

Modelos de Variabilidade
Estocástica
e
Deformação Temporal

Flávio Augusto Ziegelmann

Dissertação apresentada ao
Instituto de Matemática e Estatística
da Universidade de São Paulo
para a obtenção do grau de
Mestre em Estatística

Área de Concentração: Séries Temporais e Econometria
Orientador: Pedro L. Valls Pereira

São Paulo, setembro de 1996

Modelos de Variabilidade Estocástica
e
Deformação Temporal

Este exemplar corresponde à redação
final da dissertação devidamente corrigida
e apresentada por Flávio Augusto Ziegelmann e aprovada
pela Comissão Julgadora.

São Paulo, 23 de setembro de 1996.

Banca Examinadora:

- Prof. Dr. Pedro Luiz Valls Pereira (Orientador) - IME-USP
- Prof. Dr. Luiz Koodi Hotta - IMECC-UNICAMP
- Prof. Dr. Marcos Eugenio da Silva - FEA-USP

Sumário

Agradecimentos	iv
Resumo	vi
Introdução	1
1 Volatilidade e Características de Dados Financeiros	3
1.1 Volatilidade: Sua Importância	3
1.2 Fatos Estilizados	4
1.3 Terminologias Básicas de Análise de Séries Temporais	5
1.4 Características dos Retornos do IBOVESPA	8
2 Variabilidade Condicional Estocástica	12
2.1 Modelos ARCH	13
2.1.1 Características Gerais	14
2.1.2 Estimação	15
2.1.3 Previsão	16
2.1.4 Concordância com os Fatos Estilizados	17

2.2	Modelos GARCH	17
2.2.1	Características Gerais	18
2.2.2	Estimação	19
2.2.3	Previsão	20
2.2.4	Concordância com os Fatos Estilizados e o Formato ARCH(∞) do GARCH(1,1)	20
2.3	Outros Modelos da Família ARCH	21
2.3.1	AGARCH (GARCH Assimétrico)	21
2.3.2	EGARCH (GARCH Exponencial)	22
2.3.3	IGARCH (GARCH Integrado)	22
2.3.4	ARCH-M (ARCH em Média)	23
2.4	Modelos de Volatilidade Estocástica	23
2.4.1	Propriedades Básicas	24
2.4.2	Estimação	26
3	Modelos de Volatilidade Estocástica com Deformação Temporal	37
3.1	Por que Deformação Temporal?	37
3.2	Metodologia	38
3.2.1	Modelagem a Tempo Discreto	38
3.2.2	Modelagem a Tempo Contínuo	39
3.3	Conjuntos de Dados em Alta Frequência	47
4	Aplicação: Estimação da Volatilidade dos Retornos do IBOVESPA	48

SUMÁRIO

iii

Conclusão

60

Apêndice

62

Referências Bibliográficas

64

Índice

67

Agradecimentos

Este trabalho teve a duração de aproximadamente um ano, tendo sido iniciado após o término de meus créditos obrigatórios de mestrado. Inicialmente, houve uma dificuldade acentuada na parte de revisão bibliográfica, principalmente devido à minha inexperiência na área financeira e à vasta e recente literatura existente sobre volatilidade estocástica. Na parte aplicada, na modelagem a tempo discreto, o principal pacote computacional utilizado foi o STAMP. Já na modelagem a tempo contínuo, utilizou-se um módulo de maximização do programa GAUSS, sendo que os códigos para a estimação da volatilidade foram em parte gentilmente cedidos pela professora Joana Jasiak e em parte elaborados por mim.

Neste término de dissertação, gostaria de agradecer ao meu orientador, professor Pedro Valls, que em todos os momentos mostrou-se solícito, procurando suprir minhas deficiências sempre que necessário. Meus agradecimentos, em especial, às professoras Mariane Streibel e Clélia Tolo por sua ajuda nos momentos em que precisei. Também, a todos os professores e funcionários com quem convivi durante estes quase dois anos e meio, meu muito obrigado. Finalmente, pelo auxílio financeiro durante este trabalho, agradeço ao CNPq.

À minha noiva Patrícia por tudo que ela foi para mim neste período, pela sua ajuda, carinho e amor em todos os momentos, todo o meu amor.

Muito obrigado a meus pais e a toda minha família pelo amor que sempre me deram e continuam me dando.

Obrigado a meus amigos e colegas que estiveram e continuam a meu lado.

Abstract

Estimating and forecasting volatility of an asset is a very important task in financial markets. Our aim in this work is to cover the most used stochastic conditional variance models and to propose the concept of time deformation in this context. The idea is that the market does not change as calendar time goes by, but as new information arrives.

We also estimate the volatility of IBOVESPA returns, applying Stochastic Volatility Models both without and with Time Deformation.

Resumo

Estimar e prever a volatilidade de um ativo é uma tarefa muito importante em mercados financeiros. Nosso objetivo neste trabalho é cobrir os modelos de variância condicional estocástica mais utilizados e propor o conceito de deformação temporal neste contexto. A idéia é que o mercado modifica-se com a chegada de novas informações, e não com o decorrer do tempo de calendário.

Nós também estimamos a volatilidade dos retornos do IBOVESPA, aplicando Modelos de Volatilidade Estocástica sem e com Deformação Temporal.

Introdução

Estimar e prever de modo adequado a volatilidade de um ativo financeiro é uma tarefa de extrema importância em análise de risco. Convém ressaltar, entretanto, que o enfoque desta dissertação é estatístico, sem a intenção de explorar as aplicações financeiras específicas. Existem 3 objetivos bem delimitados. O primeiro é o de cobrir, ao menos na parte teórica, a maior parte das modelagens atuais existentes formuladas para estimação e previsão de volatilidade condicional estocástica, onde devemos interpretar volatilidade como uma medida geral de variabilidade, quer seja o desvio padrão de uma variável, quer seja sua variância, ou até mesmo o logaritmo da variância. O segundo objetivo é o de apresentar a idéia de “tempo operacional” e “deformação temporal”, incorporando estes conceitos dentro da análise de variabilidade estocástica. O terceiro e final objetivo é o de aplicar algumas das modelagens propostas, especificamente os denominados Modelos de Volatilidade Estocástica sem e com Deformação Temporal, aos dados do índice IBOVSPA (Índice de Bolsas de Valores do Estado de São Paulo), comparando os resultados obtidos.

No capítulo 1, dados financeiros são caracterizados, procurando-se dar uma idéia da importância do cálculo da volatilidade. Também são apresentados alguns conceitos básicos utilizados em Análise de Séries Temporais. No capítulo 2, introduzimos a noção de variabilidade condicional estocástica, explorando as características de dois grandes ramos de modelos, a saber, modelos da família GARCH e modelos de Volatilidade Estocástica. Estes dois ramos basicamente diferem no que diz respeito ao tipo de informação utilizado para estimar a volatilidade. O terceiro capítulo é sobre modelos de Volatilidade Estocástica com Deformação Temporal. O conceito de deformação surge com a idéia de que uma variável financeira, em nosso caso, particularmente, a volatilidade, sofre alterações à medida que novas informações chegam, e não conforme a evolução do tempo de calendário, igualmente espaçado. Dias carregados de novas informações, sejam

elas “boas” ou “ruins”, causam agitação no mercado, enquanto dias “calmos” o deixam estagnado. No último capítulo aplicam-se as modelagens de Volatilidade Estocástica sem e com Deformação Temporal aos dados do IBOVESPA, comparando-se os resultados. Através da abordagem de deformação a tempo contínuo, encontramos que aumento no volume de negociação acelera o tempo operacional, aumentando a volatilidade. Também encontramos que queda dos preços acarreta em maior aceleração do tempo operacional e mais volatilidade do que aumento dos mesmos, ou seja, mercado em queda causa maior agitação nas variáveis financeiras.

Capítulo 1

Volatilidade e Características de Dados Financeiros

Ao longo deste trabalho, estaremos preocupados em, dentro do contexto univariado, estimar e prever a volatilidade (variabilidade) dos retornos de ativos financeiros, os quais serão definidos por $y_t = \log x_t - \log x_{t-1}$, onde x_t é o preço do ativo no instante t . Para tanto, é necessário que incorporemos em nossos modelos as características deste tipo de variável. Neste capítulo, primeiramente forneceremos uma idéia bastante intuitiva da importância do cálculo da volatilidade de um ativo. Após, discutiremos as características empíricas relatadas na literatura e encontradas em dados financeiros (especificamente em retornos de ativos). A seguir, apresentaremos algumas terminologias básicas utilizadas em Análise de Séries Temporais, as quais serão úteis no decorrer do trabalho. Finalmente, verificaremos se nosso conjunto de dados, os retornos do índice de fechamento do IBOVESPA, possui tais características.

1.1 Volatilidade: Sua Importância

Prever a volatilidade de um ativo de modo preciso é a chave para obter sucesso na análise e no gerenciamento de risco. Muitos sistemas financeiros possuem em seu âmago a necessidade destas previsões (ver Alexander (1996)).

A volatilidade nos indica se o preço de um ativo está variando pouco ou muito,

ou seja, ela é uma medida da incerteza quanto às variações do preço. Épocas em que a variabilidade dos preços está muito alta são aquelas que possibilitam os maiores lucros ou, também, as maiores perdas; isto é, são aquelas em que o risco é maior. Já quando a volatilidade é baixa, o risco é menor. A relação é óbvia e, naturalmente, um investidor somente assumiria uma posição de alto risco caso houvesse a possibilidade de um grande retorno. Assim, numa fase de alta volatilidade, entre aqueles que assumem riscos, os “bons” tornam-se melhores, enquanto os “ruins” tornam-se piores.

1.2 Fatos Estilizados

Dados financeiros observados ao longo do tempo apresentam certas características peculiares (ver, por exemplo, Taylor (1986) e Pagan (1993)). Este é o caso dos retornos de ativos financeiros, variável que será utilizada em nosso trabalho. Os principais fatos estilizados (características) encontrados na literatura são os seguintes:

- i) os retornos não são autocorrelacionados;
- ii) os quadrados dos retornos são autocorrelacionados, com uma pequena autocorrelação de primeira ordem e uma subsequente queda bastante lenta;
- iii) existem agrupamentos com diferentes níveis de volatilidade dos retornos ao longo do tempo;
- iv) a distribuição não-condicional dos retornos possui caudas pesadas em relação à distribuição normal;
- v) para alguns tipos de retornos, a volatilidade reage de maneira diferenciada se os preços estão aumentando ou declinando. Sabe-se que declínio no mercado causa maior volatilidade do que ascendência.

Note que estas características guardam muitas relações entre si, dando origem a conclusões semelhantes. Por exemplo, se os retornos tivessem distribuição normal, a característica i) inviabilizaria a ii), pois não-correlação para variáveis gaussianas implica independência de quaisquer funções disjuntas dessas variáveis. Do mesmo modo, a característica iv) aponta diretamente para a não-normalidade ou para a não-independência. De fato, a autocorrelação dos quadrados tem forte relação com as caudas pesadas, pois, quando surgirem valores extremos, quer positivos, quer negativos (ou seja, valores extremos nos quadrados), devido à autocorrelação dos quadrados, haverá tendência de que os próximos quadrados

também sejam extremos, originando um excesso destes extremos, isto é, caudas pesadas.

De forma análoga, as características ii) e iii) estão intimamente relacionadas. Como falamos no parágrafo anterior, se os quadrados dos retornos são autocorrelacionados, isto significa que esperamos que valores grandes (em módulo) dos retornos sejam seguidos por valores também grandes, enquanto valores pequenos situem-se próximos a outros valores pequenos, causando assim os agrupamentos, em diversos níveis, da volatilidade.

Assim, a partir da década passada, muitos modelos estatísticos foram desenvolvidos para obter a volatilidade, incorporando em suas estruturas estes fatos estilizados. Os principais tipos de modelos são os ARCH (“Autoregressive Condicional Heteroskedasticity”), GARCH (“Generalized ARCH”) e de Volatilidade Estocástica (V.E.).

1.3 Terminologias Básicas de Análise de Séries Temporais

A fim de explorarmos as propriedades dos modelos que serão utilizados no trabalho, é conveniente definirmos alguns conceitos básicos empregados em Análise de Séries Temporais. Assim, apresentamos as seguintes definições:

Definição 1.1 O processo y_t será um **ruído branco (RB)** se satisfizer as seguintes condições:

- i) $E[y_t] = \mu$, para $\forall t$;
- ii) $Var[y_t] = \sigma^2$, para $\forall t$;
- iii) $Cov(y_t, y_{t+s}) = 0$, para $\forall t$ e $\forall s \neq 0$.

Neste texto, μ será tomado igual a zero. Se tivermos a condição mais forte de independência entre y_t e y_{t+s} , então y_t será **ruído branco independente**. Ainda, se $y_t \sim N(0; \sigma^2)$, então y_t será **ruído branco gaussiano**.

