

## **Ainda os modelos GARCH\***

Rodrigo De Losso da Silveira Bueno<sup>§</sup>

### **RESUMO**

Este artigo resume de maneira sistemática a vasta literatura sobre modelos GARCH. Trata-se de uma compilação sobre o assunto, com a intenção de divulgar e difundir tais modelos para aqueles fluentes na língua portuguesa. Estende-se os modelos GARCH em relação à introdução feita por Issler (1999) em duas dimensões: generalizando a especificação univariada e também apresentando os modelos multivariados.

**Palavras-chave:** Modelos GARCH.

### **ABSTRACT**

This article summarizes the huge literature on GARCH Models. It is a survey on this subject to spread those models throughout the Portuguese readers. Issler (1999) briefly introduces univariate GARCH models, and provides results of several Brazilian univariate financial series. We extend such study in terms of specification of the univariate model and by presenting the Multivariate GARCH Models.

**Key words:** GARCH Models.

**JEL classification:** C5, C52, C32.

---

\* O autor agradece as sugestões de Denisard C. O. Alves, Vera L. Fava e Fabiana Tito. Erros remanescentes são de sua responsabilidade.

§ IPE/USP e Doutorando The University of Chicago. [delosso@usp.br](mailto:delosso@usp.br)

Recebido em agosto de 2001. Aceito em fevereiro de 2002.

## 1 Introdução<sup>1</sup>

Há duas décadas assistiu-se ao surgimento dos modelos auto-regressivos de heterocedasticidade condicional, ARCH, formalizados por Engle (1982), e posteriormente generalizados por Bollerslev (1986), GARCH, dando origem a uma vasta literatura sobre o assunto. Antes disso, os modelos econométricos de séries de tempo, financeiros e macroeconômicos, focalizavam essencialmente a questão do primeiro momento condicional, ou seja, os modelos ARIMA. Dependências temporais de ordem superior eram simplesmente tratadas como perturbações aleatórias.

O desenvolvimento desses modelos ocorreu por causa do aumento crescente da importância do risco e da incerteza na teoria econômica moderna, e pelo fato de modelos como o CAPM de Sharpe (1964) e Lintner (1965) não funcionarem tão bem empiricamente. Assim, em virtude das dificuldades encontradas em testá-los e procurando-se estender as fronteiras da econometria de séries temporais, desenvolveram-se técnicas que permitem a modelagem temporal de variâncias e covariâncias.

O uso dos modelos de heterocedasticidade condicional é análogo ao de modelos lineares de séries de tempo, ARIMA( $p, d, q$ ), que se baseiam na estimação da média condicional. De modo semelhante, os modelos GARCH baseiam-se na estimação da variância condicional em vez de considerá-la constante ao longo do tempo. Por este fato, pode-se dizer que a variância em determinado instante depende de variâncias passadas.

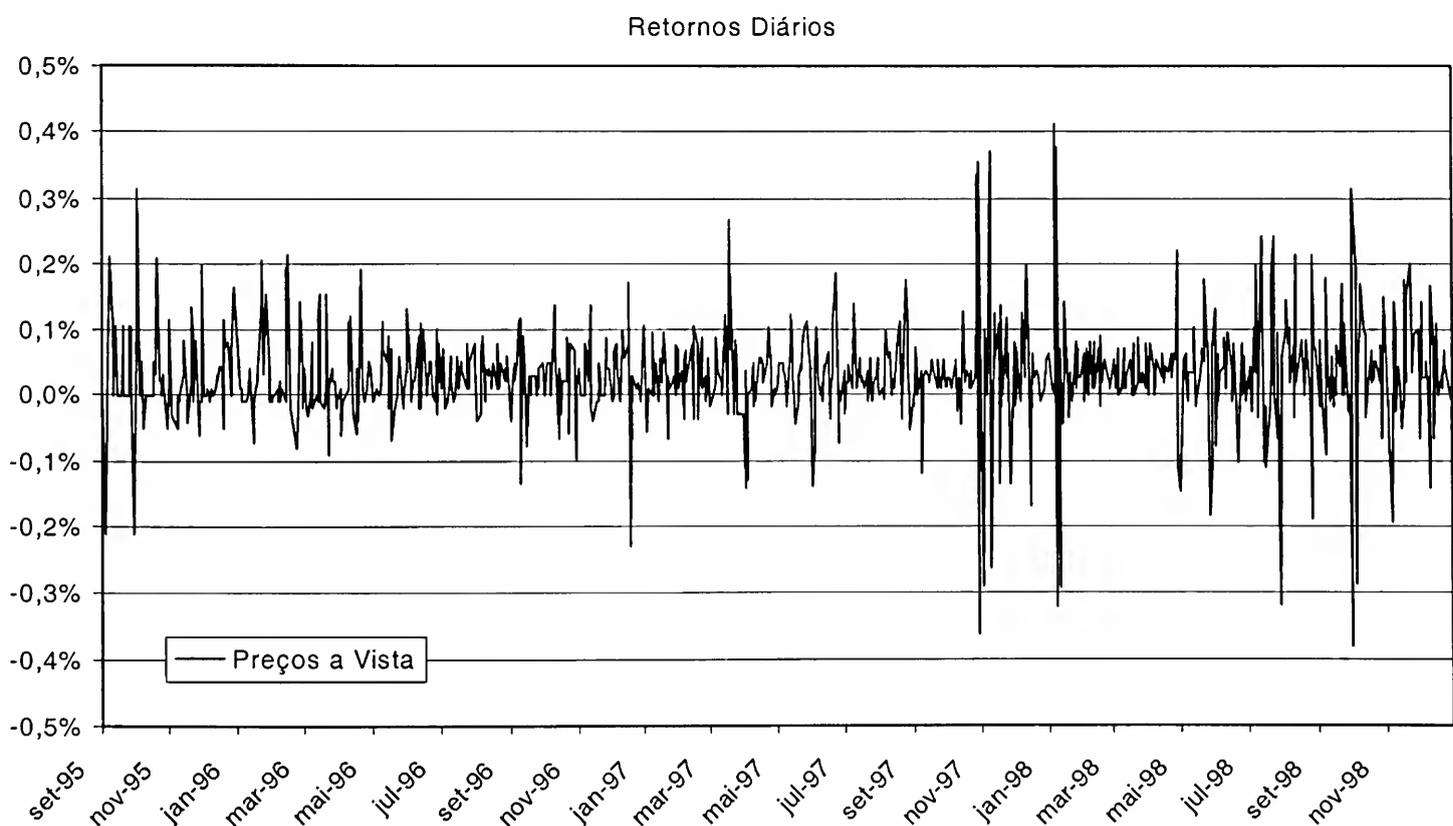
A grande contribuição de seus criadores está na distinção de uso entre momentos de segunda ordem condicionais e não condicionais. Enquanto a matriz de covariância não condicional para as variáveis de interesse pode ser invariante no tempo, a matriz de covariância condicional freqüentemente depende de estados passados da natureza. Ora, o entendimento dessa dependência temporal é fundamental para muitas questões em macroeconomia e finanças, como, por exemplo, a natureza da volatilidade de retornos observada em relação à prevista por modelos de apreçamento de ativos. Além disso, da perspectiva de inferência econométrica, a perda de eficiência assintótica, por se desconsiderar a natureza condicional da heterocedasticidade, pode ser extremamente significativa. Dessa maneira, condicionando as previsões às informações presentes, geralmente obtêm-se estimativas mais precisas da incerteza do erro de previsão como resultado da estimação mais acurada do intervalo de confiança dessa previsão e, potencialmente, pode-se simular a volatilidade observada dos retornos (ver Bansal e Yaron, 2000).

---

1 Baseada em Bollerslev, Engle e Nelson (1994).

Empiricamente, observa-se que as séries financeiras não têm, necessariamente, distribuição Normal-padrão, dada a elevada probabilidade de eventos extremos, tanto positivos quanto negativos. Quando a probabilidade de eventos extremos é maior do que na Normal-padrão, diz-se que a série tem caudas gordas, chamadas de leptocúrticas. De fato, há evidências empíricas de que grandes oscilações são seguidas por outras significativas (no Gráfico 1, vê-se um exemplo desses fatos para a variação de preço a vista do dólar). Várias são as razões para esse comportamento. Suponha-se, por exemplo, que uma má notícia qualquer seja divulgada. Isso deverá provocar muito mais oscilações nos preços dos ativos do que antes, quando o mercado estava tranqüilo. Essas oscilações geralmente perduram por um certo tempo, até que se retorne ao comportamento anterior. Além disso, períodos em que não há negociações podem resultar em aumento de volatilidade. Por exemplo, os agentes estão absorvendo informações durante o final de semana, e o reflexo pode ser uma elevação da volatilidade na abertura das transações na segunda-feira.

**Gráfico 1**  
**Retornos do Dólar a Vista**



Como esses, outros fenômenos que se dão em determinado instante podem alterar a volatilidade da série, tornando importante que se consiga prever o comportamento da variância ao longo do tempo. Isso indica que considerar a variância condicional **constante** ao

longo do tempo é uma hipótese discutível. Daí a relevância dos modelos de heterocedasticidade condicional que, como será mostrado, permitem captar a dinâmica da variância ao longo do tempo e também abarcar casos em que a curtose da série é alta (ver Bollerslev, Engle e Nelson, 1994 e Nelson, 1989).

Uma vez reconhecidas as vantagens desses modelos, uma vasta literatura empírica e teórica vem se acumulando. O interesse deste artigo é simplesmente compilar parte dessa literatura, com a intenção de divulgar e difundir tais modelos para aqueles fluentes na língua portuguesa, à semelhança de trabalhos como os de Bollerslev, Engle e Nelson (1994) e Bera e Higgins (1995), porém num menor grau de formalidade e complexidade.

Estende-se os modelos GARCH em relação à introdução feita por Issler (1999) em duas dimensões: generalizando a especificação univariada, e também apresentando os modelos multivariados.

O artigo está organizado da seguinte forma: a seção 2 apresenta os modelos GARCH univariados que foram desenvolvidos nos últimos 20 anos; a seção 3 apresenta os testes para a detecção de modelos GARCH univariados; as seções 4, 5, 6 e 7 discutem, respectivamente, a identificação, a estimação, a inferência e a previsão em modelos GARCH univariados; a seção 8 generaliza os modelos GARCH univariados; a seção 9 apresenta os modelos GARCH multivariados, discutindo testes, estimação e inferência; a seção 10 conclui o trabalho.

