

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Edison Americo Huarsaya Tito

**Abordagens de Inferência Evolucionária em Modelos
Adaptativos**

Tese de Doutorado

Tese apresentada como requisito parcial para
obtenção do título de Doutor pelo Programa de Pós-
Graduação em Engenharia Elétrica da PUC-Rio.

Orientadores: Marco Aurélio Cavalcanti Pacheco
Marley Maria Bernardes Rebuszi Vellasco

Rio de Janeiro
Março de 2003



Edison Americo Huarsaya Tito

Abordagens de Inferência Evolucionária em Modelos Adaptativos

Tese apresentada como requisito parcial para obtenção do título de Doutor pelo Programa de Pós-Graduação em Engenharia Elétrica da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Marco Aurélio Cavalcanti Pacheco
Orientador, PUC-Rio

Marley Maria Bernardes Rebuzzi Vellasco
Orientador, PUC-Rio

Cristiano Augusto Coelho Fernandes
PUC-Rio

Carlos Roberto Hall Barbosa
PUC-Rio

Hedibert Freitas Lopes
UFRJ

Valmir C. Barbosa
UFRJ

Leandro dos Santos Coelho
PUC-PR

Maria Luiza F. Velloso
UERJ

Ney Augusto Dumont
Coordenador(a) Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 12 de Março de 2003

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Edison Americo Huarsaya Tito

Graduou-se em Engenharia Eletrônica na Universidad Nacional de San Agustín de Arequipa - Perú em 1995. Obteve o título de Mestre em Engenharia Elétrica pela Pontifícia Universidade Católica de Rio de Janeiro em 1999, tendo como área de concentração: Métodos de Apoio à Decisão e como linha de Pesquisa: Inteligência Computacional. Desenvolveu junto com os seus orientadores diversos projetos para a indústria.

Ficha Catalográfica

Tito, Edison Americo Huarsaya

Abordagens de inferência evolucionária em modelos adaptativos / Edison Americo Huarsaya Tito; orientadores: Marco Aurélio Cavalcanti Pacheco, Marley Maria Bernardes Rebuzzi Vellasco. – Rio de Janeiro : PUC, Departamento de Engenharia Elétrica, 2003.

[15], 107 f. : il. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Inclui referências bibliográficas.

1. Engenharia elétrica – Teses. 2. Estatística Bayesiana. 3. Simulações estocásticas. 4. Algoritmos genéticos. I. Pacheco, Marco Aurélio Cavalcanti. II. Vellasco, Marley Maria Bernardes Rebuzzi. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

A meus pais e meus irmãos

Agradecimentos

Aos meus orientadores, Dr. Marco Aurélio C. Pacheco e Dra. Marley Vellasco pelo estímulo e parceria para a realização deste trabalho.

Ao CNPq e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Ao Dr. Arnaud Doucet do laboratório de processamento de sinais da universidade de Cambridge; ao Dr. Hedibert Lopes do DME, UFRJ e ao Dr. Cristiano Fernandes do DEE, PUC/RJ, pelas críticas e sugestões realizadas ao trabalho.

Aos professores, pesquisadores, funcionários e colegas do Departamento de Engenharia Elétrica da PUC-Rio.

Aos professores que participaram da Comissão examinadora.

Aos meus pais, pela educação, atenção e carinho de todas as horas.

A Deus, por ter iluminado meu caminho.

Resumo

Huarsaya Tito, Edison Americo. **Abordagens de Inferência Evolucionária em Modelos Adaptativos**. Rio de Janeiro, 2003. 107p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Em muitas aplicações reais de processamento de sinais, as observações do fenômeno em estudo chegam sequencialmente no tempo. Conseqüentemente, a tarefa de análise destes dados envolve estimar quantidades desconhecidas em cada observação concebida do fenômeno.

Na maioria destas aplicações, entretanto, algum conhecimento prévio sobre o fenômeno a ser modelado está disponível. Este conhecimento prévio permite formular modelos Bayesianos, isto é, uma distribuição *a priori* sobre as quantidades desconhecidas e uma função de verossimilhança relacionando estas quantidades com as observações do fenômeno. Dentro desta configuração, a *inferência Bayesiana* das quantidades desconhecidas é baseada na distribuição *a posteriori*, que é obtida através do teorema de Bayes.

Infelizmente, nem sempre é possível obter uma solução analítica exata para esta distribuição *a posteriori*. Graças ao advento de um formidável poder computacional a baixo custo, em conjunto com os recentes desenvolvimentos na área de simulações estocásticas, este problema tem sido superado, uma vez que esta distribuição *a posteriori* pode ser aproximada numericamente através de uma distribuição discreta, formada por um conjunto de amostras.

