FUNDAÇÃO
GETULIO VARGAS

EPGE
Escola de Pós-Graduação em Economia

SEMINÁRIOS
DE PESQUISA
ECONÔMICA

Coordenação: Prof. Rubens Penha Cysne
Diretor de Pesquisa da EPGE/FGV
Email: rubens@fgv.br - ☎ (021) 536-9245

"MODELOS DINÂMICOS NÃO LINEARES E APLICAÇÕES DE QUADRATURA GAUSSIANA"

HELIO MIGON
(COPPE - UFRJ)

LOCAL
Fundação Getulio Vargas
Praia de Botafogo, 190 - 10° andar - Auditório

DATA
08/05/97 (5ª feira)

HORÁRIO
16:00h

SIMULATION AND COMPUTATION

Volume 25 Number 2 1996

COMMUN. STATIST.—SIMULA., 25(2), 369-380 (1996)

LONGITUDINAL DATA ANALYSIS OF ANIMAL GROWTH VIA MULTIVARIATE DYNAMIC MODELS

Emanuel P. Barbosa
Instituto de Ciências Exatas - UFMG
CP 702 - 30161-970 - Belo Horizonte, MG - Brazil
e-mail: emanuel at oraculo.lcc.ufmg

Helio S. Migon
Instituto de Matemática - UFRJ
CP 68530 - 20944-970 - Rio de Janeiro, RJ - Brazil
e-mail: mad02014 at ufrj.bitnet

Key Words & Phrases: Longitudinal data; Animal growth; Dynamic hierarchical models; Biological parameters.

ABSTRACT

We propose models to analyze animal growth data with the aim of estimating and predicting quantities of biological and economical interest such as the *maturing rate* and *asymptotic weight*. It is also studied the effect of environmental factors of relevant influence in the growth process.

The models considered in this paper are based on an extension and specialization of the dynamic hierarchical model (Gamerman & Migon, 1993) to a non-linear growth curve setting, where some of the growth curve parameters are considered exchangeable among the units. The inference for these models are approximate conjugate analysis based on Taylor series expansions and linear Bayes procedures.

1. INTRODUCTION

Longitudinal data of growth usually present non-linear behavior in relation to the assessment condition, as for instance the time, and are observed for more than one individual or experimental unit. A typical case is the study of animal growth, where the mechanism of growth of individuals of a certain race are characterized by biological factors of genetic nature, and environmental factors (see for instance, Fitzhugh,1976).

In this paper, models are proposed to analyze growth data of cattle (weight versus age, corresponding to the monitoring of cows of a certain race), with the aim of estimating and predicting quantities of biological and economical interest such as *maturing rate* (measure of how quick the animal reaches its final weight) and *asymptotic weight*. The information contained in these parameters is important for the design of animal improvement programs to increase the efficiency of the production process. It is also studied the effect of environmental factors of relevant influence in the growth process such as the animal birth season.

A traditional univariate modeling of these data through classical growth curves suffers from the so called aggregation dilemma (Blattberg & George, 1991): a single curve for the entire data set ('pooled model') sin for excess of aggregation and consequent loss of information and, on the other hand, an independent model for each individual can (and in fact it does) present instability in the parameter estimates: such problem is approached through shrinkage procedures based on hierarchical and shared parameter models. Since the initial presentation of the so called hierarchical models by Lindley & Smith (1972) as a flexible Bayesian modelling structure for data analysis, its practical use has been frequent, as for instance, in Racine-Poon (1985), Blattberg & George (1991), Lange, Carlin & Gelfand (1992) and many others.

The class of models considered in this paper can be interpreted as a multivariate extention of the models introduced by Migon & Gamerman (1993) or, alternatively, as a non-linear extention of the models presented by Barbosa & Harrison (1992). The former is based on an extension and specialization of the dynamic hierarchical model (Gamerman & Migon, 1993) to a growth curve setting, where all the parameters are considered exchangeable among the units of same birth season.

The hierarchical dynamic models proposed in this paper are defined in three parts or stages. In the first one, a 3-parameter growth curve is defined for each unit (where the parameters have the interpretation of maturing rate, growth factor & asymptotic weight), and stochastic constraints (structural information) are imposed on these parameters in the second stage according to three different hypothesis: model 1 - the maturing rate parameter exchangeable among the units, model 2 - maturing rate and

growth factor exchangeable and model 3 - all the parameters exchangeable; the third part of the model (evolution equation) expresses the dynamic evolution of the parameters.

The inference for these models are approximate conjugate analysis based on 1^{st} order Taylor Series expansions and linear Bayes procedures; the implementation of these models are made considering flat priors and some practical procedures such as: discount factors for the dynamic evolution, variance law (proportional to the level) for the observational variance, variances in the second stage estimated off-line from corresponding univariate models, etc. Models fitting and predictive performance are assessed through MAE - mean absolute error for all models, including a degenerate case of model 3 where the variance in the second stage is zero, which we call 'pooled model'.

This paper is organized as follows. In Section 2 a general formulation and analysis of the models proposed in this paper are presented. These models are adapted to a growth curve context in Section 3, resulting in three classes of hierarchical and sharing parameters models. Fitting results and models comparison are discussed in section 4, followed by some final remarks in section 5.

2. HIERARCHICAL NON-LINEAR DYNAMIC MODELS

2.1 Modeling Structure

The dynamic hierarchical model is composed by three parts: the observation equation, including a link function, the structural equation and the evolution equation, describing, respectively, the observational distribution, the parameter hierarchy structure and the parameter dynamics, as follows:

(i) - Observational distribution:

$$(y_t|\eta_t, \sigma_t^2) \sim N[\eta_t, V_{1,t}\sigma_t^2] \tag{1}$$

where: $V_{1,t} = diag(v(\eta_{1,t}), \cdots, v(\eta_{r,t}))$, with $v(.)$ defining a variance law to be specified; σ_t^2 is a scale factor, and y_t is a $r \times 1$ vector of observation units taken at time t.

The observational process mean η_t is related to the state parameters of the first level of the hierarchy $\theta_{1,t}$ through the inversible link function $h(.)$ such that: $\lambda_t = h(\eta_t) = F_{1,t}\theta_{1,t}$, where $F_{1,t}$ is a known regression matrix of proper dimensions.

(ii) - Structural equation:

$$(\theta_{1,t}|\theta_{2,t}, \sigma_t^2, D_{t-1}) \sim [F_{2,t}\theta_{2,t}, V_{2,t}\sigma_t^2] \tag{2}$$

where $F_{2,t}$ is a known contraction matrix and $\theta_{2,t}$ is a vector of structural parameters.

(iii) - State parameters evolution:

$$(\theta_{2,t}|\theta_{2,t-1},\sigma_t^2,D_{t-1}) \sim [g_t(\theta_{2,t-1}),W_t\sigma_t^2] \tag{3}$$

where $g_t(.)$ is the process evolution function and [ ] denotes a distribution specified only by the first two moments.

There is no theoretical restrictions to consider additional stages in the parameters hierarchy. It is worth to mention that the dimension of the parameters reduce from stage to stage. Following, inferential aspects of the non-linear hierarchical models with two stages are discussed.

2.2 - Inference procedure

Inference is performed sequentially in time as follows. At time t the prior moments for the first and second stage state parameters $\theta_{1,t}$ and $\theta_{2,t}$ are obtained from the posterior moments of $\theta_{2,t-1}$ through the structural and evolution equations with a first order Taylor series expansion of $g(.)$ about the posterior mean of $\theta_{2,t-1}$.

The prior moments of the observational mean η_t are matched to those obtained from a local linear approximation to the link function equation, and a Bayesian conjugate analysis for (η_t,σ_t^2) with normal-inverse gamma prior distribution is used to obtain the posterior distribution for these parameters. Linear Bayes methods are then used in conjunction with the link function equation to derive the posterior moments for $\theta_{1,t}$ and $\theta_{2,t}$. It is worth mention that the inferential procedure described above is conditional on the value of the matrices $V_{2,t}$ and W_t. The values of W_t are determined considering the discount factor method (West & Harrison, 1989) and the elements of $V_{2,t}$, in a empirical Bayes fashion. Further details about the inferential procedure can be found in Migon & Barbosa (1994).

3. A CLASS OF GROWTH MODELS

A large class of growth models is presented combining the models developed in Section 2 with the univariate models introduced by Migon & Gamerman (1993) which are described in this section.

3.1 - Univariate Growth Models

Let $\eta_t = E[y_t|D_{t-1},\beta]$ be the expected response function of the univariate time series y_t, where D_{t-1} represent all the past information available before time t, and β is the adopted global parametrization. A general family of growth curves is defined by $h(\eta_t) = \beta_1 + \beta_2\beta_3^t$, where $h(\eta) = \eta^\tau$ if $\tau \neq 0$ and $\log(\eta)$ if $\tau = 0$, which relates the data mean to the parameters used to describe it.

Some important examples are the modified exponential ($\tau = 1$), the Gompertz ($\tau = 0$) and the logistic ($\tau = -1$). It should be noted that the

transformation is applied to the process mean, with the data in its original scale, and also, that if $\beta_3 > 0$ and $\tau \geq 0$ the data grows without limit, and if $\beta_3 < 1$ its grows under an asymptotic value.

