

Modelos GARCH Bayesianos: Métodos Aproximados e Aplicações

Helio S. Migon*

Josmar Mazucheli**

Resumo

Os modelos GARCH são resumidamente revisados e reformulados na classe dos modelos dinâmicos Bayesianos. Procedimentos de quadratura Gaussiana e aproximações de Laplace são utilizados na estimação de modelos GARCH, estáticos e dinâmicos, de baixa dimensionalidade. Os algoritmos desenvolvidos são validados a partir de dados artificialmente gerados. Quatro séries reais de retornos financeiros são analisados. A qualidade de diferentes modelos e métodos de estimação são acessadas através de sua capacidade preditiva e, também, comparando-se as volatilidades previstas, não observáveis, com sua proxie natural, o quadrado dos retornos.

Abstract

The class of GARCH models is briefly revised and those models reformulated as dynamic Bayesian models. Gaussian quadrature technique and Laplace approximation are used to estimate both static and dynamic GARCH models of moderate dimension. The implemented algorithms are validated using artificially generated data. Four real return time series were analyzed and the performance of the different models and estimation methods accessed through their predictive capability and, also, comparing the non-observed volatilities with the square of the returns, its natural proxy.

Palavras-chave: Modelos GARCH dinâmicos; método de quadratura gaussiana aproximações de Laplace; capacidade preditiva.

Códigos JEL: C1; C11 e C22.

*UFRJ.

**DEs/UEM.

1. Introdução

Os modelos GARCH, introduzidos por Bollerslev (1986), vêm sendo exaustivamente utilizados na modelagem da volatilidade de séries financeiras. Uma cuidadosa descrição de suas propriedades teóricas é apresentada em Bollerslev et al (1992).

Neste artigo, nossa contribuição consiste em considerar como os modelos GARCH podem ser formulados na classe dos modelos dinâmicos Bayesianos e apresentar técnicas numéricas para sua inferência. É útil relembrar que modelar a variância observacional é uma prática comum na área de previsões Bayesianas. Se a dimensão do espaço paramétrico é moderada podemos utilizar com vantagens procedimentos de quadratura Gaussiana e aproximações de Laplace. Este é exatamente o caso quando $p, q < 2$. Modelos de séries temporais de componentes não observáveis com erro ARCH (Harvey, Ruiz e Sentana (1992)) são reconsiderados do ponto de vista Bayesiano, neste artigo. Métodos de quadratura gaussiana no contexto de modelos GARCH dinâmicos são discutidos em Valle (1997).

Outras contribuições nesta direção são os trabalhos de Muller e Pole (1994), os quais discutem modelos dinâmicos multivariado, com erros ARCH, através de um algoritmo tipo Metropolis-Hasting. Esta abordagem embora muito geral é, ainda, pouco operacional o que nos motiva a investigar formas aproximadas de análise, como os métodos de quadratura gaussiana e de Laplace.

Os resultados que apresentamos neste artigo fazem parte de um projeto mais abrangente, consistindo em estudar modelos univariados de volatilidade determinística e estocástica no contexto Bayesiano e explorar aproximações numéricas nesses modelos incluindo mudança de regime e o teste de diferentes hipóteses sobre as distribuições observacionais.

Apresentaremos alguns exemplos com dados artificialmente gerados de forma a validar a implementação do algoritmo proposto e, a seguir, analisaremos quatro séries reais. A performance dos diferentes modelos será acessada através de sua capacidade preditiva e, também, comparando-se as volatilidades previstas, não observáveis, com uma proxy natural, o quadrado dos retornos, como proposto em Hamilton e Susmel (1994).

Este artigo está organizado em cinco seções e dois apêndices. Na próxima seção resumimos os principais fatos teóricos associados aos modelos GARCH. Os modelos dinâmicos GARCH são apresentados na seção 3 e, sua inferência, na 4. Finalmente, propomos modelos alternativos e comparamos suas eficiências através de diferentes funções de perda.

2. Modelos GARCH e Propriedades

O modelo GARCH introduzido por Bollerslev (1986), permite que os erros num modelo de séries temporais tenham uma variância que varie no tempo. Especificamente a variância está relacionada aos quadrados dos valores prévios e aos valores defasados da variância. Seja $\{\nu_t\}$ um processo estocástico com média zero e variância $V_t = \alpha_0 + \sum_{i=1}^p \alpha_i \nu_{t-i}^2 + \sum_{j=1}^q \delta_j V_{t-j}$.

Esta é a classe dos modelos GARCH(p,q), com $p, q > 0$ e inteiros, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\delta_j \geq 0$, $i = 1, \dots, p$, e $j = 1, \dots, q$. Quando $q = 0$ temos os modelos ARCH(p). A região de estacionariedade desses modelos é descrita por $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \delta_j < 1$. Vale notar que $\nu_t^2 = \alpha_0 + \sum_{i=1}^m (\alpha_i + \delta_i) \nu_{t-1}^2 + \epsilon_t - \sum_{i=1}^q \delta_i \epsilon_{t-i}$, onde $m = \max(q, p)$, o qual é um processo ARMA(m,q). Em resumo, o quadrado da série dos retornos são distribuídos segundo um processo ARMA. Embora muitos autores adotem a hipótese de normalidade para descrever a distribuição condicional dos retornos, $\nu_t \sim N[0, V_t]$, várias outras distribuições têm sido discutidas na literatura. Bollerslev (1992),

sugeriu a distribuição t-Student, Jorion (1988) utilizou uma mistura normal-Poisson e Hsieh (1989) uma mistura de normal-lognormal.

Sob as condições de estacionariedade mencionadas a variância não condicional num modelo GARCH(p,q) é dada por $var(\nu_t) = \alpha_o / (1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \delta_j)$. Além disto, o coeficiente de curtose para os casos particulares ARCH(1) e GARCH(1,1) são, respectivamente, $E(\nu_t^4) / E^2(\nu_t^2) = 3(1 - \alpha_1) / (1 - 3\alpha_1^2)$ e $6(1 - \alpha_1) / (1 - \delta_1^2 - 2\alpha_1\delta_1 - 3\alpha_1^2)$. Uma medida de persistência da volatilidade, no modelo GARCH(1,1), é dada pela solução da equação de diferenças de primeira ordem $V_{t+m} = \alpha_0 + (\alpha_1 + \delta_1)V_{t+m-1}$ para o horizonte m . O parâmetro de decaimento exponencial, $\lambda = \alpha_1 + \delta_1$, define a medida de persistência.