Definição 1.2 O processo y_t será uma **diferença martingala (DM)** quando satisfizer as seguintes condições:

- i) $\mathbb{E}[|y_t|] < \infty$, para $\forall t$;
 ii) $\mathbb{E}[y_t | f(y_{t-1}, y_{t-2}, \dots)] = 0$, para $\forall t$.

Todas DM's têm média zero e são serialmente não-correlacionadas. Note que uma DM impõe uma condição mais forte do que não-autocorrelação. Uma seqüência serialmente não-correlacionada não pode ser prevista com base em uma função linear de seus valores passados. No caso de uma seqüência DM, nenhuma função de valores passados, linear ou não-linear, pode prevê-la.

Definição 1.3 y_t será fracamente estacionário ou estacionário em covariância se

- i) $\mathbb{E}[y_t] = \mu$, para $\forall t$;
 ii) $\mathbb{E}[|y_t^2|] < \infty$, para $\forall t$;
 iii) $Cov(y_t, y_s) = C_y(t, s) = C_y(t - s) = C_y(m)$, onde $m = t - s$, para $\forall t$ e $\forall s$.

A condição i) estabelece que, independentemente de onde estivermos no tempo, a média permanece a mesma, enquanto a condição iii) diz que a covariância só depende da distância temporal entre os pontos observados. Note também que $C_y(m) = C_y(-m)$.

Definição 1.4 O processo y_t será estritamente ou fortemente estacionário se

$$F(y_{t+h}, y_{t+h+1}, \dots, y_{t+h+p}) = F(y_t, y_{t+1}, \dots, y_{t+p})$$

para $\forall h$ e $\forall p$.

Ou seja, um processo é dito ser estritamente estacionário se a distribuição conjunta de $(y_t, y_{t+j_1}, y_{t+j_2}, \dots, y_{t+j_n})$ depende somente dos intervalos separando os tempos e não dos tempos propriamente ditos. Veja que estacionariedade estrita, aliada à condição de segundo momento finito, implica em estacionariedade em covariância.

Definição 1.5 y_t será um processo de Médias Móveis de Ordem q , denotado por $MA(q)$, quando satisfizer a seguinte relação:

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (1.1)$$

ou ainda

$$y_t - \mu = \theta(L)\varepsilon_t \quad (1.2)$$

onde ε_t é ruído branco com $\text{Var}[\varepsilon_t] = \sigma^2$, $\mathbb{E}[y_t] = \mu$, $L\varepsilon_t = \varepsilon_{t-1}$ e

$$\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$$

A função de autocovariância de ordem j é dada por

$$\gamma_j = \begin{cases} (-\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \dots + \theta_q\theta_{q-j})\sigma^2 & , j = 0, 1, 2, \dots, q \\ 0 & , j > q \end{cases}$$

onde $\theta_0 = -1$.

Definição 1.6 y_t será um processo **Autorregressivo de Ordem p** , denotado por $AR(p)$, quando satisfizer a relação

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (1.3)$$

ou

$$\phi(L)(y_t - \mu) = \varepsilon_t \quad (1.4)$$

onde ε_t é ruído branco com $\text{Var}[\varepsilon_t] = \sigma^2$, $\mathbb{E}[y_t] = \mu$ e $\mu = \frac{c}{\phi(1)}$, tal que $\phi(1) = 1 - \phi_1 - \dots - \phi_p$.

As funções de autocovariância são dadas por

$$\gamma_j = \begin{cases} \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p} & , j = 1, 2, \dots \\ \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma^2 & , j = 0 \end{cases}$$

Definição 1.7 y_t será um processo **Autorregressivo e de Médias Móveis de Ordem p, q** , denotado por $ARMA(p, q)$, quando tiver a forma

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (1.5)$$

ou

$$\phi(L)(y_t - \mu) = \theta(L)\varepsilon_t \quad (1.6)$$

onde, novamente, $\mathbb{E}[y_t] = \mu = \frac{c}{\phi(1)}$.

A função de autocovariância, para $j > q$, é dada por

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}$$

assumindo uma forma mais complicada para $j \leq q$, onde aparecerão os termos de médias móveis.

1.4 Características dos Retornos do IBOVESPA

Nesta seção, nossa proposta é a de averiguarmos se os retornos do índice de fechamento do IBOVESPA, no período de 2 de janeiro de 1986 a 4 de dezembro de 1995, possuem as características estilizadas descritas na seção anterior. Temos um total de 2.426 observações (dias úteis) que cobrem um período de aproximadamente 10 anos.

Primeiramente, olhando para o gráfico dos retornos ao longo do tempo, mostrado na figura 1.1, tentou-se localizar alguns agrupamentos de volatilidade. Vemos que 4 agrupamentos foram escolhidos para fins de avaliação, c1, c2, c3 e c4.

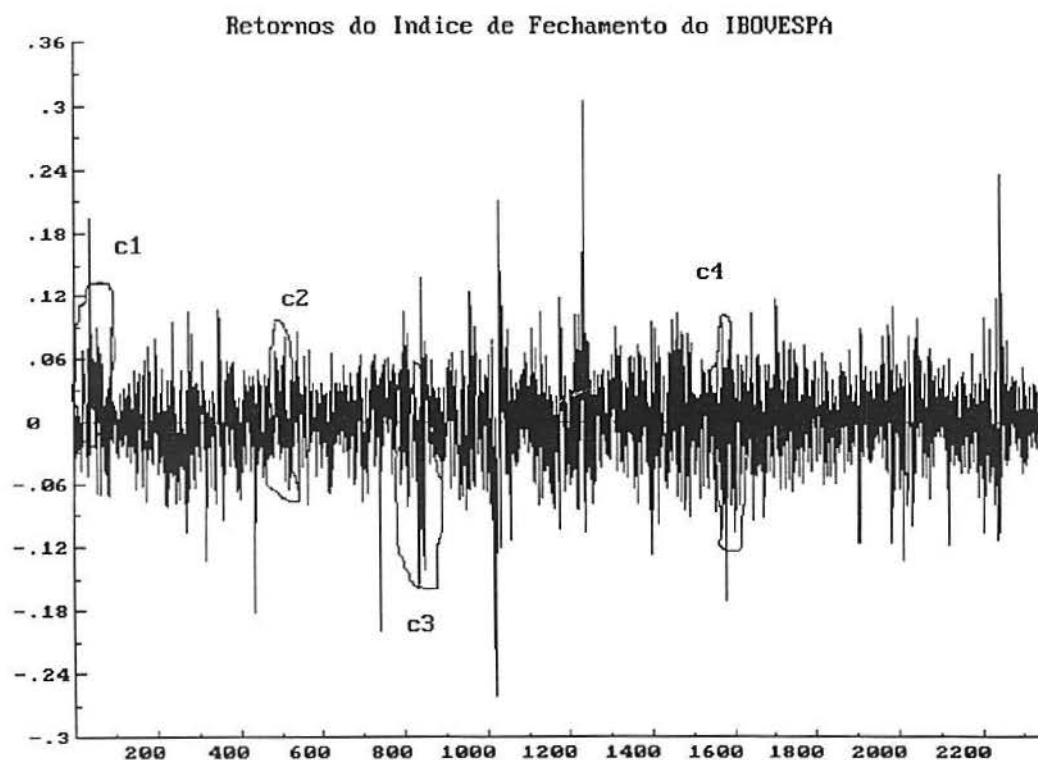


Figura 1.1: Retornos

Note que os agrupamentos c2 e c4 expressam de modo razoável a condição de erraticidade para os retornos e autocorrelação para seus quadrados, esta última devido à amplitude de variação dos retornos “manter-se” dentro do agrupamento, considerando tanto retornos positivos como negativos. No entanto, se olharmos para os agrupamentos c1 e c3, veremos que a amplitude de variação fica restrita preponderantemente a retornos com o mesmo sinal, indicando uma possível presença de autocorrelação nos próprios retornos. Lembre-se de que um dos fatos estilizados é que os retornos não são autocorrelacionados.

Obviamente, esta distinção entre os agrupamentos não é muito perceptível e, em nosso caso, ela foi sugerida após observarmos os gráficos das funções de autocorrelação dos retornos e de seus quadrados, dados na figura 1.2.

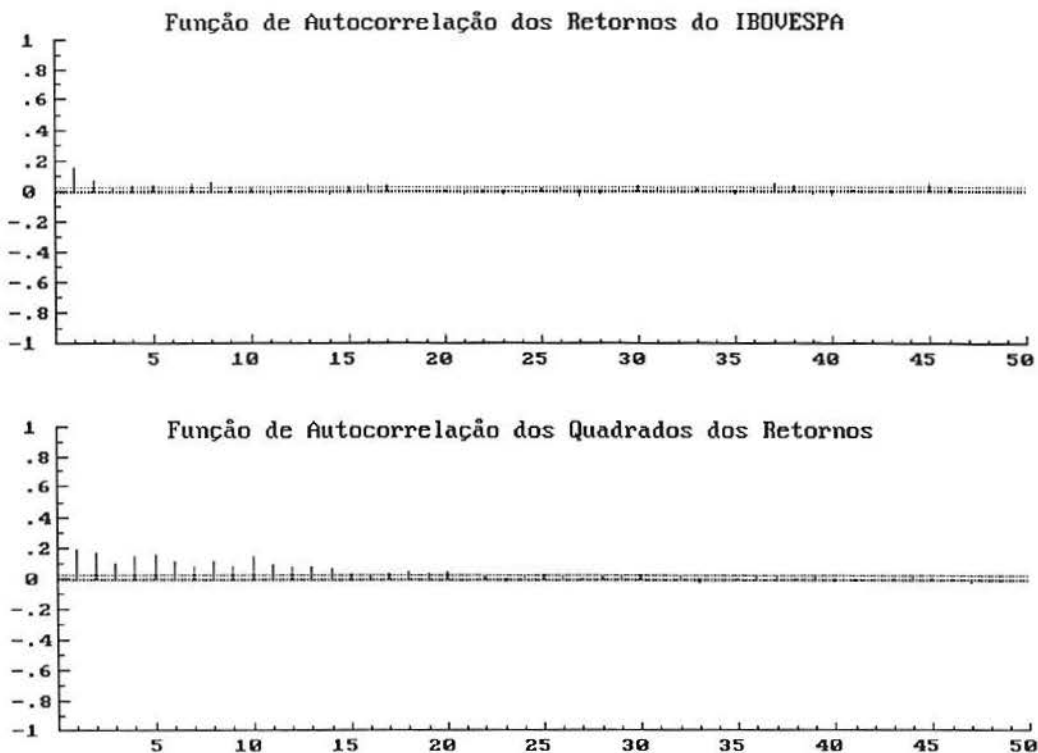


Figura 1.2: Autocorrelações

Nesta figura, vemos claramente que a autocorrelação de ordem 1 dos retornos é significativa, comparando com os limites de 2 desvios (se o processo x_t é ruído branco, então a distribuição assintótica de sua função de autocorrelação amostral é $N(0; 1/T)$, onde T é o tamanho da amostra), rejeitando-se a hipótese de ruído branco e, portanto, violando o primeiro dos fatos estilizados. Seu valor é de 0,1477, enquanto o limite de 2 desvios é de 0,0203. Por outro lado, podemos notar que a autocorrelação dos quadrados é significativa; além disso, seu valor de ordem 1 é pequeno, aproximadamente 0,2, e, após esta ordem, a função decaí lentamente, confirmando o segundo dos fatos estilizados.

A seguir, observemos a figura 1.3, a qual nos mostra o histograma dos retornos padronizados e uma densidade não-paramétrica ajustada a eles.

