UMA APLICAÇÃO DE SISTEMAS DE INFERÊNCIA NEBULOSOS PARA A IDENTIFICAÇÃO AUTOMÁTICA DE MODELOS ARMA SAZONAIS E/OU NÃO SAZONAIS

Luiza Maria Oliveira da Silva, MsC
Faculdades Ibmec – Avenida Rio Branco, 108/ 5º andar Rio de Janeiro – Brasil
cep. 20040-001
luiza.maria@ibmecrj.br

Maria Augusta Soares Machado, DSc
Faculdades Ibmec – Avenida Rio Branco, 108/ 5º andar Rio de Janeiro – Brasil
cep. 20040-001
mmachado@ibmecrj.br

Reinaldo Castro Souza, PhD
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 – Rio de Janeiro – Brasil – cep. 22453-900
reinaldo@ele.puc-rio.br

Resumo

Este trabalho apresenta uma metodologia usando sistemas de inferência nebulosos para a identificação automática de estruturas Box & Jenkins sazonais e/ou não-sazonais.

Palavras-chaves: Identificação de Estruturas Box & Jenkins, Sistemas de Inferência Nebulosos, Inteligência Computacional Aplicada.

Abstract

This paper presents a methodology using a fuzzy inference systems to identify automatically Box & Jenkins seasonals and/or not seasonals structures.

Key words : Box and Jenkins model identification, Fuzzy Inference Systems, Soft Computing.

1 - INTRODUÇÃO

A metodologia Box & Jenkins (Box & Jenkins,1976) tem sido utilizada para fazer previsões com resultados melhores que outros métodos de previsão utilizados até então. Entretanto, alguns analistas têm relutado em usar esta metodologia, em parte, porque a identificação da estrutura adequada é uma tarefa bastante difícil. Por este motivo , a tecnologia de sistemas especialistas tem sido utilizada nesta identificação, por exemplo, em problemas de classificação de séries temporais (Reynolds , Stevens, Mellichamp & Smith,1995 e Machado,M.A,2000).
O presente trabalho tem como objetivo apresentar uma metodologia que utiliza técnicas de Inteligência Artificial, combinando conceitos de regras e conjuntos nebulosos, para a identificação automática de Estruturas Box & Jenkins sazonais.
Existe um sistema especialista utilizando *neuro-fuzzy* para identificação de séries temporais não-sazonais – modelos AR(1), AR(2), MA(1), MA(2), e ARMA(1,1) – (Machado, M. A., 2000) onde se

identifica apenas o filtro não-sazonal. O que será feito neste trabalho será a identificação dos modelos ARMA simples e/ou sazonais, identificando os filtros sazonal e linear da série.

2. METODOLOGIA BOX & JENKINS

A análise de séries temporais, segundo Box & Jenkins (1976), tem como objetivo principal a realização de previsão (Makridakis et al, 1998). Essa metodologia permite que valores futuros de uma série sejam previstos tomando por base apenas seus valores presentes e passados. Isso é feito através da correlação temporal existente entre os valores existentes.

A realização do processo temporal pelo método de Box & Jenkins é representada por um conjunto de processos estocásticos denominados modelos ARIMA (*autoregressive integrated moving average*) onde em cada instante de tempo t, existe um conjunto de valores que a série pode assumir, aos quais estão associadas possibilidades de ocorrência (Tápia, M., 2000).

Para cada instante de tempo t, é possível que exista uma função de densidade de probabilidade logo, cada variável aleatória $Z_t,\ t = t_1, t_2, ...$ pode ter média e variância específicas.

O trabalho consiste em descobrir qual é o processo que gera a série em estudo, isto é, qual o modelo que repesenta melhor a série.

Os modelos ARIMA simples de ordem (p, d, q) são definidos como

$$\phi(B)\ \nabla^d\ Z_t = \theta(B)\, a_t$$

e resultam da combinação de três componentes também denominados “filtros” que são:

- O componente auto-regressivo de ordem p – AR(p) (*autoregressive*);
- O filtro de integração de ordem d – I (*integrated*);
- O componente de médias móveis de ordem q – MA(q) (*moving average*).