## 2 Modelos GARCH<sup>2</sup>

### 2.1 ARCH( $q$ )

O primeiro modelo de variância condicional foi proposto por Engle (1982):

**Definição 1: ARCH** - Seja  $\varepsilon_t$  um processo estocástico real em tempo discreto e  $I_{t-1}$  a  $\sigma$ -álgebra de toda a informação até  $t-1$  e suponha que se queira estimar o modelo ARIMA( $p, 0, q$ ); então, define-se o processo ARCH como sendo:

---

2 Neste trabalho assume-se prévio conhecimento dos modelos tradicionais de séries de tempo, em especial da metodologia de Box-Jenkins (ver Box, Jenkins e Reinsel, 1994; Hamilton, 1994; Brockwell e Davis, 1991; Enders, 1995; Spanos, 1986 e Mills, 1990, 1999).

$$y_t = \frac{\Theta(L)}{\Phi(L)} \varepsilon_t \quad (1)$$

onde:

$\Phi(L) = (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)$  é a polinomial de auto-regressão<sup>3</sup>,

$L$  denota o operador defasagem, isto é,  $Ly_t = y_{t-1}$ ,

$\Theta(L) = (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)$  é a polinomial de médias móveis;

$$\varepsilon_t = \sigma_t u_t$$

sendo  $u_t \sim i.i.d.(0,1)$ ,

$$\sigma_t^2 = \alpha + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2, \text{ independente de } u_t.$$

A distribuição do erro,  $u_t$ , não precisa ser necessariamente Normal, basta ser independentemente e identicamente distribuída para que os momentos assintóticos da série sejam iguais aos da Normal. O fato de  $\text{Var}(u) = 1$ , ou seja, da variância ser fixa, não implica perda de generalidade. Qualquer outro número poderia ser usado, e o ajuste se daria nos coeficientes de  $\sigma_t^2$ . Com isso, é fácil verificar que a variância não condicional é dada por:

$$E(\varepsilon_t^2) = E(\sigma_t^2) \cdot E(u_t^2) = \sigma_\varepsilon^2 = \frac{\alpha}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_q} \quad (2)$$

3 Às vezes usa-se a letra  $B$  no lugar de  $L$ .

É preciso formular algumas restrições ao modelo para garantir a estacionariedade e positividade da variância condicional. Em primeiro lugar, uma condição suficiente para que a variância condicional seja sempre positiva é de que  $\alpha > 0$ ,  $\alpha_i \geq 0$ ,  $\forall i, i = 1, 2, \dots, q$ . Além disso,  $\sum_{i=1}^q \alpha_i < 1$  é condição necessária para que a série seja estacionária. Formalmente tem-se:

**Teorema 1.** O  $q$ -ésimo processo linear ARCH, com  $\alpha > 0$ ,  $\alpha_i \geq 0$ ,  $\forall i$ , é estacionário se, e somente se, as raízes da função característica associada estiverem fora do círculo unitário. Além disso, a variância estacionária é dada por (2).

Prova: Ver Engle (1982) ou Gouriéroux (1997).

Da mesma maneira que nos modelos ARIMA, o modelo condicional no segundo momento pode ser assim expresso:

$$E(\sigma_t^2 | I_{t-1}) = \alpha + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad (3)$$

Para verificar, intuitivamente, que a curtose da série é plenamente considerada em modelo ARCH, enuncia-se o seguinte teorema para o caso particular em que  $q = 1$ .

**Teorema 2:** Seja o processo ARCH  $\varepsilon_t = \sigma_t u_t$ , sendo  $u_t \sim i.i.d.(0,1)$ , onde  $\sigma_t^2 = \alpha + \alpha_1 \varepsilon_{t-1}^2$ ,

independente de  $u_t$ . Então a curtose do modelo é dada por:  $K(\varepsilon_t) = 3 \left[ \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} \right]$

Prova.  $E(\varepsilon_t^4) = E(\sigma_t^4) E(u_t^4) = 3E(\sigma_t^4)$  Mas,  $E(\sigma_t^4) = E\left[\left(\alpha^2 + 2\alpha\alpha_1\varepsilon_t^2 + \alpha_1^2\varepsilon_t^4\right)\right]$

Portanto:  $E(\varepsilon_t^4) = 3 \left[ \alpha^2 + \frac{2\alpha^2\alpha_1}{1-\alpha_1} + \alpha_1^2 E(\varepsilon_t^4) \right] \Rightarrow (1-3\alpha_1^2) E(\varepsilon_t^4) = \frac{3\alpha^2(1+\alpha_1)}{1-\alpha_1}$

o que resulta, após algumas poucas manipulações algébricas, em:

$$E(\varepsilon_t^4) = \left[ \frac{3\alpha^2}{(1-\alpha)^2} \right] \left[ \frac{1-\alpha^2}{1-3\alpha^2} \right] \tag{4}$$

Logo:

$$K(\varepsilon_t) = 3 \left[ \frac{1-\alpha^2}{1-3\alpha^2} \right] \tag{5}$$

Q.E.D.

O primeiro fator entre colchetes de (4) é três vezes a variância ao quadrado. Analisando a expressão entre colchetes de (5) é fácil verificar que se  $\alpha = 0$ , a série é mesocúrtica. Por outro lado, para que a curtose seja finita é preciso que  $3\alpha^2 < 1 \Rightarrow \alpha^2 < \frac{1}{3}$ . Além disso,  $0 < \alpha < 1 \Rightarrow 1 - \alpha^2 > 1 - 3\alpha^2$ . Portanto, a curtose da série pode ser maior do que 3, e o modelo ARCH é capaz de incorporar esse fato.

## 2.2 GARCH(p, q)

Um dos problemas da formulação inicial de Engle, verificado empiricamente, é que os modelos ARCH(q) demandam muitos parâmetros para serem estimados adequadamente. Seguindo a idéia dos modelos ARIMA, em que todo MA( $\infty$ ) pode ser representado por um AR(p), p finito, Bollerslev (1986) propôs o modelo de heterocedasticidade condicional autorregressivo generalizado, GARCH, cuja definição é a seguinte:

**Definição 2: GARCH** – Seja  $\varepsilon_t$  um processo estocástico real em tempo discreto e  $I_{t-1}$  a  $\sigma$ -álgebra de toda a informação até t-1 e suponha que se queira estimar o modelo (1). Então, define-se processo GARCH como tendo a seguinte equação da variância condicional:

$$\sigma_t^2 = \alpha + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \tag{6}$$

onde:

$p$  - ordem do processo auto-regressivo na variância condicional;

$q$  - ordem do processo de "médias móveis";

$\alpha > 0$ ,  $\alpha_i \geq 0$ ,  $\beta \geq 0$  - condição suficiente para que a variância seja positiva.

Uma condição suficiente para a estacionariedade é que as raízes da polinomial da variância estejam fora do círculo unitário, como no caso dos modelos ARIMA. Uma condição necessária é:  $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$ . Se  $p = 0$ , volta-se ao modelo proposto por Engle. Formalmente tem-se:

**Teorema 3:** O processo GARCH ( $p$ ,  $q$ ), dado pela Definição 2, é estacionário com

$$E(\varepsilon_t) = 0, \text{var}(\varepsilon_t) = \frac{\alpha}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j} \text{ e } \text{cov}(\varepsilon_t, \varepsilon_j) = 0 \text{ se } i \neq j \text{ se, e somente se, as raízes}$$

da função característica associada estiverem fora do círculo unitário.

Prova: Ver Bollerslev (1986).

Pode-se interpretar o modelo GARCH( $p$ ,  $q$ ) como um processo auto-regressivo em  $\varepsilon_t^2$ . Para ver isso defina-se:

$$v_t = \varepsilon_t^2 - \sigma_t^2 = (u_t^2 - 1)\sigma_t^2 \quad (7)$$

É fácil verificar que  $v_t$  tem média zero e é não autocorrelacionado. A seguir, substituindo em (6) obtém-se:

$$\varepsilon_t^2 = \alpha + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \varepsilon_{t-j}^2 - \sum_{j=1}^p \beta_j v_{t-j}^2 + v_t \quad (8)$$

Ou seja, pode-se reescrever (8) como:

$$\varepsilon_t^2 = \alpha + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{j=1}^p \beta_j v_{t-j}^2 + v_t \quad (9)$$

A equação (9) é muito parecida com um processo ARMA[ $\max(p, q), p$ ], sugerindo que a identificação do modelo GARCH pode seguir a mesma metodologia proposta por Box-Jenkins. No entanto, deve-se ter cautela com isso, pois a metodologia será válida sob a hipótese de se tratar efetivamente de um modelo GARCH( $p, q$ ) (ver Mills, 1999, para considerações adicionais).

Da mesma maneira que no processo ARCH, é possível mostrar que o modelo GARCH implica a seguinte equação para a curtose no GARCH(1, 1) (ver Bollerslev, 1986):

$$K(\varepsilon_t) = \frac{6\alpha_1^2}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2} \quad (10)$$

O advento do modelo GARCH propiciou uma maior eficiência na estimação e uma economia de parâmetros a serem estimados. Na prática, é muito usado o modelo GARCH(1, 1), no qual são estimados apenas três parâmetros na equação da variância. Se em vez disso fosse usado o ARCH( $q$ ), o número de parâmetros a serem estimados poderia ser maior. Assim, o modelo GARCH pode ser considerado um avanço, pois permite menor perda de graus de liberdade.

As restrições impostas sobre os coeficientes dos modelos GARCH, de que eles devem ser todos maiores do que zero, podem ser relaxadas. Na verdade, as condições gerais de estacionariedade que devem ser impostas sobre os coeficientes não foram ainda derivadas. Para casos específicos, no entanto, existem alguns resultados importantes. Por exemplo, Nelson e Cao (1992) dão as condições de estacionariedade para os modelos GARCH(1,  $q$ ) e GARCH(2,  $q$ ). Em particular, o GARCH(1, 2) será estacionário se:

$$\alpha \geq 0;$$

$$0 \leq \beta_1 < 1,$$

$$\beta_1\alpha_1 + \alpha_2 \geq 0; \text{ e}$$

$$\alpha_1 \geq 0$$

$\alpha_2$  pode ser menor do que zero. Analogamente o GARCH (2, 1) será estacionário caso:

$$\alpha \geq 0,$$

$$\alpha_1 \geq 0,$$

$$\beta_1 \geq 0,$$

$$\beta_1 + \beta_2 < 1, \text{ e}$$

$$\beta_1^2 + 4\beta_2 \geq 0$$

Nada impede que  $\beta_2$  seja menor do que zero. As duas condições acima são muito importantes porque o relaxamento da hipótese de que todos os coeficientes do modelo devem ser positivos permite resultados mais próximos da prática e permitem a imposição de restrições de estimação mais adequadas.

### 2.3 EGARCH( $p, q$ )

Apesar dos modelos anteriores terem se constituído em um grande avanço na modelagem econométrica, eles supõem que os impactos dos choques aleatórios são simétricos. Contudo, há muitas evidências empíricas de que isso não é verdade. Choques no retorno de ações, por exemplo, são fortemente assimétricos: choques negativos são seguidos por maiores aumentos na volatilidade do que choques positivos.

O modelo GARCH captura, com sucesso, os fatos de a série ter caudas espessas e de um aumento de volatilidade persistir por algum tempo, isto é, os chamados de períodos de aglomerações de alta volatilidade. Mas não está, por exemplo, bem adaptado para captar efeitos de alavancagem. Ou seja, os modelos GARCH são incapazes de captar assimetrias nos retornos. Além disso, os parâmetros do modelo têm que ser positivos para que a estacionariedade das séries seja mantida.

