VIII ERMAC-R3

5º Encontro Regional de Matemática Aplicada e Computacional

20 a 22-Novembro-2008

Universidade federal do Rio Grande do Norte - Natal/RN

André Gustavo Campos Pereira,
Depto. de Matemática, UFRN,
59.072-970, Natal, RN
E-mail: andre@ccet.ufrn.br.

Cátia Regina dos Santos Silva **Marconio Silva dos Santos**,
Curso de Matemática, UFRN,
59.072-970, Natal, RN
E-mail: cirai@terra.com.br, marcofratello@libero.it,

Judinor Aguirres Guerra Junior (*in memorien*)
Depto. de Engenharia Elétrica, UFRN,
59.072-970, Natal, RN.

Resumo

Neste trabalho, são apresentados alguns conceitos básicos sobre algoritmos genéticos e sua modelagem por Cadeias de Markov finitas. Esse processo é feito em três operações: seleção, cruzamento e mutação.

Palavras-chave

Algorítmo genético, Cadeia de Markov.

Introdução

Algoritmos genéticos (AG's) são freqüentemente usados para atacar problemas de otimização estáticos do tipo

$$\max\{f(b)|b \in B^l\}, \tag{1}$$

em que $0 < f(b) < \infty$, para todo $b \in B^l = \{0,1\}^l$ e f não é a função constante. As cadeias de Markov oferecem um modelo apropriado para analisar AG's.

Um algoritmo genético consiste em uma n-upla de strings binárias b_i de comprimento l, onde os bits de cada string são considerados genes do cromossomo de um indivíduo e onde a n-upla de cromossomos constitui uma população. Seguindo a terminologia da evolução orgânica, as operações executadas na população são chamadas de *seleção*, *cruzamento* e *mutação*. Cada indivíduo b_i representa uma possível solução do problema (1) e o valor da sua função objetivo $f(b_i)$ é dito ser sua adaptabilidade (fitness) a qual será maximizada.

Quando é usada a chamada seleção proporcional, a população da próxima geração é determinada por n experimentos randômicos independentes. A probabilidade que o indivíduo b_i seja selecionado da n-upla $(b_1, b_2, ..., b_n)$ para ser um membro da próxima geração a cada experiência é dada por

$$P(b_i \text{ ser selecionado}) = \frac{f(b_i)}{\sum_{j=1}^{n} f(b_j)} > 0 \tag{2}$$

O cruzamento permite a troca de material genético entre os indivíduos da população. O operador de cruzamento é aplicado com alguma probabilidade $p_c \in [0,1]$ a fim de construir uma string a partir de pelo menos duas outras strings randomicamente escolhidas. Se, por exemplo, o cromossomo tem tamanho l um ponto dentre $1, ..., l-1$ é escolhido aleatoriamente, e a partir desse ponto as informações genéticas dos pais serão trocadas dando origem a descendentes que carregam as informações genéticas dos pais.

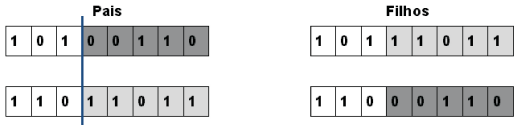

Figura 1: Exemplo de um cruzamento.

A mutação opera independentemente em cada indivíduo perturbando probabilisticamente cada bit na string. O evento de que o j-ésimo bit do i-ésimo indivíduo seja mudado é estocasticamente independente e acontece com probabilidade $p_m \in (0,1)$. Por exemplo, a probabilidade de que a string $b = (00000)$ transite para a string $b' = (10110)$ através da mutação é

$$p_m \cdot (1-p_m) \cdot p_m \cdot p_m \cdot (1-p_m) = p_m^3 \cdot (1-p_m)^2. \tag{3}$$

Cadeias de Markov finitas

Definição 1 *Chamamos* processo estocástico *a uma seqüência de variáveis aleatórias* $X_n : \Omega \longrightarrow \mathbb{R}$

definidas sobre um mesmo espaço de probabilidade.

O conjunto de todos os valores possíveis de um processo estocástico é chamado de espaço de estados. Uma cadeia de Markov é um processo estocástico que modela os fenômenos que possuem a propriedade de Markov.

Definição 2 *Um processo estocástico* $\{X_n\}_{n\in\mathbb{N}}$ *possui a propriedade de Markov se*

$$P(X_n = i_n | X_{n-1} = i_{n-1}, ..., X_0 = i_0) = \\ = P(X_n = i_n | X_{n-1} = i_{n-1}),$$

para todo n.