Os modelos ARIMA sazonais de ordem $(P, D, Q)_s$ são definidos como

$$\Phi(B^s)\,\nabla_s^D\ Z_t = \Theta(B^s)\ a_t$$

e resultam da combinação de três componentes também denominados “filtros” que são:

- O componente auto-regressivo de ordem P – AR(P) (*autoregressive*);
- O filtro de integração de ordem D – I (*integrated*);
- O componente de médias móveis de ordem Q – MA(Q) (*moving average*).

Os modelos ARIMA multiplicativos de ordem $(p, d, q) \times (P, D, Q)_s$ são aplicados a séries que apresentam correlação serial ‘dentro’ e ‘entre’ períodos sazonais e são definidos como

$$\Phi(B^s)\,\phi(B)\,\nabla_s^D\ \nabla^d\ Z_t = \Theta(B^s)\,\theta(B)\,a_t$$

A figura abaixo é uma representação em diagramas do modelo ARIMA $(p, d, q) \times (P, D, Q)_s$ que ilustra a seqüência de filtragens aplicadas ao ruído branco (a_t).

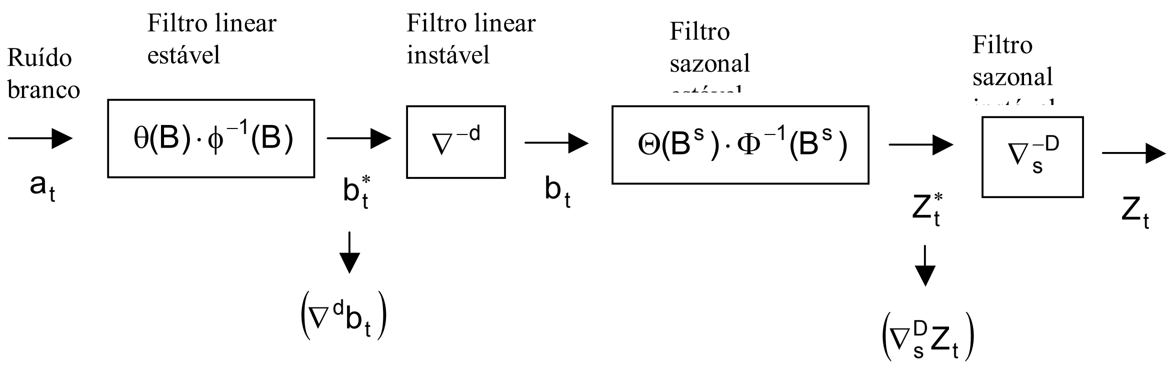

Figura 1 – Representação do modelo ARIMA $(p, d, q) \times (P, D, Q)_s$
Fonte: Souza e Camargo (2004:151)

Numa série temporal pode-se encontrar os três filtros ou um subconjunto deles, resultando daí, vários modelos na metodologia Box & Jenkins.
Uma condição que tem que ser colocada no processo estocástico é que este tem que ser estacionário. Um processo estocástico é dito estacionário de segunda ordem quando as seguintes condições forem satisfeitas para qualquer instante de tempo t:

$$E[z_t] = E[z_{t+k}] = \mu \,,$$
$$Var[z_t] = E[(z_t - \mu)^2] = \sigma^2$$
$$Cov[z_t, z_{t+k}] = E[(z_t - \mu)\cdot(z_{t+k} - \mu)]$$

As duas primeiras condições indicam que a média e a variância de Z_t não variam com o tempo e a terceira, indica que as autocovariâncias não dependem do tempo e sim em relação à distância k que separa as observações.
Quando a série recebe a influência de fatores sazonais, outro tipo de correlação passa a ter importância: a correlação entre os instantes de tempo distantes entre si por s ou múltiplos de s, onde s representa o período da sazonalidade.
A tabela a seguir apresenta as propriedades e características para a identificação teórica dos parâmetros p, q, P e Q dos modelos AR(p), MA(q), ARMA(p,q), SAR(P), SMA(Q) e SARMA(P,Q).