Tentando superar essas dificuldades, Nelson (1991) propôs o modelo de heterocedasticidade condicional auto-regressivo generalizado exponencial, EGARCH, que permite a absorção de impactos assimétricos e a possibilidade irrestrita de que alguns coeficientes sejam negativos.

$$\ln \sigma_t^2 = \alpha + \sum_{j=1}^p \beta_j \ln \sigma_{t-j}^2 + \sum_{i=1}^q \alpha_i \left\{ \phi \left[ \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| - E \left( \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| \right) - \gamma \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right] \right\} \quad (11)$$

Por causa da especificação logarítmica, não é possível que a variância seja negativa, portanto não há a necessidade de restringir os parâmetros ao  $\mathcal{R}_+^{\max[p,q]+3}$ . O parâmetro  $\gamma$  permite que os efeitos sejam assimétricos. Se ele for igual a zero, então um choque positivo tem o mesmo efeito na volatilidade que um choque negativo de mesma magnitude. Se  $0 < \gamma < 1$ , um choque negativo aumenta a volatilidade mais do que um positivo. Se o parâmetro for maior do que 1, um choque positivo reduz a volatilidade mais do que proporcionalmente, enquanto a volatilidade aumenta em caso contrário. (Hamilton, 1994) Ademais, os erros não são elevados ao quadrado, de forma a medir a diferença entre choques positivos e negativos. É preciso, no

entanto, que as raízes da polinomial  $1 - \sum_{j=1}^p \beta_j L^j$  estejam fora do círculo unitário a fim de que a variância seja estacionária. A estacionariedade estrita é dada, segundo Nelson (1991),

$$\text{se } \sum_{i=1}^q \alpha_i^2 < \infty$$

Nelson (1991) propõe a função de Distribuição Erro Generalizado, GED, para estimar o EGARCH:

$$f(u_t) = \frac{\nu \exp\left(-\frac{1}{2} \left| \frac{\varepsilon_t}{\lambda} \right|^\nu\right)}{\lambda 2^{\frac{\nu+1}{\nu}} \Gamma\left(\frac{1}{\nu}\right)} \quad (12)$$

onde:

$\Gamma(\cdot)$  é a função gama;

$$\lambda \equiv \left[ \frac{2^{-\frac{2}{\nu}} \Gamma\left(\frac{1}{\nu}\right)}{\Gamma\left(\frac{3}{\nu}\right)} \right]^{\frac{1}{2}} \text{ é uma constante;}$$

$0 < \nu \leq \infty$  é um parâmetro positivo determinando o grau de curtose da distribuição;

$$-\infty < \varepsilon_t < \infty$$

Em particular, quando  $\nu = 2$  e  $\lambda = 1$ , tem-se a distribuição normal. Se  $\nu < 2$ , a distribuição tem caudas gordas.

## 2.4 ARCH-M

Muitas teorias financeiras exigem uma explícita relação entre média e risco, ou variância. Com o advento dos modelos ARCH tornou-se mais fácil obter esse tipo de especificação, bastando acrescentar à equação da média a variável explicativa  $\sigma_t$  ou seu quadrado. Esse modelo foi inicialmente proposto por Engle, Lillien e Robins (1987), podendo ser facilmente estendido a casos mais gerais, como GARCH-M, EGARCH-M, e assim por diante. Formalmente, o modelo tem a seguinte configuração:

$$y_t = \frac{\Theta(L)}{\Phi(L)} \varepsilon_t + \delta \sigma_t^a \quad (13)$$

onde  $\varepsilon_t = \sigma_t u_t$

sendo  $u_t \sim i.i.d.(0,1)$

onde:

$$\sigma_t^a = f(\varepsilon_{t-j}, \sigma_{t-i}), i, j = 1, 2, \dots, \max[p, q];$$

$$a = 1, 2.$$

Se  $y_t$  representa os retornos de uma série financeira, intuitivamente espera-se que  $\delta$  seja positivo.

## 2.5 TGARCH (p, q) e GJR GARCH (p, q)

As formulações econométricas para a captação de impactos assimétricos têm sido bastante estudadas, como no caso do EGARCH. Zakoian (1990) propôs o modelo de heterocedasticidade condicional auto-regressivo generalizado truncado, TGARCH, da seguinte forma:

$$\sigma_t^2 = \alpha + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{i=1}^q \left[ \alpha_i d_{(\varepsilon_{t-i} \leq 0)} |\varepsilon_{t-i}|^a - \gamma \varepsilon_{t-i}^a \right] \quad (14)$$

onde  $d(\ )$  é a variável “dummy” igual a zero caso o erro não satisfaça a condição imposta entre parênteses, e 1, em caso contrário;

$a = 1$ .

Glosten, Jagannathan e Runkle (1993) propuseram a estimação desse modelo com a seguinte modificação:  $a = 2$ .

## 2.6 NGARCH

Higgins e Bera (1992) propuseram o modelo ARCH não linear, NARCH, mas que pode ser estendido para um NGARCH da seguinte maneira:

$$\sigma_t^a = \alpha + \sum_{j=1}^p \beta_j \sigma_{t-j}^a + \sum_{i=1}^q \left[ \alpha_i |\varepsilon_{t-i}|^a \right] \quad (15)$$

Esse modelo tem a vantagem de deixar o parâmetro  $a$  livre para ser estimado e, portanto, é bem geral; porém, não incorpora impactos assimétricos.

## 2.7 A-PGARCH

Combinando as idéias dos dois últimos modelos apresentados, Ding, Granger e Engle (1993) propuseram o seguinte modelo assimétrico, cuja assimetria é captada pelo coeficiente  $\gamma$ .

$$\sigma_t^a = \alpha + \sum_{j=1}^p \beta_j \sigma_{t-j}^a + \sum_{i=1}^q \left[ \alpha_i (|\varepsilon_{t-i}| - \gamma \varepsilon_{t-i})^a \right] \quad (16)$$

## 2.8 STARCH - Modelos de volatilidade estocástica

Harvey, Ruiz e Sentana (1992) propuseram um modelo de variância condicional diferente do tradicional, pois usam o chamado modelo estrutural, que é estimado por meio do Filtro de Kalman (ver detalhes adicionais em Harvey, 1989 e Harvey, Ruiz e Shephard, 1994).

A aplicação desse modelo no caso de variância condicional exige que a série seja um ruído branco. Em geral, se o mercado é eficiente, os retornos da série costumam ser um ruído branco. No entanto, é difícil encontrar séries que satisfaçam estritamente a condição do modelo.

Formalmente, o modelo de volatilidade estocástica ou ARCH estrutural é o seguinte:

$$y_t = \mu v_t e^{\sigma_t^2/2} \quad (17)$$

onde:

$\mu$  é constante;

$$v_t \xrightarrow{d} N(0,1),$$

$$\sigma_t^2 = \alpha + \phi \sigma_{t-1}^2 + \eta_t,$$

$$\text{Var}(\eta_t) = \sigma_\eta^2,$$

$$\text{Cov}(v_t, \eta_{t-s}) = 0, \forall s,$$

$\xrightarrow{d}$  denota convergência em distribuição.

Elevando ao quadrado e linearizando (17) obtém-se:

$$\ln y^2 = \ln \mu^2 + \ln v_t^2 + \sigma_t^2 \quad (18)$$

e

$$\sigma_t^2 = \alpha + \phi \sigma_{t-1}^2 + \eta_t \quad (19)$$

Este é um típico modelo de séries de tempo estrutural que pode ser estimado utilizando-se o Filtro de Kalman, o qual permite que a variância varie ao longo do tempo não apenas por ser temporalmente dependente, mas também por incluir um termo estocástico em sua especificação. Esta é uma diferença fundamental em relação aos modelos apresentados anteriormente, nos quais não havia termos estocásticos na equação da variância que não fossem dependentes da equação da média. Pode-se observar, ainda, que o termo estocástico da equação da média não é  $\sigma_t^2$ , embora, por influência deste, a equação da média incorpore o termo estocástico da equação da variância. O modelo tem uma diferença adicional: o erro da equação da média não está presente na equação da variância.

### 3 Testes para detecção de modelos GARCH

Existem vários testes para a identificação de modelos GARCH. Apresentam-se, a seguir, os mais comuns. Outras possibilidades podem ser encontradas em Bollerslev, Engle e Nelson (1994).

#### 3.1 Teste ACF e PACF

O modelo GARCH é parecido com um modelo ARMA ( $p, q$ ). Assim, suas funções de autocorrelação, ACF, e autocorrelação parcial, PACF, devem sugerir se a série é heterocedástica, da mesma maneira que dão uma idéia das ordens  $p$  e  $q$  de um modelo ARMA. Para estimar essas funções, o primeiro passo é obter os quadrados dos erros estimados pela regressão do modelo estabelecido como se a variância condicional fosse constante:

$$\bar{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t^2 \quad (20)$$

onde  $T$  - número de resíduos.

Em seguida, calcula-se a função de autocorrelação amostral para o quadrado dos resíduos e representa-a graficamente. A função é dada por:

$$\hat{\rho}_i = \frac{\sum_{t=i+1}^T (\hat{\varepsilon}_t^2 - \bar{\sigma}^2)(\hat{\varepsilon}_{t-i}^2 - \bar{\sigma}^2)}{\sum_{t=1}^T (\hat{\varepsilon}_t^2 - \bar{\sigma}^2)^2} \quad (21)$$

Em grandes amostras, o desvio padrão de  $\hat{\rho}_i$  pode ser aproximado por  $T^{-0.5}$  (ver Enders, 1995, p. 148 ou Mills, 1999).<sup>4</sup> Se os valores de  $\hat{\rho}_i$  forem estatisticamente diferentes de zero, então diz-se que algum modelo ARCH está presente.

A função de autocorrelação parcial é obtida colocando-se num gráfico o coeficiente  $\phi_{s,s}$  contra  $s$  estimado a partir da seguinte equação:

$$\varepsilon_t^2 = \phi_s + \phi_{s,1}\varepsilon_{t-1}^2 + \phi_{s,2}\varepsilon_{t-2}^2 + \dots + \phi_{s,s}\varepsilon_{t-s}^2, \quad s = 1, 2, \dots, q \quad (22)$$

Da mesma forma, em grandes amostras o desvio padrão de  $\hat{\phi}_{s,s}$  pode ser aproximado por  $T^{-0.5}$ .

No caso deste teste, diferentemente dos modelos ARMA, a ACF dá a ordem da autoregressão do GARCH e a PACF dá a ordem das “médias móveis” do GARCH.

Porém, se tais modelos não forem do tipo GARCH, obviamente tais testes não servem para definir  $p$  e  $q$ , mas sim para identificar a existência de heterocedasticidade condicional.

### 3.2 Teste de Ljung-Box

A estatística de Ljung-Box,<sup>5</sup>  $Q$ , pode ser usada para testar a presença de heterocedasticidade condicional. Sua hipótese nula é de que a soma das autocorrelações é estatisticamente diferente de zero. Isto é:

4 Na verdade, isso não é estritamente verdadeiro do ponto de vista estatístico (ver Priestley, 1981). Mas usa-se tal método por questões práticas.

5 Esta estatística deve ter preferência sobre a anterior.

$$H_0 \sum_{i=1}^n \rho_i = 0 \text{ contra } H_a : \rho_i \neq 0 \text{ para algum } i$$

A estatística de Ljung-Box é dada por:

$$Q = T(T+2) \sum_{i=1}^n \frac{\hat{\rho}_i^2}{T-i} \xrightarrow{d} \chi_n^2 \quad (23)$$

onde  $\chi_n^2$  - distribuição qui-quadrada com  $n$  graus de liberdade.