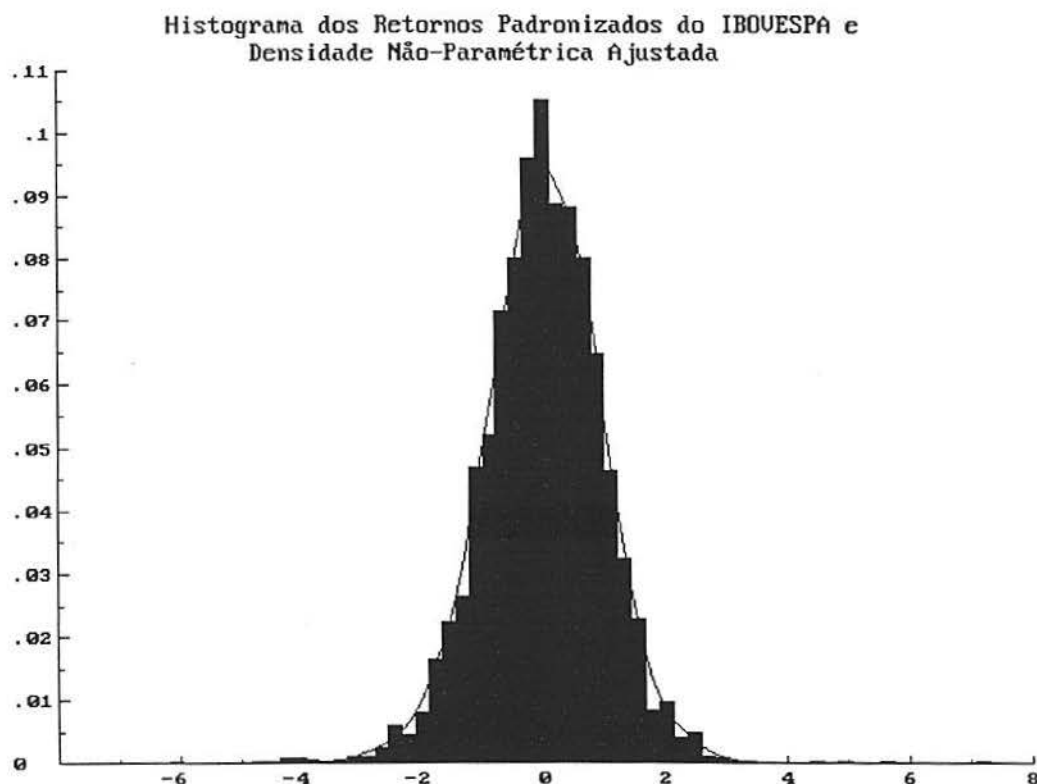


Figura 1.3: Histograma e Densidade

Grosseiramente, percebe-se que as caudas do histograma são pesadas e que ele não apresenta assimetria. Tais percepções visuais são confirmadas nos testes Qui-quadrado para curtose e assimetria (ver apêndice). Aceita-se a hipótese de simetria com nível descritivo amostral de aproximadamente 18%, enquanto rejeita-se a hipótese de não haver excesso de curtose a um nível descritivo inferior a 0,01%.

Para concluirmos, devemos enfatizar que vários fatos estilizados foram confirmados. Entretanto, os retornos são autocorrelacionados. Esta característica, como veremos no decorrer do texto, fará com que sejamos obrigados a transformar nossos dados. Voltaremos a falar sobre isto no capítulo 4.

Capítulo 2

Variabilidade Condicional Estocástica

Neste capítulo, estudaremos alguns dos modelos formulados para a estimação e previsão da volatilidade dos retornos. Como vimos no capítulo 1, uma característica das séries de retornos de ativos financeiros é a formação de agrupamentos de volatilidade, ou seja, destacam-se grupos de observações com diferentes níveis de variabilidade ao longo do tempo. Sendo assim, um “bom” modelo deveria incorporar esta característica de variação da volatilidade. Um modelo que não incorporasse esta idéia “permitiria” o uso de medidas de variabilidade amostral tradicionais, como, por exemplo, a variância não-condicional, a qual seria uma medida demasiadamente globalizadora, muito influenciada por observações extremas e incapaz de detectar as mudanças de volatilidade ocorridas, perdendo assim valiosa informação contida nos dados. Mesmo uma medida de variabilidade com uma “janela móvel” ao longo do tempo, apesar de oferecer algumas vantagens em relação à anterior, seria influenciada por observações extremas durante um período de tempo igual ao tamanho da janela utilizada, possuindo, digamos assim, uma “memória indesejável”, ou seja, o efeito de observações “incomuns” é propagado.

Os modelos que apresentaremos têm por objetivo estimar e prever a volatilidade condicional. As principais formas utilizadas na literatura são os modelos ARCH, GARCH (e derivados) e de Volatilidade Estocástica (V.E.). Para estes tipos de modelos, a variância em um determinado instante de tempo, condicional ao conjunto de informações até o instante anterior, passa a ser considerada

uma variável aleatória, formando, ao longo do tempo, um processo estocástico. Formalmente, utilizando uma abordagem paramétrica e uma suposição de gaussianidade para fins de entendimento, temos

$$y_t | I \sim N(0; \sigma_t^2) \quad (2.1)$$

onde y_t é a variável de interesse,
 I é um conjunto de informações
e σ_t^2 é a variância condicional ao conjunto I .

A diferença entre os modelos da família GARCH e os modelos de V.E. basicamente encontra-se no conjunto de informações utilizado em cada uma das modelagens. Nos modelos GARCH, I é composto por observações passadas da variável y_t , ou seja,

$$I = Y_{t-1} = \{y_1, \dots, y_{t-1}\}$$

Já nos modelos de V.E., $I = Y_t$ é formado por um processo latente (não-observável), o logaritmo da volatilidade condicional até o instante t , o qual evolui com a chegada de novas informações ao mercado financeiro, oriunda de variáveis também não-observáveis.

2.1 Modelos ARCH

Engle (1982) introduziu a classe de modelos ARCH como um modo conveniente para modelar a variância condicional. Sua forma, no caso geral (ARCH(p)), é dada por

$$y_t = \mathbf{x}_t' \boldsymbol{\beta} + e_t \quad (2.2)$$

$$e_t = \varepsilon_t \sigma_t \quad (2.3)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \dots + \alpha_p e_{t-p}^2 \quad (2.4)$$

para $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, 2, \dots, p$, onde y_t é o retorno no instante t , σ_t^2 é a variância no instante t condicional a Y_{t-1} , $\varepsilon_t \sim NIID(0; 1)$ e \mathbf{x}_t é um processo determinístico ou estocástico (podendo até mesmo ser endógeno) que representa a média condicional de y_t . Note que as condições $\alpha_0 > 0$ e $\alpha_i \geq 0$ asseguram que $\sigma_t^2 \geq 0$, para todo t .

2.1.1 Características Gerais

Com o intuito de explorar as propriedades desta classe de modelos, nos concentraremos no modelo mais simples, o ARCH(1), onde

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 \quad (2.5)$$

Assumindo que $\beta = 0$, tem-se, então,

$$y_t = e_t = \varepsilon_t \sigma_t \quad (2.6)$$

Assim, vemos facilmente que $y_t \mid Y_{t-1} \sim N(0; \sigma_t^2)$. Já os momentos não-condicionais de y_t podem ser obtidos das propriedades da distribuição normal de ε_t e do processo autorregressivo para σ_t^2 . Portanto,

$$\mathbb{E}[y_t] = \mathbb{E}[\varepsilon_t \sigma_t] = \mathbb{E}[\varepsilon_t] \mathbb{E}[\sigma_t] = 0$$

se $\mathbb{E}[|\sigma_t|] < \infty$.

Para calcularmos $\mathbb{E}[y_t^2]$, notemos que o modelo dado pelas equações (2.5) e (2.6) pode ser escrito como

$$y_t^2 = \sigma_t^2 + (y_t^2 - \sigma_t^2) = \alpha_0 + \alpha_1 y_{t-1}^2 + \nu_t \quad (2.7)$$

onde

$$\nu_t = \sigma_t^2 (\varepsilon_t^2 - 1) \quad (2.8)$$

Deste modo, ν_t é uma DM (mantido que $\mathbb{E}[|\sigma_t^2|] < \infty$), já que $\mathbb{E}[\nu_t \mid Y_{t-1}] = 0$. Com isso, obtemos de (2.7) que

$$\sigma^2 = \mathbb{E}[y_t^2] = \frac{\alpha_0}{1 - \alpha_1} = \mathbb{E}[\sigma_t^2] \quad (2.9)$$

para $\alpha_1 < 1$.

Agora,

$$\mathbb{E}[y_t^3] = \mathbb{E}[\varepsilon_t^3 \sigma_t^3] = \mathbb{E}[\varepsilon_t^3] \mathbb{E}[\sigma_t^3] = 0$$

assim como todos outros momentos de ordem ímpar (garantido que o momento de ordem par superior seguinte exista), já que ε_t tem densidade simétrica em relação a zero.

Já,

$$\mathbb{E}[y_t^4] = \mathbb{E}[\varepsilon_t^4] \mathbb{E}[\sigma_t^4] = 3 \mathbb{E}[\sigma_t^4] = 3 \frac{\alpha_0^2 (1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}$$

pois

$$\phi = \mathbb{E}[\sigma_t^4] = \mathbb{E}[(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \sigma_{t-1}^2)^2] = \alpha_0^2 + 2\alpha_0 \alpha_1 \sigma^2 + 3\alpha_1^2 \phi$$

que, usando (2.9), resulta em

$$\phi = \frac{\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}$$

para $3\alpha_1^2 < 1$.

Assim, com base nos resultados acima, obtemos as seguintes conclusões:

- y_t é uma diferença martingala se $\mathbb{E}[|\sigma_t|] < \infty$. Ainda, se $0 \leq \alpha_1 < 1$, y_t é ruído branco.
- y_t^2 segue um processo autorregressivo. Para que y_t^2 seja estacionário em covariância, necessitamos $3\alpha_1^2 < 1$.
- A medida de curtose de y_t , se $3\alpha_1^2 < 1$, é dada por

$$c = \frac{\mathbb{E}[y_t^4]}{(\mathbb{E}[y_t^2])^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2}$$

Note que, exceto para o caso trivial em que $\alpha_1 = 0$ e conseqüentemente $y_t \sim NIID(0; \alpha_0)$, quando c existe, ele é maior do que 3, ou seja, y_t tem leptocurtose: caudas pesadas.

- Se y_t^2 for estacionário em covariância, temos, diretamente do processo autorregressivo, que $\rho_{y^2}(s) = \alpha_1^s$ é o coeficiente de autocorrelação de ordem s de y_t^2 .
- Pelas duas primeiras conclusões, têm-se que y_t é ruído branco para $0 \leq \alpha_1 < 1$ e y_t^2 é um autorregressivo, caracterizando agrupamentos de volatilidade (em diferentes níveis) ao longo do tempo.

2.1.2 Estimação

Os modelos ARCH foram construídos de modo a permitir que a verossimilhança seja facilmente encontrada. Utilizando a decomposição dos erros de predição (e

ignorando constantes), obtemos

$$\begin{aligned} \log f(y_1, \dots, y_T | y_0; \boldsymbol{\theta}) &= \sum_{t=1}^T \log f(y_t | Y_{t-1}; \boldsymbol{\theta}) \\ &= -\frac{1}{2} \sum_{t=1}^T \log \sigma_t^2 - \frac{1}{2} \sum_{t=1}^T \frac{y_t^2}{\sigma_t^2} \end{aligned} \quad (2.10)$$

onde $\boldsymbol{\theta} = (\alpha_0, \alpha_1)'$.

Assim, obtêm-se os estimadores tradicionais de máxima verossimilhança para α_0 e α_1 . O comportamento assintótico desses estimadores foi estudado por Weiss (1986), que mostrou sua normalidade sob a condição de y_t ter quarto momento finito. Posteriormente, Lee e Hansen (1994) e Lumsdaine (1996), ainda no contexto assintótico, abordaram a não gaussianidade de ε_t , tratando (2.10) como uma função de quase-verossimilhança¹. Note que também poderíamos assumir outras distribuições a priori para ε_t , como é o caso, por exemplo, da distribuição t , que possui caudas mais pesadas do que a normal. Neste caso, também obteríamos uma verossimilhança condicional exata a ser maximizada.

Em relação à não-normalidade de ε_t e conseqüente não-normalidade condicional de y_t , tem havido interesse em procedimentos não-paramétricos para a estimação da densidade de ε_t (Engle e Gonzales-Rivera (1991)).

2.1.3 Previsão

Note que, nos modelos ARCH, para $s > 0$, a previsão

$$\begin{aligned} \hat{y}_{T+s} &= \mathbb{E}[y_{T+s} | Y_T] \\ &= \mathbb{E}[\varepsilon_{T+s} | Y_T] \mathbb{E}[\sigma_{T+s} | Y_T] = 0 \end{aligned}$$

já que $\mathbb{E}[\varepsilon_{T+s} | Y_T] = \mathbb{E}[\varepsilon_{T+s}] = 0$.

Agora,

$$\begin{aligned} \text{Var}[y_{T+s} | Y_T] &= \mathbb{E}[y_{T+s}^2 | Y_T] = \mathbb{E}[\sigma_{T+s}^2 \varepsilon_{T+s}^2 | Y_T] \\ &= \mathbb{E}[\sigma_{T+s}^2 | Y_T] \mathbb{E}[\varepsilon_{T+s}^2 | Y_T] = \mathbb{E}[\sigma_{T+s}^2 | Y_T] \end{aligned}$$

¹Esta abordagem trata a perturbação como se ela fosse gaussiana, obtendo assim estimadores de quase-máxima verossimilhança

Ainda, para o modelo ARCH(1),

$$\mathbb{E}[\sigma_{T+s}^2 | Y_T] = \mathbb{E}[(\alpha_0 + \alpha_1 y_{T+s-1}^2) | Y_T]$$

que, aplicando o operador Esperança, substituindo $\mathbb{E}[y_{T+s-1}^2 | Y_T]$ por $\mathbb{E}[\sigma_{T+s-1}^2 | Y_T]$ e recursivamente utilizando (2.5), facilmente verifica-se resultar em

$$\mathbb{E}[\sigma_{T+s}^2 | Y_T] = \alpha_0 \sum_{i=0}^{s-1} \alpha_1^i + \alpha_1^s y_T^2$$

Perceba que, quando y_t não é um processo estacionário, por exemplo $\alpha_1 = 1$, $\text{Var}[y_{T+s} | Y_T]$ explode à medida que s cresce. Entretanto, se não há muita persistência no processo, Baille e Bollerslev (1992) concluem que uma aproximação normal baseada em $y_{T+s} | Y_T \sim N(0; \mathbb{E}[y_{T+s}^2 | Y_T])$ pode ser satisfatória.