To make the above formulation more flexible we reparametrize the model defining the parameters recursively in order to account for local behavior of the growth curve as opposed to the former global description. Also we permit that the parameters evolve smoothly in time. The dynamic evolution is:

$$\theta_t = g(\theta_{t-1}) + w_t \tag{4}$$

where $\theta_t = (\mu, \gamma, \phi)_t^T$ is the vector of parameters with $\mu_t = h(\eta_t)$ the process level, γ_t the increment from time $t-1$ to t and ϕ_t, a dump factor ($\phi < 1$). The non-linear evolution function is defined as $g(\theta_{t-1}) = (\mu_{t-1} + \gamma_{t-1}, \gamma_{t-1}\phi_{t-1}, \phi_{t-1})$ and ω_t is a random component.

Note that the case $\phi_t = 1, \forall t$, corresponds to the linear growth model. The local parameters relate with the global ones through: $\beta_1 = \mu_t + \frac{\gamma_t}{1-\phi_t}$, $\beta_2 = -\frac{\gamma_t}{1-\phi_t}$ and $\beta_3 = \phi_t$.

The model specification is completed with the observational distribution defined by $y_t|\eta_t \sim N[\eta_t, v(\eta_t)]$, where the inference procedure involve the modeling of $v(\eta)$, the use of discount factors associated to ω_t and Taylor series approximations to coupe with the existing non-linearity.

3.2 - Hierarchical Growth Curve Models

The basic structure of the models presented in this section is built from the combination of the models introduced in sections 2.1 and 3.1, with the adoption of exchangeability assumptions among the units, concerning one or more parameters.

Such hypothesis, called structural equations in the hierarchical model context, can be defined in at least 3 alternative forms for the generalized growth models presented in this paper. For each of these alternatives, we present the regressors F_1 and F_2 and the non-linear evolution function $g(.)$. Among these models, just the one where the three parameters are exchangeable in relation to the units, is in fact hierarchical; the other two models have structural equations redundant.

Also, for a particular observational unit i ($i = 1, 2, \cdots, r$) we have:

$$(y_{t,i}|\eta_{t,i}, \sigma_t^2) \sim N[\eta_{t,i}, \sigma_t^2 V_{t,i}] \tag{5}$$

with $V_{t,i} = v(\eta_{t,i})$ for a given variance law $v(.)$ and a link function described by $\eta_t = \mu_{t,i}^{\tau}$, $\tau = -1, 0, 1$.

Model 1: One parameter exchangeable (dump factor ϕ).

In a biological context of growth curves the dump factor ϕ is associated to the animal maturing rate. It is supposed that the maturing rate is a race characteristic and therefore the parameter associated to it should be

exchangeable among the animals, but each animal develops its own individual weight level. Under these hypothesis, the parametrization adopted is, $\theta_1 = (\mu_1, \gamma_1, \phi_1, \cdots, \mu_r, \gamma_r, \phi_r)^T$ and $\theta_2 = (\mu_1, \gamma_1, \cdots, \mu_r, \gamma_r, \phi)^T$.

It should be noted that $r_1 = 3r$ and $r_2 = 2r + 1$, where r is the number of units considered; using the Kronecker-product notation, the regressors are defined by $F_1 = (1,0,0) \otimes I_r$ and $F_2 = (F_2^1, \cdots, F_2^r)^T$, where

$$F_2^i = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

The non-linear evolution function $g(.)$ is defined for this model by $g(\theta_{2,t-1}) = (\mu_1 + \gamma_1,\ \gamma_1\phi, \cdots,\ \mu_r + \gamma_r,\ \gamma_r\phi,\ \phi)_{t-1}^T$ and the structural variance has the form $V_{2,t} = I_r \otimes V$, with $V = diag(0, 0, V_\phi)$, since with the exception of the parameter ϕ, the others two are free of structural restrictions.

Although the formulation presented is correct, it is not parsimonious since the structural equation is redundant; this happens because given $\theta_{2,t}$ in the hierarchical model, there is no transference of information from y_t to $\theta_{1,t}$, since $y_t \perp \theta_{1,t}|\theta_{2,t}$, i.e., $cov(y_t, \theta_{1,t}|\theta_{2,t}) = 0$.

The proof is immediate: for instance, in the case of an identity link function, it is easily obtained from the observation and structural equations that, $cov(y_t, \theta_{1,t}|\theta_{2,t}) = cov(F_1 v_{2,t} + v_{1,t},\ v_{2,t}) = \sigma_t^2 F_{1,t} V_{2,t} = 0$ where $v_{i,t} \sim N[0,\ \sigma_t^2 V_{i,t}]$, $i = 1, 2$.

Model 2: Two parameters exchangeable (maturing rate & growth factor)

In this model only the process level μ is supposed distinct for each individual. The parametrization adopted is $\theta_1 = (\mu_1, \gamma_1, \phi_1, \cdots, \mu_r, \gamma_r, \phi_r)^T$ and $\theta_2 = (\mu_1, \cdots, \mu_r, \gamma, \phi)^T$. Note that, $r_1 = 3r$ and $r_2 = r + 2$, which characterizes a large reduction in parameter dimension.

The asymptotes are determined uniquely for each individual. The regressor F_1 is defined as in model 1, but the components of F_2 are:

$$F_{2,t}^i = \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & \cdots & 0 & 1 \end{pmatrix}$$

The non-linear evolution function $g(.)$ is defined in this model by $g(\theta_{2,t-1}) = (\mu_1 + \gamma, \cdots, \mu_r + \gamma,\ \gamma\phi,\ \phi)_{t-1}^T$ and the structural variance $V_{2,t} = I_r \otimes V$, with

$$V = \begin{pmatrix} 0 & 0 & 0 \\ 0 & V_\gamma & V_{\gamma\phi} \\ 0 & V_{\gamma\phi} & V_\phi \end{pmatrix}$$

Model 3: All parameters exchangeable

This model can be considered canonical since all the parameters are exchangeable and the reduction in dimension is maximum, i.e., $r_2 = 3$. Also, there is a unique asymptote common to all individuals. The parametrization adopted is $\theta_1 = (\mu_1, \gamma_1, \phi_1, \cdots, \mu_r, \gamma_r, \phi_r)^T$ and $\theta_2 = (\mu, \gamma, \phi)^T$. The F_1 regressor is the same as in the other two previous models but F_2 is defined by $F_2 = I_3 \otimes 1_r$, where $1_r = (1, \cdots, 1)^T$. The non-linear function g(.) is, $g(\theta_{2,t-1}) = (\mu + \gamma,\ \gamma\phi,\ \phi)^T_{t-1}$.

For each one of the three models presented, the linearization matrix $G_t = g'(m_{2,t})$, necessary to the description of the state parameters evolution, is easily obtained.

4. FITTING RESULTS AND MODELS COMPARISON

4.1- Introduction

From an initial data set of cattle weights and other variables collected at EMBRAPA - The Brazilian Institute for Agricultural and Cattle-Raising Research, in São Carlos, SP, corresponding to the monitoring of cows of Canchim race from its birth till its 30 months old at intervals of 3 months, were considered the data of 227 females which information of weight, birth season, etc. were available. The regularity of 3 months in the observation interval is guaranteed by standardization via linear correction in the original data, as recommended by the *Beef Improvement Federation.*

From this animal population it was considered a stratified random sample of 20 animals, with 10 units of them at each of two opposite birth seasons. The importance of considering birth season data corresponding to each animal is clear since this variable is a very representative environmental factor and the animal growth process depends on the environment..

Before the presentation of the data analysis in the next section using the proposed multivariate models, it is presented here some preliminary results corresponding to the fitting of the univariate models introduced in section 3.1 using individual data of each animal as well as average weights for the animals in each birth season.

Three types of growth curves were sequentially fitted: logistic, Gompertz and modified exponential, and in each case, two versions were considered, static and dynamic, where the distinction is made by the value of the discount factor. The analyses were considered using well centered prior distributions with reasonably vague variances, characterized by their first two moments, denoted by m_0 and C_0 (diagonal matrix), as shown below.

The prior moments for the modified exponential growth model were specified such that when the animal borns, its level of weight is in the interval (30; 56) kg with 95% of probability, which is in accordance with

TABLE I: Prior Moments

	M-Exponential	Gompertz	Logistic
m_0	$(43.5, 60, .95)$	$(3.8, .90, .95)$	$(.023, -.015, .95)$
C_0	$(43.5, 60, .05)$	$(1.0, 1.0, .05)$	$(.05, .05, .05)$

reality. The prior moments for the other growth models were specified in a similar way. The variance law considered is given by $V_t \propto \mu_t$. The discount factors for the system evolution were fixed at 0.95 for the level and growth factor and at 0.98 for the maturing rate ϕ, which corresponds to the expectation that the latter parameter should be more stable than the others.

Since the dynamic models have in general more flexibility for fitting purposes, we consider them from now on, even though there is no expressive advantages in predictive terms. The logistic model will be deleted from subsequent analyses since it has shown poor fitting.