Uma cadeia de Markov é dita ser finita quando o espaço de estados E é finito. A probabilidade $p_{ij}(t)$ é chamada de probabilidade de transição de i para j no passo t. Se as probabilidades de transição são independentes de t, isto é, $p_{ij}(t) = p_{ij}(s)$ para todo $i, j \in E$ e para todo $s, t \in \mathbb{N}$, então a cadeia de Markov é dita homogênea.

As probabilidades de transição de uma cadeia de Makov homogênea finita podem ser representadas por uma matriz $P = (p_{ij})$, onde $p_{ij} \in [0, 1]$ e

$$\sum_{j=1}^{|E|} p_{ij} = 1,$$

para todo $i \in E$.

Modelagem do algoritmo genético por cadeias de Markov

O AG pode ser descrito como uma cadeia de Markov: O estado do AG na próxima etapa do algorítmo só depende dos genes dos indivíduos no instante anterior ao passo. Cada elemento do espaço de estados E pode ser considerado como a representação binária de um número de inteiro não-negativo.

Teorema 1 *As mudanças probabilísticas dos genes dentro da população causadas pelos operadores genéticos são capturadas pela matriz de transição* P *que pode ser decomposta em um produto de matrizes estocásticas* $P = SCM$*, onde* S*,* C *e* M *descrevem as transições intermediárias causadas pela seleção, cruzamento e mutação, respectivamente.*

Demonstração. Tendo em vista que o algoritmo genético realiza sobre uma população, em cada geração e nesta ordem, uma seleção, um cruzamento e uma mutação, cujas matrizes de transição serão representadas, respectivamente, por S, C e M, conseqüentemente, a matriz de transição P, resultante de uma geração, será representada pelo produto SCM. Sendo assim, o elemento $P(i, j)$ será dado pela probabilidade de chegada ao estado j dado que se encontrava no estado i e é obtido pelo produto $\sum_{k\in E}\sum_{t\in E} S(i, k)C(k, t)M(t, j)$ cujos fatores são respectivamente as probabilidades de seleção, cruzamento e mutação, de chegada ao estado correspondente à segunda componente dado que partiu do estado correspondente à primeira componente.

Para se obter a matriz P, define-se as variáveis aleatórias X_s, X_c, e X_m que representam os estados de seleção, cruzamento e mutação respectivamente. Considerando E o espaço de estados desse processo, X_0 a variável aleatória que representa o estado inicial do processo, e os operadores de seleção, de cruzamento e de mutação, tem-se

$$\begin{aligned}
&P(X_m = j | X_0 = i) = \\
&= \frac{P(X_m = j, X_0 = i)}{P(X_0 = i)} \\
&= \sum_{k\in E} \frac{P(X_m = j, X_s = k, X_0 = i)}{P(X_0 = i)} \\
&= \sum_{k\in E}\sum_{t\in E} \frac{P(X_m = j, X_c = t, X_s = k, X_0 = i)}{P(X_0 = i)} \\
&= \sum_{k\in E}\sum_{t\in E} P(X_m = j | X_c = t, X_s = k, X_0 = i) \\
&\quad P(X_c = t | X_s = k, X_0 = i) P(X_s = k | X_0 = i) \\
&= \sum_{k\in E}\sum_{t\in E} S(i, k)C(k, t)M(t, j) \\
&= SCM(i, j).
\end{aligned}$$

Então, a condição de que $P = SCM$ é satisfeita. ■

Definição 3 *Dizemos que uma matriz é* positiva *quando todos os seus elementos são positivos.*

Definição 4 *Dizemos que uma matriz quadrada é* matriz estocástica *quando todos os elementos são não negativos e a soma dos elementos de cada linha é igual a 1.*

Teorema 2 *A matriz* P *obtida no teorema anterior é positiva.*

Demonstração. Por M ser positiva e SC ser estocástica temos que $P = SCM$ é positiva. ■

Um resultado bastante conhecido na teoria de cadeias de Markov é o seguinte.

Teorema 3 *Toda cadeia de Markov homogênea irredutivel, aperiódica com espaço de estados finito é convergente.*

Teorema 4 *A cadeia de Markov representada pela matriz* P *acima é convergente.*

Demonstração. Como P é positiva, temos que ela é irredutível e aperiódica. Como ela é finita,

irredutível e aperiódica, o teorema anterior nos garante que ela é convergente. ■

A pergunta que surge naturalmente é: ela converge quase certamente para o ponto de máximo procurado? Mostraremos através de exemplos, utilizando funções simples e um número pequeno de pontos que tal situação não ocorre.