	AR(p)	MA(q)	ARMA(p,q)	SAR(P)	SMA(Q)	SARMA(P,Q)
Modelo expresso em termos dos w_t's anteriores	$\phi(B)\,w_t = a_t$	$\theta^{-1}(B)\,w_t = a_t$	$\theta^{-1}(B)\phi(B)w_t = a_t$	$\Phi(B^s)\,w_t = a_t$	$\Theta^{-1}(B^s)\,w_t = a_t$	$\Theta^{-1}(B^s)\Phi(B^s)w_t = a_t$
Modelo expresso em termos dos a_t's anteriores	$w_t = \phi^{-1}(B)a_t$	$w_t = \theta(B)a_t$	$w_t = \phi^{-1}(B)\theta(B)a_t$	$w_t = \Phi^{-1}(B^s)a_t$	$w_t = \Theta(B^s)\,a_t$	$w_t = \Phi^{-1}(B^s)\Theta(B^s)a_t$
Função de Autocorrelação ρ_k	Infinita (exponenciais amortecidas e/ou senóide amortecido). Não se anulam bruscamente.	Finita. Anulam-se bruscamente no lag k.	Infinita (exponenciais amortecidas e/ou senóide amortecidos para k > q-p). Não se anulam bruscamente.	Infinita (exponenciais amortecidas e/ou senóide amortecido). Não se anulam bruscamente.	Finita. Anulam-se bruscamente no lag k.	Infinita (exponenciais amortecidas e/ou senóide amortecidos para k > Q - P). Não se anulam bruscamente.
Função de autocorrelação parcial ϕ_{kk}	Finita. Anulam-se bruscamente no lag k.	Infinita (dominada por exponenciais amortecidas e/ou senóide). Não se anulam bruscamente.	Infinita (dominada por exponenciais amortecidas e/ou senóide amortecidos para k > q-p). Não se anulam bruscamente.	Finita. Anulam-se bruscamente no lag k.	Infinita (domonada por exponenciais amortecidas e/ou senóide). Não se anulam bruscamente.	Infinita (exponenciais amortecidas e/ou senóide para k > Q - P). Não se anulam bruscamente.

Tabela 1 – Comportamento teórico dos modelos AR(p), MA(q), ARMA(p,q), SAR(P), SMA(Q) e SARMA(P,Q)
Fonte: Souza e Camargo (2004:68)

3. MATEMÁTICA NEBULOSA (FUZZY SETS)

A Teoria dos Conjuntos Nebulosos foi desenvolvida por Lotf Zadeh da Universidade da Califórnia, em Berkeley, nos anos 60 e vem sido utilizada nas áreas de classificação, mineração de dados, previsão de séries temporais e sistemas especialistas entre outras.
O objetivo desta teoria é de modelar as informações de caráter impreciso ou vago. Humanos geralmente avaliam vários conceitos de forma diferente. Por exemplo, as expressões da linguagem natural como ‘pequeno’, ‘médio’ e ‘grande’ podem ser transformadas em uma linguagem matemática através de conjuntos nebulosos.
A lógica nebulosa (fuzzy logic) é tida como uma generalização da lógica clássica (Booleana), expandida para manipular conceitos de “verdades parciais”, entendidos como valores que se situam entre o “completamente verdadeiro” e o “completamente falso” (Zadeh,1979 ; Jang & Mizutani,1997).
Da mesma forma que existe uma estreita relação entre o conceito de conjunto e a lógica clássica , há uma estreita relação entre a lógica nebulosa e a teoria dos conjuntos nebulosos.
Na teoria dos conjuntos nebulosos, o raciocínio exato corresponde a um caso limite do raciocínio aproximado e é interpretado como um caso particular de uma quantidade nebulosa. O valor verdade de uma proposição pertence a um conjunto parcialmente ordenado, ao contrário dos sistema lógicos binários (conjuntos *crisp*), onde o valor verdade só pode assumir dois valores: verdadeiro (1) ou falso (0).
O grau de pertinência dos elementos de um conjunto é especificado por um número: 1 para os estritamente membros, 0 para os não-membros e valores do intervalo (0, 1) para representar a transição entre esses extremos. A
O conhecimento do fenômeno é expresso através de afirmativas do tipo: ‘se (um conjunto de condições é satisfeito) então (pode-se inferir um conjunto de conseqüências)’.
As regras são decididas ‘a priori’, baseadas no conhecimento do sistema em estudo (conhecimento do especialista).
O raciocínio nebuloso é formado pelos seguintes eventos:

- transformação de variáveis numéricas em variáveis nebulosas usando o processo de fuzzificação
- criação de uma base de regras nebulosas expressas por declarações do tipo ‘ se... ....então’
- criação de um sistema de inferência que mapeia conjuntos nebulosos em conjuntos nebulosos (ou não). Este manipula o caminho no qual as regras são combinadas
- obtenção do resultado utilizando-se um defuzzificador que mapeia o conjunto nebuloso de saída em um número real no caso de saída nebulosa.

O diagrama a seguir ilustra um sistema de inferência nebuloso genérico com saída nebulosa.

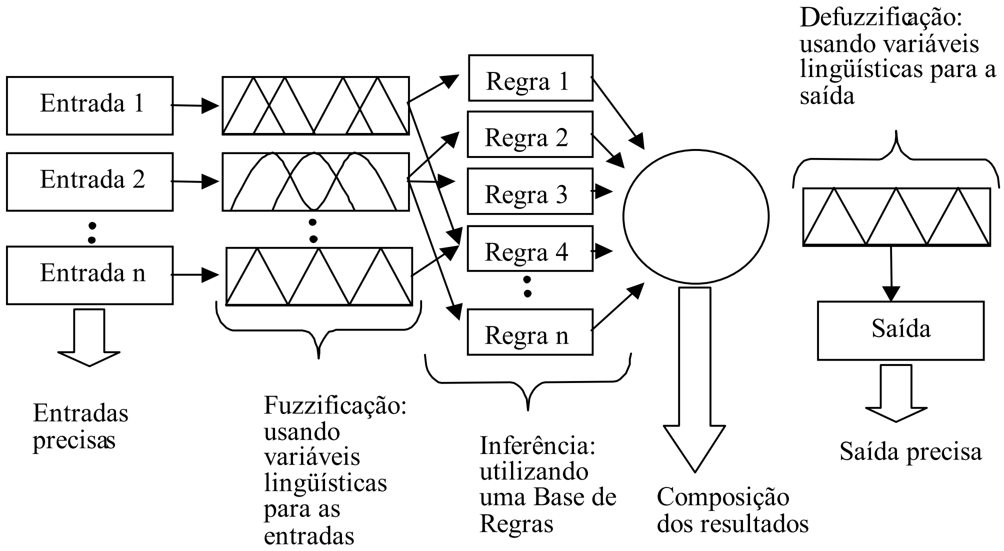

Figura 2- Sistema de inferência nebuloso
Fonte: Baptista, 2003.

A matemática nebulosa permite a criação de sistemas de inferências nebulosos e representa o conhecimento de forma explícita, através de regras nebulosas, possibilitando, facilmente , o entendimento do sistema em estudo.

4. SISTEMA PROPOSTO

Na tomada de decisão sobre a estrutura Box & Jenkins adequada à modelagem de um determinado processo estocástico, é grande a dose de julgamento subjetivo a ser realizado por um especialista de análise de séries temporais. É proposta deste trabalho, usar a flexibilidade para tratamento de situações desta natureza possibilitada pela modelagem de problemas por meio da lógica nebulosa.
O sistema especialista proposto é composto de bases de regras nebulosas para as estruturas Box & Jenkins de modelos sazonais cuja especificação geral consiste: de variáveis de entrada caracterizando os valores das funções de autocorrelação e autocorrelação parcial.
O procedimento para a realização deste trabalho é:

- Passo 1:

Variávies de entrada: para os modelos sazonais de período 12, as variávies de entrada são vetores formados pelos estimadores das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) nos *lags* 12, 24, 36 e 48. Estes estimadores são gerados através de simulações de séries dos modelos SAR(1), SAR(2), SMA(1), SMA(2) e SARMA(1,1) com 100, 200, 300, 400 e 500 observações para cada um deles.