### 3.3 Teste multiplicador de Lagrange ou de Engle

Outro teste possível é o do Multiplicador de Lagrange, LM (ver White, 1980; Bollerslev, Engle e Nelson, 1994 e Johnston e Dinardo, 1997 para outros detalhes), inicialmente proposto por Engle (1982) para o processo ARCH, e depois estendido por Lee (1991) para o GARCH. O teste é definido como:

$$H_0 : \alpha_i = 0, \text{ para todo } i \times H_a . \alpha_i \neq 0 \text{ para algum } i$$

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2 \quad (24)$$

cuja estatística é dada por:

$$TR^2 \xrightarrow{d} \chi_q^2$$

onde  $R^2$  representa o coeficiente de determinação de (24).

Se os coeficientes são conjuntamente próximos de zero, a regressão terá pouco poder explicativo, logo, o  $R^2$  deverá ser baixo. Assim, se a hipótese nula não é rejeitada, diz-se que não há presença de modelos ARCH.

Em caso de rejeição da hipótese nula os testes não determinam que tipo de modelo ARCH deverá ser usado. Na verdade, os testes apenas indicam se a covariância condicional é, ou não, constante. Segundo Bollerslev, Engle e Nelson (1994), mesmo em modelos lineares a seleção do modelo apropriado não é trivial. Com relação às inúmeras possibilidades que os modelos ARCH oferecem, o problema torna-se ainda mais complicado. Convém notar que esses testes podem indicar erroneamente a presença de modelos ARCH, porque são sensíveis a especificações incorretas na equação da média. Ou seja, os erros estimados pelo modelo considerando variância não condicional podem indicar a presença de modelos ARCH simplesmente por terem sido especificados de forma errônea, cometendo-se o erro do tipo I.

#### 4 Identificação de modelos GARCH

Já foi mencionado que, detectada a existência de heterocedasticidade condicional, não é possível verificar que tipo de modelo GARCH poderia ser estimado. Existem alguns testes para verificar se o modelo GARCH é simétrico ou assimétrico (Hegerud, 1997), mas tais testes só conseguem separar um modelo GARCH simétrico de outro assimétrico. Uma vez que existem inúmeros modelos assimétricos, restaria saber de qual deles se originam os erros, mantendo-se a dificuldade de especificação.

Uma alternativa usual na literatura é escolher o modelo utilizando-se os critérios de Akaike ou Schwarz. Porém suas propriedades estatísticas na presença de modelos GARCH são desconhecidas. No entanto, esse é um critério muito fácil de ser usado, por isso é largamente preferido. A fórmula do Schwarz, *BIC*, é dada por:

$$BIC(m) = \ln \sigma^2 + \frac{m}{T} \ln T \quad (25)$$

onde  $m$  é o número de parâmetros estimados do modelo.

Outra alternativa proposta é definir um critério de validação dentro da amostra ou, preferencialmente, fora da amostra, escolhendo-se uma função-erro - tipicamente o erro quadrático médio. No caso da equação da variância, a função-erro pode ser (Bollerslev, Engle e Nelson, 1994) a seguinte:

$$E_1 = \sum_{t=1}^T (\varepsilon_t^2 - \sigma_t^2)^2 \quad (26)$$

A função-erro definida em (26) prejudica estimativas de variância condicional que são diferentes do resíduo quadrado realizado de maneira simétrica. Não considera, porém, nenhuma punição para variâncias negativas ou zero que são claramente não intuitivas. Por este critério, regressões de mínimos quadrados dos resíduos quadrados sobre as informações passadas terão a menor perda dentro da amostra.

Uma alternativa mais razoável seria a função-erro definida em termos percentuais dos erros quadrados:

$$E_2 = \sum_{t=1}^T \frac{(\varepsilon_t^2 - \sigma_t^2)^2}{\sigma_t^4} \quad (27)$$

Outra possibilidade é calcular a função-erro implícita na função gaussiana:

$$E_3 = \sum_{t=1}^T \left( \ln \sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2} \right) \quad (28)$$

Ou mais simplesmente prejudicando fortemente a função-erro quando a variância é próxima de zero:

$$E_4 = \sum_{t=1}^T \left( \ln \frac{\varepsilon_t^2}{\sigma_t^2} \right) \quad (29)$$

Apesar das possibilidades apresentadas, a mais natural função-erro deve ser baseada nos objetivos do trabalho na qual deverá ser aplicada (ver Bollerslev, Engle e Nelson, 1994, e suas referências).

## 5 Estimação de modelos GARCH

O método geralmente usado para estimar esses modelos é o de máxima verossimilhança. Esse método exige que se formule uma hipótese sobre a verdadeira distribuição da série *a priori*.

Geralmente escolhe-se a distribuição Normal, em função das propriedades convenientes que encerra. Segundo Hentschel (1995), alguns autores escolhem a distribuição *t-Student* ou Distribuição Erro Generalizado para fazer a estimação quando as provas empíricas mostram a presença de leptocurtose. No entanto, posto que a GED considera a distribuição normal como um caso especial, suas caudas não são suficientemente espessas para abarcar o grande número de observações extremas que a série financeira pode Ter. (Bera e Higgins, 1995)

Na falta de normalidade dos erros, uma alternativa seria escolher a distribuição Normal para a estimação, entendendo-se isso como uma estimação de quase-máxima-verossimilhança. E dependendo do modelo, os resultados pouco se modificarão com a escolha dessa função. (Hentschel, 1995)

Admitindo-se que a distribuição dos erros seja Normal,<sup>6</sup> a função de verossimilhança é a seguinte:

$$L = L(\Theta; y_t) = \prod_{t=\mu+1}^T \frac{1}{(2\pi \sigma_t^2)^{\frac{1}{2}}} \exp \left\{ \frac{\varepsilon_t^2}{-2\sigma_t^2} \right\} \quad (30)$$

onde:

$\Theta$  é o vetor de parâmetros a serem estimados da equação da média e da variância;

$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_\mu^2 = \frac{\sum_{t=1}^T \varepsilon_t^2}{T}$  são as variâncias iniciais do processo de otimização, necessárias ao algoritmo;

$\mu = \max[p, q]$ ,  $p$  e  $q$  da equação da variância, é o número de variâncias para iniciar o processo de otimização.

Supondo-se que a distribuição dos resíduos siga uma *t-Student*, como em Bollerslev (1987), a função de verossimilhança deve ser:

6 Os erros são não correlacionados.

$$L = L(\Theta; \sigma) = \prod_{t=\mu}^T \left[ \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{(\nu-2)\pi\sigma_t^2}} \left(1 + \frac{\varepsilon_t^2}{(\nu-2)\sigma_t^2}\right)^{-\frac{\nu+1}{2}} \right] \quad (31)$$

onde:

$\Gamma(\cdot)$  é a função gama;

$\nu > 2$ .

Nessa função, quando  $\nu \rightarrow \infty$ , a distribuição converge para uma Normal. Entretanto, para  $4 < \nu < \infty$ , a curtose,  $k = 3 \frac{\nu-2}{\nu-4} > 3$ , excede a curtose da Normal.

Um possível método a ser usado na estimação é o algoritmo de Berndt, Hall, Hall e Hausman (1974) – BHHH,<sup>7</sup> dado pela seguinte expressão:

$$\Theta^{i+1} = \Theta^i + \lambda_i \left( \sum_{t=1}^T \frac{\partial \ln L_t}{\partial \Theta} \frac{\partial \ln L_t}{\partial \Theta'} \right)^{-1} \sum_{t=1}^T \frac{\partial \ln L_t}{\partial \Theta'} \quad (32)$$

onde:

$\Theta^i$  - vetor de parâmetros estimados após a *i-ésima* iteração

$\lambda_i$  - tamanho do passo escolhido para maximizar a função de verossimilhança em dada direção.

<sup>7</sup> Ver, ainda, Mills (1999) e Judge *et alii* (1988).

Para a obtenção dos parâmetros iniciais pode-se fazer a regressão por mínimos quadrados ordinários e utilizar os valores dos parâmetros estimados como valores iniciais para o método iterativo. Outra maneira de obter tais parâmetros é ajustar um modelo AR(I)MA e utilizar os valores obtidos como valores iniciais. Os valores dos parâmetros dados pelo teste LM podem servir também como parâmetros iniciais para a estimação.

Contudo, é preciso observar para a variância condicional que:

- (a) As condições de estacionariedade e de não negatividade são difíceis de serem observadas na prática; mas existem métodos de reparametrização que permitem assegurar a hipótese de não negatividade (Bollerslev, Engle e Nelson, 1994);
- (b) Como o procedimento descrito é numérico, é possível que haja problemas de convergência se o valor inicial não for adequado. Não obstante, a estimação por máxima verossimilhança mantém as propriedades assintóticas usuais como consistência e distribuição Normal assintótica;
- (c) Um possível método para assegurar a convergência é estimar os parâmetros por mínimos quadrados ordinários, calcular os erros, calcular  $\sigma_t$ , reestimar os parâmetros por mínimos quadrados generalizados, e assim sucessivamente até a convergência dos coeficientes. Entretanto, hoje esse procedimento não tem tanto sentido, dado o estágio computacional disponível.

De fato, somente é possível proceder como sugerido no item (c) se os coeficientes puderem ser consistentemente estimados pela equação da média. Engle (1982) apresenta as condições segundo as quais isso pode ser feito para o modelo ARCH( $q$ ). Bollerslev, Engle e Nelson (1994), entretanto, sustentam que a perda de eficiência assintótica pode ser muito significativa quando existe heterocedasticidade. Os coeficientes podem ser não viesados, mas terão variância maior em relação a um modelo em que tudo é estimado simultaneamente. Além disso, para esse tipo de procedimento é necessário que a matriz de covariância condicional seja bloco-diagonal e que os erros sejam supostos gaussianos; do contrário, não se pode estimar em separado a equação da média e da variância (ver Gouriéroux, 1997). Por exemplo, os modelos EGARCH( $p, q$ ) e GARCH-M( $p, q$ ) não são bloco-diagonais, logo, é preciso que as equações da média e variância sejam estimadas simultaneamente.

Existe ainda outra razão para estimar ambas as equações simultaneamente. Suponha-se que se queira especificar um modelo ARMA-GARCH sobre os retornos de uma série que, sob a hipótese de eficiência de mercado, teria apenas uma constante na equação da média. Segundo

Engle,<sup>8</sup> se a estimação fosse em separado, possivelmente seriam encontrados parâmetros estatisticamente diferentes de zero, além da constante, na equação da média. Todavia, a estimação simultânea poderia resultar em parâmetros não significativos para a equação da média. Bera e Higgins (1995, p. 227) afirmam que freqüentemente a estimação do processo GARCH muda a especificação do modelo na equação da média. Ou seja, após o processo de identificação do modelo via equação da média, procede-se a estimação e verificação da presença de modelos GARCH. Em caso de rejeição da hipótese de não existência de modelos GARCH, sua estimação subsequente pode alterar a especificação do modelo efetuado na equação da média.