Exemplo da construção das matrizes de transição de um AG

Nesta seção apresentamos os gráficos da probabilidade de o ponto de máximo da função objetivo pertencer a população ao longo do tempo. Veremos que mesmo nos utilizando de funções simples a probabilidade da população conter o ponto de máximo converge, mas não converge para 1.

Os algoritmos para a construção da matriz de seleção, cruzamento e mutação, bem como os gráficos abaixo foram executados no programa MATLAB 6.5. Deixamos os algorítmos das matrizes disponíveis no site http://br.geocities.com/algenetico a fim de não ocupar muito espaço neste resumo.

Nossos exemplos são compostos por uma população de 3 elementos onde o espaço de busca é composto de 8 pontos.

Agradecimentos

Agradecemos a colaboração da Profa. Dra. Viviane Simioli Medeiros Campos (UFRN).

Referências

[1] Cavicchio, D. J., "Adaptive Search Using Simulated Evolution", Ph.D. Thesis, University of Michigan, Ann Arbor, MI, 1970.

[2] Coello, Carlos A. C., Introducción a los Algoritmos Genéticos, Soluciones Avanzadas. *Tecnologías de Información y Estrategias de Negocios*, Año 3, Número 17 (1995) 5-11.

[3] Coelho, L. S. e Coelho, A.A., Algoritmos Evolutivos em Identificação e Controle de Processos: Uma Visão Integrada e Perspectivas, *SBA Controle & Automação*, Vol. 10 No. 01 / Jan., Fev., Mar., Abril., 1999.

[4] Davis, T. E., "Toward an extrapolation of the simulated annealing convergence theory onto the simple genetic algorithm". Tese de PhD, Universidade da Florida, Gainesville, 1991.

[5] Davis, T. E. e Principe, J., "A Markov chain framework for the simple genetic algorithm". Evolutionary Computation, 1993.

[6] De Jong, K. A., "An Analysis of the Behavior of a Class of a Genetic Adaptive System", Phd. Tesis, University of Michigan, Ann Arbor, MI, 1975.

[7] Dorea, C. C. Y., Pereira, A G. C. e Rojas Cruz, J. A., "Metropolis Hastings Algorithm: the Non-Homogeneus Case", Stochatic Modelling and Aplications, Vol. 5, No 2, (2002) 10-20.

[8] Dorea, C. C. Y., Martins Neto, D. S. B. e Pereira, A G. C., "Suficient Conditions for Egordicity and Convergence of MH, AS and EM Algorithms". Stochastic Models, Vol. 20, No 2, (2004) 193-204.

[9] Eiben, A. E. Aarts, E.H.L. and Van Hee, K.M., "Global convergence of genetic algorithms: A markov chain analysis", in Parallel Problem Solving from Nautre, H.-P. Schwefel and R. Männer (Eds.), Berlim and Heidelberg: Springer, 1991.

[10] Goldberg, D. E., "Genetic Algorithm in Search & Machine Learning", Addison Wesley, 1989.

[11] Guerra Jr, J. A., "Auto-Sintonia de Controlador PID Digital usando Algoritmos Genéticos para Controle de Velocidade de Motor de Indução trifásico". XIII CBA (XIII Congresso Brasileiro de Automática), pp. 879 - 884, 2000.

[12] Günter, R., "Convergence Analysis of Canonic Genetic Algorithms", IEEE Transaction on Neural Networks, 1994.

[13] Günter, R., *Finite Markov Chain in Evolutionary Computation: A Tour d'Hrizon*, Fundamenta Informaticae, 1-22, 1998.

[14] Holland, J. H., "Adaptation in natural and artifitial systems", Ann Arbor: The University of Michigan Press, 1975.

[15] Lukas, E., "Stochastic Convergence", Academic Press, New York, 1975.

[16] Mahfoud, S. W. e Goldberg, D. E., *Parallel Recombinative Simulated Annealing: A Genetic Algorithm.* Parallel Computing, 21(1): 1 - 28, 1995.

[17] Michalewicz, Z., "Genetic Algorithms + Data Structures = Evolutions Programs". Springer-Verlag, 1992.

[18] Ochi, L. S., "Algoritmos Genéticos: Origem e Evolução", 1999.

[19] Pereira, A G. C., "Analisys of Non-Homogeneus Markov Chain Egordicity via a Doeblin Type Condition". Tese de Doutorado do Departamento de Matemática da UnB (Universidade de Brasília), 2002.

---

[20] Whitley D., "A genetic Algorithm Tutorial", Colorado state university, 1993.