4.2 - Models Implementation: numerical results

The hierarchical and shared parameters models for growth data analysis presented in this paper are fitted to the data described in 4.1, according to the inferential procedures proposed, considering the Gompertz and modified exponential forms, in their dynamic versions. Since models 1 and 2 have their structural equations redundant (they are a kind of degenerated hierarchical) we call them shared parameters models. Model 3 is considered in two versions: with $V_2 \neq 0$ (hierarchical) and with $V_2 = 0$ ('pooled model').

Model assessment is based on the predictive Mean Absolute Error - MAE. The models implementation makes use of the same discount factors and prior distributions considered in the univariate case; also, the second level state parameter variances are obtained off-line from the corresponding univariate fittings. The MAE values are presented for each fitted model considering not the individual units but just their quantiles.

From Table II and III, corresponding to two different birth seasons, it is clear that the hierarchical model has a better performance than the pooled model almost independently of the functional form considered, what suggests the usefulness of the proposed hierarchical methodology. It should be noted that this model permits to summarize the information contained in the data in a unique growth curve, but keeping the capability of making inferences of each individual separately.

TABLE II: MAE for Models 1, 2, 3 (Season I)

Gompertz	*Model* 1	*Model* 2	*Model* 3 − *hier.*	*Model* 3 − *pooled*
Q1	18.2	17.3	28.3	29.4
Q2	24.1	23.6	38.6	40.1
Q3	44.2	37.0	45.1	44.0
M-Exp.				
Q1	25.9	18.5	20.5	33.1
Q2	30.1	23.6	36.9	34.4
Q3	36.4	44.8	48.2	51.5

TABLE III: MAE for Models 1, 2, 3 (Season II)

Gompertz	*Model* 1	*Model* 2	*Model* 3 − *hier.*	*Model* 3 − *pooled*
Q1	36.8	32.8	37.4	36.8
Q2	40.1	42.5	41.8	53.7
Q3	53.1	47.7	49.7	60.5
M-Exp.				
Q1	34.7	38.3	31.6	37.0
Q2	38.3	48.1	35.2	45.3
Q3	44.1	52.0	52.1	50.1

These tables also show that, in general, models 1 and 2 have better predictive performance than the pooled model and, in some cases slightly better than the hierarchical model. It should be noted that these shared parameters models expresses the assumption that certain parameters (as for instance, maturing rate) are a race characteristic and therefore should be shared by the units. A model that absorb this knowledge a priori like models

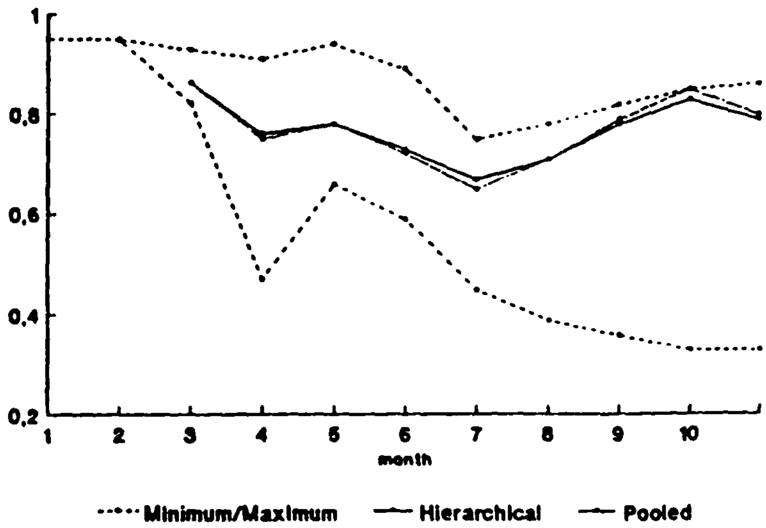


FIG.1 - Posterior Mean for the Third Parameter (Season I)

1 and 2, makes a better use of experimental information and therefore, it is not surprising their superior predictive performance.

It should be stressed the fact that the hierarchical model ($V_2 \neq 0$) and the shared parameter models in fact are useful to approach complementary questions. The shared parameter models are more effective for prediction and the hierarchical model is more adequate to summarize the information related to all individuals (of a certain race).

In order to exemplify the aggregation dilemma it is presented in figure 1, the graphic of the posterior mean of the ϕ parameter evaluated from the univariate models and from the hierarchical and pooled models. The dotted lines represent the minimum and maximum observed values from the univariate analysis for the 10 units considered in the season 1. The others represent the posterior mean evolution, respectively, for the hierarchical model (full line) and the pooled model (broken line). It is clear, as expected, that the estimates from the multivariate models are more stable and smooth than the ones obtained from the univariate models.

5. FINAL REMARKS

It is worth to stress the main findings and conclusions we can get from this application and to state some relevant points for further investigation.

One of the main aspects of the class of proposed models considered in this paper is it great flexibility. First, this class of models has a variety of particular and useful models embedded in it, defined by different functional forms and different exchangeability assumptions. Also, for a chosen functional form and a given hierarchical structure, the time varying parameters can accommodate different quantitative aspects of the growth process. Another important aspect is the fact that the parameters have a biological interpretation.

As expected all models perform better than the pooled model for data sets corresponding to each birth season. For our objectives, we can say that hierarchical models and shared parameter models are in some sense, complementary, since the two versions of model 3 are appropriate for data description and comparison (between seasons), and models 1 and 2 are more adequate for prediction.

One of the points for further work is the inclusion of an extra level of hierarchy in the model to contemplate the effects of different birth seasons.

We should note that the proposed models, conditionally on the value of the ϕ parameter would be linear and consequently an alternative form of analysis could make use of the Gauss-Hermite numerical integration over (ϕ, V_2). However, this approach would be more sensible only for models where V_2 has few non-zero elements. Alternatively, the estimation of V_2 could be carried out through Markov chain Monte Carlo technology, in the dynamic context as considered here.

BIBLIOGRAPHY

Barbosa, E. P. & Harrison, P. J. (1992) - Variance Estimation for Multivariate Dynamic Linear Models. Journal of Forecasting, vol 11, 7, 621-628.

Blattberg, R.C & George, E.I (1991), Shrinkage Estimation of Price and Promotion Elasticities: seemingly unrelated equations. Journal of the American Statistical Association, 86, 414, 304-315.

Fitzhugh, H.A (1976) - Analysis of growth curves ands strategies for altering their shapes. Journal of Animal Science, 42, 4.

Gamerman, D & Migon, H.S (1993), Dynamic Hierarchical Models. Journal of the Royal Statistical Society, B, 55, 3, 629-642.

Lange, N., Carlin, B.P & Gelfand, A.E (1992), Hierachical Bayes Model for the Progression of HIV Infection Using Longitudinal CD4 T-Cell Numbers. Journal of the American Statistical Association, 87, 419.

Lindley, D.V & Smith, A.M.F (1972), Bayes Estimates for the Linear Model (with discussion). Journal of the Royal Statistical Society, B, 34, 1-41.

Migon, H.S & Barbosa, E.P. (1994), Non-linear Dynamic Models with Application to Growth Curves. Rel.Tec., LES, UFRJ.

Migon, H.S & Gamerman, D (1991), Generalized Exponential Growth Models - a Bayesian approach. **Journal of Forecasting**, 12, 573-584 .

Racine-Poon, A (1985), A Bayesian Approach to Non-Linear Random Effect Models, **Biometrics**, 41, 1015-1023.

West, M & Harrison, P.J. (1989), **Bayesian Forecasting and Dynamic Models**, Springer-Verlag.

Received April, 1995; Revised July, 1995.

Uma Aplicação de Métodos de Monte-Carlo em Modelos de Volatilidade

Hedibert Freitas Lopes (IM-UFRJ)
Hélio S. Migon (IM/COPPE-UFRJ)

Abstract

Os modelos para séries temporais heterocedásticas têm sido explorados exaustivamente na última década. Engle (1982) e Bollerslev (1986) observaram que as mudanças na volatilidade de várias séries tendiam a ser serialmente correlacionadas e propuseram, respectivamente, os modelos ARCH (Autoregressive Conditional Heteroscedasticity) e GARCH (Generalized ARCH).

O mais popular desses modelos é o GARCH(1,1), onde a série de interesse $\{y_t\}$ é um processo ruído branco gaussiano com variância unitária, $\{\epsilon_t\}$, multiplicado por um fator de escala $\{\sigma_t\}$, isto é,

$$y_t = \sigma_t \epsilon_t, \quad t = 1, \cdots, T, \quad \epsilon_t \sim NID(0,1) \tag{1}$$

$$\sigma_t^2 = \gamma + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, \quad \gamma > 0, \quad \alpha + \beta < 1. \tag{2}$$

Este modelo pode ser estendido adicionando-se mais defasagens de y_t e σ_t^2 na equação (2), embora seja pouco comum nas aplicações. Quando $\beta = 0$ temos um ARCH(1).

Os principais fatores que podem explicar o crescente interesse nesta área são:[1]

1. O reconhecimento de que muitos modelos teóricos em finanças estão relacionados a volatilidade; por exemplo, os retornos de investimentos em ações e as taxas de câmbio exibem, em geral, mudanças na volatilidade ao longo do tempo.
2. A modelagem da volatilidade é um excelente campo para o desenvolvimento e aplicação de novas técnicas estatísticas para modelos não-lineares e não-gaussianos[2].

