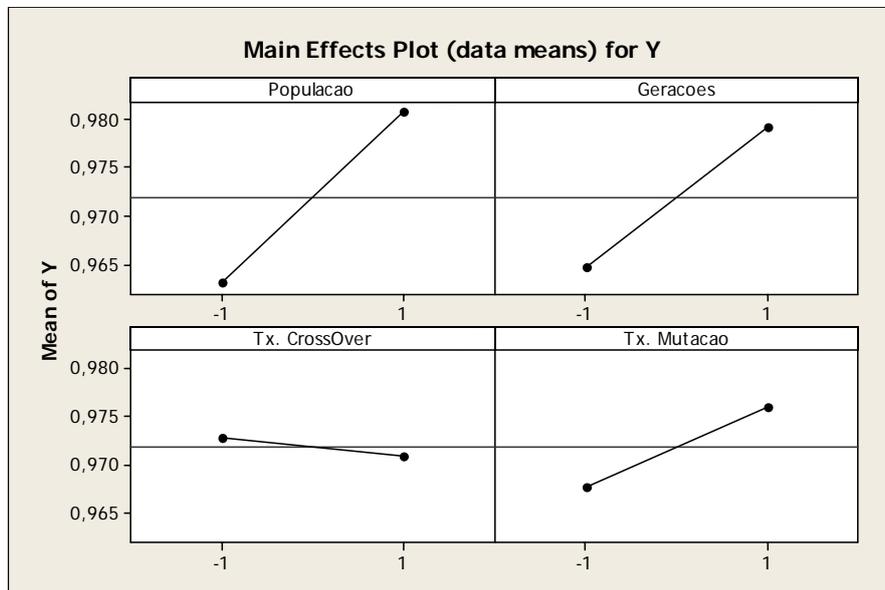


Figura 4: Gráfico dos efeitos principais

6.7. CONCLUSÕES E RECOMENDAÇÕES

Com base nos resultados mostrados no item anterior concluí-se que os fatores principais A (Tamanho da População) e B (Número de Gerações), assim como a interação entre os fatores BCD (número de gerações * taxa de cruzamento * taxa de mutação) são o que mais influencia significativamente a variável resposta.

Observa-se que existe uma interação que influencia mais significativamente a resposta final que os fatores Taxa de Cruzamento e Taxa de Mutação isoladamente.

7. CONCLUSÕES

Este trabalho evidencia que a escolha dos parâmetros do algoritmo genético não deve ser feita de forma genérica. Uma escolha mal feita pode prejudicar o desempenho do algoritmo na obtenção dos resultados esperados.

Conforme citado na introdução deste artigo, não existe um consenso na definição destes parâmetros. Desta forma, buscou-se apresentar um método, para uma melhor configuração destes parâmetros, através da utilização de projeto de experimentos.

Na ilustração apresentada neste trabalho, a utilização do Projeto de Experimentos nos parâmetros dos algoritmos genéticos permitiu uma análise mais eficiente entre estes parâmetros. Além disso, ficou comprovada a existência de interações significativas entre estes fatores, auxiliando o processo de tomada de decisão de quem está planejando o experimento. Desta forma, o reconhecimento do efeito individual e combinado dos fatores favorece a elaboração de um melhor projeto de experimento para que se possa atingir melhores resultados na otimização de funções.

Cabe ressaltar, como proposta de trabalhos futuros, que a metodologia desenvolvida neste artigo será utilizada em uma pesquisa de doutorado. Nesta pesquisa pretende-se verificar como a utilização dos algoritmos genéticos, em conjunto com projeto de experimentos, podem melhorar a simulação de eventos discretos para sistemas de manufatura.

AGRADECIMENTOS

Aos programas de Pós-Graduação da Unesp (Guaratinguetá-SP) e Unifei (Itajubá-MG) que, através de um convênio, possibilitaram a realização desta pesquisa, tema de tese de doutorado.

A FAPEMIG por apoio a pesquisa.

REFERÊNCIAS BIBLIOGRÁFICAS

Bäck, T; Fogel, H. P. e Michalewicz, Z. **Evolutionary Computation 2: Advanced Algorithm and Operators**. Bristol, UK, Institute of Physics, 2000.

Bäck, T; Schwefel, H. P. An Overview of Evolucionary Algorithms for Parameter Optimization. **Evolucionary Computation**, v.1, n. 1, p. 1-23. The MIT Press, 1993.

Barbosa, A. O. **Simulação e Técnicas da Computação Evolucionaria Aplicadas a Problemas de Programação Linear Inteira Mista**. Tese de Doutorado. Universidade Tecnológica Federal do Paraná: pós-graduação em Engenharia Elétrica e Informática Industrial. Curitiba, 2005.

Box, G. E. P., Hunter, W. G., Hunter, J. S. **Statistics for Experimenters: An Introduction to Design, Data Analysis and Model Building**. Wiley, New York, 1978.