## 6 Inferência em modelos univariados

Utilizar a estimação por máxima verossimilhança, dadas determinadas condições de regularidade, garante a consistência e a normalidade assintótica dos estimadores. Nos modelos GARCH, no entanto, a verificação de tais condições de regularidade tem se mostrado extremamente difícil. (Bollerslev, Engle e Nelson, 1994) Apenas em modelos específicos obtêm-se provas da validade dessas condições. Por exemplo, Lumsdaine (1995) mostra que a seguinte condição sobre o GARCH (1, 1) é suficiente para assegurar sua ergodicidade<sup>9</sup> e estacionariedade:

$$E \ln \left[ \alpha_1 u_t^2 - \beta_1 \right] < 0 \quad (33)$$

De fato, pela desigualdade de Jensen:

$$E \ln \left[ \alpha_1 u_t^2 + \beta_1 \right] < \ln E \left[ \alpha_1 u_t^2 + \beta_1 \right] = \ln \left[ \alpha_1 + \beta_1 \right] \quad (34)$$

Assim, a superfície dos parâmetros cobre o interessante caso do IGARCH (1, 1), no qual  $\alpha_1 + \beta_1$ .<sup>10</sup> Ou seja, não obstante sua semelhança com o modelo integrado de ordem 1, sua interpretação assintótica é completamente diferente.

8 Correspondência pessoal.

9 O conceito de ergodicidade e suas conseqüências podem ser encontrados em Hamilton (1994) e White (2000).

10 O modelo IGARCH( $p, q$ ) é aquele em que os coeficientes da equação da variância somam um, de modo que a variância não condicional passa a tender ao infinito.

Esse fato sugere que se questione se as inferências sobre os parâmetros de um modelo GARCH qualquer são válidas. Na verdade, a teoria assintótica sobre tais especificações é extremamente difícil. (Bollerslev, Chou e Kroner, 1992) Mas, garantida aquela condição de regularidade, mesmo na presença de IGARCH (1, 1), as inferências estatísticas convencionais continuam válidas, embora no IGARCH(1, 1) haja a necessidade de um considerável tamanho amostral para que as distribuições assintóticas convirjam adequadamente. Portanto, pode-se admitir como razoável as inferências tradicionais na presença de modelos GARCH, e se existe a suspeita da existência de um IGARCH (1, 1) é preciso que o tamanho da amostra seja bastante grande. (Lumsdaine, 1995) Na prática, os estudos empíricos simplesmente assumem que as condições de regularidade são satisfeitas. (Bollerslev, Engle e Nelson, 1994, p. 2980)

## 7 Previsão de modelos GARCH

Em modelos econométricos convencionais, com variância condicional constante ao longo do tempo, a incerteza do erro de predição é uma função crescente do horizonte de predição e não depende da origem da previsão. Na presença de modelos ARCH, entretanto, a acuidade das previsões dependerá do conjunto de informações correntes não trivialmente. A construção do intervalo de confiança do erro e os testes de estabilidade estrutural amostrais requerem, por isso, a avaliação das variâncias condicionais do erro futuro.

O problema reside no fato de que a distribuição dos erros é leptocúrtica, alterando sensivelmente a maneira de se construir o intervalo de confiança. Na realidade, ainda estão sendo feitas pesquisas com vistas a encontrar as formulações corretas para a determinação dos intervalos de confiança e da própria função de predição. Heynen e Kat (1994) apresentam funções para os casos de EGARCH (1, 1) e GARCH (1, 1) que, aliás, são os modelos cujas propriedades assintóticas já são conhecidas.

Não obstante, a modo de fazer previsões segue os mesmos passos do modelo ARIMA. Prevê-se um passo à frente; com o resultado obtido, faz-se o passo seguinte, e assim sucessivamente.

Heynen e Kat (1994) realizaram um estudo de previsão usando o índice de ações e câmbio. Seus resultados podem ser sintetizados no seguinte: em geral, a volatilidade é mais previsível no longo prazo do que no curto prazo. Além disso, o desempenho do modelo de previsão depende sensivelmente do ativo com que se está trabalhando. Ainda segundo os autores, índices de ações são melhor previstos utilizando-se o modelo de volatilidade estocástica, enquanto que para o câmbio o modelo GARCH (1, 1) apresentou melhores resultados.

## 8 Generalização dos modelos GARCH: mudança na especificação

Há dois tipos de generalizações possíveis para os modelos GARCH: podem ser diretamente generalizados para o caso multivariado; e podem ser especificados de maneira mais geral. A diferença básica é que o número de parâmetros a serem estimados aumenta exponencialmente para o primeiro caso, o que implica a necessidade de restrições adicionais (ver Engle e Kroner, 1995).

Na presente seção apresenta-se a generalização quanto à especificação do modelo. Esse tipo de modelo é devido a Hentschel (1995), que propõe estimar a seguinte equação de variância:

$$\frac{\sigma_t^\lambda - 1}{\lambda} = \alpha + \alpha_1 \sigma_{t-1}^\lambda f^v(\varepsilon_{t-1}) + \beta_1 \frac{\sigma_{t-1}^\lambda - 1}{\lambda} \quad (35)$$

onde  $f(\varepsilon_t) = |\varepsilon_t - b| - c(\varepsilon_t - b)$

A generalização proposta contempla o caso de modelos GARCH (1, 1), por questões de facilidade de exposição e notação. Naturalmente isso se estenderia para o caso geral, colocando-se os somatórios nos lugares apropriados. Para o caso em que  $\lambda = 1$ , o autor supracitado batizou o modelo de GARCH valor absoluto, AGARCH.

É notório o uso da transformação de Box-Cox (ver Davidson e MacKinnon, 1993) para determinar de que tipo de não-linearidade trata o modelo. Além disso, o autor introduziu uma pequena modificação nas parcelas em que aparecem os erros, multiplicando-as pelo desvio padrão elevado ao respectivo expoente e uma constante. Na prática, isso não altera os modelos apresentados anteriormente.

O autor mostra que com as restrições apropriadas seu modelo equivale aos estudados anteriormente no caso  $p = q = 1$ . Isto é:

**Tabela 1**  
**Parâmetros e Modelo**

$\lambda$	$\nu$	$B$	$c$	Modelo
0	1	0	Livre	EGARCH
1	1	0	$ c  \leq 1$	TGARCH
2	2	0	0	GARCH
2	2	Livre	0	NGARCH (Engle e Ng)
2	2	0	Livre	GJR GARCH
Livre	$\lambda$	0	0	NGARCH
Livre	$\lambda$	0	$ c  \leq 1$	A-P GARCH

Substituindo as restrições na equação (35), chega-se aos modelos apresentados anteriormente, à exceção dos fatores multiplicativos do erro.

A generalização do modelo permite que haja dois tipos possíveis de assimetria: o parâmetro  $b$  controla a magnitude de assimetria na curva de variância com relação à sua mudança horizontal (ver figuras a seguir). O parâmetro  $c$ , por sua vez, permite que a assimetria se dê com relação à inclinação da curva de variância. Além disso, o modelo permite que a potência da especificação seja livre.

Para entender melhor, considerem-se os seguintes casos:

#### 8.1.1 Caso 1. $\nu = 1, c = b = 0$

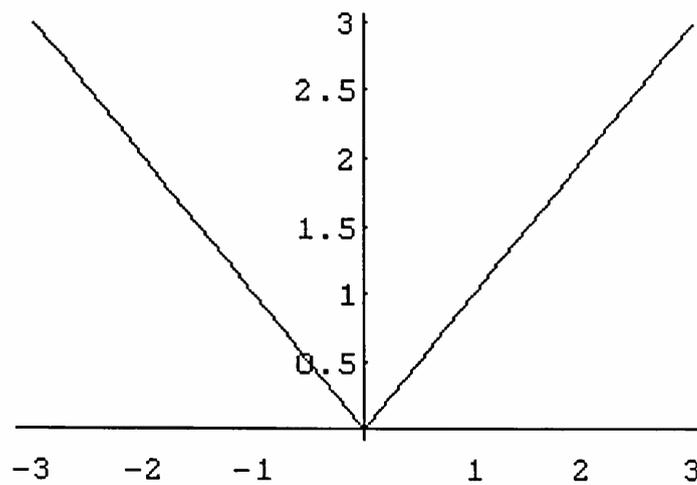
Neste caso, é fácil ver que:

$$\varepsilon_t > 0 \Rightarrow f(\varepsilon_t) = \varepsilon_t$$

$$\varepsilon_t < 0 \Rightarrow f(\varepsilon_t) = -\varepsilon_t$$

Duas retas de 45° ortogonais, com origem em zero são apresentadas no Gráfico 2. Este é o caso trivial em que impactos positivos e negativos têm o mesmo efeito sobre a variância.

**Gráfico 2**  
 $v = 1, c = b = 0$



### 8.1.2 Caso 2: $v = 1, c = 0, b = 1$

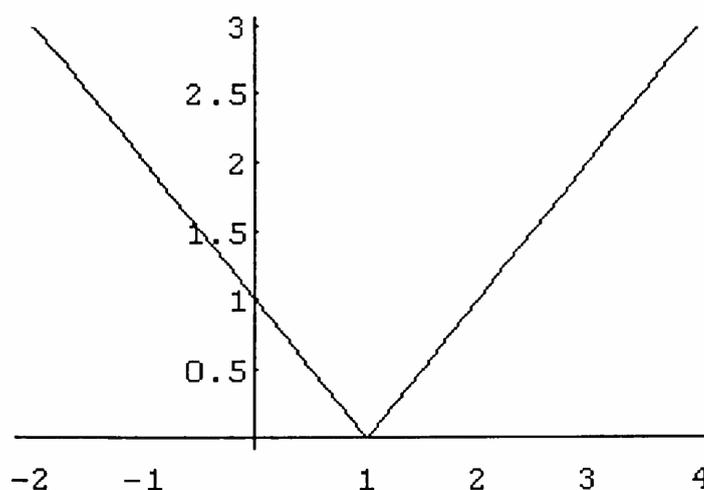
Agora tem-se:

$$\varepsilon_t - b > 0 \Rightarrow f(\varepsilon_t) = \varepsilon_t - b$$

$$\varepsilon_t - b < 0 \Rightarrow f(\varepsilon_t) = -\varepsilon_t + b$$

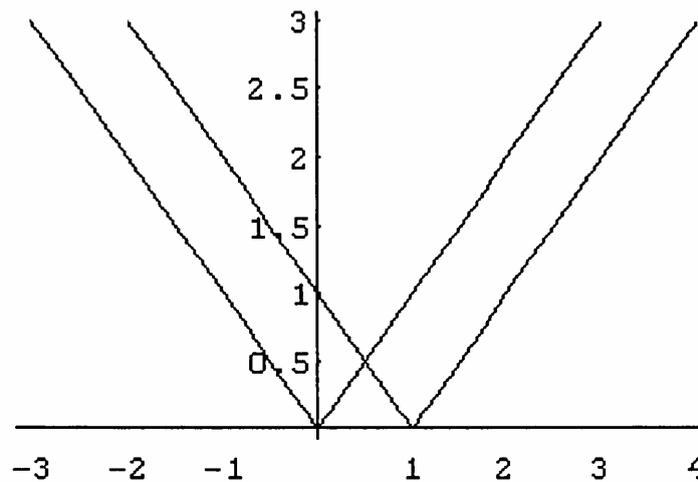
Simplemente há um deslocamento para a direita das curvas do exemplo anterior. Em outras palavras, **choques próximos a zero aumentam a volatilidade** e choques próximos a um não a alteram significativamente. Este caso é interessante, porque nenhum modelo contemplou casos assim anteriormente.

**Gráfico 3**  
 $v = 1, c = 0, b = 1$



O caso em que  $b < 0$  é simétrico ao anterior e não será discutido aqui. O Gráfico 4 junta os dois anteriores e permite que se observe a assimetria horizontal.