- Passo 2:

Funções de pertinência das variáveis de entrada do modelo: são ajustadas funções de pertinência, *Generalized bell,* antecedentes das regras nebulosas para representar a significância dos picos para os valores positivos e negativos das autocorrelações e das autocorrelações parciais estimados com a finalidade de criar uma base de regras característica para cada estrutura estudada.

- Passo 3:

Regras do modelo que serão do tipo

Se ACF no *lag* 12 é A_1 e ACF no *lag* 24 é A_2 e ... e PACF no *lag* 48 é A_8 então o modelo é {SAR(1) ou SAR(2) ou SMA(1) ou SMA(2) ou SARMA(1,1)), usando antecedentes do tipo *AND*.

Para determinar as regras foram utilizados os seguintes passos:
1º) Estimadores necessários, ACF's e PACF's, para a identificação do modelo Box & Jenkins, gerados no passo 1.
2º) A partir desses dados, foram criadas regras para o sistema FUZZY.

- Passo 4:

Agregação das regras:
O sistema desenvolvido consiste de 114 regras para o modelo $SAR(1)_{12}$ e 174 regras para o modelo $SMA(1)_{12}$. O restante das regras para os demais modelos estão sendo elaboradas.
Cada regra determina a variável de saída e esta, identificará o tipo de modelo Box & Jenkins.

- Passo 5:

Defuzzificação da agregação das regras através do centróide.

5- RESULTADOS OBTIDOS E CONCLUSÕES

Para realizar a validade do sistema foram geradas novas séries representativas dos modelos Box & Jenkins com 100, 200, 300, 400 e 500 observações para cada um deles e cada série gerada possui. valor aleatório dos parâmetros Φ e Θ (parâmetros utilizados na construção dos modelos Box & Jenkins SAR(1) e SMA(1), respectivamente). Estas séries foram testadas no sistema FUZZY e no software FPW – versão 3.5 – para fazer a identificação e fazer uma comparação entre o número de acertos no sistema FUZZY e no software FPW.

Nas tabelas a seguir, tem-se o número de acertos para as séries com 100, 200, 300, 400 e 500 observações utilizando o sistema de inferência nebuloso (FUZZY) e o FPW para determinados intervalos de Φ, para o modelo SAR(1) e Θ para o modelo SMA(1) e o número de acertos no global para cada tamanho de série.
Pode-se observar que no geral a identificação através do sistema de inferência nebuloso (FUZZY) é tão bom quanto a identificação através do FPW. Para valores de Φ entre: (-1; -0,9) a relação de acerto foi equivalente; (-0,9; -0,1), (-0,1; 0,1) e (0,1; 0,9) para as séries com 100 observações o sistema FUZZY identificou melhor e para as demais séries, o resultado foi equivalente e, (0,9; 1) o sistema FUZZY identificou melhor.

Número de observações da série	100		200		300		400		500	
Sistema	FUZZY	FPW	FUZZY	FPW	FUZZY	FPW	FUZZY	FPW	FUZZY	FPW
Valores de Φ	Número de séries com determinado valor de Φ									
(-1; -0,9)	2 séries		8 séries		4 séries		5 séries		11 séries	
	100%	100%	100%	75%	100%	100%	100%	100%	100%	91%
(-0,9; -0,1)	32 séries		37 séries		40 séries		38 séries		38 séries	
	75%	47%	78%	65%	73%	80%	84%	71%	74%	82%
(-0,1; 0,1)	12 séries		10 séries		12 séries		10 séries		8 séries	
	25%	0%	0%	0%	17%	8%	50%	30%	38%	38%
(0,1; 0,9)	50 séries		42 séries		40 séries		43 séries		39 séries	
	62%	36%	95%	71%	80%	83%	70%	63%	85%	79%
(0,9; 1)	4 séries		3 séries		4 séries		4 séries		4 séries	
	100%	0%	100%	0%	100%	25%	75%	25%	75%	25%
Porcentagem total de acertos	64	35	80	60	71	71	75	63	78	76