3. O Modelo Dinâmico GARCH

Um modelo dinâmico GARCH univariado é descrito por:

$$\begin{aligned}
 y_t \mid \theta, \beta_t &\sim N(\mathbf{F}'_t(\theta)\beta_t, V_t(\theta)) \\
 \beta_t \mid \theta, \beta_{t-1} &\sim N(\mathbf{G}_t(\theta)\beta_{t-1}, \mathbf{W}_t(\theta)) \\
 V_t(\theta) &= \alpha_o + \sum_{i=1}^p \alpha_i (y_{t-i} - F'_{t-i}(\theta)\beta_{t-i})^2 + \sum_{j=1}^q \delta_j V_{t-j}(\theta) \quad (1)
 \end{aligned}$$

onde θ é o vetor de hiperparâmetros, β_t são os parâmetros do modelo, F_t é um vetor de regressores e G_t é uma função que introduz a dinâmica dos parâmetros. Assumiremos que y_o, \dots, y_{-p+1} , V_o, \dots, V_{-q+1} , W_t são conhecidos, onde $\alpha_o > 0$, $\alpha_i \geq 0$ e $\delta_j \geq 0$ são os hiperparâmetros do modelo, os quais são desconhecidos e restritos à região de estacionariedade. A dependência dos hiperparâmetros é dada através de $V_t(\theta)$.

Condiciona ao vetor de hiperparâmetros, $\theta' = (\alpha_o, \alpha_1, \dots, \alpha_p, \delta_1, \dots, \delta_q)$ e a V_o, \dots, V_{-q+1} , o modelo proposto nesta seção

é não linear, fazendo com que seu tratamento analítico seja difícil. Para solucionar este problema será necessário fazer aproximações, as quais serão discutidas na próxima seção. Este modelo corresponde à abordagem Bayesiana de um caso especial da proposta de Harvey, Ruiz e Sentana (1992) e é, também, uma extensão usual da prática Bayesiana de se modelar a variância observacional (West e Harrison, 1997).

Suponhamos que a informação disponível no instante t é $D_t = \{y_t, D_{t-1}\}$ e que a distribuição inicial $(\beta_o | \theta, D_o) \sim N(m_o, C_o)$. Assim as distribuições para qualquer tempo $t \geq 1$ são dadas por:

- Priori em t : $\beta_t | \theta, D_{t-1} \sim N[a_{t-1}, R_{t-1}]$
- Preditiva a 1 passo: $y_t | \theta, D_{t-1} \sim N[f_t, Q_t]$
- Posteriori em t : $\beta_t | \theta, D_t \sim N[m_t, C_t]$

onde: $m_t = a_t + A_t(y_t - f_t)$, $A_t = R_t F_t Q_t^{-1}$, $C_t = R_t - A_t A_t' Q_t$, $a_t = G_t m_{t-1}$, $R_t = G_t C_{t-1} G_t' + W_t$, $f_t = F_t' a_t$ e $Q_t = F_t' R_t F_t + V_t$

4. Inferência em MD com Erro GARCH(p,q)

Na análise de modelos dinâmicos não lineares (MDNL) esquemas de integração são necessários para a estimação das características das distribuições a posteriori a cada etapa do processo. Vários artigos apresentam técnicas numéricas para a estimação e previsão em modelos dinâmicos não lineares. Pole (1988) e Pole e West (1988) usam quadratura Gaussiana com grade dinâmica, obtendo as distribuições a cada tempo. Muller (1992), em MDNL, propõe o uso do algoritmo Metropolis-Hasting, para propagar as amostras obtidas via Monte Carlo, simuladas da distribuição a priori em cada etapa do modelo dinâmico. O algoritmo requer a reconstrução das distribuições cor-

respondentes ao período $t - 1$, o que é feito através de uma mistura de processos de Dirichlet.

Neste artigo propomos o uso de quadratura Gaussiana com grade fixa e baseada em toda a informação disponível e, também, o uso da forma exponencial da aproximação de Laplace. Nossa hipótese básica consiste em admitir que os hiperparâmetros da estrutura GARCH sejam invariantes no tempo, o que tem se mostrado válido em alguns exemplos reais como será visto.

É importante notar que, mesmo condicionalmente aos hiperparâmetros, esta classe de modelos é ainda intratável, pois os V_t são funções dos β_{t-i} , $i = 1, \dots, p$. Como os métodos explorados neste artigo perdem eficiência na razão da dimensão do vetor paramétrico, nossas aplicações se restringirão aos casos em que $p, q < 2$. A alternativa em modelos de maior dimensão paramétrica é o uso de técnicas de simulação estocástica.

A inferência em modelos dinâmicos univariados com erros observacionais segundo um processo GARCH(p,q) pode ser condicional ou não ao vetor de hiperparâmetros θ . Mesmo condicionalmente aos hiperparâmetros teremos de fazer uso de linearização para viabilizar as inferências, como mostraremos. A análise condicional não leva em conta a variabilidade derivada da estimativa dos hiperparâmetros.

Para melhor compreender a afirmação acima é útil verificar o que ocorre com a distribuição preditiva, por exemplo. Esta tem forma intratável como pode ser visto, pois depende dos β_{t-i} , $i = 1, \dots, p$.

$$\begin{aligned}
 p(y_t | \theta, D_{t-1}) &\propto \int p(y_t | \beta_{t-1}, \theta) p(\beta_{t-1} | \theta, D_{t-1}) d\beta_{t-1} \\
 &\propto \int (\mathbf{F}'_t \mathbf{W}_t \mathbf{F}_t + \mathbf{V}_t)^{-1/2} \exp[\mathbf{m}'_{t-1} \mathbf{C}_{t-1}^{-1} \mathbf{m}_{t-1}] \\
 &\times \exp\left[-\frac{(y_t - \mathbf{F}'_t \mathbf{G}_t \mathbf{m}_{t-1})^2}{2Q_t}\right] \exp\left[-\frac{U(\beta)}{2}\right] d\beta_{t-1} \quad (2)
 \end{aligned}$$

com $U(\beta) = (\beta_{t-1} - \mathbf{d}_{t-1})' \mathbf{S}_{t-1}^{-1} (\beta_{t-1} - \mathbf{d}_{t-1})$, onde $\mathbf{d}'_{t-1} = [(y_t \mathbf{F}'_t \mathbf{G}_t / K_t) + \mathbf{m}'_{t-1} \mathbf{C}_{t-1}^{-1}] \mathbf{S}_{t-1}$, $\mathbf{S}_{t-1} = [\mathbf{C}_{t-1}^{-1} + (\mathbf{G}'_t \mathbf{F}_t \mathbf{F}'_t \mathbf{G}_t / K_t)]^{-1}$, $K_t = V_t + \mathbf{F}'_t \mathbf{W}_t \mathbf{F}_t$ e $Q_t = V_t + \mathbf{F}'_t \mathbf{R}_t \mathbf{F}_t$. Esta integral não tem uma solução analítica fechada pois V_t é uma função dos β_{t-i} anteriores.