**Gráfico 4**  
**Assimetria Horizontal**



### 8.1.3 Caso 3: $v = 1$ , $b = 0$ , $0 < c < 1$

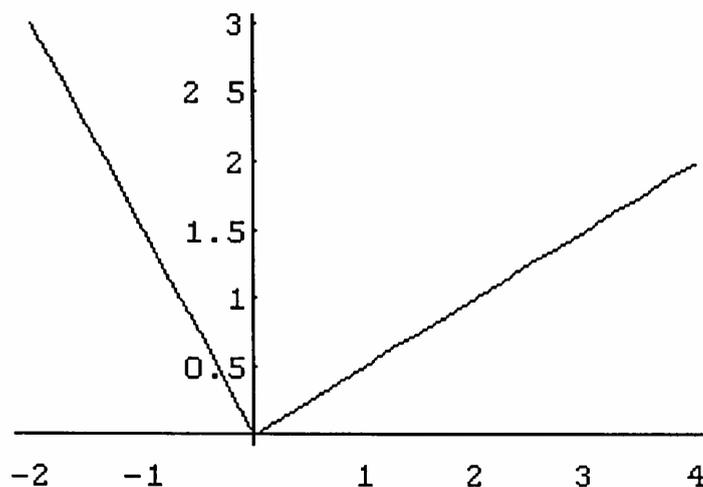
Neste caso obtém-se:

$$\varepsilon_t > 0 \Rightarrow f(\varepsilon_t) = \varepsilon_t - c\varepsilon_t = (1-c)\varepsilon_t$$

$$\varepsilon_t < 0 \Rightarrow f(\varepsilon_t) = -\varepsilon_t - c\varepsilon_t = (-1-c)\varepsilon_t$$

Está havendo uma mudança na inclinação da curva de variância, o que significa que choques negativos têm seu efeito aumentado sensivelmente no que refere à variância. Choques positivos, por sua vez, pouco aumentam a variância. O gráfico a seguir mostra isso.

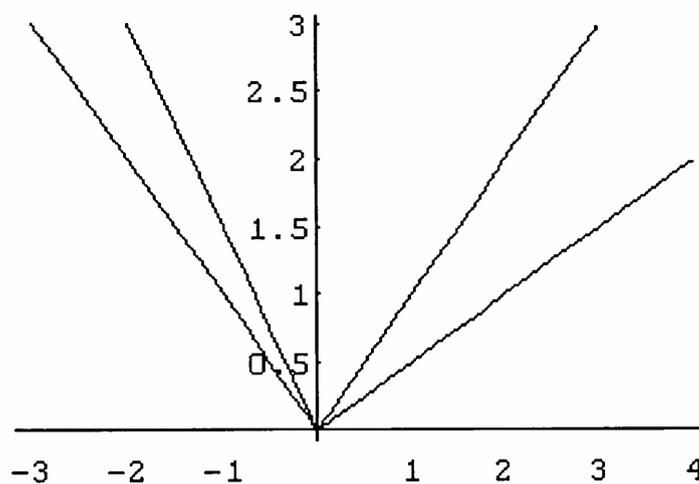
**Gráfico 5**  
 $v = 1, b = 0, 0 < c < 1$



Se  $-1 < c < 0$ , o efeito é simétrico e não será discutido aqui.

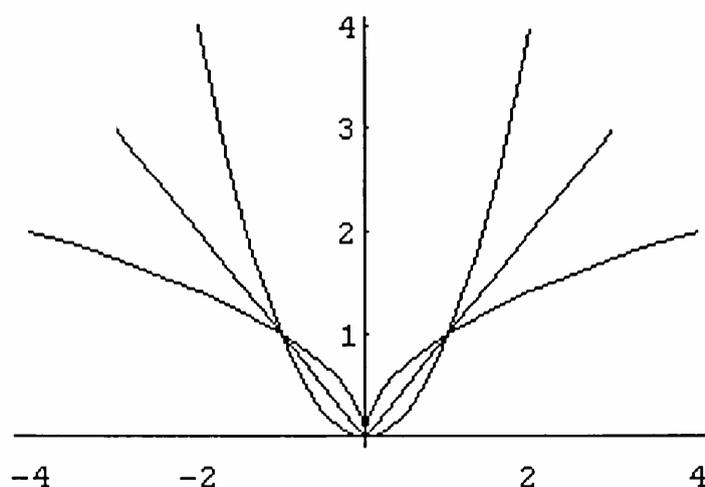
O Gráfico 6 mostra os Gráficos 2 e 5 no mesmo plano, podendo-se perceber que a rotação se dá em sentido horário quando isso acontece:

**Gráfico 6**  
**Assimetria Angular**



É possível verificar que se  $v = 1$ , a curva de variância será uma reta; se  $v = 2$ , a curva de variância será convexa; e se  $v = 0,5$ , a curva de variância será côncava.

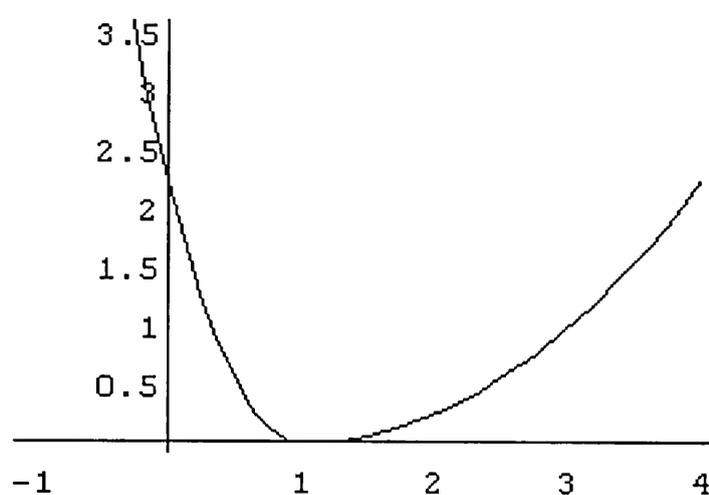
**Gráfico 7**  
Vários Valores para  $\nu$



O uso desse modelo mais geral deve ser estudado porque os modelos GARCH apresentados anteriormente não consideram simultaneamente assimetrias na origem e na inclinação. Pode-se mostrar que os modelos anteriores ou calculam um tipo ou outro de assimetria. Por exemplo, Hentschel mostra que ao deixar o parâmetro  $b$  livre o EGARCH captaria os dois tipos de assimetria.

Para efeitos de ilustração, considere-se o modelo com  $\nu = 2$ ,  $b > 0$ ,  $0 < c < 1$ .

**Gráfico 8**  
 $\nu = 2$ ,  $b > 0$ ,  $0 < c < 1$



A interpretação é a mesma para o caso linear, ou seja, choques menores do que um aumentam a volatilidade exponencialmente. Choques positivos maiores do que 1 têm um efeito menor no aumento da volatilidade.

## 9 Generalização dos modelos GARCH para o contexto multivariado

A segunda generalização possível dos modelos GARCH é adaptá-los a espaços de dimensão maior do que um. São os modelos GARCH multivariados. O modelo GARCH Multivariado é denotado pelo prefixo “M”

Nesta seção apresentam-se as três possibilidades mais importantes. Trata-se do modelo proposto por Bollerslev, Engle e Wooldridge (1988), que estenderam o modelo GARCH-M para o contexto multivariado usando o operador Vech, denotado por MGARCH-M. O outro possível modelo multivariado é o Bekk, devido a Baba, Engle, Kraft e Kroner (ver Engle e Kroner, 1995). Finalmente, apresenta-se o modelo de correlação condicional constante, devido a Bollerslev (1990) e Baillie e Bollerslev (1990).

### 9.1 O Modelo Vech

A matriz de covariância dos retornos é simétrica. Logo, os elementos diferentes entre si estão no triângulo, digamos, inferior dessa matriz. A técnica de estimação consiste, pois, em aplicar um operador nessa matriz de tal sorte que se fique apenas com os elementos do triângulo inferior em forma de vetor. Este é o chamado operador Vech.<sup>11</sup> Por exemplo, a matriz de covariância de um modelo com  $n$  variáveis explicativas no tempo  $t$  é assim descrita:

$$\Sigma_t = \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} & \cdots & \sigma_{1n,t} \\ \sigma_{21,t} & \sigma_{22,t} & \cdots & \sigma_{2n,t} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1,t} & \sigma_{n2,t} & \cdots & \sigma_{nn,t} \end{bmatrix} \quad (36)$$

Como a matriz de covariância é simétrica, pode-se aplicar o operador Vech para empilhar a porção triangular inferior em um vetor de dimensão  $\frac{n(n+1)}{2} \times 1$ . Ou seja, coloca-se a primeira coluna sobre a segunda; o resultado sobre a terceira; e assim sucessivamente:

<sup>11</sup> O operador Vech procede da mesma maneira, mas transforma todos os elementos da matriz em um único vetor, ou seja, não desconsideraria o triângulo superior da matriz de covariância (ver detalhes em Lütkepohl, 1991)

$$Vech(\Sigma_t) = \begin{bmatrix} \sigma_{11,t} \\ \sigma_{21,t} \\ \vdots \\ \sigma_{n1,t} \\ \sigma_{22,t} \\ \sigma_{32,t} \\ \vdots \\ \sigma_{n2,t} \\ \vdots \\ \sigma_{nn,t} \end{bmatrix} \quad (37)$$

O modelo MGARCH-M em Vech pode ser assim formulado:

$$y_t = b + \Sigma_t \delta + \varepsilon_t \quad (38)$$

$$Vech(\Sigma_t) = C + \sum_{i=1}^q A_i Vech(\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{j=1}^p B_j Vech(\Sigma_{t-j}) \quad (39)$$

onde:

$y_t$  é um vetor  $n \times 1$ ,

$b$  é um vetor de médias (ou constantes)  $n \times 1$ ,

$\delta$  é um vetor  $n \times 1$ ,

$\varepsilon_t | I_{t-1} \sim N(0, \Sigma_t)$  é um vetor  $n \times 1$  de perturbações aleatórias, que segue um processo MGARCH;

$C$  é um vetor  $\frac{n(n+1)}{2} \times 1$ ,

$A_i$  é uma matriz  $\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}$ ,

$B_i$  é uma matriz  $\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}$ .