2.1.4 Concordância com os Fatos Estilizados

Pelo exposto das características gerais do modelo ARCH(1), vimos que ele incorpora vários dos fatos estilizados empíricos encontrados na literatura. Entretanto, quando y_t^2 é estacionário, sua função de autocorrelação é dada por $\rho_{y^2}(s) = \alpha_1^s$, o que não reflete a baixa autocorrelação de primeira ordem e posterior queda muito lenta para zero encontradas empiricamente. Veja que se α_1 for pequeno teremos baixa autocorrelação de primeira ordem, mas o decaimento será muito rápido. No entanto, se α_1 for grande (próximo de 1), o decaimento será mais lento, mas a autocorrelação de primeira ordem será alta, contradizendo os resultados empíricos.

Sendo assim, tal inadequação motiva à busca de outros modelos que capturem melhor as características empíricas das funções de autocorrelação de y_t^2 . Uma útil extensão dos modelos ARCH são os modelos GARCH (ARCH Generalizado) introduzidos por Bollerslev (1986).

2.2 Modelos GARCH

Os modelos ARCH têm sido estendidos em diversas direções. Uma das mais importantes extensões são os modelos GARCH, que incluem volatilidade condicional defasada na equação da volatilidade condicional σ_t^2 .

2.2.1 Características Gerais

O exemplo mais simples e também mais utilizado é o GARCH (1,1), dado por

$$y_t = \varepsilon_t \sigma_t \quad (2.11)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (2.12)$$

onde $\varepsilon_t \sim NIID(0; 1)$, $\alpha_0 > 0$, $\alpha_1 > 0$ e $\beta_1 > 0$.

Note que estamos assumindo novamente $\beta = 0$, assim como fizemos para o modelo ARCH(1). Este modelo pode ser generalizado para um GARCH(p, q), permitindo p defasagens de y_t^2 e q de σ_t^2 .

Em termos dos quadrados de y_t , podemos escrever as equações (2.11) e (2.12) como

$$y_t^2 = \alpha_0 + (\alpha_1 + \beta_1) y_{t-1}^2 + \nu_t - \beta_1 \nu_{t-1} \quad (2.13)$$

onde $\nu_t = \sigma_t^2(\varepsilon_t^2 - 1)$ é novamente uma DM.

Por (2.13), observa-se que o modelo GARCH inclui uma parte de médias móveis para y_t^2 , ou seja, y_t^2 segue um ARMA(1,1). Perceba que, se $\alpha_1 = 0$, então as partes autorregressiva e de médias móveis têm β_1 como raiz comum. Assim, formalmente, obtemos

$$(1 - \beta_1 L) y_t^2 = \alpha_0 + (1 - \beta_1 L) \nu_t \quad (2.14)$$

que se reduz a

$$y_t^2 = \alpha'_0 + \nu_t$$

onde $\alpha'_0 = \frac{\alpha_0}{1 - \beta_1}$.

Assim, se $\alpha_1 = 0$, β_1 não é identificável e a variância condicional dada por (2.12) torna-se uma constante. No entanto, se $\alpha_1 \neq 0$, o processo y_t^2 é um ARMA(1,1), com funções de autocorrelação dadas por

$$\rho_1 = \frac{[1 - \beta_1(\alpha_1 + \beta_1)]\alpha_1}{1 + \beta_1^2 - 2\beta_1(\alpha_1 + \beta_1)} \quad (2.15)$$

$$\rho_s = (\alpha_1 + \beta_1)^{s-1} \rho_1, \quad s \geq 2 \quad (2.16)$$

Lembrando da falta de adequamento do modelo ARCH(1) em relação à função de autocorrelação encontrada empiricamente para os quadrados dos retornos, a

qual apresenta baixa autocorrelação de primeira ordem e decaimento lento para as demais, vemos que o modelo GARCH(1,1) é capaz de suprir tal deficiência, bastando que α_1 seja pequeno e $\alpha_1 + \beta_1$ grande, que é o que acontece frequentemente com dados reais.

Agora, analisando os momentos não-condicionais de y_t , Bollerslev (1986) mostra que a condição necessária e suficiente para que o momento de ordem $2r$ (ordem par) exista é

$$\mu(\alpha_1, \beta_1, r) = \sum_{j=0}^r \binom{r}{j} a_j \alpha_1^j \beta_1^{r-j} < 1 \quad (2.17)$$

onde $a_0 = 1$, $a_j = \prod_{i=1}^j (2i - 1)$ ($j = 1, 2, \dots$)

Ele também mostra que os momentos de ordem $2r$ são dados por

$$\mathbb{E}[y_t^{2r}] = a_r \left[\sum_{k=0}^{r-1} a_k^{-1} \mathbb{E}[y_t^{2k}] \alpha_0^{r-k} \binom{r}{r-k} \mu(\alpha_1, \beta_1, k) \right] [1 - \mu(\alpha_1, \beta_1, r)]^{-1} \quad (2.18)$$

enquanto os momentos de ordem $(2r-1)$ (ordem ímpar) são iguais a zero (supondo a existência do momento de ordem $2r$).

Assim,

$$\begin{aligned} \mathbb{E}[y_t^2] &= \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \\ \mathbb{E}[y_t^4] &= 3\alpha_0^2 (1 + \alpha_1 + \beta_1) [(1 - \alpha_1 - \beta_1)(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)]^{-1} \end{aligned}$$

O coeficiente de curtose é igual a

$$c = \frac{\mathbb{E}[y_t^4]}{(\mathbb{E}[y_t^2])^2} = 3(1 - (\alpha_1 + \beta_1)^2)(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)^{-1}$$

Note que o coeficiente de curtose será maior do que 3 quando $\alpha_1^2 > 0$, ou seja, se α_1^2 não for muito pequeno, então y_t terá caudas pesadas. Também percebe que, conforme a condição (2.17), y_t^2 será estacionário em covariância se e somente se $(\beta_1^2 + 2\alpha_1\beta_1 + 3\alpha_1^2) < 1$.

2.2.2 Estimação

Para os modelos GARCH, a função de verossimilhança pode ser facilmente encontrada, de modo análogo ao descrito para os modelos ARCH, obtendo-se os

estimadores de máxima verossimilhança para os parâmetros envolvidos no modelo, $\theta = (\alpha_0, \alpha_1, \beta_1)'$.

A diferença em relação aos modelos ARCH é que, quando derivamos o logaritmo da função de verossimilhança em relação aos parâmetros, nos deparamos com um processo recursivo, o qual exigirá o estabelecimento não só de valores prévios para y_t como para σ_t^2 , para $t \leq 0$.

2.2.3 Previsão

Como para os modelos ARCH, temos que, para $s \geq 1$,

$$\begin{aligned}\mathbb{E}[y_{T+s} | Y_T] &= 0 \\ \text{Var}[y_{T+s} | Y_T] &= \mathbb{E}[y_{T+s}^2 | Y_T] \\ &= \mathbb{E}[\sigma_{T+s}^2 | Y_T]\end{aligned}$$

que, para o modelo GARCH(1,1), é dada por

$$\begin{aligned}\mathbb{E}[\sigma_{T+s}^2 | Y_T] &= \mathbb{E}[\alpha_0 + \alpha_1 y_{T+s-1}^2 + \beta_1 \sigma_{T+s-1}^2 | Y_T] \\ &= \alpha_0 + (\alpha_1 + \beta_1) \mathbb{E}[y_{T+s-1}^2 | Y_T]\end{aligned}$$

Resolvendo recursivamente, obtemos

$$\mathbb{E}[y_{T+s}^2 | Y_T] = \alpha_0 \sum_{i=0}^{s-1} (\alpha_1 + \beta_1)^i + (\alpha_1 + \beta_1)^s y_T^2$$

2.2.4 Concordância com os Fatos Estilizados e o Formato ARCH(∞) do GARCH(1,1)

Vimos que o modelo GARCH(1,1) é capaz de representar um número maior dos fatos estilizados empíricos do que o modelo ARCH(1), principalmente no que diz respeito à função de autocorrelação dos quadrados dos retornos. Mostraremos agora como o modelo GARCH(1,1) pode ser escrito na forma ARCH(∞). Considere a equação

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Podemos escrevê-la como

$$(1 - \beta_1 L) \sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2$$

Ou ainda

$$\sigma_t^2 = \alpha'_0 + \frac{\alpha_1}{1 - \beta_1 L} y_{t-1}^2 = \alpha'_0 + \sum_{j=0}^{\infty} \beta_1^j L^j \alpha_1 y_{t-1}^2 \quad (2.19)$$

para $|\beta_1 L| < 1$, onde $\alpha'_0 = \frac{\alpha_0}{1 - \beta_1}$.

Expandindo-se (2.19), chega-se a

$$\sigma_t^2 = \alpha'_0 + \alpha_1 y_{t-1}^2 + \beta_1 \alpha_1 y_{t-2}^2 + \beta_1^2 \alpha_1 y_{t-3}^2 + \dots \quad (2.20)$$

a qual tem o formato de um ARCH(∞).

Vemos que para os valores geralmente encontrados empiricamente para o GARCH(1,1), α_1 pequeno e $(\alpha_1 + \beta_1)$ grande, a volatilidade condicional no instante t depende de uma combinação linear dos quadrados dos retornos, a qual dá pesos que vão decaindo lentamente à medida que o valor observado se distancia temporalmente de t . Observe que o peso da observação referente ao instante $t-1$, o mais próximo, já não é grande. Recorde-se que, para o modelo ARCH(1), σ_t^2 somente dependia de y_{t-1}^2 , o qual recebia um peso grande, pois α_1 , para o ARCH(1), é normalmente encontrado grande (o que não acontece para o GARCH(1,1)). Como resultado, teremos que a volatilidade fornecida pelo GARCH(1,1) deverá ser mais suave do que a fornecida pelo ARCH(1).

2.3 Outros Modelos da Família ARCH

Há um grande número de extensões dos modelos ARCH (ver Bollerslev, Chou e Kroner (1992)). Cada uma delas sempre procura representar de maneira mais adequada algum dos fatos estilizados encontrados empiricamente. Apresentaremos sucintamente alguns destes modelos em suas formas mais simples.

2.3.1 AGARCH (GARCH Assimétrico)

Para o AGARCH(1,1), a equação da volatilidade condicional² pode ser dada por

$$\sigma_t^2 = \alpha_0 + \alpha_1 (y_{t-1} - \xi)^2 + \beta_1 \sigma_{t-1}^2 \quad (2.21)$$

²Para variações e detalhes, ver Pagan (1993).

onde $\alpha_0, \alpha_1, \beta_1, \xi > 0$.

Neste modelo, retornos negativos induzirão a uma volatilidade condicional maior do que retornos positivos, refletindo uma das características estilizadas dadas no capítulo 1, a saber, mercado em queda apresenta mais volatilidade do que mercado em ascensão.

2.3.2 EGARCH (GARCH Exponencial)

As restrições de não-negatividade para os parâmetros dos modelos GARCH podem restringir excessivamente a dinâmica das volatilidades condicionais obtidas. Nelson (1991) eliminou a necessidade de tais restrições, modelando o logaritmo da volatilidade. O EGARCH(1,1) é dado por

$$\log \sigma_t^2 = \alpha_0 + \beta_1 \log \sigma_{t-1}^2 + \gamma_0 z_{t-1} \quad (2.22)$$

onde $z_t = [|\varepsilon_t| - (2/\pi)^{1/2}] + \delta \varepsilon_t$ e $\varepsilon_t \sim NIID(0;1)$

Obviamente, o modelo EGARCH(1,1) sempre produz uma variância condicional σ_t^2 positiva para quaisquer valores de $\alpha_0, \gamma_0, \beta_1$ (necessita apenas que $|\beta_1| < 1$ para ser estacionário). Este modelo também permite que σ_t^2 seja não-simétrica em relação à ε_t , e, para δ negativo, exibirá volatilidade maior quando ε_t for grande e negativo.

2.3.3 IGARCH (GARCH Integrado)

Quando $\alpha_1 + \beta_1 = 1$,³ nós podemos reescrever o modelo GARCH(1,1) como

$$\sigma_t^2 = \alpha_0 + (1 - \lambda)y_{t-1}^2 + \lambda\sigma_{t-1}^2 \quad (2.23)$$

para $0 \leq \lambda \leq 1$.