Geweke (1989) e Müller & Pole (1994) estudam os modelos GARCH sobre o enfoque bayesiano; Geweke utiliza técnicas de integração Monte Carlo com *Importance Sampling*, enquanto Müller & Pole utilizam Métodos Monte Carlo via Cadeias de Markov.

Uma alternativa aos modelos ARCH é permitir que σ_t^2 dependa de algum componente não-observável ou variável latente modelando seu logaritmo (isto é, $h_t = \log \sigma_t^2$), através de um processo estocástico linear. Modelos desse tipo são comumente chamados de *Modelos de Volatilidade Estocástica (SV)*. Esses modelos são uma versão, em tempo-discreto dos modelos que governam a maior parte da teoria moderna de finanças (Taylor (1993)).

[1] Existem centenas de artigos publicados sobre o tema.

[2] Carlin, Polson & Stoffer (1992) desenvolveram métodos de estimação de modelos dinâmicos não-lineares e não-normais usando *Gibbs Sampling*.

Uma versão simples de modelos de volatilidade estocástica foi proposta por Taylor (1986):

$$y_t = exp\{h_t/2\}\varepsilon_t, \quad t = 1, \cdots, T, \quad \varepsilon_t \sim NID(0,1) \tag{3}$$
$$h_t = \gamma_0 + \gamma_1 h_{t-1} + \nu_t, \quad \nu_t \sim (0, \sigma_\nu^2) \tag{4}$$

Uma interpretação da variável latente h_t é representar a variação irregular e aleatória de novas observações, o que é muito difícil de se modelar diretamente em mercados financeiros [3].

O modelo SV pode ser re-escrito como

$$\log\{y_t^2\} = h_t + \log\{\varepsilon_t^2\} \tag{5}$$
$$h_t = \gamma_0 + \gamma_1 h_{t-1} + \nu_t \tag{6}$$

o qual pode ser pensado como um modelo de espaços de estado.

A facilidade de avaliar a função de verossimilhança e sua adaptação para muitas séries temporais econômicas tornou os modelos ARCH muito populares. Por outro lado, a difícil avaliação da função de verossimilhança dos modelos de volatilidade estocástica dificultaram sua aplicação empírica[4].

Recentes esforços têm sido feitos para reverter esse quadro, uma vez que, como já foi dito, os modelos SV têm boas propriedades, pelos menos quando se trata da sua adaptação a situações econômicas empíricas. Jarquier, Polson & Rossi (1994) abordam o problema sobre o enfoque de Modelos Hieráquicos Bayesiano, isto é: $p(y_t \mid h_t)$, $p(h_t \mid \omega)$ e $p(\omega)$. O exemplo apresentado em (3) e (4) corresponde a

$$(y_t \mid \sigma_t^2) \sim N[0, \sigma_t^2] \tag{7}$$
$$(h_t \mid \omega) \sim N[\mu_t, \sigma_\nu^2] \tag{8}$$
$$(\omega) \sim p(\omega) \tag{9}$$

onde: $\sigma_t^2 = exp(h_t)$, $\mu_t = \gamma_0 + \gamma_1 h_{t-1}$ e $\omega = (\gamma_0, \gamma_1, \sigma_\nu)$.

Este trabalho tem como primeiro objetivo fazer uma comparação das representações mais simples vistas acima, onde a estimação é feita através de dois procedimentos de integração : (i) Amostrador de Gibbs e (ii) *Sampling-Resampling* que gera amostras de (γ, α, β) ou $(\gamma_0, \gamma_1, \sigma_\nu)$ da priori, amostras estas que são reponderadas pela função de verossimilhança e tranformadas em amostras da posteriori, onde toda análise é então concluída. Essa última técnica (chamada de *Sampling-Importance Resampling (SIR)*, quando ligada a uma *Importance Function*), tem um grande apelo pedagógico e sugere uma estratégia de cálculo fácil de implementar (Smith & Gelfand (1992). Fazemos esses exercícios com uma série simulada com uma estrutura na volatilidade e com algumas séries da economia brasileira.

[3] O interesse em modelar volatilidade já vem de longa data, com o trabalho de Clark (1973) que propôs um modelo de misturas para a distribuição das mudanças nos preços de ações.

[4] Kim & Shepard (1994) e Shepard (1995) fazem extensas revisões dos principais conceitos, técnicas de estimação e previsão e desafios que estão por traz dos modelos ARCH e dos modelos SV.

UMA APLICAÇÃO DE MÉTODOS DE MONTE CARLO EM MODELOS DE VOLATILIDADE ESTOCÁSTICA

Helio S. Migon e **Hedibert F. Lopes**

Universidade Federal do Rio de Janeiro - IM & COPPE

e-mail: migon at pep.ufrj.br

CONTEÚDO

0. INTRODUÇÃO

Motivação

- Teoria moderna de finanças propõe modelos de decisão relacionados com a volatilidade.
- Modelos de volatilidade são excelentes exemplos de modelos não-lineares e não-normais.
- Em modelos dinâmicos é comum modelar-se a variância observacional.

Objetivos

- Investigar uma ampla classe de modelos de volatilidade estocástica – uni e multivariada.
- Propor métodos de estimação Bayesiana e comparar resultados.
- Discutir tópicos de seleção de modelos via fator de Bayes e medidas de não linearidade.
- Avaliar influência usando medidade de Kullback-Leibner

MODELO ARCH(1)

$$y_t = h_t^{1/2}\epsilon_t, \qquad \epsilon_t \sim n[0,1]$$

$$h_t = \alpha + \delta\, y_{t-1}^2, \quad \alpha > 0$$

Alternativamente

$$y_t^2 = \alpha + \delta\, y_{t-1}^2 + \nu_t, \qquad \nu_t = h_t(\epsilon_t^2 - 1)$$

Propriedades

$$K = \frac{E[y_t^4]}{E^2[y_t^2]} = \frac{3(1-\delta)}{1-3\delta^2}, \quad \text{se } \ 3\delta^2 < 1$$

- y_t^2 é covariância-estacionário com função de autocorrelação:

$$\rho_{y_t^2}(s) = \delta^s, \quad \text{se} \quad \delta^2 < 1/3$$

- y_t é ruído branco se: $\delta < 1$

$$y_t = h_t^{1/2}\,\epsilon_t, \qquad \epsilon_t \sim n[0,1]$$

$$h_t = \alpha + \delta\, y_{t-1}^2 + \phi\, h_{t-1}, \quad \alpha > 0$$

Alternativamente

$$y_t^2 = \alpha + (\delta + \phi)\, y_{t-1}^2 + \nu_t - \phi\, \nu_{t-1}, \quad \nu_t = h_t(\epsilon_t^2 - 1)$$

Propriedades

$$K = \frac{E[y_t^4]}{E^2[y_t^2]} = 3 + \frac{3\ var[E[y_t^2|y_{t-1}]]}{E[y_t^2]^2},$$

- y_t^2 é covariância-estacionário com função de autocorrelação:

$$\rho_{y_t^2}(1) = \frac{1 - \delta(\delta+\phi)}{1 + \delta^2 - 2\delta(\alpha+\phi)}$$

$$\rho_{y_t^2}(s) = \rho_{y_t^2}(1)^{s-1}, \quad s \geq 2$$

- y_t é ruído branco se: $\delta + \phi < 1$

MODELO DE VOLATILIDADE ESTOCÁSTICA LOG-AR(1)

VEST - LOG-AR(1)

$$y_t = h_t^{1/2}\ \epsilon_t, \qquad \epsilon_t \sim n[0,1]$$

$$\Psi_t = log(h_t)$$

$$\Psi_t = \alpha + \delta\ \Psi_{t-1} + \sigma_\nu\ \nu_t$$

Alternativamente

$$log(y_t^2) = \Psi_t + a_t, \qquad a_t = log(\epsilon_t^2) \sim Log - \chi^2$$

$$\Psi_t = \alpha + \delta\ \Psi_{t-1} + \sigma_\nu\ \nu_t$$

MODELO DINÂMICO NÃO NORMAL E NÃO LINEAR

com hiperparâmetros $\omega = (\alpha, \delta, \sigma_\nu^2)$

Propriedades

$$K = \frac{E[y_t^4]}{E^2[y_t^2]} = 3\exp(\sigma_\Psi^2, \quad \mu_\Phi = \frac{\alpha}{1-\delta} \text{ e } \sigma_\Phi^2 = \frac{\sigma_\nu^2}{1-\delta^2}$$

- y_t^2 é similar a ARMA(1,1) com função de autocorrelação:

$$\rho_{y_t^2}(s) = \frac{\exp(\sigma_\Phi^2) - 1}{3\exp(\sigma_\Phi^2) - 1}\delta^s$$

VEST é similar ao GARCH(1,1)

- y_t é ruído branco se: $\delta + \phi < 1$

2. UMA CLASSE GERAL DE MODELOS DE VOLATILIDADE ESTOCÁSTICA

$$
\begin{gathered}
(y_t|h_t) \sim EF[0,h_t] \\
(h_t|\theta,\sigma_t^2,D_{t-1}) \sim [\mu(\theta,D_{t-1}),\sigma_\nu^2] \\
p(\theta|\sigma_\nu^2) \\
p(\sigma_\nu^2)
\end{gathered}
$$

onde:

- [ ] – é qualquer distribuição sobre R^+, especificada pela média e variância, EF- familia exponencial e
- $D_{t-1} = (y_1,\cdots,y_{t-1},h_1,\cdots,h_{t-1})$

Se desejável θ pode variar no tempo

MODELO HIERÁRQUICO

EXEMPLOS

- **ARCH(1)**

$$\mu(\theta, D_{t-1}) = \alpha + \delta\ y_{t-1}^2$$

$$\sigma_\nu^2 \Longrightarrow 0$$

- **VEST-AR(1)**

$$\mu(\boldsymbol{\theta}, D_{t-1}) = \alpha + \delta\ log(h_{t-1}), \quad \forall \sigma_\nu^2$$

$$(h_t | \boldsymbol{\theta}, h_{t-1}) \sim LN[\mu, \sigma_\nu^2]$$

- **VEST–NORMAL/GAMA AR(1)**

$$(y_t | h_t) \sim N[0, h_t]$$

$$(h_t | h_{t-1}) \sim Ga[\mu(\boldsymbol{\theta}, D_{t-1}),\ \sigma_\nu^2]$$

$$\mu(\boldsymbol{\theta}, D_{t-1}) = \alpha + \delta\ log(h_{t-1}), \quad \forall\ \sigma_\nu^2$$

$$a_t = \frac{(\alpha + \delta\ h_{t-1})^2}{\sigma_\nu^2}$$

$$b_t = \frac{\alpha + \delta\ h_{t-1}}{\sigma_\nu^2}$$

Figura 1:

(a) Série Original (T=250) e Função de autocorrelação (FAC)

(b) Série Transformada (quadrado) e Função de autocorrelação (FAC)

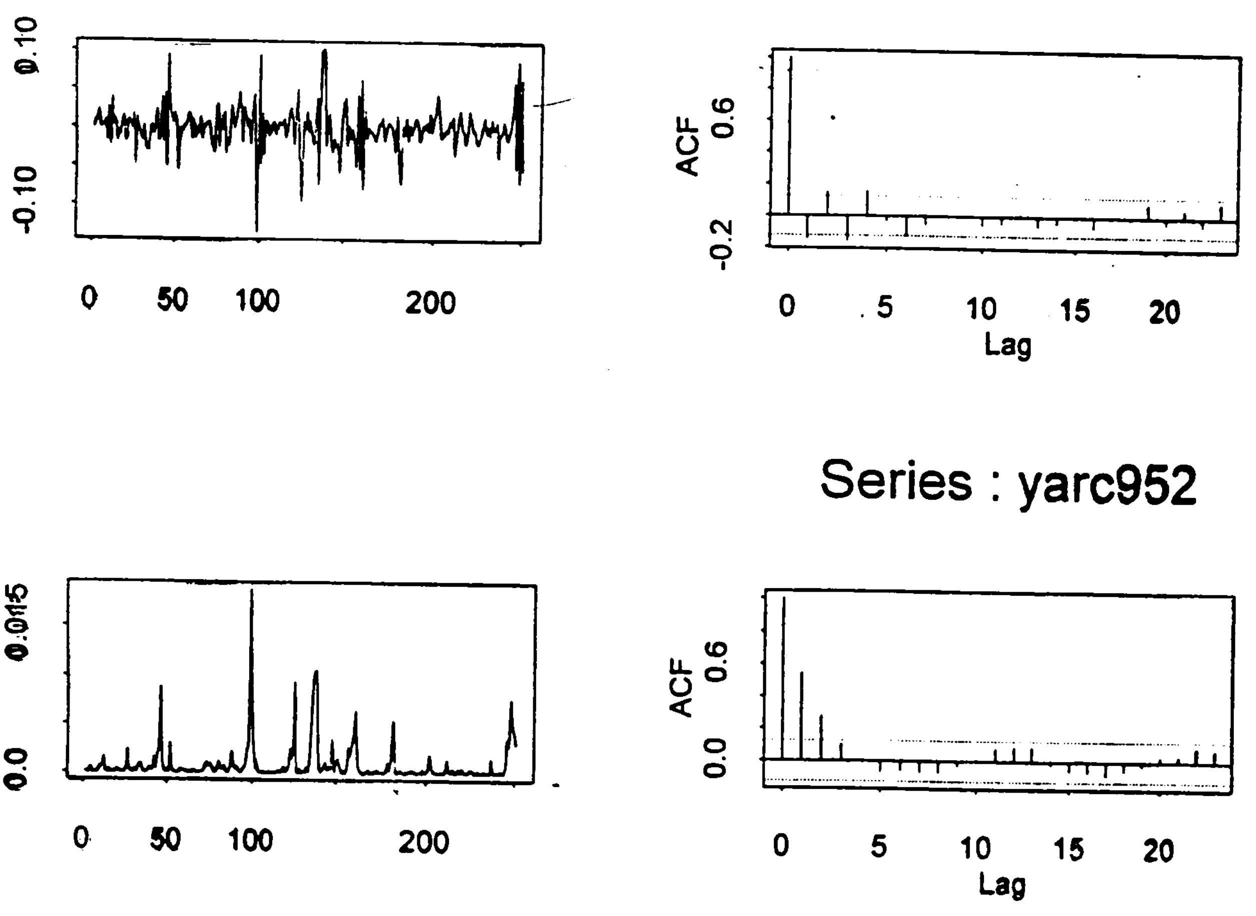
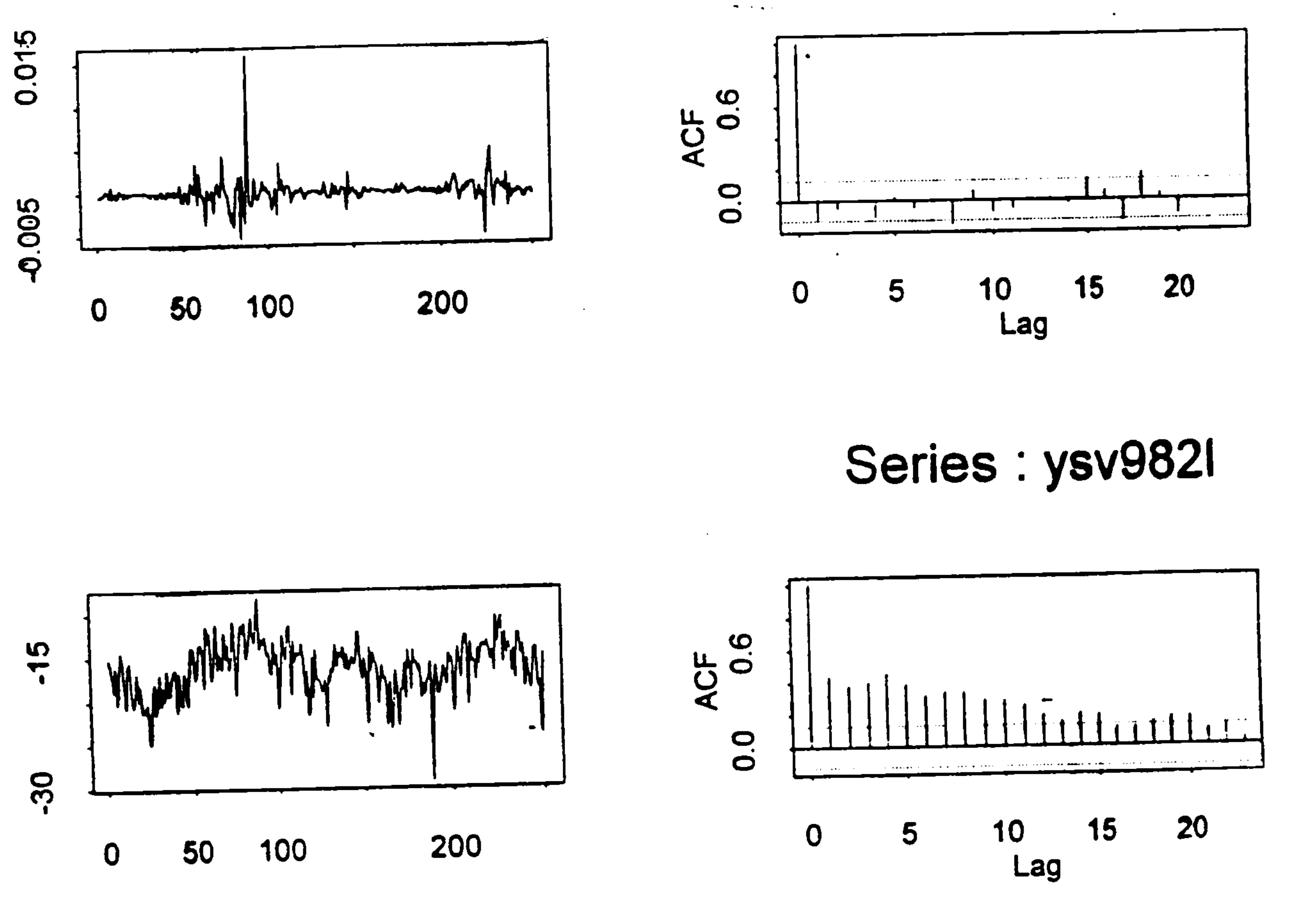