A proposta deste artigo é prosseguir fazendo mais uma aproximação, o que corresponde a substituir β_{t-i} por $E(\beta_{t-i} | D_{t-1}) = \mathbf{m}_{t-i}^{(i-1)}$, $i = 1, \dots, p$, e $\mathbf{m}_{t-1}^{(0)} = \mathbf{m}_{t-1}$, médias das distribuições suavizadas. As variâncias V_t e Q_t serão assim aproximadas por $V_t = \alpha_o + \sum_{i=1}^p \alpha_i (y_{t-i} - \mathbf{F}'_{t-i} \mathbf{G}_{t-i} \mathbf{m}_{t-i}^{(i-1)})^2 + \sum_{j=1}^q \delta_j V_{t-j}$ e $Q_t = \mathbf{F}'_t \mathbf{R}_t \mathbf{F}_t + V_t$. Substituindo estes resultados em (2) tem-se que $y_t | \theta, D_{t-1} \sim N[f_t, \hat{Q}_t]$, onde $f_t = \mathbf{F}'_t \mathbf{a}_t$ e $\hat{Q}_t = \mathbf{F}'_t \mathbf{R}_t \mathbf{F}_t + V_t$.

Esta aproximação deverá ser feita a cada etapa do modelo dinâmico, de forma que as distribuições envolvidas serão “linearizadas”. Supondo θ conhecido as inferências condicionais são imediatas. Para obtermos as inferências marginais em modelos dinâmicos univariado podemos usar quadratura Gaussiana ou o método de Laplace como descritos nos anexos. É conveniente se proceder a reparametrização: $\phi_o = \log(\alpha_o)$, $\phi_i = \log(\alpha_i / (1 - \sum_{k=1}^i \alpha_k))$, $i = 1, \dots, p$, e $\psi_j = \log(\delta_j / (1 - \sum_{k=1}^p \alpha_i - \sum_{l=1}^j \delta_l))$, $j = 1, \dots, q$ antes de realizar as integrais mencionadas.

As principais características das distribuições a posteriori condicionadas podem ser obtidas via teorema de Bayes e as distribuições marginais seguem da integração segundo $p(\theta | D_{t-1})$, onde $\theta' = (\phi_o, \phi_1, \dots, \phi_p, \psi_1, \dots, \psi_q)$. Como não temos solução analítica, propomos uma nova aproximação, isto é.: $p(\theta | D_{t-1})$ e $p(\theta | D_t)$ serão substituídos por $p(\theta | D_T)$. Assim

$$\begin{aligned}
 p(y_t | D_{t-1}) &= \int_{\theta} p(y_t | \theta, D_{t-1}) p(\theta | D_T) d\theta \\
 p(\beta_t | D_t) &= \int_{\theta} p(\beta_t | \theta, D_t) p(\theta | D_T) d\theta.
 \end{aligned}
 \tag{3}$$

Implícita nesta aproximação está a suposição de que θ é invariante no tempo. Assim é usado $p(\theta|D_T)$ em lugar de $p(\theta|D_t)$, $\forall t$. A distribuição de $\theta|D_T$ pode ser obtida pela aproximação de Laplace ou pela quadratura Gaussiana, onde:

$$p(\theta | D_T) \propto p(\theta | D_0) \prod_{t=1}^T p(y_t | \theta, D_{t-1})$$

As distribuições marginais de interesse serão misturas avaliadas na grade disponível, sendo aproximadas por,

$$p(y_t | D_{t-1}) = \sum w_j p(y_t | \theta_j, D_{t-1}) p(\theta_j | D_T)$$

$$p(\beta_t | D_t) = \sum w_j p(\beta_t | \theta_j, D_t) p(\theta_j | D_T), \quad \forall t$$

Os momentos da distribuição a posteriori, bem como os da priori e da preditiva, podem ser obtidos como:

$$E(\beta_{t-1} | D_{t-1}) = \sum_j m_{t-1}(\theta_j) w_j p(\theta_j | D_T)$$

$$E(\beta_{t-1}^2 | D_{t-1}) = \sum_j (m_{t-1}^2(\theta_j) + C_{t-1}(\theta_j)) w_j p(\theta_j | D_T)$$

Na aproximação de Laplace, ao invés de substituímos os β 's pelas médias suavizadas, usamos derivadas analíticas da função log-verossimilhança. Note que esta pode ser descrita pela soma das contribuições de cada tempo através das distribuições preditivas. Assim $L(\theta) = \sum_1^T l_t(\theta)$, onde

$$l_t = -\log(V_t)/2 - (y_t - \mu_t)^2 / (2V_t)$$

Como descrito no apêndice, a solução via método de Laplace envolverá as derivadas analíticas de l_t (Fiorentini e Calzolari, 1996) dadas por:

$$\begin{aligned} \frac{\partial l_t}{\partial \theta} &= -\frac{1}{2} \frac{1}{V_t} \frac{\partial V_t}{\partial \theta} + \frac{1}{2} \frac{\partial V_t}{\partial \theta} \frac{\epsilon_t^2}{h_t^2} = \frac{1}{2} \frac{1}{V_t} \frac{\partial V_t}{\partial \theta} \left(\frac{\epsilon_t^2}{V_t} - 1 \right) \\ \frac{\partial^2 l_t}{\partial \theta \partial \theta'} &= \left(\frac{\epsilon_t^2}{V_t} - 1 \right) \frac{\partial}{\partial \theta'} \left(\frac{1}{2} \frac{1}{V_t} \frac{\partial V_t}{\partial \theta} \right) - \frac{1}{2} \frac{1}{h_t^2} \frac{\partial V_t}{\partial \theta} \frac{\partial V_t}{\partial \theta'} \frac{\epsilon_t^2}{h_t^2} \end{aligned} \quad (4)$$

onde: $\epsilon_t = (y_t - \mu_t)$.

5. Aplicações

Iniciaremos apresentando alguns resultados obtidos com séries artificialmente geradas, com o objetivo de validar os algoritmos desenvolvidos. O primeiro conjunto de dados foi gerado através de modelos dinâmico de primeira ordem com erros GARCH e a estimação feita via quadratura Gaussiana como descrito no anexo 1. O segundo exemplo consistirá na aplicação do método de Laplace a dados artificiais de um modelo GARCH estático. Este conjunto de dados será analisado também via MCMC utilizando o algoritmo de Metropolis implementado no software BUGS, o qual está relacionado com o artigo de Bawens e Lubrano (1996).

Na segunda parte apresentamos os resultados obtidos para quatro séries reais do mercado brasileiro. Estas foram analisadas por modelos GARCH estáticos e dinâmicos e também por modelos dinâmicos de primeira ordem com estimação de variância. Todas as análises são realizadas sob a suposição de normalidade.

5.1. Dados artificialmente gerados

Este primeiro exemplo corresponde a um MD de primeira ordem com erros GARCH(1,1), o qual foi cuidadosamente analisado em

Valle (1997). Um conjunto de duzentas observações foi gerado pelo MD de primeira ordem cujo mecanismo de geração é dado por:

$$\begin{aligned}
 y_t &= \mu_t + \nu_t & \nu_t &\sim N(0, V_t) \\
 \mu_t &= \mu_{t-1} + \omega_t & \omega_t &\sim N(0, W_t), & W_t &\ll V_t \\
 V_t &= 0.2 + 0.2(y_{t-1} - \mu_{t-1})^2 + 0.6V_{t-1} & & t = 1, \dots, 200
 \end{aligned}$$

Podemos observar, na figura abaixo, o gráfico dos dados gerados, destacando-se sua alta volatilidade. Examinando-se as funções de autocorrelação e de autocorrelações parciais para a série dos quadrados ν_t^2 , onde $\nu_t = y_t - \mu_t$, confirmou-se que estes se comportam como um ARMA(1,1) não Gaussiano, sugerindo que o processo gerador dos dados é aproximadamente um GARCH(1,1).