Perceba que os momentos não-condicionais deste modelo não estão definidos, pois estamos tratando de um GARCH(1,1) não-estacionário.

³Ver Bollerslev e Engle (1986).

2.3.4 ARCH-M (ARCH em Média)

Engle, Lilien e Robins (1987) estenderam o modelo ARCH(p) de modo a permitir que a média dos retornos dependesse de sua própria volatilidade condicional. Sua aplicação é direcionada para ativos de risco. O modelo ARCH-M(1) é dado por

$$y_t = g(\sigma_t^2; \theta) + \varepsilon_t \sigma_t \quad (2.24)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 [y_{t-1} - g(\sigma_{t-1}^2; \theta)]^2 \quad (2.25)$$

onde normalmente é utilizada uma parametrização linear para $g(\sigma_t^2; \theta)$.

As propriedades estatísticas deste modelo foram estudadas por Hong (1991).

2.4 Modelos de Volatilidade Estocástica

Uma alternativa aos modelos da família ARCH é permitir que σ_t^2 dependa não de observações passadas, mas sim de componentes não-observáveis, ou seja, de uma estrutura latente. Tais modelos são denominados de Modelos de Volatilidade Estocástica. A forma mais utilizada destes modelos é dada por

$$y_t = \mathbf{x}_t' \boldsymbol{\beta} + e_t \quad (2.26)$$

$$e_t = \sigma \varepsilon_t \exp\left\{\frac{h_t}{2}\right\} \quad (2.27)$$

$$h_{t+1} = \gamma_0 + \gamma_1 h_t + \eta_{t+1} \quad (2.28)$$

onde

y_t é a variável de interesse,

\mathbf{x}_t é um vetor $k \times 1$ de variáveis explicatórias,

$\boldsymbol{\beta}$ é um vetor $k \times 1$ de coeficientes,

ε_t é ruído branco com média zero e variância σ_ε^2 ,

σ é um fator de escala positivo e

η_t é outro ruído branco com média zero, variância σ_η^2 e independente de ε_t .⁴

⁴A suposição de dependência entre as perturbações é tratada em Harvey e Shephard(1993b).

2.4.1 Propriedades Básicas

Com o objetivo de explorarmos algumas propriedades deste modelo, adicionaremos as suposições de gaussianidade sobre ε_t e η_t , considerando ainda $\text{Var}[\varepsilon_t] = 1$, $\beta = \mathbf{0}$ e $\sigma = 1$, obtendo o modelo formado pelas equações

$$y_t = \varepsilon_t \exp\left\{\frac{h_t}{2}\right\} \quad (2.29)$$

$$h_{t+1} = \gamma_0 + \gamma_1 h_t + \eta_{t+1} \quad (2.30)$$

À medida que h_t é gaussiano, temos que ele é um AR(1) gaussiano padrão. Este processo será estacionário se $|\gamma_1| < 1$, com

$$\mu_h = \mathbb{E}[h_t] = \frac{\gamma_0}{1 - \gamma_1} \quad (2.31)$$

$$\sigma_h^2 = \text{Var}[h_t] = \frac{\sigma_\eta^2}{1 - \gamma_1^2} \quad (2.32)$$

Como y_t é o produto dos processos ε_t e $\exp\{\frac{h_t}{2}\}$, onde ε_t é estacionário por suposição, então ele será estacionário se e somente se o processo h_t for estacionário.

Assim, sob a suposição de estacionariedade, utilizando as propriedades da função geradora de momentos da distribuição normal (ver, por exemplo, Murteira (1990)), os momentos de ordem par de y_t serão dados por

$$\begin{aligned} \mathbb{E}[y_t^r] &= \mathbb{E}[\varepsilon_t^r] \mathbb{E}[\exp\{\frac{r}{2}h_t\}] \\ &= \frac{r!}{2^{\frac{r}{2}}(\frac{r}{2})!} \exp\left\{\frac{r}{2}\mu_h + \frac{r^2\sigma_h^2}{8}\right\} \end{aligned} \quad (2.33)$$

enquanto, devido à simetria de ε_t em relação a zero, os momentos de ordem ímpar serão iguais a zero.

Para calcularmos a medida de curtose c , temos de (2.33) que

$$\mathbb{E}[y_t^2] = \exp\left\{\mu_h + \frac{\sigma_h^2}{2}\right\} \quad (2.34)$$

$$\mathbb{E}[y_t^4] = 3 \exp\{2\mu_h + 2\sigma_h^2\} \quad (2.35)$$

Então,

$$c = \frac{\mathbb{E}[y_t^4]}{(\mathbb{E}[y_t^2])^2} = 3 \frac{\exp\{2\mu_h + 2\sigma_h^2\}}{[\exp\{\mu_h + \frac{\sigma_h^2}{2}\}]^2} = 3 \exp\{\sigma_h^2\} \geq 3$$

o que mostra que este modelo de V.E. implica que y_t tenha caudas mais pesadas do que as da distribuição normal.

Quanto às propriedades dinâmicas de y_t , temos as seguintes:

- Como ε_t é ruído branco e η_t e ε_t são independentes, então, se $|\gamma_1| < 1$ e portanto h_t for estacionário, y_t será ruído branco.
- Para h_t estacionário,

$$\text{Cov}(y_t^2, y_{t-s}^2) = \mathbb{E}[y_t^2 y_{t-s}^2] - (\mathbb{E}[y_t^2])^2$$

onde $\mathbb{E}[y_t^2]$ é dada por (2.33), enquanto, para $s \geq 1$,

$$\begin{aligned} \mathbb{E}[y_t^2 y_{t-s}^2] &= \mathbb{E}[(\varepsilon_t^2 e^{h_t})(\varepsilon_{t-s}^2 e^{h_{t-s}})] \\ &= \mathbb{E}[\varepsilon_t^2] \mathbb{E}[\varepsilon_{t-s}^2] \mathbb{E}[e^{h_t+h_{t-s}}] \\ &= \mathbb{E}[\exp\{h_t + h_{t-s}\}] \end{aligned}$$

Agora, já que h_t é um AR(1) gaussiano, supondo-se a normalidade multivariada, obtemos

$$h_t + h_{t-s} \sim N(2\mu_h; 2\sigma_h^2(1 + \gamma_1^s))$$

Então, novamente utilizando os resultados da função geradora de momentos da distribuição normal, chegamos a

$$\mathbb{E}[\exp\{h_t + h_{t-s}\}] = \exp\{2\mu_h + \sigma_h^2(1 + \gamma_1^s)\}$$

e, portanto,

$$\begin{aligned} \text{Cov}(y_t^2, y_{t-s}^2) &= \exp\{2\mu_h + \sigma_h^2 + \sigma_h^2 \gamma_1^s\} - \exp\{2\mu_h + \sigma_h^2\} \\ &= \exp\{2\mu_h + \sigma_h^2\} (\exp\{\sigma_h^2 \gamma_1^s\} - 1) \end{aligned}$$

conduzindo a

$$\rho_{y^2}(s) = \frac{\exp\{\sigma_h^2 \gamma_1^s\} - 1}{3 \exp\{\sigma_h^2\} - 1}$$

já que

$$\begin{aligned} \text{Var}[y_t^2] &= \mathbb{E}[y_t^4] - (\mathbb{E}[y_t^2])^2 \\ &= 3 \exp\{2\mu_h + 2\sigma_h^2\} - \exp\{2\mu_h + \sigma_h^2\} \\ &= \exp\{2\mu_h + \sigma_h^2\} (3 \exp\{\sigma_h^2\} - 1) \end{aligned}$$

Ainda, se γ_1 é grande, utilizando a expansão de Taylor para o numerador da expressão anterior, encontramos

$$\rho_{y^2}(s) \simeq \frac{\exp\{\sigma_h^2\} - 1}{3 \exp\{\sigma_h^2\} - 1} \gamma_1^s \quad (2.36)$$

Olhando para esta função, vemos que ela tem o formato da função de autocorrelação de um processo ARMA(1,1), estabelecendo deste modo uma forte relação com o modelo GARCH(1,1).

Considere agora a equação (2.29). Elevando todos os termos ao quadrado e posteriormente tomando logaritmos, obtemos

$$\log y_t^2 = h_t + \log \varepsilon_t^2 \quad (2.37)$$

$$= w + h_t + \xi_t \quad (2.38)$$

$$h_{t+1} = \gamma_0 + \gamma_1 h_t + \eta_{t+1} \quad (2.39)$$

onde ξ_t é ruído branco com média zero e $w = \mathbb{E}[\log \varepsilon_t^2]$.

Assim, temos que o processo $\log y_t^2$ é a soma de um processo autorregressivo com um ruído branco. Podemos escrever, em função dos processos corrigidos pelas médias,

$$\log y_t^2 = \frac{\eta_t}{1 - \gamma_1 L} + \xi_t$$

ou

$$(1 - \gamma_1 L) \log y_t^2 = (1 - \gamma_1 L) \xi_t + \eta_t$$

Portanto, $\log y_t^2$ segue um ARMA(1,1). A função de autocorrelação deste processo, para $s \geq 1$, é dada por

$$\rho_{\log y^2}(s) = \frac{\gamma_1^s}{1 + (1 - \gamma_1^2)(\sigma_\xi^2/\sigma_\eta^2)}$$

2.4.2 Estimação

Nesta seção, inicialmente olharemos para o modelo sem variáveis explanatórias. Abordaremos, de forma bastante superficial, o Método dos Momentos Generalizados, o qual não será utilizado na parte prática de nosso trabalho. A seguir, veremos com mais detalhes o Método de Quase-Máxima Verossimilhança, que,

juntamente com o Filtro de Kalman, fornece estimativas para os parâmetros e para a volatilidade condicional. Também será sugerida, seguindo Harvey e Shephard (1993a), a correção das observações através da eliminação da heteroscedasticidade, podendo-se obter mais informações sobre a distribuição de ε_t e um estimador e preditor para a variância dos retornos. Finalmente, indicaremos o caminho a ser tomado no caso da presença de variáveis explanatórias no modelo.

Método dos Momentos Generalizados (GMM)

O Método dos Momentos Generalizados, na forma em que será sugerido aqui, embora tenha versões muito anteriores, foi desenvolvido recentemente por Hansen (1982). A principal vantagem deste método é que ele requer somente a especificação de certas condições sobre os momentos da variável, não exigindo a descrição completa de sua distribuição. A sua idéia, já bastante conhecida, é estimar os parâmetros desconhecidos do modelo através da solução de relações de igualdade entre os momentos amostrais e os momentos populacionais que dependem destes parâmetros.

No caso de Modelos de Volatilidade Estocástica, há muitos momentos que podemos usar para a estimação de seus parâmetros. Assim, pode existir um número maior de condições (igualdades) a serem satisfeitas do que parâmetros a estimar. Deste modo, não podemos escolher um único estimador que satisfaça às várias condições. Entretanto, podemos escolher os estimadores, de modo que eles satisfaçam da melhor maneira possível às diferentes condições, através da minimização de uma função critério, tal como

$$q(\boldsymbol{\theta}; y_1, \dots, y_T) \equiv \mathbf{g}' \mathbf{W} \mathbf{g} \quad (2.40)$$

onde \mathbf{g} é um vetor $k \times 1$ representando as condições a serem satisfeitas, como por exemplo,

$$\mathbf{g} = \begin{bmatrix} \frac{1}{T} \sum y_t^2 - \mathbb{E}[y_t^2] \\ \frac{1}{T} \sum y_t^4 - \mathbb{E}[y_t^4] \\ \frac{1}{T} \sum y_t^2 y_{t-1}^2 - \mathbb{E}[y_t^2 y_{t-1}^2] \\ \vdots \\ \frac{1}{T} \sum y_t^2 y_{t-k+2}^2 - \mathbb{E}[y_t^2 y_{t-k+2}^2] \end{bmatrix}$$

e W é uma matriz $k \times k$ de pesos, simétrica definida positiva, refletindo a importância de satisfazer cada uma das relações de igualdade.

Hansen (1982) estabeleceu a caracterização mais geral deste método, permitindo a não-independência entre as variáveis, e derivou as propriedades assintóticas destes processos serialmente dependentes. Uma descrição detalhada é dada em Hamilton (1994).

Algumas das principais desvantagens deste método em relação a nosso modelo, como relatado em Shephard (1994), são as seguintes:

- GMM somente pode ser usado se h_t for estacionário. Assim, quando γ_1 está próximo de 1, GMM não deve funcionar bem.
- Os estimadores dos parâmetros não têm a propriedade de invariância.
- GMM não oferece estimadores para a volatilidade condicional, necessitando assim de um método complementar para estimá-la.