Figura 2:

(a) Série Original (T=250) e Função de autocorrelação (FAC)

(b) Série Transformada (log-quadrado) e Função de autocorrelação (FAC)



1

ABORDAGEM BAYESIANA

SAMPLING IMPORTANCE RESAMPLING

Um algorítimo aplicável tanto em modelos ARCH como VEST é descrito por

i) Gerar N pontos da distribuição a priori

$$\boldsymbol{\theta}_i \sim p(\boldsymbol{\theta}), i = 1, \cdots, N$$

ii) Gerar, no caso VEST, N pontos da distribuição de evolução das volatilidade

$$\boldsymbol{\nu}_i \sim N[0,1], i = 1, \cdots, N$$

iii) Calcular as realizações da volatilidade

$$h_{t,i} = \alpha_i + \delta_i h_{t-1,i} + \sigma_\nu \epsilon_{t,i}$$

iv) Obter os pesos gerados pela verossimilhança

$$w_i = \frac{\exp(w_i^*)}{\Sigma \exp(w_i^*)}, \quad \text{onde}$$

$$w_i^* = -[\Sigma\ log(\sigma_{t,i}^2) + \Sigma \frac{y_t^2}{\sigma_{t,i}^2}]$$

v) Gerar M pontos a partir da distribuição discreta $\{\boldsymbol{\theta}_i, \omega_i\}$

$$\theta_j^* \sim p(\theta_1, \cdots, \theta_N), \quad i = 1, \cdots, N \text{ e } j = 1, \cdots, M$$

AMOSTRADOR DE GIBBS

Modelo de Volatilidade Estocástica

Posteriori Conjunta

$$p(h,\omega|y) \propto p(y|h)p(h|\omega)p(\omega)$$

Condicionais Completas

$$p(\omega|h,y) \Longrightarrow \text{Modelo Linear Bayesinao}$$
$$p(h|\omega,y) \Longrightarrow \text{Fazer Passo a Passo}$$

$$\begin{aligned} p(h_t|h_{t-1},h_{t+1},\omega,y_t) &= p(y_t|h_t)p(h_t|h_{t-1})p(h_{t+1}|h_t) \\ &= h_t^{-1/2}\exp[-y_t^2/2h_t]1/h_t\exp[-(log(h_t)-\mu_t)^2/(2log(\sigma^2)] \end{aligned}$$

$$\begin{aligned} \text{onde :} \quad \mu &= [\alpha(1-\delta)+\delta(log(h_{t+1}+log(h_{t-1})]/(1+\delta^2) \\ \sigma^2 &= \sigma_\omega^2/(1-\delta^2) \end{aligned}$$

2

4.1. DADOS ARTIFICIAIS

Modelo ARCH(1)

- Série gerada com parâmetros $\alpha = 0.00015$, $\delta = 0.95$ e $T = 250$.

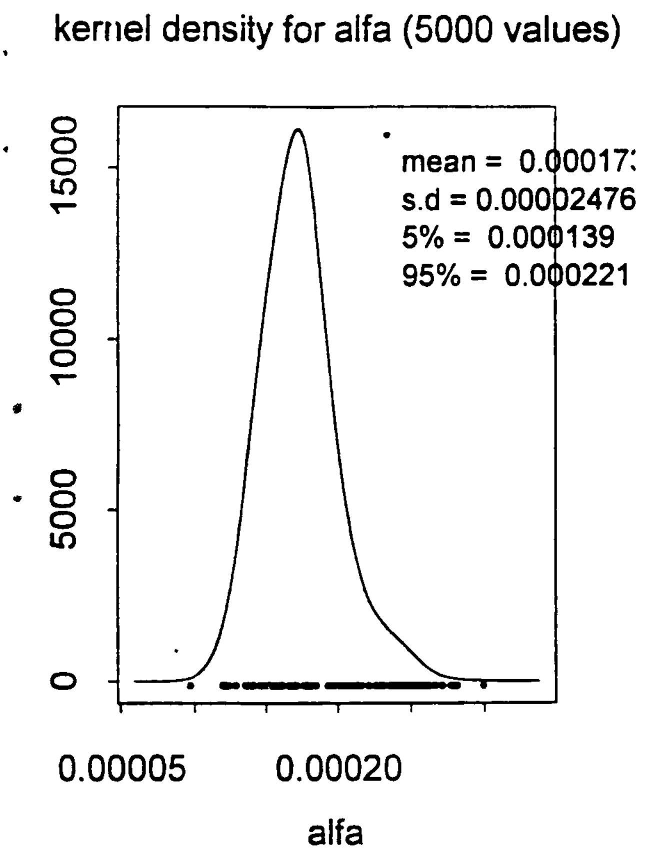
  - **SIR**, com priori não informativa (bem escolhida!), fornece resultados mais razoáveis que o método de máxima verossimilhança, com inicialização $(0.00015, 0.95)$ extremamente precisa.

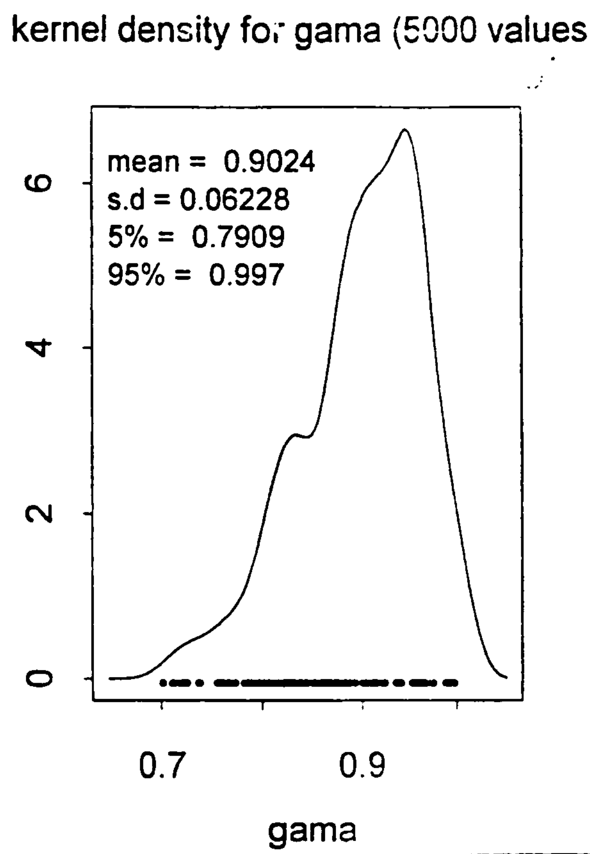
RESULTADOS OBTIDOS

	SIR	MLE
$\hat{\alpha}$	0.000173 (0.000025)	0.00016 (0.0000089)
$\hat{\delta}$	0.902386 (0.062271)	1.0316 (0.16250)

  - Observe que $\delta_{MLE} > 1$, implicando em processo não estacionário, além da convergência ser fraca.

  - Gibbs Sample: o programa utilizado, BUGS, não consegue verificar a log-concavidade, embora esta seja garantida.





Modelo VEST–LOG-AR(1)

- Série gerada com parâmetros $\alpha = -0.245$, $\delta = 0.98$ e $\sigma^2_{nu} = 0.5$, com $h_0 = 1.0$ e $T = 250$.

RESULTADOS OBTIDOS

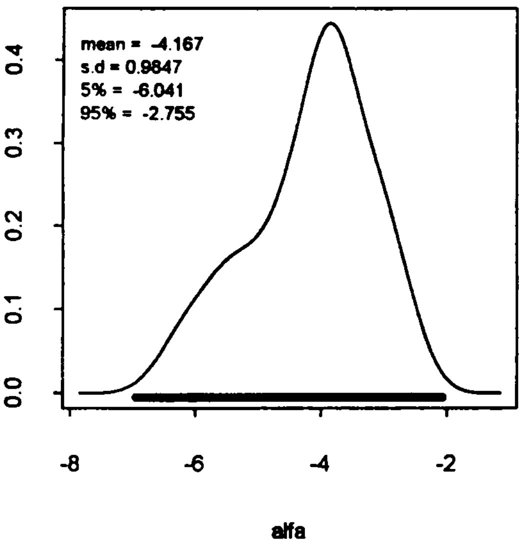
- **SIR**, com priori não informativa (bem escolhida!) e supondo σ^2_ν conhecido
- MCMC via Bugs, com 3000 iterações.