Tabela 2 – Resultados obtidos para a série $SAR(1)_{12}$

Número de observações da série	100		200		300		400		500	
Sistema	FUZZY	FPW	FUZZY	FPW	FUZZY	FPW	FUZZY	FPW	FUZZY	FPW
Valores de Φ	Número de séries com determinado valor de Φ									
(-1; -0,9)	8 séries		5 séries		4 séries		4 séries		6 séries	
	100%	88%	100%	100%	100%	100%	100%	75%	100%	100%
(-0,9; -0,1)	39 séries		36 séries		40 séries		37 séries		36 séries	
	85%	51%	83%	64%	90%	85%	97%	84%	92%	81%
(-0,1; 0,1)	9 séries		11 séries		12 séries		13 séries		9 séries	
	0%	0%	1%	1%	33%	17%	38%	23%	55%	11%
(0,1; 0,9)	39 séries		47 séries		41 séries		41 séries		48 séries	
	92%	44%	96%	79%	93%	80%	95%	90%	100%	96%
(0,9; 1)	5 séries		1 série		3 séries		5 séries		1 série	
	100%	80%	100%	100%	100%	100%	100%	100%	100%	100%
Porcentagem total de acertos	85	82	48	82	66	85	76	89	79	93

Tabela 3 – Resultados obtidos para a série $SAR(1)_{12}$

6.REFERÊNCIAS BIBLIOGRÁFICAS

BAPTISTA, JOSÉ MAURÍCIO, Utilizando Lógica Fuzzy em um Modelo Macroeconômico, Dissertação de mestrado profissionalizante, Faculdades Ibmec. Rio de Janeiro, 2003.

BARRETO, M. J., *Inteligência Artificial no Limiar do Século XXI*, Duplic-Prestação de Serviços, 1997.

BRAGA, M.J.,BARRETO, J.M., MACHADO, M.A., *Conceitos da Matemática Nebulosa na Análise de Risco*, Artes & Rabiscus, 1995.

BOX, P.E. G., JENKINS, M.G., *Time Series Analysis Forecasting and Control*, Holden - Day Inc., 1976.

COSTA, ALEX et al, Lógica Fuzzy: Conceitos e Aplicações http://www.inf.unisinos.br/~cazella/dss/fuzzy_relatorio.pdf - acesso em 26/04/2004.

HAMILTON, JAMES D., *Time Series Analysis*, Princeton University Press, New Jersey, 1994.

HARVEY, ANDREW C., *Time Series Models*, , MIT Press, Great Britain 2sd ed. , pp. 22/28, 1993.

MACHADO, MARIA AUGUSTA S., Identificação das Estruturas Box & Jenkins não Sazonais usando Redes Neurais Nebulosas, Tese de Doutorado, PUC-RJ, 2000.

MORETTIN, PEDRO A., TOLOI, CÉLIA M., *Previsão de Séries Temporais*, Atual, 1987.

REYNOLDS, B., STEVENS T., MELLICHAMP R., SMITH M. J., *Box-Jenkins Forecast Model Identification*, A.I. Expert June 1995.

SHUMWAY, ROBERT H., STOFFER, DAVID S., *Time Series Analysis and Its Applications*, Springer, 2000.

SOUZA, C.R., CAMARGO, M.E., *Análise e Previsão de Séries Temporais: os Modelos ARIMA*, s/e, 2ª edição, Rio de Janeiro, 2004.

TANSCHEIT, RICARDO, Fundamentos de lógica Fuzzy e controle Fuzzy http://www.ica.ele.puc-rio.br/cursos/download/SI-Logica_Controle_Fuzzy.pdf - acesso em 26/04/2004.

ZADEH, L.A., *A Theory of Approximate Reasoning*, Machine Intelligence, Vol. 9, HAYES, J., MICHIE, D., MIKULICH, L. I., editors Halstead Press 1979.

ZIMMERMANN, H.J., *Fuzzy Set Theory and its Applications*, Kluwer-Nijhoff, 1980.