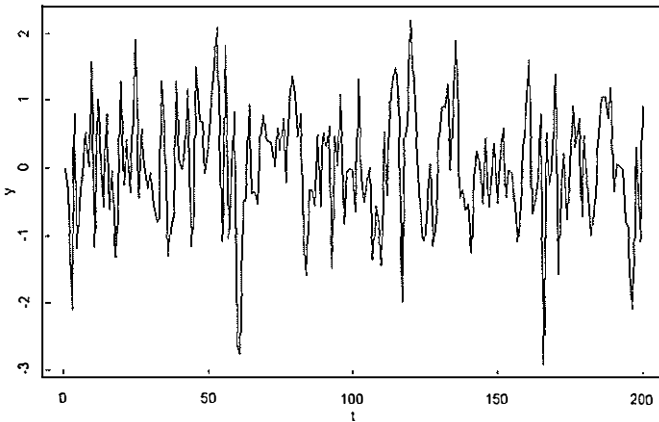


Figura 1

MD de primeira ordem com erros GARCH(1,1)
 $\alpha_0 = .2, \alpha_1 = .2, \delta_1 = .6, Y_0 = 0, \mu_0 = 0$ e $V_0 = 1$

Neste exemplo assumiremos também que os valores de $y_0 = 0$, $\mu_0 = 0$ e $V_0 = 0.1$ são conhecidos e todas as inferências serão feitas

condicionadas a estes valores. Para a aplicação do nosso algoritmo as transformações sugeridas serão usadas e as distribuições a priori para os hiperparâmetros convenientemente escolhidas. Um fator de desconto igual a .99 em correspondência com a razão de variâncias usadas na geração dos dados será utilizada.

Com estes valores iniciais o algoritmo de quadratura gaussiana foi executado, usando a princípio uma grade de três pontos para cada hiperparâmetro e fazendo só três iterações com a finalidade de reposicioná-la. A seguir, uma grade de 7 pontos é aplicada e iterada até a convergência. O tempo gasto neste exemplo foi de aproximadamente 2 minutos usando-se um programa desenvolvido em OX (Doornik, 1998). As estimativas $\hat{\alpha}_0 = E(\alpha_0 | D_T)$, $\hat{\alpha}_1 = E(\alpha_1 | D_T)$ e $\hat{\delta}_1 = E(\delta_1 | D_T)$ foram calculadas como funções dos parâmetros $\theta' = (\phi_0, \phi_1, \psi_1)$. Na tabela abaixo os resultados comparativos para dois valores do fator de desconto são apresentados.

Tabela 1
Resultados Obtidos para o MD com
Erros Observacionais GARCH(1,1)

Parâmetro	Valor Gerado	Estimativa ($\rho = 0.99$)	Estimativa ($\rho = 0.98$)
α_0	0.2	0.145(0.08)	0.148(0.07)
α_1	0.2	0.393(0.14)	0.402(0.14)
δ_1	0.6	0.515(0.43)	0.508(0.43)

Os desvios padrões da distribuição a posteriori são apresentados entre parênteses. Salienta-se que as estimativas de forma considerável variando-se o fator de desconto, embora ocorra ligeiro aumento nos desvios padrões. Isto se deve ao aumento da incerteza, pois a perda de informação muda de quase 1% para 2%.

No próximo exemplo um conjunto de 500 observações de um GARCH(1,1) foi gerado com hiperparâmetros $\theta = (0.2, 0.2, 0.6)'$ e esperança condicional nula para todo o tempo. A tabela abaixo mostra os resultados da estimação via máxima verossimilhança e pelo método de Laplace. Na última linha da tabela são apresentados resultados de estimação via MCMC. Esses resultados são indistinguíveis do ponto de vista prático. Os resultados descritos foram obtidos a partir de rotinas desenvolvidos em SAS/IML, as quais podem ser obtidas com o segundo autor. As modas e as matrizes Hensianas, quando necessário, foram calculadas numericamente. As soluções de MCMC foram obtidas através do algoritmo de Metropolis implementado no *software* BUGS - *Bayesian Analysis Using Gibbs Sampler*.

Tabela 2
Modelos GARCH(1,1) Estático

Método	α_0	α_1	δ	Correlação	
MLE	0.255 (0.076)	0.337 (0.076)	0.456 (0.102)	0.368	-0.865 -0.702
Laplace	0.302 (0.084)	0.352 (0.074)	0.417 (0.091)	0.187	-0.857 -0.584
MCMC	0.294 (0.072)	0.332 (0.073)	0.432 (0.088)		
REAL	0.2	0.2	0.6		

Na próxima figura observamos a série gerada artificialmente. As densidades a posteriori para os parâmetros da estrutura GARCH são ligeiramente assimétricas e concentram massa de probabilidade na região de inversibilidade como era de se esperar. Na figura 3, as estimativas de volatilidade baseadas na máxima verossimilhança e no método de Laplace são comparadas com a volatilidade artificialmente gerada. As três séries são indistinguíveis evidenciando a boa qualidade preditiva.