	SIR	MCMC
$\hat{\alpha}$	−0.4723 (0.3381)	−0.4155 (0.1548)
$\hat{\delta}$	0.8851 (0.0565)	0.9042 (0.0364)
$\hat{\sigma}_\nu^{-2}$	2.00 fixo	3.19 (1.1030)

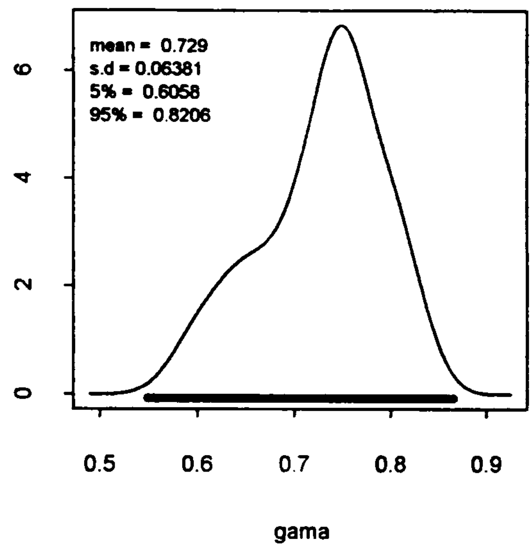
- Esses resultados, embora preliminares são animadores.
- Os métodos tem comportamento similar.
- Os intervalos de probabilidade de 95% praticamente contém o verdadeiro valor dos parâmetros.

Volatilidade Estocastica

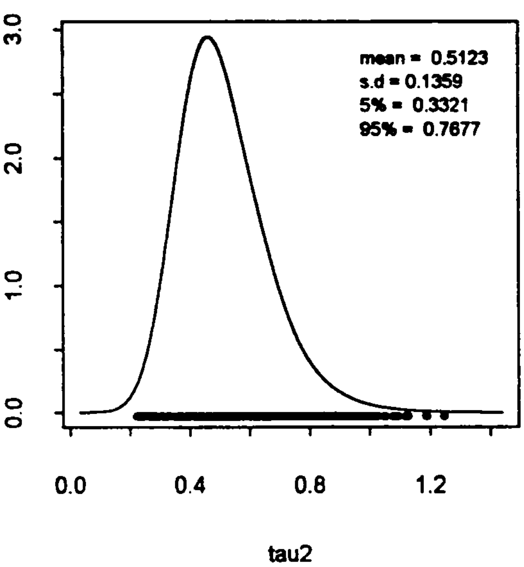
kernel density for alfa (3000 values)



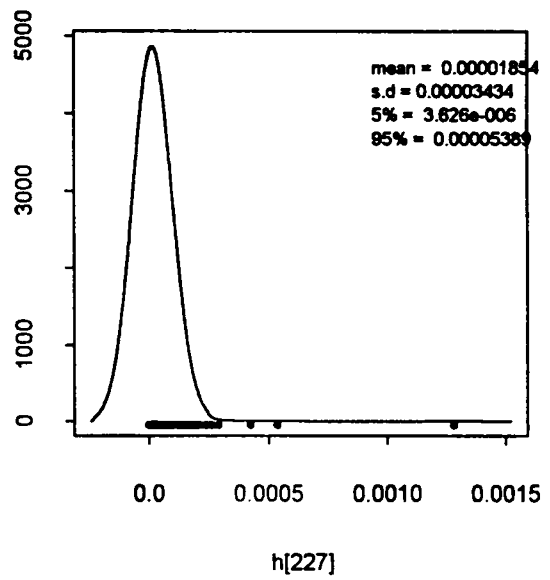
kernel density for gama (3000 values)

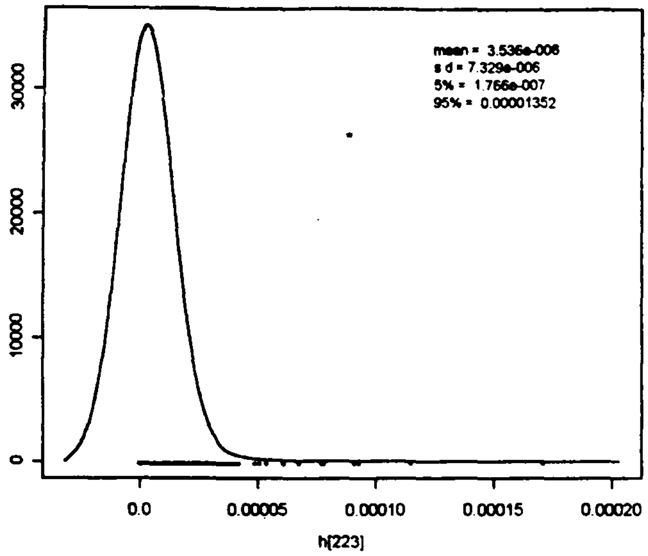


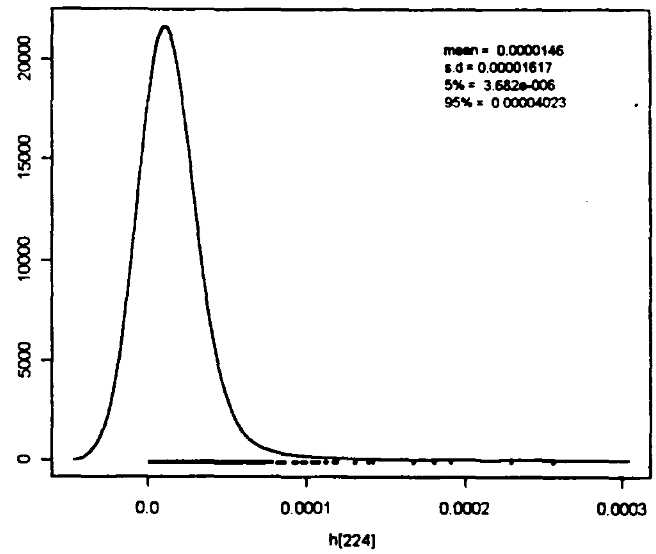
kernel density for tau2 (3000 values)



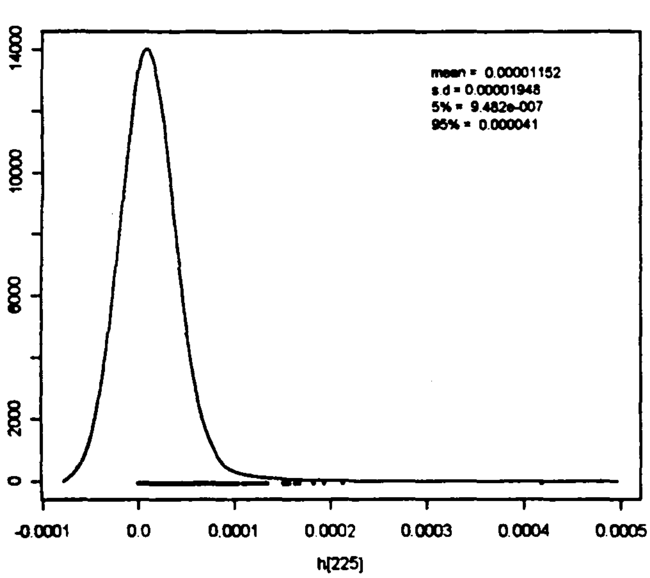
kernel density for h[227] (3000 values)



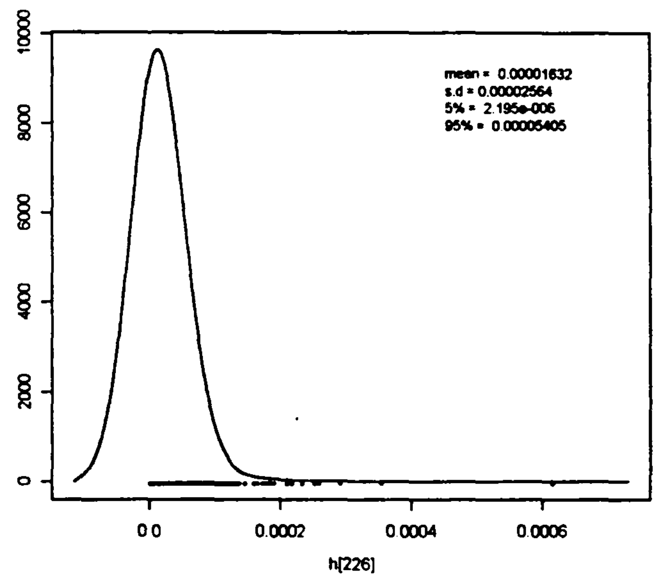




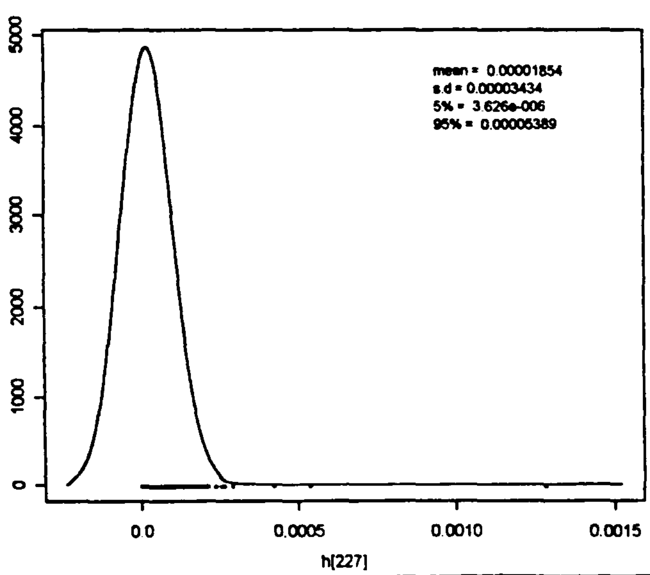
kernel density for h[225] (3000 values)