Abrindo as equações matriciais anteriores tem-se:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{n,t} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} + \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} & \cdots & \sigma_{1n,t} \\ \sigma_{21,t} & \sigma_{22,t} & \cdots & \sigma_{2n,t} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1,t} & \sigma_{n2,t} & \cdots & \sigma_{nn,t} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{n,t} \end{bmatrix} \tag{40}$$

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{21,t} \\ \vdots \\ \sigma_{n1,t} \\ \sigma_{22,t} \\ \sigma_{32,t} \\ \vdots \\ \sigma_{n2,t} \\ \vdots \\ \sigma_{nn,t} \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \\ c_{22} \\ c_{32} \\ \vdots \\ c_{n2} \\ \vdots \\ c_{nn} \end{bmatrix} + \sum_{i=1}^q \begin{bmatrix} a_{11,i} & a_{12,i} & \cdots & a_{1 \frac{n(n+1)}{2},i} \\ a_{21,i} & a_{22,i} & \cdots & a_{2 \frac{n(n+1)}{2},i} \\ \vdots & \vdots & & \vdots \\ a_{\frac{n(n+1)}{2} 1,i} & a_{\frac{n(n+1)}{2} 2,i} & \cdots & a_{\frac{n(n+1)}{2} \frac{n(n+1)}{2},i} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-i}^2 \\ \varepsilon_{2,t-i} \varepsilon_{1,t-i} \\ \vdots \\ \varepsilon_{n,t-i} \varepsilon_{1,t-i} \\ \varepsilon_{2,t-i}^2 \\ \varepsilon_{3,t-i} \varepsilon_{2,t-i} \\ \vdots \\ \varepsilon_{n,t-i} \varepsilon_{2,t-i} \\ \vdots \\ \varepsilon_{n,t-i}^2 \end{bmatrix} +$$

$$\sum_{j=1}^p \begin{bmatrix} b_{11,j} & b_{12,j} & \cdots & b_{1 \frac{n(n+1)}{2},j} \\ b_{21,j} & b_{22,j} & \cdots & b_{2 \frac{n(n+1)}{2},j} \\ \vdots & \vdots & & \vdots \\ b_{\frac{n(n+1)}{2} 1,j} & b_{\frac{n(n+1)}{2} 2,j} & \cdots & b_{\frac{n(n+1)}{2} \frac{n(n+1)}{2},j} \end{bmatrix} \begin{bmatrix} \sigma_{11,t-j} \\ \sigma_{21,t-j} \\ \vdots \\ \sigma_{n1,t-j} \\ \sigma_{22,t-j} \\ \sigma_{32,t-j} \\ \vdots \\ \sigma_{n2,t-j} \\ \vdots \\ \sigma_{nn,t-j} \end{bmatrix} \tag{41}$$

Na formulação acima o número de parâmetros a serem estimados, apenas na equação da variância, é de  $\left[ \frac{n(n+1)}{2} \right]^2 (p+q) + \frac{n(n+1)}{2}$ . Assim, por serem muitos, geralmente impõe-se que os elementos das diagonais secundárias das matrizes  $A_i$  e  $B_i$  sejam iguais a zero.

Para se ter uma idéia do número de parâmetros que devem ser estimados quando não se impõe a restrição sugerida, se  $n = 4$  e  $p = q = 1$ , devem ser estimados 210 parâmetros na equação da variância. Impondo-se 0 nas diagonais secundárias, então o número de parâmetros

a serem estimados reduz-se drasticamente para  $\frac{n(n+1)}{2}(p+q+1)$  (30 no exemplo anterior), proporcionando uma economia de  $\frac{n(n+1)}{2}(p+q) \left[ \frac{n(n+1)}{2} - 1 \right]$  parâmetros (ou seja, 180 parâmetros deixam de ser estimados, no exemplo dado).

No caso bivariado, usando-se o MGARCH (1, 1) o número de parâmetros reduz-se de 21 para 9. As equações da covariância para esse modelo são:

$$\begin{cases} \sigma_{11,t} = c_{11} + a_{11,1}\varepsilon_{1,t-1}^2 + b_{11,1}\sigma_{11,t-1} \\ \sigma_{12,t} = c_{12} + a_{12,1}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + b_{12,1}\sigma_{12,t-1} \\ \sigma_{22,t} = c_{22} + a_{22,1}\varepsilon_{2,t-1}^2 + b_{22,1}\sigma_{22,t-1} \end{cases} \quad (42)$$

Ou seja, a variância de uma variável é explicada apenas pelo quadrado de seus erros passados e de sua variância passada, enquanto que a covariância das equações é explicada pelas covariâncias passadas e pelo produto cruzados dos erros passados.

Engle e Kroner (1995) afirmam que a imposição dessas restrições parece ser plausível intuitivamente porque as informações sobre as variâncias são usualmente reveladas nos resíduos quadrados passados e nas variâncias passadas; e se as variâncias alteram-se vagarosamente, o resíduos quadrados passados e as variâncias passadas devem ser capazes de prever as variâncias futuras. O mesmo comportamento aconteceria no caso das covariâncias. Isso significa afirmar que as informações sobre as variâncias dadas pelas covariâncias, e sobre estas dadas por aquelas, são negligenciáveis.

## 9.2 O modelo Bekk

Além do problema do excessivo número de parâmetros a serem estimados, é preciso garantir que a matriz de covariância seja positiva semidefinida em cada momento do tempo (ver outros questionamentos em Gouriéroux, 1997). Mas isso não é possível por meio da formulação anterior de maneira prática. (Engle e Kroner, 1995) Condições que crescem na proporção do crescimento do número de parâmetros têm que ser impostas.

Engle e Kroner (1995) impõem a restrição de matrizes positivas definidas. Esse tipo de imposição é pouco restritivo, pois pouquíssimos modelos Vech não poderiam ser substituídos pelo Bekk. Formalmente:

$$\Sigma_t = F'F + \sum_{i=1}^q D_i' \varepsilon_{t-i} \varepsilon_{t-i}' D_i + \sum_{j=1}^p E_j' \Sigma_{t-j} E_j \quad (43)$$

onde  $F, D_i, E_j$  são matrizes  $n \times n$ .

Exemplificando-se para  $n = 2, p = q = 1$ , tem-se:

$$\begin{aligned} \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} &= \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}' \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} + \\ &+ \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}' \begin{bmatrix} \varepsilon_{1,t-1}^2 & \varepsilon_{1,t-1} \varepsilon_{2,t-1} \\ \varepsilon_{2,t-1} \varepsilon_{1,t-1} & \varepsilon_{2,t-1}^2 \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} + \\ &+ \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}' \begin{bmatrix} \sigma_{11,t-1} & \sigma_{12,t-1} \\ \sigma_{21,t-1} & \sigma_{22,t-1} \end{bmatrix} \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}. \end{aligned} \quad (44)$$

Nessa formulação o número de parâmetros a serem estimados reduz-se de 21 para 12, garantindo-se que a matriz de covariância será positiva semidefinida em cada momento do tempo. Genericamente, o número de parâmetros a serem estimados nesse modelo é

de  $n^2(p + q + 1)$ , o que proporciona uma economia de  $\frac{n(n-1)}{4} [(p+q)n(n+3) - 2]$  parâmetros que não serão estimados, relativamente ao modelo Vech.<sup>12</sup>

### 9.3 O modelo de correlação condicional constante

As dificuldades para se estimar os modelos Vech e Bekk são inúmeras, particularmente em termos computacionais. Bollerslev (1990) propôs um modelo em que as covariâncias condicionais são parametrizadas de modo a serem proporcionais ao produto dos desvios padrões correspondentes. (Bollerslev, Engle e Nelson, 1994) Isso diminui a intensidade computacional requerida para estimar o modelo e torna mais fácil impor as condições necessárias para garantir que a matriz de covariância seja positiva semidefinida.

Considere-se a matriz diagonal  $D_t$ ,  $n \times n$ , com covariâncias condicionais ao longo de sua diagonal. Ou seja,  $\{D_t\}_{ii} = \{\Sigma_t\}_{ii}$  e  $\{D_t\}_{ij} \equiv 0, \forall i \neq j, i, j = 1, 2, \dots, n$ . Seja  $\Gamma_t$  a matriz de correlação condicional, isto é  $\{\Gamma_t\}_{ij} = \{\Sigma_t\}_{ij} \left[ \{\Sigma_t\}_{ii} \{\Sigma_t\}_{jj} \right]^{-\frac{1}{2}}, i, j = 1, 2, \dots, n$ .

O modelo de correlação condicional constante assume, então, que  $\Gamma_t = \Gamma$  é invariante no tempo, de tal modo que a variação temporal em  $\Sigma_t$  é determinada somente pelas variâncias condicionais. Dessa maneira:

$$\Sigma_t = D_t^{\frac{1}{2}} \Gamma D_t^{\frac{1}{2}} \quad (45)$$

Considerando que a variância condicional ao longo de  $D_t$  devem ser todas positivas e assumindo que  $\Gamma$  seja positiva definida, a seqüência de matrizes de covariância condicional,  $\{\Sigma_t\}_{t=1, T}$ , é positiva definida quase certamente para todo  $t$ . Além disso, a inversa de  $\Sigma_t$  é simplesmente:

12 As inter-relações entre os modelos Vech e Bekk estão formuladas em Engle e Kroner (1995).

$$\Sigma_t^{-1} = D_t^{-\frac{1}{2}} \Gamma^{-1} D_t^{-\frac{1}{2}} \quad (46)$$

Com essa formulação é possível concentrar a função de verossimilhança em  $\Gamma$ , simplificando bastante a estimação. Além disso, não é mais preciso inverter a matriz  $\Sigma_t$   $T$  vezes; basta apenas a inversão da matriz  $\Gamma$ , haja vista que as matrizes  $D_t$  são diagonais. (Gouriéroux, 1997) Por exemplo, no caso bivariado tem-se a seguinte formulação:

$$\Sigma_t = \begin{bmatrix} \sqrt{\sigma_{11,t}} & 0 \\ 0 & \sqrt{\sigma_{22,t}} \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \sqrt{\sigma_{11,t}} & 0 \\ 0 & \sqrt{\sigma_{22,t}} \end{bmatrix} \quad (47)$$

Segundo Bera e Higgins (1995), é óbvio que a principal restrição a essa formulação é a forte hipótese de que a correlação é constante. Bera e Roth (1991, *apud* Bera e Higgins, 1995) sugerem um teste para verificar se a correlação é constante e rejeitam esta hipótese nula em muitas séries financeiras.

Outra formulação possível são os modelos fatoriais (ver Bera e Higgins, 1995; Bollerslev, Engle e Nelson, 1994; Gouriéroux, 1997 e suas referências).

#### 9.4 Estimação

Definida a formulação do modelo, é preciso estimá-lo. No caso de se supor normalidade multivariada para os resíduos, a função de verossimilhança linearizada condicionada às  $\mu$  matrizes de covariância iniciais assim se apresenta (ver Bollerslev, Engle e Nelson, 1994):

$$l(\Theta; \Psi) = -n \frac{T - \mu}{2} \ln 2\pi - \frac{1}{2} \sum_{t=\mu+1}^T \ln |\Sigma_t| - \sum_{t=\mu+1}^T \frac{\varepsilon_t' \Sigma_t^{-1} \varepsilon_t}{2} \quad (48)$$

onde:

$\Psi$  é o vetor de parâmetros a serem estimados nas equações de covariância;

$$\Sigma_1 = \Sigma_2 = \dots = \Sigma_\mu = \frac{\sum_{t=1}^T \varepsilon_t' \varepsilon_t}{T},$$

$$\mu = \max[p, q]$$

Conforme atestam Bollerslev, Engle e Wooldridge (1988), a estimativa de máxima verossimilhança para os parâmetros, dadas algumas condições de regularidade, será assintoticamente normal e não-viesada com matriz de covariância igual à inversa da matriz de informação de Fischer (ver Ramanathan, 1993) e, portanto, os procedimentos tradicionais de inferência podem ser imediatamente procedidos.