Neste contexto, este trabalho aborda o campo de simulações estocásticas sob a ótica da genética Mendeliana e do princípio evolucionário da “sobrevivência dos mais aptos”. Neste enfoque, o conjunto de amostras que aproxima a distribuição *a posteriori* pode ser visto como uma população de indivíduos que tentam sobreviver num ambiente Darwiniano, sendo o indivíduo mais forte, aquele que possui maior probabilidade. Com base nesta analogia, introduziu-se na área de simulações estocásticas (a) novas definições de *núcleos de transição* inspirados nos operadores genéticos de *cruzamento* e *mutação* e (b) novas definições para a *probabilidade de aceitação*, inspirados no *esquema de seleção*, presente nos Algoritmos Genéticos.

Como contribuição deste trabalho está o estabelecimento de uma equivalência entre o teorema de Bayes e o princípio evolucionário, permitindo,

assim, o desenvolvimento de um novo mecanismo de busca da solução ótima das quantidades desconhecidas, denominado de *inferência evolucionária*. Destacam-se também: (a) o desenvolvimento do *Filtro de Partículas Genéticas*, que é um algoritmo de aprendizado *online* e (b) o *Filtro Evolutivo*, que é um algoritmo de aprendizado *batch*. Além disso, mostra-se que o *Filtro Evolutivo*, é em essência um *Algoritmo Genético* pois, além da sua capacidade de convergência a distribuições de probabilidade, o *Filtro Evolutivo* converge também a sua moda global. Em conseqüência, a fundamentação teórica do *Filtro Evolutivo* demonstra, analiticamente, a convergência dos Algoritmos Genéticos em espaços contínuos.

Com base na análise teórica de convergência dos algoritmos de aprendizado baseados na *inferência evolucionária* e nos resultados dos experimentos numéricos, comprova-se que esta abordagem se aplica a problemas reais de processamento de sinais, uma vez que permite analisar sinais complexos caracterizados por comportamentos não-lineares, não-gaussianos e não-estacionários.

Palavras-chave

Estatística Bayesiana; Simulações Estocásticas; Algoritmos Genéticos.

Abstract

Huarsaya Tito, Edison Americo. **Evolutionary Inference Approaches for Adaptive Models**. Rio de Janeiro, 2003. 107p. Ph.D. Thesis - Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro.

In many real-world signal processing applications, the phenomenon's observations arrive sequentially in time; consequently, the signal data analysis task involves estimating unknown quantities for each phenomenon observation.

However, in most of these applications, prior knowledge about the phenomenon being modeled is available. This prior knowledge allows us to formulate a Bayesian model, which is a prior distribution for the unknown quantities and the likelihood functions relating these quantities to the observations. Within these settings, the *Bayesian inference* on the unknown quantities is based on the posterior distributions obtained from the Bayes' theorem.

Unfortunately, it is not always possible to obtain a closed-form analytical solution for this posterior distribution. By the advent of a cheap and formidable computational power, in conjunction with some recent developments in stochastic simulations, this problem has been overcome, since this posterior distribution can be obtained by numerical approximation.

Within this context, this work studies the stochastic simulation field from the Mendelian genetic view, as well as the evolutionary principle of the "survival of the fittest" perspective. In this approach, the set of samples that approximate the *posteriori* distribution can be seen as a population of individuals which are trying to survive in a Darwinian environment, where the strongest individual is the one with the highest probability. Based in this analogy, we introduce into the stochastic simulation field: (a) new definitions for the *transition kernel*, inspired in the genetic operators of crossover and mutation and (b) new definitions for the *acceptation probability*, inspired in the *selection scheme* used in the Genetic Algorithms.

The contribution of this work is the establishment of a relation between the Bayes theorem and the evolutionary principle, allowing the development of a new optimal solution search engine for the unknown quantities, called *evolutionary inference*. Other contributions: (a) the development of the *Genetic Particle Filter*,

which is an evolutionary online learning algorithm and (b) the *Evolution Filter*, which is an evolutionary batch learning algorithm. Moreover, we show that the *Evolution Filter* is a Genetic algorithm, since, besides its capacity of convergence to probability distributions, it also converges to its global modal distribution. As a consequence, the theoretical foundation of the *Evolution Filter* demonstrates the convergence of Genetic Algorithms in continuous search space.

Through the theoretical convergence analysis of the learning algorithms based on the *evolutionary inference*, as well as the numerical experiments' results, we verify that this approach can be applied to real problems of signal processing, since it allows us to analyze complex signals characterized by non-linear, non-gaussian and non-stationary behaviors.

Keywords

Bayesian Statistics; Stochastic Simulation; Genetic Algorithms.