Método de Quase-Máxima Verossimilhança (QML)

Forma de Espaço de Estados e Filtro de Kalman O filtro de Kalman é um procedimento recursivo utilizado para calcular o estimador ótimo do vetor de estados no tempo t , baseado na informação disponível até este instante. Para sua aplicação, o modelo deve ser escrito na forma de espaço de estados, que geralmente é expressa, para o caso univariado, como

$$y_t = z_t \alpha_t + d_t + \varepsilon_t \quad (2.41)$$

$$\alpha_t = T_t \alpha_{t-1} + c_t + R_t \eta_t \quad (2.42)$$

onde

$$E[\varepsilon_t] = 0, \text{Var}[\varepsilon_t] = \sigma_\varepsilon^2,$$

$$E[\eta_t] = 0, \text{Var}[\eta_t] = Q_t,$$

para $t = 1, 2, \dots, T$. Também, tanto ε_t , como η_t , são perturbações serialmente não-correlacionadas.

Quanto as dimensões dos vetores e matrizes, temos que z_t é $1 \times m$, α_t é $m \times 1$, T_t é $m \times m$, c_t é $m \times 1$, R_t é $m \times g$, η_t é $g \times 1$ e Q_t é $g \times g$.

Além disso, $\mathbb{E}[\alpha_0] = \mathbf{a}_0$, $\text{Var}[\alpha_0] = \mathbf{P}_0$ e $\mathbb{E}[\varepsilon_t \eta_{si}] = 0$, $i = 1, 2, \dots, g$, para $\forall s$, $t = 1, 2, \dots, T$, e as perturbações também são não-correlacionadas com o vetor inicial α_0 .

A equação (2.41) é chamada de equação de medida, enquanto (2.42) é a equação de transição. Os parâmetros desconhecidos σ_ε^2 e aqueles das matrizes \mathbf{z}_t , \mathbf{T}_t , \mathbf{R}_t e \mathbf{Q}_t são chamados de hiperparâmetros.

Derivando o filtro de Kalman, obtêm-se as equações de predição e atualização, que podem ser escritas em único conjunto de recursões, dado por

$$\begin{aligned} \mathbf{a}_{t+1/t} &= (\mathbf{T}_{t+1} - \mathbf{K}_t \mathbf{z}_t) \mathbf{a}_{t/t-1} + \mathbf{K}_t \mathbf{y}_t + (\mathbf{c}_{t+1} - \mathbf{K}_t \mathbf{d}_t) \\ \mathbf{P}_{t+1/t} &= \mathbf{T}_{t+1} (\mathbf{P}_{t/t-1} - \mathbf{P}_{t/t-1} \mathbf{z}'_t f_t^{-1} \mathbf{z}_t \mathbf{P}_{t/t-1}) \mathbf{T}'_{t+1} + \mathbf{R}_{t+1} \mathbf{Q}_{t+1} \mathbf{R}'_{t+1} \end{aligned}$$

para $t = 1, \dots, T$, onde

$$\mathbf{K}_t = \mathbf{T}_{t+1} \mathbf{P}_{t/t-1} \mathbf{z}'_t f_t^{-1} \quad (2.43)$$

$$f_t = \mathbf{z}_t \mathbf{P}_{t/t-1} \mathbf{z}'_t + \sigma_\varepsilon^2 \quad (2.44)$$

\mathbf{K}_t é conhecida como a matriz ganho do filtro.

Se as perturbações e o vetor de estados inicial forem normalmente distribuídos, então $\mathbf{a}_{t/t-1}$ é um estimador ótimo de α_t , no sentido em que minimiza o erro quadrático médio (ver Harvey (1989)). Se a suposição de normalidade é retirada, ele ainda será um estimador ótimo (no sentido anterior) dentre os estimadores lineares de α_t . Como subproduto do filtro de Kalman, nós também obtemos estimadores suavizados do vetor de estados, $\mathbf{a}_{t/T}$, os quais utilizam toda informação amostral para a estimação no instante t , sendo, portanto, melhores que os estimadores filtrados que só utilizam a informação disponível até o instante $t - 1$.

Paralelamente, em um procedimento iterativo, obtêm-se as estimativas de máxima verossimilhança dos parâmetros do modelo, através da função de verossimilhança normal (se a suposição de gaussianidade não é assumida, os estimadores dos parâmetros resultarão em estimadores de quase-máxima verossimilhança), extraída via decomposição dos erros de predição ($\nu_t = y_t - \mathbb{E}[y_{t/t-1}]$), onde $\nu_t \sim NID(0; f_t)$. O logaritmo da verossimilhança é dado por

$$\log L = -\frac{T}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log f_t - \frac{1}{2} \sum_{t=1}^T \frac{\nu_t^2}{f_t} \quad (2.45)$$

Implementação para o Modelo de Volatilidade Consideremos novamente o modelo (2.37), incluindo o fator de escala σ^2 ,

$$\log y_t^2 = \log \sigma^2 + h_t + \log \varepsilon_t^2 \quad (2.46)$$

ou ainda

$$\log y_t^2 = w + h_t + \xi_t \quad (2.47)$$

$$h_{t+1} = \gamma_0 + \gamma_1 h_t + \eta_{t+1} \quad (2.48)$$

onde

$$w = \log \sigma^2 + \mathbb{E}[\log \varepsilon_t^2] \text{ e } \mathbb{E}[\xi_t] = 0.$$

À medida que ξ_t é ruído branco, (2.47) e (2.48) são uma representação em espaço de estados linear (susceptível ao tratamento através do Filtro de Kalman), onde (2.47) é a equação de medida e (2.48) é a equação de transição. A perturbação $\log \varepsilon_t^2$ em (2.46) tem momentos que dependem da distribuição de ε_t . Se ε_t for normal (suposição freqüentemente adotada), então $\mathbb{E}[\log \varepsilon_t^2] = -1,27$ e $\text{Var}[\log \varepsilon_t^2] = 4,93$ (ver Abramowitz e Stegun (1970)).

Como, em nossa representação de espaço de estados, a perturbação ξ_t dificilmente terá distribuição normal (isso somente ocorreria caso ε_t^2 tivesse distribuição lognormal), então o modelo será não-gaussiano e não poderemos obter os estimadores de máxima verossimilhança tradicionais. O que faremos é, conforme falamos anteriormente, tratar as perturbações como se elas tivessem distribuição normal e maximizar a verossimilhança dos erros de predição obtidos através do Filtro de Kalman. Os estimadores de quase-máxima verossimilhança obtidos têm, sob certas condições de regularidade, distribuição assintótica normal (Dunsmuir (1979)), isto é

$$T^{1/2}(\hat{\theta} - \theta) \sim NM(0; C(\theta))$$

onde

$$C(\theta) = A^{-1}(2A + B)A^{-1} = 2A^{-1} + A^{-1}BA^{-1}$$

com A e B envolvendo integrais de funções das derivadas da função geradora espectral do processo em relação a θ e um coeficiente de excesso de curtose do processo de ruído (da forma reduzida do modelo - forma de modelos ARIMA) em relação à distribuição normal.

Este resultado assume que η_t e ξ_t têm quartos momentos finitos e que os parâmetros não estão sobre a fronteira do espaço paramétrico. Veja que não foi

assumida uma distribuição para ε_t , abordagem que é chamada de QML irrestrita. Se este tipo de estimação é realizado, pode-se inferir, através do estimador da variância de ξ_t , uma distribuição aproximada para ε_t . Assim, o fator de escala σ^2 poderá ser estimado posteriormente através do estimador de w , já que $w = \log \sigma^2 + \mathbb{E}[\log \varepsilon_t^2]$. Também é comum assumir que o fator σ^2 seja a própria variância de $\sigma\varepsilon_t$, o que implicaria que em nosso modelo $\text{Var}[\varepsilon_t] = 1$. Quanto ao estimador de w , ele pode ser obtido de duas formas, conjuntamente com os demais na maximização da função de quase-verossimilhança ou, olhando para a equação (2.47), através da média de $\log y_t^2$.

Outra opção é especificarmos completamente a distribuição de ε_t , por exemplo, assumirmos que $\varepsilon_t \sim NIID(0; 1)$. Tal atitude elimina a necessidade de estimação da variância de ξ_t . O fator de escala σ^2 será novamente estimado através do estimador de w .

Ressaltemos que em qualquer das duas abordagens, se não supusermos σ conhecido, então não conseguiremos estimar γ_0 , ou seja, estaremos impondo que o processo h_t , a longo prazo, tenha média zero.

Ainda é importante salientar que, como falamos anteriormente, ao se aplicar o Filtro de Kalman, obtêm-se tanto as estimativas filtradas quanto suavizadas da volatilidade condicional. As estimativas filtradas utilizam somente a informação até o instante anterior ao que está sendo estimado, enquanto as estimativas suavizadas utilizam toda informação amostral, conduzindo a erros quadráticos médios inferiores e, portanto, sendo preferíveis para fins de inferência em relação à volatilidade condicional.

Observações Corrigidas por Heteroscedasticidade e Previsão

Utilizando QML irrestrita, obtemos uma sugestão para a distribuição de ε_t através da variância estimada de ξ_t . Entretanto, como veremos agora, utilizando as observações corrigidas por heteroscedasticidade, podemos ter bases mais fortes para a determinação da distribuição de ε_t , além de obtermos um útil estimador para a variância de previsão de y_t . Genericamente, estas observações corrigidas são definidas por

$$\tilde{y}_t = y_t e^{-0.5h_{t|T}} \quad , \quad t = 1, \dots, T \quad (2.49)$$

onde y_t é a variável observada e $h_{t|T}$ é o estimador suavizado da volatilidade, que possui a propriedade de erro quadrático médio mínimo entre os estimadores

lineares.⁵

O estimador $h_{t|T}$, na abordagem clássica de extração de sinal (suavizamento), é modelado como uma combinação linear das observações, ou seja, a volatilidade pode ser vista como uma ponderação dos logaritmos dos quadrados dos retornos, podendo-se atribuir importâncias a eles. Para usarmos este estimador, a série deve ser suficientemente grande e t não deve estar próximo do início ou do final da série. Nós o utilizaremos para descrevermos as propriedades das observações corrigidas.

Então, considerando a equação (2.47), temos

$$h_{t|T} = w(L)(\log y_t^2 - w) = \sum_{j=-\infty}^{\infty} w_j L^j (\log y_t^2 - w) \quad (2.50)$$

tal que (ver Harvey (1989), pg. 227)

$$w(L) = \frac{|\theta_1(L)|^2 \sigma_1^2 |\Delta_1(L)\gamma_1(L)|^{-2}}{\sum_{m=0}^M |\theta_m(L)|^2 \sigma_m^2 |\Delta_m(L)\gamma_m(L)|^{-2}} \quad (2.51)$$

onde

$\theta_m(L)$ é um polinômio de médias móveis;

$\gamma_m(L)$ é um polinômio autorregressivo;

$\Delta_m(L)$ é um polinômio de diferenças

e $|\theta_m(L)|^2 = \theta_m(L)\theta_m(L^{-1})$.

Os polinômios acima são definidos quando o modelo descrito pelas equações (2.47) e (2.48), onde cada um dos processos é corrigido pela sua média, é colocado na forma

$$\log y_t^2 = \frac{\eta_t}{1 - \gamma_1 L} + \xi_t \quad (2.52)$$

denominada UCARIMA⁶. Em nosso caso, não temos polinômios não-estacionários, do tipo $\Delta_m(L)$, nem polinômios de médias móveis $\theta_m(L)$. O valor de M é igual a 1, pois, além do ruído ξ_t , o único componente não-observável é a volatilidade condicional. Deste modo, olhando para (2.52), vemos que

⁵No caso em que as perturbações são gaussianas, $h_{t|T}$ é o estimador de erro quadrático médio mínimo de h_t dentre todos os estimadores, lineares ou não.

⁶Forma ARIMA (ARMA não-estacionário), onde os componentes são não-observáveis.

$$\theta_m(L) = \Delta_m(L) = 1, \quad m = 0, 1;$$

$$\gamma_0(L) = 1;$$

$$\gamma_1(L) = 1 - \gamma_1 L;$$

$$\sigma_0^2 = \sigma_\xi^2;$$

$$\sigma_1^2 = \sigma_\eta^2;$$

$$q = \sigma_\eta^2 / \sigma_\xi^2.$$

Substituindo os valores acima em (2.51), encontramos

$$\begin{aligned} w(L) &= \frac{\sigma_\eta^2}{|\gamma_1(L)|^2} / \frac{\sigma_\eta^2 + |\gamma_1(L)|^2 \sigma_\xi^2}{|\gamma_1(L)|^2} \\ &= \frac{\sigma_\eta^2}{\sigma_\eta^2 + |\gamma_1(L)|^2 \sigma_\xi^2} \\ &= \frac{q}{q + |1 - \gamma_1 L|^2} \end{aligned} \tag{2.53}$$

Agora, substituindo (2.46) em (2.50), chegamos a

$$h_{t|T} = w(L) \log \varepsilon_t^2 + w(L) h_t - w(1) \mathbb{E}[\log \varepsilon_t^2]$$

que, substituindo em (2.49) e após algum desenvolvimento, resulta em

$$\tilde{y}_t = \sigma \varepsilon_t \prod_{j=-\infty}^{\infty} (\varepsilon_{t+j}^2)^{-0.5w_j} \exp\{0.5[(1 - w(L))h_t + w(1)\mathbb{E}[\log \varepsilon_t^2]]\}$$

onde os w_j são determinados por algoritmos apropriados (ver Whittle (1983, pg.59)).