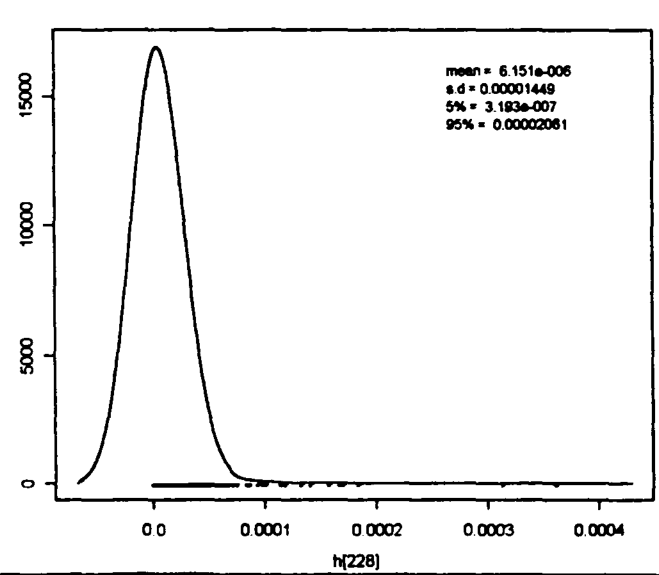
kernel density for h[226] (3000 values)



kernel density for h[227] (3000 values)



kernel density for h[228] (3000 values)



4.2. DADO REAL – TAXA DE JUROS

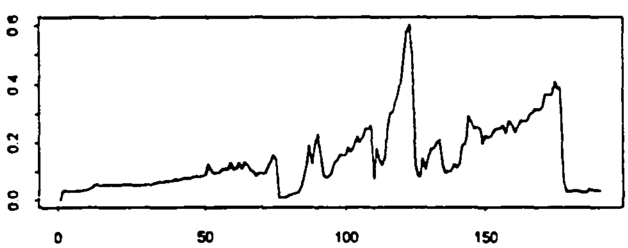
Esse conjunto de dados refere-se a taxa de juros de CDB pré-fixado (tres dias úteis), isto é $log(1+i_t/100)$. Apresentamos abaixo uma análise usando SIR e m modelos ARCH com uma estrutura AR(1) para cuidar da não estacionariedade da série original. Algumas comparações são apresentadas.

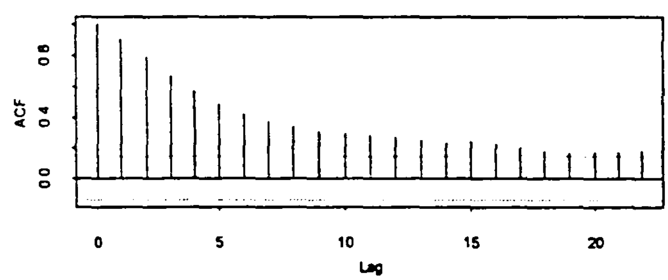
- **MODELO UTILIZADO**

$$y_t = \rho y_{t-1} + h_t^{1/2}\epsilon_t, \quad \epsilon_t \sim N[0,1]$$

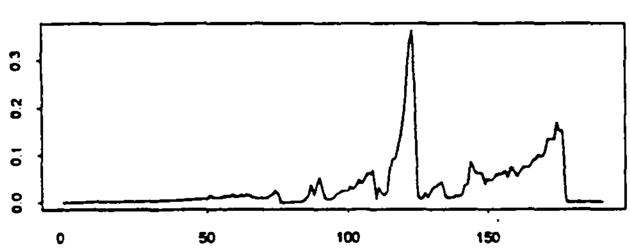
$$h_t = \alpha + \delta y_{t-1}^2, \quad y_0 = 0$$

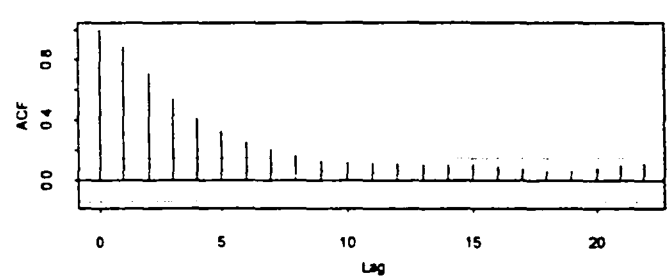
- **SIR**, com priori não informativa (bem escolhida!), fornece resultados mais razoáveis que o método de máxima verossimilhança, com inicialização obtida de uma rodada de SIR que degenera, $(\rho, \alpha, \delta) = (0.736, 0.00156, 0.3473)$





Series : juros2





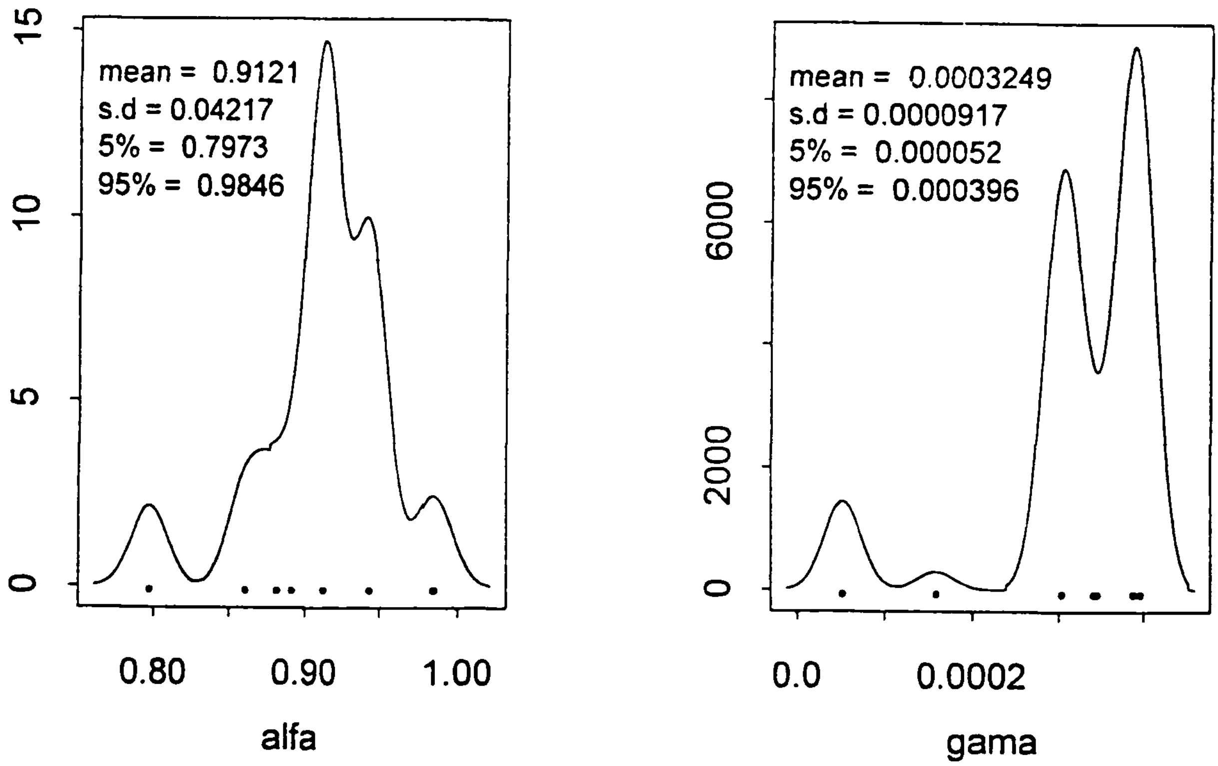
Series : juros2l[2:189]

RESULTADOS OBTIDOS

	SIR	MLE
$\hat{\rho}$	0.9121 (0.042)	0.7636 (0.010)
$\hat{\alpha}$	0.000325 (0.000092)	0.000576 (0.000121)
$\hat{\delta}$	0.84499 (0.0184)	1.40061 (0.2601)

- Observe que $\delta_{MLE} > 1$, implicando em processo não estacionário, embora a convergência seja forte.
- Outras inicializações do método de máxima verossimilhança não convergiram.

- Resultados de VEST foram obtidos no BUGS para um modelo sem considerar a componente de nível e, também, para um modelo com ρ fixo. Esses resultados, embora razoáveis precisam ser melhor analizados.



5. EXTENÇÕES E CONCLUSÕES

- Métodos de Monte Carlo são efetivos.
- Classe ampla de modelos a investigar.
- Implementar técnicas de diagnóstico e seleção de modelos via Fator de Bayes, medidas de não linearidade e influência através de Kullback-Leibner
- Representar modelos ARCH em estrutura linear dinâmica.

6. REFERÊNCIAS BÁSICAS

Jacquier, E, Polson, N.G & Rossi, P.E (1994) - Bayesian analysis of stochastic volatility models, JBES.

Mills, T.C (1993) - The econometric modelling of fiancial time series, CUP.

Muller, P (1991) - A dynamic vector ARCH model for exchange rate data, ISDS Tec.Rep.

Taylor, S (1986) Modelling financial time series.

Shephard, N (1995) - statistical aspects of ARCH and Stochastic Volatility, Disc.Paper 94.