Sumário

1	Introdução	16
1.1.	Abstração do Modelo de Espaço de Estados	18
1.2.	As Estratégias de Aprendizado	19
1.3.	Objetivos e Contribuições deste Trabalho	20
1.4.	Organização da Tese	22
1.5.	Resumo	23
2	Aprendizado por Métodos Bayesianos	24
2.1.	A Estratégia Bayesiana de Aprendizado	25
2.2.	O Filtro Bootstrap	28
2.3.	O Algoritmo MCCM	31
2.4.	Resumo	34
3	Aplicação do Princípio Evolucionário em Otimização	35
3.1.	Os Algoritmos Genéticos	37
3.2.	Fundamentos Teóricos dos Algoritmos Genéticos	40
3.3.	Resumo	42
4	Novas Aplicações do Princípio Evolucionário	43
4.1.	A Estratégia Evolucionária de Aprendizado	43
4.2.	O Filtro de Partículas Genéticas	44
4.3.	O Filtro Evolutivo	53
4.4.	Resumo	61
5	Estudo de Casos	62
5.1.	Filtragem dos Estados de um Sistema Não-Linear Unidimensional	62
5.2.	Otimização de Funções	76
5.3.	Aprendizado de um Sistema de Volatilidade Estocástica	86
5.4.	Resumo	100

6 Conclusões e Trabalhos Futuros	101
7 Referências Bibliográficas	104

Lista de figuras

Figura 1 - Representação gráfica do algoritmo do FPB	30
Figura 2 - Representação gráfica do algoritmo do FPG	52
Figura 3 - Equivalência entre o princípio evolucionário e o teorema de Bayes	60
Figura 4 - Simulação do modelo não-linear unidimensional	63
Figura 5 - Desempenho do FPB e do FPG na média dos estados do Modelo	64
Figura 6 - Desempenho do FPB e do FPG na MAP dos estados do Modelo	64
Figura 7 - Desempenho do FPB e do FPG na média da saída do Modelo	65
Figura 8 - Desempenho do FPB e do FPG na MAP da saída do Modelo	65
Figura 9 - Desempenho do FPB na evolução da densidade dos estados	66
Figura 10 - Desempenho do FPG na evolução da densidade dos estados	66
Figura 11 - Desempenho do FPB na evolução da densidade das observações	67
Figura 12 - Desempenho do FPG na evolução da densidade das observações	67
Figura 13 - Desempenho da média dos estados obtido pelo FPB	68
Figura 14 - Desempenho da média dos estados obtido pelo FPG	68
Figura 15 - Desempenho da MAP dos estados obtido pelo FPB	69
Figura 16 - Desempenho da MAP dos estados obtido pelo FPG	69
Figura 17 - Desempenho da média das observações obtido pelo FPB	70
Figura 18 - Desempenho da média das observações obtido pelo FPG	70
Figura 19 - Desempenho da MAP das observações obtido pelo FPB	71
Figura 20 - Desempenho da MAP das observações obtido pelo FPG	71
Figura 21 - Desempenho da média dos estados obtido pelo FPB	72
Figura 22 - Desempenho da média dos estados obtido pelo FPG	72
Figura 23 - Desempenho da MAP dos estados obtido pelo FPB	73
Figura 24 - Desempenho da MAP dos estados obtido pelo FPG	73
Figura 25 - Desempenho da média das observações obtido pelo FPB	74
Figura 26 - Desempenho da média das observações obtido pelo FPG	74
Figura 27 - Desempenho da MAP das observações obtido pelo FPB	75
Figura 28 - Desempenho da MAP das observações obtido pelo FPG	75
Figura 29 - Função de custo a ser maximizada	77
Figura 30 - Função de custo a ser minimizada	77