Assumindo η_t gaussiano, h_t estacionário, ε_t simétrico e a existência de seu k -ésimo momento, temos que todos os momentos de ordem ímpar de \tilde{y}_t são zero. Ainda, sob a suposição de independência entre ε_t e η_s para todo t e s , notando que

$$(1 - w(L))h_t = \frac{1 - w(L)}{1 - \gamma_1 L} \eta_t$$

tem distribuição normal com média zero e variância denotada por ν (considerando o processo h_t corrigido pela média), os momentos de ordem par serão dados por

$$\mathbb{E}[\tilde{y}_t^k] = \sigma^k \mathbb{E}[|\varepsilon_t|^k] \prod_{j=-\infty}^{\infty} \mathbb{E}[|\varepsilon_{t+j}|^{-kw_j}] e^{\frac{\nu k^2}{8}} e^{0.5kw(1)\mathbb{E}[\log \varepsilon_t^2]} \tag{2.54}$$

Perceba que as observações corrigidas, sob as suposições dadas acima, serão simétricas e não-correlacionadas. Além disso, utilizando a teoria probabilística, podemos obter os momentos de \tilde{y}_t para algumas distribuições padrão de ε_t .

Note que o que foi desenvolvido supõe o conhecimento dos parâmetros e das volatilidades. Entretanto, utilizaremos estimadores consistentes dessas quantidades, calculando assim momentos amostrais que, comparados com os momentos populacionais teóricos, poderão nos dar mais informações sobre a distribuição de ε_t do que aquelas fornecidas somente pela estimativa da variância de ξ_t .

Tendo em mãos as observações corrigidas, podemos calcular a variância amostral

$$\tilde{\sigma}_*^2 = T^{-1} \sum_{t=1}^T \tilde{y}_t^2 \quad (2.55)$$

que é um estimador para $\sigma_*^2 = \sigma^2 \text{Var}[\varepsilon_t]$.

Este estimador $\tilde{\sigma}_*^2$ será utilizado para o cálculo do estimador da variância da previsão de y_t . Ou seja, o estimador da variância de y_t , dada a amostra inteira, será dado por

$$\hat{\text{Var}}[y_t | Y_T] = \tilde{\sigma}_*^2 \exp(h_{t|T})$$

para $t = 1, \dots, T$.

que, levado a previsões futuras, resulta em

$$\hat{\text{Var}}[y_{T+j} | Y_T] = \tilde{\sigma}_*^2 \exp(\hat{\gamma}_1^j h_{T|T})$$

para $j = 1, 2, \dots$

Harvey e Shephard (1993a) mostram que, para a normal padrão e para alguns graus de liberdade da distribuição t , as observações corrigidas fornecem estimadores com vício pequeno.

Já,

$$\begin{aligned} \mathbb{E}[y_{T+j} | Y_T] &= \sigma^2 \mathbb{E}[\varepsilon_{T+j} e^{\frac{h_t}{2}} | Y_T] \\ &= \sigma^2 \mathbb{E}[\varepsilon_{T+j}] \mathbb{E}[e^{\frac{h_t}{2}} | Y_T] \\ &= 0 \end{aligned}$$

Para concluirmos, é interessante percebermos que, na dedução do estimador da variância, a única condição colocada sobre a distribuição de ε_t é a de possuir segundo momento finito.

Variáveis Explanatórias

Nesta subseção, indicaremos resumidamente um dos possíveis procedimentos a serem seguidos no caso de haver variáveis explanatórias incluídas no modelo (Harvey e Shephard (1993a)). Ou seja, estamos tratando do modelo descrito pelas equações (2.26), (2.27) e (2.28), dadas no início da seção e repetidas aqui por conveniência:

$$y_t = \mathbf{x}'_t \boldsymbol{\beta} + e_t \quad (2.56)$$

$$e_t = \sigma \varepsilon_t \exp\left\{\frac{h_t}{2}\right\} \quad (2.57)$$

$$h_{t+1} = \gamma_0 + \gamma_1 h_t + \eta_{t+1} \quad (2.58)$$

Assume-se a existência do segundo momento de ε_t , $\mathbb{E}[\varepsilon_t] = 0$ e que h_t é estacionário. Os passos são os seguintes:

1. Conduz-se uma regressão de mínimos quadrados ordinários (OLS) de y_t em relação a \mathbf{x}_t .
2. Tomam-se os resíduos obtidos no primeiro passo e aplica-se a modelagem descrita nas subseções anteriores, obtendo-se estimadores de QML para os parâmetros e o estimador suavizado de h_t , $h_{t|T}$.
3. Multiplica-se a equação (2.56) por $\exp\{-\frac{1}{2}h_{t|T}\}$, resultando no modelo

$$\tilde{y}_t = \tilde{\mathbf{x}}'_t \boldsymbol{\beta} + \tilde{u}_t \quad , \quad t = 1, \dots, T$$

onde \tilde{u}_t tem todas as propriedades das observações corrigidas descritas na subseção anterior, ou seja, tem média zero, variância constante e são serialmente não-correlacionadas.

4. Obtêm-se os estimadores de mínimos quadrados generalizados (GLS) de $\boldsymbol{\beta}$ dados por

$$\tilde{\boldsymbol{\beta}} = \left[\sum_{t=1}^T e^{-h_{t|T}} \mathbf{x}_t \mathbf{x}'_t \right]^{-1} \sum_{t=1}^T e^{-h_{t|T}} \mathbf{x}_t y_t$$

Note que este estimador difere do estimador GLS tradicional em que não conhecemos os verdadeiros valores dos parâmetros envolvidos na matriz de correlação dos u_t 's, pois o termo h_t é estocástico.

A matriz de covariâncias de $\tilde{\beta}$, assim como suas propriedades assintóticas e eficiência podem ser encontradas em detalhe em Harvey e Shephard (1993a). Os resultados dependerão dos tipos de variáveis explanatórias colocadas no modelo; elas podem ser determinísticas, estocásticas exógenas ou até mesmo endógenas.

Capítulo 3

Modelos de Volatilidade Estocástica com Deformação Temporal

Neste capítulo, inicialmente introduziremos a noção de Deformação Temporal dentro do contexto de Análise de Séries Temporais Financeiras, justificando seu uso e importância. Após, abordaremos uma metodologia adequada para seu tratamento dentro de Modelos de Volatilidade Estocástica, primeiramente com uma rápida abordagem a tempo discreto e posteriormente, seguindo o trabalho de Ghysels e Jasiak (1994), com uma descrição detalhada a tempo contínuo.

3.1 Por que Deformação Temporal?

Variáveis financeiras, tais como preço de uma ação, são afetadas com a chegada de novas informações ao mercado financeiro. Dias em que informações não surgem, sejam elas de qualquer tipo, “boas” ou “ruins”, são caracterizados por pouca movimentação ou alteração nos preços das ações. Entretanto, dias carregados de novas informações têm efeito justamente oposto, ou seja, apresentam alta variação nos preços das ações. Tal constatação conduziu Mandelbrot e Taylor (1967) e Clark (1973) a modelarem preços de ações como processos estocásticos subordinados, ou seja, ao invés de modelarem os preços como função do tempo de calendário, quer diário, semanal, mensal, etc., eles os modelaram como função

da chegada de nova informação ao mercado, a qual é um processo que flui aleatoriamente ao longo do tempo. Assim, segundo esta abordagem, não olhamos mais para nosso processo de interesse evoluindo através de tempos de calendário igualmente espaçados, e sim, evoluindo de acordo com a chegada de nova informação, originando uma nova dimensão temporal, a qual chamaremos de tempo operacional s . Desta maneira, fica caracterizado o que denominamos de Deformação Temporal.

3.2 Metodologia

Acabamos de ver dentro de um contexto geral a idéia de Deformação Temporal e sua motivação. Entretanto, nosso interesse é o de aplicarmos este conceito dentro do tratamento de Modelos de Volatilidade Estocástica. Inicialmente, abordaremos superficialmente o caso discreto. A seguir, veremos detalhadamente a modelagem a tempo contínuo, sua aproximação a tempo discreto e estimação baseada no Filtro de Kalman e Quase-Máxima Verossimilhança.

3.2.1 Modelagem a Tempo Discreto

Considere o modelo a tempo discreto dado por

$$y_t = \sigma \varepsilon_t \exp\left\{\frac{h_t^+}{2}\right\} \quad (3.1)$$

$$h_t^+ = h_t + \mathbf{x}_t' \boldsymbol{\beta} \quad (3.2)$$

$$h_t = \gamma h_{t-1} + \eta_t \quad (3.3)$$

Perceba que a volatilidade não é mais descrita por um simples processo autorregressivo de ordem 1, AR(1), como foi feito no capítulo anterior. Lá, o que acontecia, era que a volatilidade condicional evoluía latentemente através do tempo de calendário, sem depender, ao menos explicitamente, de outros processos que determinassem sua evolução, embora, implicitamente, através do Filtro de Kalman, sua estimação estivesse intimamente ligada ao processo dos retornos, y_t .

Nesta nova abordagem, a volatilidade condicional depende diretamente dos processos representados pelas variáveis explanatórias, ou seja, sua evolução é

\exists uma variável aleatória U , tal que

$$\lim_{i \rightarrow \infty} \mathbb{E}[(U_i - U)^2] = 0$$

que será escrito como

$$l.i.m._{i \rightarrow \infty} U_i = U$$

Definição 3.2 O processo $X(t)$ é estocasticamente contínuo em $t = t_0$ se

$$l.i.m._{t \rightarrow t_0} X(t) = X(t_0)$$

isto é, se e somente se

$$\lim_{t \rightarrow t_0} \mathbb{E}[\{X(t) - X(t_0)\}^2] = 0$$

Definição 3.3 $X(t)$ é dito ser estocasticamente diferenciável em $t = t_0$, com derivada

$$\left[\frac{dX(t)}{dt} \right]_{t=t_0} = \dot{X}(t_0)$$

se e somente se

$$l.i.m._{h \rightarrow 0} \left\{ \frac{X(t_0 + h) - X(t_0)}{h} \right\}$$

existe e é igual a $\dot{X}(t_0)$, isto é, se e somente se

$$\lim_{h \rightarrow 0} \mathbb{E} \left[\left\{ \frac{X(t_0 + h) - X(t_0)}{h} \right\}^2 \right]$$

existe.

Definição 3.4 $W(t)$ é um processo de Wiener ou movimento Browniano se satisfaz as seguintes condições:

- i) $W(0) = 0$;
- ii) $\mathbb{E}[W(t)] = 0$, para $\forall t > 0$;
- iii) $W(t)$ é normal, para cada $t > 0$;
- iv) $W(t)$ tem incrementos independentes e estacionários.

Definição 3.5 *Uma equação diferencial estocástica de primeira ordem, para o caso particular em que $W(t)$ é um processo de Wiener, é dada pela seguinte relação:*

$$dX(t) = a(t, X(t), \theta)dt + b(t, X(t), \theta)dW(t)$$

onde a e b são funções limitadas e mensuráveis para $t \in [0, T]$, θ é o vetor de parâmetros e $X(t)$ é a solução da equação.

Se a e b forem funções lineares de $X(t)$, então a equação acima será chamada de equação diferencial estocástica linear de primeira ordem.

Definição 3.6 $X(t)$ é dito ser um **AR(1) a tempo contínuo** se satisfizer a seguinte equação diferencial estocástica linear de primeira ordem:

$$dX(t) + \alpha X(t)dt = dW(t)$$

onde $W(t)$ é um processo de Wiener.

Convém salientar que, devido à natureza de $W(t)$, o processo $X(t)$ dado pela equação acima é contínuo mas não é diferenciável (ver Priestley (1981)).

Agora, mais familiarizados com estas definições, utilizando operações de limite para os processos de preços e volatilidade, uniremos os trabalhos de Mandelbrot e Taylor (1967) e Clark (1973) sobre processos estocásticos subordinados aplicados a finanças e os trabalhos de Hull e White (1987), entre outros, que modelaram as variações na volatilidade como uma equação diferencial estocástica (a qual não tinha, ao menos explicitamente, relação com a chegada de informações), para construirmos nosso Modelo de Volatilidade Estocástica (a tempo contínuo) com Deformação Temporal.

Formulação do Modelo

Consideremos o Modelo de Volatilidade Estocástica a tempo contínuo sem deformação dado pelas equações diferenciais estocásticas abaixo:

$$dx(t) = \mu x(t)dt + \sigma(t)x(t)dW_1(t) \quad (3.10)$$

$$d \log \sigma^2(t) = a(\log \sigma^2(t) - b)dt + cdW_2(t) \quad (3.11)$$

onde $W_1(t)$ e $W_2(t)$ são dois processos de Wiener ou Movimentos Brownianos padrão.