No caso de se considerar como distribuição condicional a *t-Student* multivariada, assim será a sua função de verossimilhança:

$$l(\Theta; \Psi) = -\frac{T-\mu}{2} \ln \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{(\nu-2)\pi}} - \frac{1}{2} \sum_{t=\mu+1}^T \ln |\Sigma_t| - \frac{\nu+n}{2} \sum_{t=\mu+1}^T \ln \left( 1 + \frac{\varepsilon_t' \Sigma_t^{-1} \varepsilon_t}{\nu-2} \right) \quad (49)$$

### 9.5 Restrições e inferência

Para que uma das funções acima seja maximizada é preciso impor restrições para garantir que a matriz de covariância seja positiva definida. Por isso, estima-se a função de verossimilhança sujeita à restrição de que os autovalores da matriz de covariância condicional sejam simultaneamente maiores do que zero.

Para garantir que a série seja estacionária é preciso que as raízes da seguinte polinomial estejam fora do círculo unitário:

$$\left| I - \sum_{i=1}^{\max[p,q]} (\alpha_i + \beta_i) \lambda^i \right| = 0 \quad (50)$$

onde  $\lambda$  é o autovalor associado a cada dos autovetores.

Como os parâmetros são restritos, surgem sérias dificuldades para a inferência estatística. O método usual de inferência estatística, o cálculo da matriz de covariância e construção da estatística *t*, a partir dos desvios padrões dos parâmetros, não funciona quando as restrições

envolvem desigualdades. Isso se deve ao fato de que as regiões de confiança não serão simétricas em relação às estimativas. Para superar essa dificuldade pode-se computar os limites de confiança por intermédio de simulações (ver Schoenberg, 1997).

## 9.6 Conclusão

Além das objeções levantadas anteriormente para a formulação do MGARCH, Bollerslev, Engle e Nelson (1994) adicionam as seguintes: é preciso reconhecer como se dá a causalidade de Granger na variância, *i.e.*, verificar se existe alguma formulação passada que prediz melhor do que outra. A outra consideração é verificar se as correlações dos coeficientes de regressão são variantes no tempo e, se são, verificar se têm as mesmas propriedades de persistência como as variâncias. Uma terceira preocupação deve ser com relação à combinação linear de variáveis ou carteiras com menor persistência do que as séries ou ativos individualmente. Finalmente, a mais natural questão trata de como se propagam os efeitos assimétricos nesse espaço multivariado.

A definição do modelo GARCH multivariado a ser usado pode seguir a mesma formulação da seção 4, considerando-se os determinantes da variância e o produto de  $\varepsilon_t$  quando pós-multiplicado pelo seu transposto. Por exemplo, a equação (26) fica, no contexto multivariado, da seguinte forma:

$$E_1 = \sum_{t=1}^T |\varepsilon_t \varepsilon_t' - \Sigma_t| \quad (51)$$

## 10 Considerações finais

Neste artigo procurou-se apresentar sistematicamente os modelos GARCH, inclusive mostrando-se duas extensões importantes relativas à especificação univariada mais geral e à multivariada. O propósito foi facilitar a compreensão de seu funcionamento e ressaltar alguns problemas teóricos e práticos que surgem quando se trabalha com dados empíricos. Detalhes a respeito de especificação, verificação e estimação foram abordados.

Acredita-se que a contribuição deveu-se também ao fato de trazer para língua portuguesa, pela primeira vez, as diversas variantes dos modelos GARCH. Em língua inglesa, trabalhos desse tipo encontram-se em Bera e Higgins (1995) e Bollerslev, Engle e Nelson (1994), além de inúmeros capítulos de livros dedicados especialmente a este assunto.

Em termos de resultados empíricos, grande parte dos artigos citados ao longo deste trabalho mostra alguma aplicação. Adicionalmente, Issler (1999) usa os modelos GARCH univariados em séries financeiras brasileiras; Bueno (1999) aplica modelos multivariados para mercados futuros brasileiros; Bansal e Yaron (2000) apresentam um exemplo de aplicação em finanças para modelos de equilíbrio geral com agente representativo, como forma de replicar fatos estilizados observados no mercado financeiro.

Novos modelos e testes continuam sendo desenvolvidos, como, por exemplo, o modelo com variância condicional e mudança de regime (ver Franses e Dijk, 2000 e Hamilton e Susmel, 1994), e testes quando a equação da média tem sido mal especificada (ver Lumsdaine e Ng, 1999).

## Bibliografia

- Baillie, Richard T., Bollerslev, Tim. A multivariate generalized ARCH approach to modeling risk premia in forward foreign rate markets. *Journal of International Money and Finance*, v. 9, p. 309-324, 1990.
- Bansal, Ravi; Yaron, Amir. Risks for the long run: a potential resolution of asset pricing puzzles. *Working Paper*, The Wharton School, 2000.
- Berndt, E. K., Hall, B. H., Hall, R. E., Hausman, J. A. Estimation and inference in nonlinear structural models. *Journal of Economic and Social Measurement*, v. 3/4, p. 653-665, 1974.
- Bera, Anil K., Higgins, Mathew L. On ARCH models: properties, estimation and testing. In: Oxley, Les et alii. *Surveys in econometrics*. Oxford: Blackwell, 1995.
- Bollerslev, Tim. Modelling coherence in the short-run nominal exchange rates: a multivariate generalized ARCH approach. *Review of Economics and Statistics*, v. 72, p. 498-505, 1990.
- Bollerslev, Tim; Chou, Ray Y., Kroner, Kenneth F. ARCH modeling in finance. *Journal of Econometrics*, v. 52, n. 1, p. 5-59, 1992.
- Bollerslev, Tim; Engle, Robert F., Nelson, Daniel B. Arch models, ch. 49. In: Engle, Robert F.; Mcfadden, Daniel L., *Handbook of econometrics*, v. 4. Amsterdam: Elsevier Science, 1994.
- Bollerslev, Tim; Engle Robert F., Wooldridge, Jeffrey M. A capital asset pricing model with time varying covariances. *Journal of Political Economy*, v. 96, n. 1, p. 116-131, 1988.

- Bollerslev, Tim. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, v. 31, n. 3, p. 307-327, 1986.
- \_\_\_\_\_. A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, v. 69, p. 542-547, 1987
- Box, George E., Jenkins, Gwilyn M., Reinsel, Gregory C. *Time series analysis: forecasting and control*. Englewood Cliffs: Prentice Hall, 1994.
- Brockwell, Peter; Davis, Richard A. *Time series: theory and methods*. 2.<sup>nd</sup> ed. [5.<sup>th</sup> printing]. New York: Springer-Verlag, 1991.
- Bueno, Rodrigo D. L. S. *GARCH multivariado e taxa ótima de "hedge"* 1999. Dissertação (Mestrado). IPE-FEA/USP, São Paulo.
- Davidson, Russel; Mackinnon, James G. *Estimation and inference in econometrics*. New York: Oxford, 1993.
- Enders, Walter. *Applied econometrics time series*. New York: Wiley, 1995.
- Engle, Robert F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, v. 50, n. 4, p. 987-1007, 1982.
- Engle, Robert F.; Kroner, Kenneth F. Multivariate simultaneous generalized ARCH. *Econometric Theory*, v. 11, p. 122-150, 1995.
- Engle, Robert F., Lilien, David M., Robins, Russel P. P. Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica*, v. 55, n. 2, p. 391-407, 1987
- Franses, Philip H.; Dijk, Dick van. *Non-linear time series models in empirical finance*. Cambridge: Cambridge, 2000.
- Gouriéroux, Christian. *ARCH models and financial applications*. New York: Springer-Verlag, 1997.
- Hegerud, Gustaf E. Specification tests for asymmetric GARCH. Unpublished paper, Department of Finance, Stockholm School of Economics, Jan/1997.
- Hamilton, James D. *Time series analysis*. Princeton: Princeton, 1994.
- Hamilton, James D.; Susmel, Raul. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, v. 64, n. 1-2, p 307-33, 1994.

- Harvey, Andrew C. *Forecasting structural time series models and the Kalman filter*. Cambridge: Cambridge, 1989.
- Harvey, Andrew; Ruiz, Esther; Sentana, Enrique. Unobserved component time series models with ARCH disturbances. *Journal of Econometrics*, v. 52, p. 129-157, 1992.
- Harvey, Andrew; Ruiz, Esther; Shephard, Neil. Multivariate stochastic variance models. *Review of Economic Studies*, v. 61, p. 247-264, 1994.
- Hentschel, Ludger. All in the family nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics*, v. 39, p. 71-104, 1995.
- Heynen, Ronald C. Volatility prediction: a comparison of the stochastic volatility, GARCH (1, 1), and EGARCH (1, 1) models. *Journal of Derivatives*, p. 50-65, Winter 1994.
- Issler, João V. Estimating and forecasting the volatility of Brazilian finance series using ARCH models. *Revista de Econometria*, v. 19, n. 1, p. 5-56, 1999.
- Johnston, Jack; Dinardo, John. *Econometric methods*. 4.<sup>th</sup> ed. New York: McGraw-Hill, 1997.
- Judge *et alii*. *Introduction to the theory and practice of econometrics*. 2.<sup>nd</sup> ed. New York: Wiley, 1988.
- Lee, John H. H. A Lagrange multiplier test for GARCH models. *Economic Letters*, v. 37, p. 265-271, 1991.
- Lintner, John. The valuation and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economic and Statistics*, v. 47, p. 13-37, 1965.
- Lumsdaine, Robin L. Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1, 1) and covariance stationary GARCH(1, 1). *Econometrica*, v. 64, p. 561-573, 1996.
- Lumsdaine, Robin L., Ng, Serena. Testing for ARCH in the presence of a possibly misspecified conditional mean. *Journal of Econometrics*, v. 93, n. 2, p. 257-79, 1999.
- Lütkepohl, H. *Introduction to multiple time series analysis*. New York: Springer-Verlag, 1991.
- Mills, Terence C. *Time series techniques for economists*. Cambridge: Cambridge, 1990.
- \_\_\_\_\_. *The econometric modelling of financial time series*. 2.<sup>nd</sup> ed. Cambridge: Cambridge, 1999.

Nelson, Daniel B. Conditional heteroscedasticity in asset returns. *Econometrica*, v. 59, n. 2, p. 347-370, 1991.

\_\_\_\_\_. Modelling stock market volatility changes. ASA: Proceedings of the Business and Economics Statistics Section, p. 93-98, 1989. In: Rossi, Peter E. *Modelling stock market volatility: bridging the gap to continuous time*. San Diego: Academic Press, 1996.

Nelson, Daniel B., Cao, Charles Q. Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics*, v. 10, n. 2, p. 229-235, 1992.

Priestley, Maurice B. *Spectral analysis and time series*, v. 1/2. London: Academic Press, 1981.

Ramanathan, Ramu. *Statistical methods in econometrics*. San Diego: Academic Press, 1993.

Schoenberg, Ronald. Simulation of Bayesian posterior distributions of parameters of constrained models. Unpublished paper, The University of Washington, Aug/1997

Sharpe, William F. Capital asset prices: a theory of market equilibrium under conditions of risk. *The Journal of Finance*, v. 19, n. 3, p. 425-442, 1964.

Spanos, Aris. *Statistical foundations of econometric modelling*. Cambridge: Cambridge, 1986.

White, Halbert. A heteroskedasticity consistent covariance estimator and a direct test for heteroskedasticity. *Econometrica*, v. 48, n. 4, p. 817-838, 1980.

\_\_\_\_\_. *Asymptotic theory for econometricians*. 2.<sup>nd</sup> ed. San Diego: Academic Press, 2000.