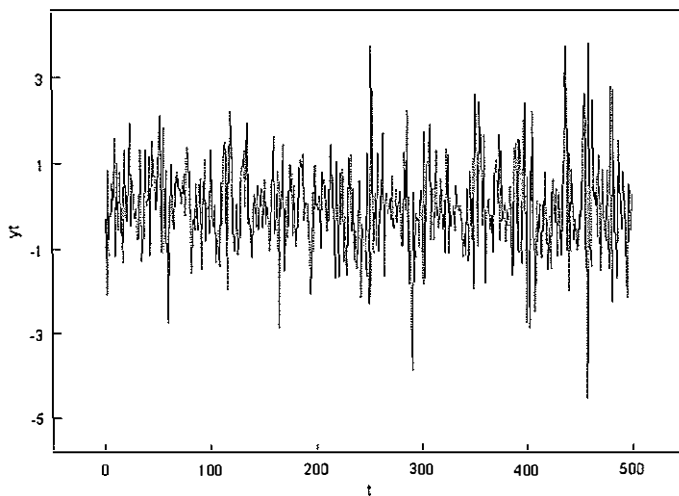


Figura 2
Série Gerada Artificialmente – GARCH(1,1) – Estático

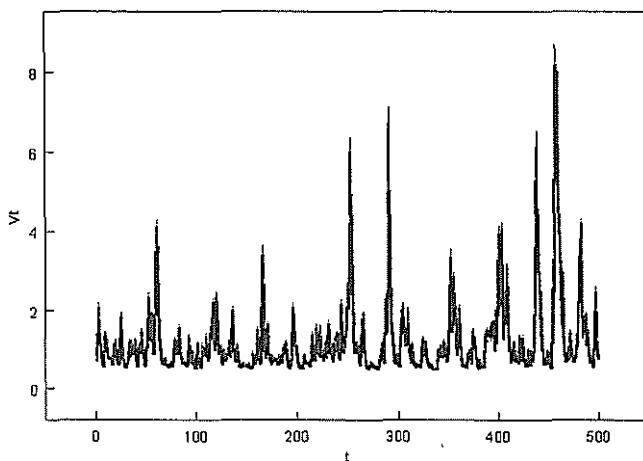


Figura 3
Volatilidade Estimada: MLE e Laplace e Original

5.2. Dados reais

Os dados analisados abaixo correspondem a cotações de diferentes ativos em períodos variados. Os retornos foram calculados através da expressão: $y_t = \log(p_t/p_{t-1})$ onde p_t é o preço observado na data t . Além das estimativas de máxima verossimilhança usando a suposição de normalidade para as quantidades observadas obtivemos também as aproximações de Laplace. Questões de ordem numérica impuseram que as séries dos retornos fossem normalizadas antes de se proceder a análise. Nos resultados apresentados a seguir além das estimativas pontuais temos os erros padrões associados. No caso do uso do método de Laplace exponencial as prioris utilizadas foram vagas e nos resultados apresentamos também as distribuições marginais de alguns dos parâmetros de interesse.

Inicialmente, são apresentadas algumas estatísticas descritivas das séries em estudo. Observa-se que o efeito autoregressivo de primeira ordem é desprezível. As quatro séries de retornos apresentam, como era de se esperar, alta curtose e assimetria moderada, exceto no caso do Câmbio com alta assimetria negativa. O período observacional iniciou-se em datas diferentes, porém todas foram observadas até 01/07/98.

Tabela 3
Descrição Preliminar dos Dados

	Câmbio	Telebras	Cbond	Cacau
Período	30/06/94	30/06/94	15/04/94	01/01/90
Número Observação	996	988	1066	2118
Média	1.46E-04	1.28E-03	6.35E-04	3.30E-04
Variância	1.26E-05	1.09E-03	3.64E-04	278E-04
Assimetria	-5.39	0.66	-0.54	0.39
Curtose	99.66	11.94	14.03	6.07
ϕ	0.13	0.005	0.00	-0.05

Os gráficos desses retornos são apresentados a seguir. As séries correspondentes aos retornos da *Telebras* e do *Câmbio* apresentam mudanças de regime. Embora as persistências possam mudar nesses momentos, nossa hipótese de trabalho é de que se mantenham estáveis, a menos de efeitos trazidos por eventuais mudanças no nível local dos retornos. Estas mudanças serão captadas através do monitoramento e intervenções Bayesianas como mostraremos mais adiante.

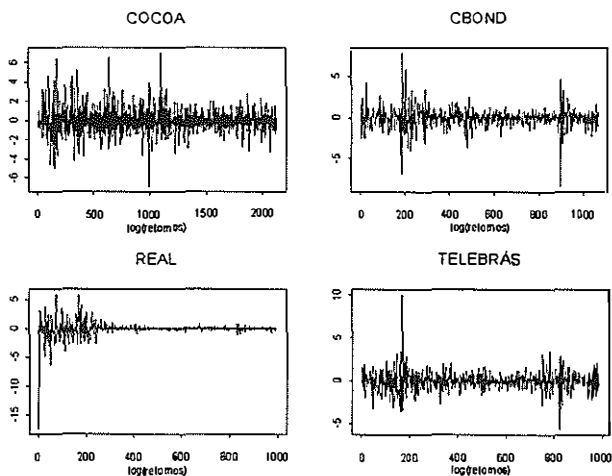


Figura 4
Taxa de Retorno do Vários Ativos

As estimativas pontuais dos parâmetros do GARCH(1,1), via método de Laplace e máxima verossimilhança, podem ser observadas na tabela abaixo. Vale observar que os resultados obtidos são virtualmente os mesmos uma vez que as prioris utilizadas no método de Laplace foram não informativas. É prudente mencionar que essas estimativas foram feitas de forma irrestrita, ou seja as condições de estacionariedade não foram impostas a priori. Além disto relembramos que todas as séries foram padronizadas para evitar dificuldades numéricas.

Tabela 4
Estimação Pontual: Laplace e Verossimilhança

Método	Câmbio			Telebras		
	α_0	α_1	δ	α_0	α_1	δ
MLE	.001 (.0001)	.091 (.017)	.895 (.014)	.032 (.010)	.180 (.029)	.791 (.032)
Laplace	.001 (.0001)	.099 (.018)	.889 (0.18)	.039 (.012)	.194 (.031)	.773 (.035)

Tabela 5
Estimação Pontual: Laplace e Verossimilhança

Método	CBOND			Cacau		
	α_0	α_1	δ	α_0	α_1	δ
MLE	.026 (.007)	.187 (.027)	.802 (.026)	.004 (.002)	.029 (.005)	.967 (.006)
Laplace	.030 (.009)	.199 (.029)	.789 (.028)	.006 (.003)	.032 (.006)	.963 (.007)

A persistência esperada para a volatilidade do Câmbio e Cbond são aproximadamente .98, enquanto que a da Telebras é .99 e no Cacau, .995. Nos gráficos a seguir são apresentadas as distribuições marginais a posteriori dos parâmetros. Todas as inferências serão baseadas nestas distribuições. Observa-se que estas são ligeiramente assimétricas e com massa de probabilidade na região de inversibilidade dos parâmetros, embora nenhuma restrição a priori tenha sido imposta. Infelizmente a metodologia de Laplace não permite obter a distribuição a posteriori de funções dos parâmetros, impedindo discutir probabilisticamente a medida de persistência.