Cheng, B.W. e Chang, C.L. A study on flowshop scheduling problem combining Taguchi experimental design and genetic algorithm. **Expert System with Applications**, volume 32, issue 2, pages 415-421, 2007.

Cox, D.R.; Reid, N. **The Theory of the Design of Experiments**. Chapman & Hall/CRC, 2000.

Cunha, Alexandre S.; Pinto, Ricardo L. U. S. Uma técnica para ajuste dos parâmetros de um Algoritmo Genético. **Anais do XXXIII Simpósio Brasileiro de Pesquisa Operacional**, Campos do Jordão, SP, p. 1105-1116, 2001.

Enns, S.T. An integrated system for controlling work load and flow in a job shop. **International Journal of Production Research**, 33, 2801–2820, 1995.

Goldberg, D. E. **Genetic Algorithms in Search, Optimization, and Machine Learning**. Addison-Wesley Publishing Company, INC. United States of America, 1989.

Iyer, S. K. e Saxena, B. Improved genetic algorithm for the permutation Flowshop scheduling problem. **Computers & Operations Research**, volume 31, issue 4, pages 593-606, 2004.

Kelton, W.D. Designing simulation experiments. **Proceedings of the 1999 Winter Simulation Conference**, 1999.

Linden, R. **Um algoritmo Híbrido para a Extração de Conhecimento em Bioinformática**. Tese (Doutorado em Engenharia Elétrica), Rio de Janeiro, Programa de Pós-graduação em Engenharia Elétrica Universidade Federal do Rio de Janeiro. 2005.

Mason, R.L.; Gunst, R.F.; Hess, J.L. **Statistical Design and Analysis of Experiments**. John Wiley & Sons Publication, 2003.

Malhotra, M.K., Ritzman, L.P. Scheduling flexibility in the service sector: a postal case study. **Production & Operations Management**, 3, 100–117, 1994.

- Mitchell, M. **An Introduction to Genetic Algorithms**. MIT Press, Massachusetts, London England, 1996.
- Montgomery, D.C. **Design and Analysis of Experiments**. 5th edition, John Wiley & Sons, Inc., 2001.
- Núñez-Letamendia, L. Fitting the control parameters of a genetic algorithm - An application to technical trading systems design. **European Journal of Operational Research**, volume 179, issue 3, pages 847-868, 2005.
- Rees, J. e Koehler, G. J. Learning genetic algorithm parameters using hidden Markov models. **European Journal of Operational Research**, 2006.
- Reeves, C. R. **Modern Heuristic Techniques for Combinatorial Problems**, McGraw-Hill. Book Company, London, 1995.
- Sadeghe, A. Scheduling problem using genetic algorithm, simulated annealing and the effects of parameter values on GA performance. **Applied Mathematical Modelling**, 2006.
- Sanchez, S.M.; Moeeni, F.; Sanchez, P.J. So many factors, so little time: Simulation experiments in the frequency domain. **International Journal of Production Economics** 103, 149–165, 2006.
- Silva E. L., Meneses E. M., **Metodologia de Pesquisa e Elaboração da dissertação**, Universidade Federal de Santa Catarina, 2001.
- Siriwardene, N. R. e Perera, B. J. C. Selection of genetic algorithm operators for urban drainage model parameter optimisation. **Mathematical and Computer Modelling**, 2006.
- Tanomaru, J. **Motivação, fundamentos e aplicações de algoritmos genéticos**. II Congresso Brasileiro de Redes Neurais. Curitiba, PR, p. 373-403, 1995.

Design of Experiments analysis applied in the definition of Genetic Algorithms parameters

Alexandre Ferreira de Pinho^{1,2}, pinho@unifei.edu.br

José Arnaldo Barra Montevechi¹, montevechi@unifei.edu.br

Fernando Augusto Silva Marins², fmarins@feg.unesp.br

¹ Universidade Federal de Itajubá (UNIFEI), Itajubá, MG, Brasil

² Universidade Estadual Paulista (UNESP), Guratinguetá, SP, Brasil

*Received: October, 2007 / Accepted: December, 2007

ABSTRACT

This article presents a Design of Experiments analysis applied in the definition of Genetic Algorithms parameters. The objective is to demonstrate that is possible to accomplish a more efficient analysis among these parameters and verify the behavior of its interactions. Initially, a bibliographical review on Genetic Algorithms will be presented, followed by a review on Design of Experiments as well. In this work it is adopted a methodology proposed by Montgomery, composed of seven steps: recognition and definition of the problem; choose of the factors, levels and variations; answer variable selection; choose of the experimental project; experiment execution; data statistical analysis; conclusions and recommendations. Finally, conclusions certifying the Design of Experiments effectiveness in order to define the Genetic Algorithms Parameters are presented.

Key words: Genetic Algorithm, Parameters, Design of Experiments.