Figura 31 – Convergência em x do FE com esquema de seleção local	78
Figura 32 – Convergência em y do FE com esquema de seleção local	78
Figura 34 – Convergência em y do FE com esquema de seleção global	79
Figura 35 – Convergência de x_t, y_t do FE com esquema de seleção local	80
Figura 36 – Convergência de x_t, y_t do FE com esquema de seleção global	80
Figura 37 – Função de custo a ser maximizada	82
Figura 38 – Função de custo a ser minimizada	82
Figura 39 – Convergência em x do FE com esquema de seleção local	83
Figura 40 – Convergência em y do FE com esquema de seleção local	83
Figura 41 – Convergência em x do FE com esquema de seleção global	84
Figura 42 – Convergência em y do FE com esquema de seleção global	84
Figura 43 – Convergência de x_t, y_t do FE com esquema de seleção local	85
Figura 44 – Convergência de x_t, y_t do FE com esquema de seleção global	85
Figura 45 – Simulação do modelo de volatilidade estocástica	86
Figura 46 - Desempenho do FPB na evolução da densidade dos estados	89
Figura 47 – Desempenho do FPB na evolução da densidade das observações	89
Figura 48 - Desempenho do FPB na estimativa dos estados	90
Figura 49 - Desempenho do FPB na estimativa das observações	90
Figura 50 - Desempenho do FPA na evolução da densidade dos estados	91
Figura 51 – Desempenho do FPA na evolução da densidade das observações	91
Figura 52 - Desempenho do FPA na estimativa dos estados	92
Figura 53 - Desempenho do FPA na estimativa das observações	92
Figura 54 - Desempenho do FPG na evolução da densidade dos estados	93
Figura 55 – Desempenho do FPG na evolução da densidade das observações	93
Figura 56 - Desempenho do FPG na estimativa dos estados	94
Figura 57 - Desempenho do FPG na estimativa das observações	94
Figura 58 – Aprendizado dos hiper-parâmetros do modelo via: FE+FPB	95
Figura 59 – Aprendizado dos hiper-parâmetros do modelo via: FE+FPA	95
Figura 60 – Aprendizado dos hiper-parâmetros do modelo via: FE+FPG	96
Figura 61 – Aprendizado dos hiper-parâmetros do modelo via: FE+FPB	96
Figura 62 – Aprendizado dos hiper-parâmetros do modelo via: FE+FPA	97
Figura 63 – Aprendizado dos hiper-parâmetros do modelo via: FE+FPG	97

Lista de tabelas

Tabela 1 - Desempenho do Erro do FPB e FPG	63
Tabela 2 - Desempenho do Erro do FPB, FPA e FPG	88
Tabela 3 - Erro da estimativa do FE+FPB para $T = 200 : 25 : 1000$	98
Tabela 4 - Erro da estimativa do FE+FPA para $T = 200 : 25 : 1000$	98
Tabela 5 - Erro da estimativa do FE+FPG para $T = 200 : 25 : 1000$	98
Tabela 6 - Erro em 50 simulações da estimativa do FE+FPB para $T = 1000$	99
Tabela 7 - Erro em 50 simulações da estimativa do FE+FPA para $T = 1000$	99
Tabela 8 - Erro em 50 simulações da estimativa do FE+FPG para $T = 1000$	99

Lista de símbolos e abreviações

$\boldsymbol{\theta}_t$	Vetor aleatório dos estados do fenômeno para o tempo t
$\boldsymbol{\theta}_t^{(i)}$	i -ésima amostra de $\boldsymbol{\theta}_t$
$\Theta_{0:t}$	Seqüência de estados $\Theta_{0:t} = \{\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_t\}$
$\Theta_{0:t}^{(i)}$	Trajatória ao longo do tempo da i -ésima amostra de $\Theta_{0:t}$
\mathbf{y}_t	Vetor aleatório das observações do fenômeno para o tempo t
$\mathbf{Y}_{1:t}$	Seqüência de observações $\mathbf{Y}_{1:t} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t\}$
$\mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Função de probabilidade com distribuição gaussiana
$\mathbf{U}(\mathbf{a}, \mathbf{b})$	Função de probabilidade com distribuição uniforme entre \mathbf{a} e \mathbf{b}
$\arg \min_{\boldsymbol{\theta}} \{\mathbf{f}(\boldsymbol{\theta})\}$	Argumento em $\boldsymbol{\theta}$ que minimiza a função $\mathbf{f}(\boldsymbol{\theta})$
$\arg \max_{\boldsymbol{\theta}} \{\mathbf{f}(\boldsymbol{\theta})\}$	Argumento em $\boldsymbol{\theta}$ que maximiza a função $\mathbf{f}(\boldsymbol{\theta})$
$\mathbf{E}(\cdot)$	Esperança matemática
$\delta(\cdot)$	Função delta de Dirac (função impulso)
$\exp(\cdot)$	Função exponencial
$\ln(\cdot)$	Função logaritmo natural
$p(\cdot)$	Função de probabilidade
$p(\cdot, \cdot)$	Função de probabilidade conjunta
$p(\cdot \cdot)$	Função de probabilidade condicional
$\mathbf{y}_t \sim p(\cdot)$	O vetor aleatório \mathbf{y}_t é distribuído de acordo com $p(\cdot)$
AGs	Algoritmos Genéticos
FPA	Filtro de Partículas Auxiliares
FPB	Filtro de Partículas Bootstrap
FPG	Filtro de Partículas Genéticas
FE	Filtro Evolutivo
PE	Programa Evolutivo
i.i.d.	Independente identicamente distribuído
MAP	Máxima a <i>posteriori</i>
MCCM	Monte Carlo via Cadeias de Markov
MEE	Modelos de espaço de estados
ASI	Amostragem Seqüencial por Importância
MCS	Monte Carlo Seqüencial