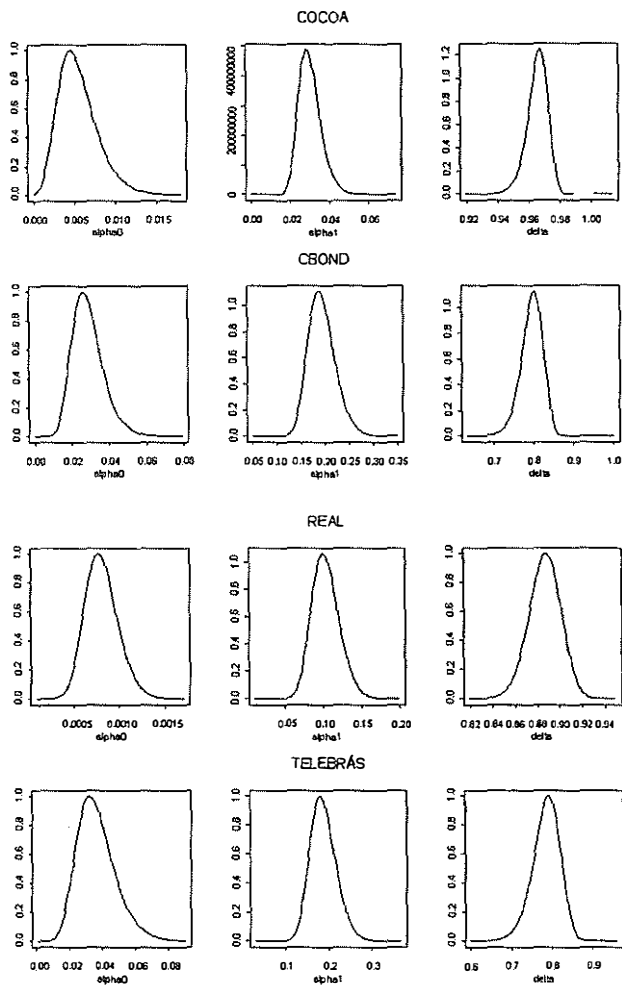


Figura 5
Distribuição a Posteriori Marginal: α_0 , α_1 e δ

Com o objetivo de comparar a performance dos modelos GARCH estimados via método de Laplace iremos utilizar as funções de perda quadrática, valor absoluto e logarítmica no módulo. Como as volatilidades são não observáveis deveremos compará-la com alguma

quantidade convenientemente escolhida. Isto será feito usando-se o quadrado dos retornos observados uma vez que estamos trabalhando com variáveis centradas na média. Assim as perdas esperadas serão dadas por:

$$EMQ = \frac{\sum_1^T (y_t^2 - \hat{V}_t)^2}{T}$$

$$EMA = \frac{\sum_1^T |y_t^2 - \hat{V}_t|}{T}$$

$$|EML| = \frac{\sum_1^T (|\log(y_t^2 / \hat{V}_t^2)|)}{T}$$

Como medida de referência avaliaremos as perdas esperadas sob a hipótese de que as variâncias condicionais são indistinguíveis. Assim, a variância não condicional, $s^2(y)$, será utilizada.

Os modelos GARCH estimados pelo método de Laplace serão comparados com um modelo dinâmico de primeira ordem com a variância estimada sequencialmente usando-se um fator de desconto. Isto corresponde a descrever-se a distribuição a posteriori da volatilidade através de, por exemplo, uma distribuição gama - $Ga(n_t, d_t)$. As seguintes equações de recorrência são facilmente obtidas

$$n_t = \delta n_{t-1} + 1 \tag{5}$$

$$d_t = \delta d_{t-1} + \frac{e_t^2}{q_t}, \tag{6}$$

onde $e_t = y_t - f_t$, $f_t = E[y_t | D_{t-1}]$ e $q_t = var[y_t | D_{t-1}]$.

A priori é elaborada de tal forma que mantem-se a média da gama e incrementa-se somente sua variância através do fator δ^{-1} . Assim este modelo aparentemente simples introduz uma dinâmica especial na volatilidade, a qual é de natureza estocástica.

O monitoramento dos erros de previsão permite acompanhar a série observada identificando, de forma cumulativa, mudanças de natureza estrutural. O monitor é baseado no fator de Bayes cumulativo

$$H_t(k) = \prod_{i=0}^{k-1} H_{t-i} = \frac{p(y_t, \dots, y_{t-k+1}) | D_{t-k}, M}{p(y_t, \dots, y_{t-k+1}) | D_{t-k}, M_A}$$

onde M_A é um modelo alternativo. A vigilância do monitor é de caráter local e baseia-se nas evidências mais recentes. Como descrito em Pole, West e Harrison (1994), o fator de Bayes local fica descrito por:

$$L_t = H_t \min(1, L_{t-1})$$

$$l_t = 1 + l_{t-1} I_{(-\infty, 1)}(L_{t-1})$$

Quando o monitor sinaliza uma ruptura estrutural, l_t indica a quantos períodos esta principiou. O monitor depende de um limite de evidência para o fator de Bayes - τ - e de um limitante para o número de períodos de acumulação - l_0 . Uma vez identificada uma mudança estrutural procede-se a uma intervenção. Esta, em sua forma mais simples, consiste em utilizar localmente fatores de descontos bem pequenos, digamos, dez vezes menores que os descontos de rotina. Este procedimento permite identificar continuamente mudanças de regimes em séries temporais.

Na literatura de modelos dinâmicos existe comprovação suficiente de que esta prática pode substituir satisfatoriamente as complexas modelagens de regimes múltiplos, usualmente denominada de multi-processo tipo 2.

Na tabela a seguir apresentamos os resultados obtidos nessas comparações.

Tabela 6

Comparação de Previsões a um Período: Volatilidade e Retorno

Modelo	\hat{V}_t				\hat{y}_t	
	EMQ	EMA	EML	LVer	EMQ	EMA
TELEBRAS						
Referência	15.93	1.29	2.52	-	-	-
Estático	16.17	1.46	2.64	-4931.45	1.005	0.639
Dinâmico-s/	14.54	1.12	2.06	-1889.15	1.018	0.647
Dinâmico-c/	17.69	1.15	2.05	-1889.46	1.069	0.662
Garch-La s/	20.69	1.30	2.06	-645.65	1.036	0.661
Garch-La c/	19.70	1.30	2.06	-670.76	1.059	0.680
CÂMBIO						
Referência	8.15	1.33	4.45	-	-	-
Estático	25.04	2.98	4.98	-4569.9	0.926	0.433
Dinâmico-s/	14.57	1.08	2.07	-841.62	0.900	0.421
Dinâmico-c/	14.49	1.08	2.07	-844.59	0.902	0.423
Garch-La s/	64.88	2.48	1.73	-271.3	1.870	0.846
Garch-La c/	48.59	1.89	1.72	-298.5	1.760	0.679
CBOND						
Referência	15.98	1.29	2.51	-	-	-
Estático	16.22	1.46	2.63	-4931.90	1.006	0.639
Dinâmico-s/	14.59	1.12	2.00	-1889.47	1.018	0.648
Dinâmico-c/	17.68	1.15	1.99	-1894.71	1.050	0.660
Garch-La s/	21.24	1.32	1.98	- 733.81	1.039	0.664
Garch-La c/	19.02	1.27	2.51	- 734.11	1.086	0.702
COCOA-final(1399)						
Referência	9.41	1.16	2.17	-	-	-
Estático	17.70	1.33	1.74	-6919.19	1.202	0.872
Dinâmico-s/	16.29	1.27	1.66	-2931.84	1.195	0.867
Dinâmico-c/	16.29	1.27	1.66	-2933.80	1.195	0.867
Garch-La s/	32.44	2.08	1.91	-5153.50	1.436	1.010
Garch-La c/	25.11	1.74	1.80	-5429.22	1.443	1.000

Como observa-se no quadro acima é difícil em várias das aplicações sustentar a hipótese de modelo GARCH. Os modelos dinâmicos produzem previsões da volatilidade superiores às das do GARCH. Entretanto, há evidências de que a estrutura GARCH produz melhores previsões dos retornos. O monitoramento e as intervenções subjetivas produzem informações qualitativas importantes, mas não necessariamente melhoram a capacidade preditiva dos modelos.

Para exemplificar as saídas típicas da análise dinâmica apresentamos o gráfico das volatilidades previstas, segundo os modelos dinâmico e o GARCH(1,1), do quadrado dos retornos observados e dos momentos de intervenção associados a análise dinâmica da série dos retornos da Telebrás.

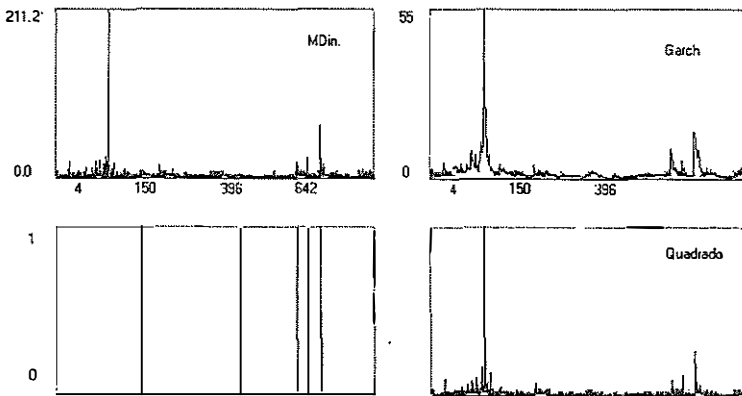


Figura 6
Volatilidade via MD – c/int. e GARCH-La

6. Conclusões e Extensões

Neste artigo uma ampla classe de modelos de volatilidade com estrutura dinâmica para a esperança condicional foi introduzida in-

cluindo aspectos da inferência, particularmente para $p, q < 2$. A proposta básica consistiu no uso de quadratura Gaussiana associado com linearização, sendo que uma alternativa promissora é o método de Laplace, o qual não utiliza a etapa de linearização.

Algumas extensões do presente trabalho envolvem: a implementação do monitoramento Bayesiano em MD univariados com erros GARCH(p,q), a comparação dos resultados obtidos com os modelos GARCH com mudança de regime (Duerker, 1997), a aplicação de Laplace em modelos dinâmicos.

Outra vertente deste trabalho pode ser a robustificação do modelo através do uso de uma t-Student com l graus de liberdade para descrever as observações do modelo dinâmico de primeira ordem com erros GARCH(p,q). No caso especial de MD de primeira ordem pode ser exemplificada substituindo nas equações da evolução do MD as funções g, G isto é:

$$m_t = a_t + C_t g(Y_t - f_t)$$

$$C_t = [R_{t-1}^{-1} + G^{-1}(y_t - f_t)]^{-1}$$

onde:

$$g(Y_t - f_t) = -\frac{\partial}{\partial(y_t - f_t)} \log p(y_t - f_t), \text{ e}$$

$$G(y_t - f_t) = -\frac{\partial^2}{\partial(y_t - f_t)^2} \log p(y_t - f_t).$$

Os modelos com mudança de regime podem ser atacados através de MCMC. Seja y_t a variável observada e s_t quantidades desconhecidas tomando valores em $1, 2, \dots, K$. Supondo que s_t pode ser descrita por uma cadeia de Markov, onde

$$Prob(s_t = j \mid s_{t-1} = i, s_{t-2} = k, \dots, y_{t-1}, y_{t-2}, \dots)$$

$$= Prob(s_t = j \mid s_{t-1} = i) = p_{ij}$$

para $i, j = 1, 2, \dots, K$. A variável s_t representa um *estado* ou *regime* no tempo t . assumindo que s_t governa os parâmetros da distribuição condicional de y_t , o objetivo é selecionar uma representação parsimoniosa para os possíveis regimes, isto é: $y_t = \sqrt{g_{s_t}} \times \tilde{u}_t$, onde $\tilde{u}_t = h_t \nu_t$ e $\nu_t \sim [0, 1]$, e a variância $h_t = a_0 + a_1 \tilde{u}_t^2 + \dots + a_q \tilde{u}_t^{2q} + \xi d_{t-1} \tilde{u}_{t-1}^2$ onde $d_{t-1} = 1$ se $\tilde{u}_{t-1} \leq 0$ e $d_{t-1} = 0$ se $\tilde{u}_{t-1} > 0$.

Submetido em março de 1999 e revisado em maio de 1999.

Referências Bibliográficas

- Bauwens, L. & Lubrano, M. 1996. "Bayesian inference on GARCH models using the Gibbs sampler". Discussion paper 96- 27, CORE, Catholic University of Louvain.
- Bollerslev, T. 1986. "Generalized autoregressive conditional heteroskedasticity". *Journal of Econometrics*, 31:307–327.
- Bollerslev, T.; Chou, R.Y. & Kroner, K.F. 1992. "Arch modelling in finance: a review of the theory and empirical evidence". *Journal of Econometrics*, 52:5–59.
- Doornik, J.A. 1998. "Object-Oriented Matrix Programming Using OX 2.0". Timberlake Consultants Ltd and Oxford.
- Duerker, M.J. 1997. "Markov switching in GARCH process and mean-reverting Stock-market volatility". *Journal of Business and Economic Statistics*, 15(1):26–34.
- Fiorentini, G. & Calzolari, G. 1996. "Analytic derivatives and the computation of GARCH estimates". *Journal of Applied Econometrics*, 11:399–417.
- Hamilton, D.J. & Susmel, R. 1994. "Autoregressive conditional heteroskedasticity and changes in regime". *Journal of Econometrics*, 64:307–333.

- Hisieh, D.A. 1989. "Modeling heteroskedacity in daily foreign exchange rates". *Journal of Business and Economic Statistics*, **7**:307–317.
- Jorion, P. 1988. "On jump processes in the foreign exchange and Stock markets". *Review of Financial Studies*, **1**:427–445.
- Muller, P. 1992. "Posterior intergration in dynamic models". *Computing Science and Statistics*, **21**.
- Muller, P. & Pole, A. 1994. "Monte Carlo posterior integration in GARCH models". Discussion paper 93-28, ISDS, Duke University.
- Naylor, J.C. & Smith, A.F.M. 1982. "Applications of a method for the efficient computation of posterior distributions". *Applied Statistics*, **31**:214–225.
- Pole, A. 1988a. "Transfer response models: a numerical approach". In: J.M. Bernardo, M.H. DeGroot, D.V. Lindley & A.F.M. Smith (eds.). *Bayesian Statistics*, **3**. Oxford University Press.
- Pole, A. & West, M. 1988. "Efficient numeric integration in dynamic models". Warwik Reseach Report 154, Departament of Statistics, University of Warwik.
- Pole, A. & West, M. 1990. "Efficient Bayesian learning in non-linear dynamic models". *Journal of Forecasting*, **9**:119–136.
- Quintana, J.M. & West, M. 1987. "Multivariate time series analysis: new techniques applied to international exchange rate data". *The Statistician*, **36**:275–281.
- SAS/IML Software: Usage and References Version 6, 1st Edition Cary NC: SAS Inc., 1989.

- Smith, A.F.M.; Skene, A.M.; Shaw, J.E.H.; Naylor, J.C. & Dransfield, M. 1985. "The implementation of the Bayesian paradigm". *Communication in Statistics*, 14:1079–1102.
- Valle, C.A.A. 1997. "Modelos Dinâmicos Bayesianos Univariados com Erros Observacionais segundo um Processo GARCH". Tese de Mestrado, UFRJ.
- West, M. & Harrison, P.J. 1997. "Bayesian Forecasting and Dynamic Models". 2nd edition. Springer-Verlag: New York.

Apêndices

O Método de Quadratura

Seja θ um parâmetro de interesse com distribuição a posterior $p(\theta | D_t)$. Nosso interesse será em integrais da forma,

$$\int g(\theta)p(\theta | D_t)d\theta_T$$

onde $D_t = (y_1, \dots, y_t)$ e o subíndice indica as componentes a serem integradas.

Smith et al (1985) e Naylor e Smith (1982) discutem experiências com métodos de integração numérica de parâmetros de distúrbio. A aproximação usa as regras da quadratura gaussiana para vetores multivariados, supondo que a densidade posterior possa ser bem aproximada pelo produto de uma distribuição multivariada e uma polinomial em $\theta_1, \dots, \theta_k$ e que os parâmetros são aproximadamente ortogonais. Para justificar a última condição, Naylor e Smith (1982) transformam os parâmetros para obter um novo conjunto de parâmetros ortogonais. No caso unidimensional, se $h(t)$ é uma função bem comportada e regular e $g(t) = h(t)(2\pi\sigma^2)^{-.5}exp\{-.5[\frac{t-\mu}{\sigma}]^2\}$ então

$$\int_{-\infty}^{+\infty} g(t)dt \approx \sum_{i=1}^n m_i g(z_i)$$

onde $m_i = \omega_i \exp(t_i^2) \sqrt{2}\sigma$, $z_i = \mu + \sqrt{2}\sigma t_i$ e $\omega_i = \frac{2^{n-1} n! \pi^{1/2} \exp\{t_i^2\}}{n^2 [H_{n-1}(t_i)]^2}$ onde os t_i são os valores do polinômio de Hermite $H_n(t)$. Para implementar esta aproximação, tabelas com os valores de t_i , ω_i , e $\omega_i \exp(t_i^2)$ são necessárias. No caso de um vetor k-dimensional, uma extensão simples de (6) é

$$g(t_1, \dots, t_k) \approx \sum_{i_k} m_{i_k}^{(k)} \dots \sum_{i_2} m_{i_2}^{(2)} \sum_{i_1} m_{i_1}^{(1)} g(z_{i_1}^{(1)}, \dots, z_{i_k}^{(k)})$$

onde $m_{i_j}^{(j)}$ e $z_{i_j}^{(j)}$ são obtidos como anteriormente, substituindo μ e σ^2 pela média e variância marginal a posteriori de θ_j .

Algoritmo

Os principais passos do algoritmo proposto para fazer as inferências marginais no MD univariado com erros GARCH(p,q) são:

- (1) Iniciar com uma grade de três pontos em cada dimensão e valores iniciais para média e variância da posteriori de θ_j ;
- (2) Iterar 3 vezes calculando a média e variância da posterior de cada uma das componentes de θ_j ;
- (3) Ampliar a grade para cinco ou mais pontos em cada dimensão, usando como grade inicial os pontos obtidos na última iteração de (2);
- (4) Calcular as médias e variâncias a posteriori e reposicionar a grade iterativamente até atingir a convergência.
- (5) Calcular as características de interesse, tais como: $E(\beta_{t-1} | D_{t-1})$, $V(\beta_{t-1} | D_{t-1})$ etc. usando as expressões descritas anteriormente.

Aproximação de Laplace

Uma aproximação útil para a avaliação de integrais do tipo

$$\int g(\theta) \exp[-nh(\theta)] d\theta$$

onde $g : \mathbf{R}^d \rightarrow \mathbf{R}^+$ e $h : \mathbf{R}^d \rightarrow \mathbf{R}$ são funções suaves de θ é dado por

$$\hat{I} = g(\hat{\theta})(2\pi/n)^{d/2} |\hat{V}|^{1/2} \exp[-nh(\hat{\theta})]$$

onde

$$\hat{V} = \left[\frac{\partial^2 h(\hat{\theta})}{\partial \theta^T \partial \theta} \right]^{-1}$$

onde $\hat{\theta}$ o valor de θ que minimiza $h(\cdot)$.

A aproximação de Laplace baseia-se na expansão de Taylor de $h(\cdot)$ e $g(\cdot)$ entorno de $\hat{\theta}$.

Nas aplicações Bayesianas, em geral:

$$-nh(\theta) = \log(q(\theta|\mathbf{x}))$$

Se $g(\theta)$ for não negativa podemos redefinir a integral de interesse por

$$I = \int \exp[-nh^*(\theta)] d\theta, \text{ onde } nh^*(\theta) = nh(\theta) - \log(g(\theta))$$

Denotando por $\hat{\theta}^*$ o valor que minimiza $h^*(\theta)$ e por $\hat{V}^* = \frac{\partial^2 h^*(\hat{\theta})}{\partial \theta \partial \theta'}$ segue que

$$\hat{I} = (2\pi)^{d/2} |\hat{V}^*| \exp[-nh^*(\hat{\theta}^*)]$$

Tierney & Kadane (1986) propõem avaliar $E[g(\theta)] = \frac{\int g(\theta)q(\theta)d\theta}{\int q(\theta)d\theta}$ aproximando separadamente o numerador e denominador. Demonstram que assim a aproximação será $o(n^{-2})$.

Relembrando que $\hat{\theta}$ é o valor de θ que maximiza $\log(q(\theta))$ e $\hat{\Sigma} = \frac{\partial^2 \log q(\hat{\theta})|x}{\partial \theta^2}$ e aplicando (2) tem-se

$$E[g(\theta)|x] = \frac{g(\hat{\theta}^*)(2\pi/n)^{d/2}|\hat{V}^*|^{1/2} \exp[-nh(\hat{\theta}^*)]}{(2\pi/n)^{d/2}|\hat{V}|^{1/2} \exp[-nh(\hat{\theta})]}$$

Usando (3) segue que:

$$E[g(\theta)] = \frac{g(\hat{\theta}^*)|\hat{V}^*|^{1/2}q(\hat{\theta}^*)}{|\hat{V}|^{1/2}q(\hat{\theta})}.$$