Note que estas equações tratam a volatilidade como um processo que evolui contínua e suavemente através do tempo. No entanto, a chegada de informação, que é o processo ao qual vinculamos a volatilidade, não evolui desta maneira contínua e suave. Além disso, o fluxo de informação é um processo latente (não-observável). Para contornarmos esta característica de não-observabilidade, nossa abordagem tratará a chegada de nova informação, que constitui o tempo operacional s , como uma função de processos observáveis no tempo de calendário t , ou seja, $s = g(t)$, onde $g(\cdot)$ será denominada de “função de conexão” (ou de deformação) entre o tempo de calendário e o tempo operacional. Assim, podemos reescrever (3.10) e (3.11) como

$$dx(t) = \mu x(t)dt + \sigma(g(t))x(t)dW_1(t) \quad (3.12)$$

$$d \log \sigma^2(s) = a(\log \sigma^2(s) - b)ds + cdW_2(s) \quad (3.13)$$

as quais vinculam a volatilidade ao tempo operacional s , causando a deformação temporal.

Veja que a equação (3.13) é uma equação diferencial estocástica linear de primeira ordem. A solução¹ é obtida facilmente, conforme segue abaixo:

$$d \log \sigma^2(s) = a(\log \sigma^2(s) - b)ds + cdW_2(s)$$

Então

$$\frac{d \log \sigma^2(s)}{ds} - a \log \sigma^2(s) = -ab + cW_2(s)$$

Multiplicando ambos os lados pelo fator integrante $\exp\{-as\}$, obtemos

$$\frac{d}{ds}(\log \sigma^2(s)e^{-as}) = e^{-as}(-ab + cW_2(s))$$

Tomando a integral entre s' e s , para $s' < s$, chegamos a

$$\log \sigma^2(s)e^{-as} - \log \sigma^2(s')e^{-as'} = b(e^{-as} - e^{-as'}) + c \int_{s'}^s e^{-ar} dW_2(r)$$

Que resulta em

$$\log \sigma^2(s) = \log \sigma^2(s')e^{a(s-s')} + (1 - e^{a(s-s')})b + c \int_{s'}^s e^{a(s-r)} dW_2(r) \quad (3.14)$$

¹Ver, por exemplo, Gihman (1972) e Wu (1985).

Definição da Função de Deformação $g(t)$

Como falamos anteriormente, o fluxo de informações que chegam ao mercado é um processo latente. Sendo assim, o caminho adotado neste trabalho foi o de estabelecer uma função de conexão entre o tempo de calendário t e o tempo operacional s , a qual é definida em termos de processos observáveis no tempo de calendário. A escolha destes processos é de importância fundamental, pois eles determinarão o ritmo de variação da volatilidade. Os processos utilizados, seguindo a abordagem de Ghysels e Jasiak (1994), serão volumes de negociação, retornos, módulos dos retornos (todos defasados em uma unidade de tempo, pois também temos como objetivo a previsão) e também algum processo que considere o efeito que dias sem negociação provocam no mercado ou que possua algum tipo de periodicidade. A inclusão dos retornos (sem módulo) defasados permite que o tempo operacional evolua em ritmos diferenciados quando o movimento dos preços é crescente ou decrescente, fazendo desta maneira com que a própria volatilidade apresente flutuações diferenciadas. Em Ghysels e Jasiak (1994), utilizando dados diários da S&P500 e NYSE de 1950 a 1987, foi encontrado que aumento em volume de negociação acelera o tempo operacional, resultando em menos persistência no processo de volatilidade com maior variância de seu ruído. Declínios nos preços, ou seja, retornos negativos, ocasionaram efeitos similares, enquanto preços crescentes, ou retornos positivos, aceleraram menos o ritmo do tempo operacional.

Formalmente, considere o Espaço de Probabilidades $(\Omega, \mathcal{F}, \mathbb{P})$ e a família $F = \{\mathcal{F}_t\}_{t=0}^{\infty}$ de sub- σ -álgebras definidas no tempo de calendário. Além disso, seja \mathbf{Z}_t um vetor m -dimensional \mathcal{F}_t -mensurável. Consideraremos a função de deformação $g(\cdot)$ uma função linear de t e assumiremos que seus incrementos, os quais estabelecem a velocidade de variação do tempo operacional, serão \mathcal{F}_{t-1} -mensuráveis através da transformação logística

$$\frac{dg(\tau; \mathbf{Z}_{t-1})}{d\tau} \equiv \dot{g}(\tau; \mathbf{Z}_{t-1}) \equiv \frac{\exp \mathbf{c}' \mathbf{Z}_{t-1}}{\frac{1}{T} \sum_{t=1}^T \exp \mathbf{c}' \mathbf{Z}_{t-1}} \quad (3.15)$$

para $t-1 \leq \tau < t$, onde \mathbf{c} é um vetor $m \times 1$.²

Note que se $\mathbf{c} = \mathbf{0}$, então $\dot{g}(\tau; \mathbf{Z}_{t-1}) = 1$, implicando na não-existência de

²Veja que o denominador de (3.15) possui uma soma que inclui variáveis que não são \mathcal{F}_{t-1} -mensuráveis, pois a soma estende-se até T . Entretanto, conforme Ghysels e Jasiak (1994), o denominador, que é um fator de escala, só está ali por razões de estabilidade numérica e sua presença não tem importância conceitual.

deformação temporal. Também

$$\Delta g(t) = g(t) - g(t-1) = \int_{t-1}^t \dot{g}(\tau; \mathbf{Z}_{t-1}) d\tau = \dot{g}(\tau; \mathbf{Z}_{t-1}) \quad (3.16)$$

onde a segunda igualdade impõe a ausência de algum possível salto de g em t , e a terceira igualdade decorre do fato de $\dot{g}(\tau; \mathbf{Z}_{t-1})$ não depender de τ .

Por construção, observamos diretamente que

$$\frac{1}{T} \sum_{t=1}^T \Delta g(t) = \frac{1}{T} \sum_{t=1}^T \dot{g}(\tau; \mathbf{Z}_{t-1}) = 1$$

Aliadas a estas condições de especificação, adicionam-se as 2 seguintes suposições:

- i) $g(0) = 0$
- ii) $0 < \dot{g}(\tau; \mathbf{Z}_{t-1}) < \infty$

Assim, as condições de especificação, aliadas a estas duas suposições, garantem que o tempo operacional evolua na mesma direção que o tempo de calendário, sem paradas ou saltos.

Nos falta ainda definir precisamente a forma funcional de \mathbf{Z}_{t-1} , a qual é dada do seguinte modo:

$$\exp\{\mathbf{c}' \mathbf{Z}_{t-1}\} \equiv \exp\{\mathbf{c}'_d \mathbf{d}_t + c_v \log Vol_{t-1} + c_r y_{t-1} + c_a | y_{t-1} |\} \quad (3.17)$$

onde

Vol_t representa o volume de negociação no instante t ;

$y_t = \Delta \log x_t$ é o retorno no instante t ;

\mathbf{d}_t é um processo que permite a inclusão de efeitos de dias sem negociação ou de outros modelos periódicos.

Perceba que uma avaliação dos coeficientes da relação (3.17) nos permite concluir como efeitos de cada um dos processos influenciam na velocidade de variação do tempo operacional e conseqüentemente na volatilidade do processo dos retornos. Para isto, basta observarmos o sinal de a na equação (3.21). Assim, perceberemos, para cada tipo de variação das variáveis incluídas em (3.17), como comportam-se a persistência e o ruído do processo de volatilidade.

Aproximação a Tempo Discreto e Estimação

Considere as equações (3.12) e (3.14), reescritas aqui novamente:

$$dx(t) = \mu x(t)dt + \sigma(g(t))x(t)dW_1(t) \quad (3.18)$$

$$\log \sigma^2(s) = \log \sigma^2(s')e^{a(s-s')} + (1 - e^{a(s-s')})b + c \int_{s'}^s e^{a(s-r)} dW_2(r) \quad (3.19)$$

Nossa intenção é deduzirmos aproximações a tempo discreto para o conjunto de equações dado acima. Tais aproximações serão necessárias para que possamos estimar, tanto os parâmetros do modelo, quanto a volatilidade condicional via Filtro de Kalman e método de QML. Assim, substituindo $s = g(t)$ e $s' = g(t-1)$ em (3.19), resulta em

$$h_t = e^{a\Delta g(t)} h_{t-1} + (1 - e^{a\Delta g(t)})b + \nu_t \quad (3.20)$$

onde $h_t = \log \sigma^2(g(t))$.

Ainda podemos escrever

$$\tilde{h}_t = e^{a\Delta g(t)} \tilde{h}_{t-1} + \nu_t \quad (3.21)$$

onde

$$\tilde{h}_t = h_t - b;$$

$$\nu_t \sim N(0; [-c^2(1 - e^{2a\Delta g(t)})]/[2a])$$

e $\Delta g(t)$ é dada por (3.16).

No que diz respeito à distribuição de ν_t , perceba que $W_2(t)$ foi definido como um Movimento Browniano padrão, ou seja, $dW_2(t) \sim NIID(0; dt)$. Assim, para

$$\nu_t = c \int_{g(t-1)}^{g(t)} e^{a(g(t)-r)} dW_2(r)$$

temos que sua distribuição é normal com $\mathbb{E}[\nu_t] = 0$ e variância dada por

$$\begin{aligned} \text{Var}[\nu_t] &= c^2 \int_{g(t-1)}^{g(t)} e^{2a(g(t)-r)} \text{Var}[dW_2(r)] \\ &= c^2 e^{2ag(t)} \int_{g(t-1)}^{g(t)} e^{-2ar} dr \\ &= -\frac{c^2 e^{2ag(t)}}{2a} [e^{-2ag(t)} - e^{-2ag(t-1)}] \\ &= -c^2(1 - e^{2a\Delta g(t)})/(2a) \end{aligned}$$

Agora, comparando (3.20) com (3.19), observamos que o processo do logaritmo da volatilidade, h_t , enquanto linear em relação ao tempo operacional, transforma-se em um modelo com coeficiente aleatório em relação ao tempo de calendário. Este tipo de processo também é chamado de processo duplamente estocástico, pois, além de seguir um AR(1), seu coeficiente é dirigido pelo processo estocástico $\Delta g(t)$.

Obtida esta aproximação para a equação (3.19), nos resta aproximarmos (3.18). Utilizando uma aproximação padrão, chegamos a

$$\log x_t = \mu + \log x_{t-1} + \sigma(g(t))\varepsilon_t$$

ou ainda

$$y_t - \mu = \sigma(g(t))\varepsilon_t \quad (3.22)$$

onde

y_t é o retorno no instante t

e $\varepsilon_t \sim NIID(0; 1)$.

Finalmente, elevando ao quadrado e tomando logaritmos em ambos os lados de (3.22), e juntando com a equação (3.21), obtemos o seguinte conjunto de equações:

$$\log(y_t - \mu)^2 = -1,27 + b + \tilde{h}_t + \xi_t \quad (3.23)$$

$$\tilde{h}_t = e^{a\Delta g(t)}\tilde{h}_{t-1} + \nu_t \quad (3.24)$$

onde $E[\xi_t] = 0$ e $\text{Var}[\xi_t] = \pi^2/2$.

Observando as equações obtidas, vemos que elas se tratam de um modelo de espaço de estados não-linear nos parâmetros, com coeficientes variando ao longo do tempo, o que as diferenciam das equações encontradas no capítulo anterior para a volatilidade sem Deformação Temporal. Entretanto, elas estão sujeitas ao tratamento via Filtro de Kalman e estimação de QML.³ Para tanto, seguindo Stock (1988), utilizaremos as seguintes condições iniciais:

- $\Delta g(1) = 1$

³O parâmetro μ poderá ser estimado através da média amostral dos retornos em uma etapa anterior, utilizando-se então os desvios da média amostral na modelagem posterior. Veremos que, na parte aplicada de nosso trabalho, utilizamos os resíduos de um AR(1) como entrada neste modelo.

- $\tilde{h}_{1/0} = 0$
- $\text{Var}[\tilde{h}_{1/0}] = \frac{\sigma_v^2}{1 - \exp\{2a\}} = -\frac{c^2}{2a}$

3.3 Conjuntos de Dados em Alta Frequência

Mercados financeiros operam, durante seu funcionamento, continuamente, em uma base de dados de alta frequência, chegando a milhares de cotações por dia. Entretanto, até bastante recentemente, os conjuntos de dados de praticamente todos os mercados vinham sendo armazenados em uma base discreta, ou seja, obtendo-se medidas resumo que raramente originavam uma frequência maior do que diária.

Atualmente, o desenvolvimento da tecnologia computacional tem permitido o armazenamento destes imensos bancos de dados, originando séries de dados com milhões de observações, as quais, conforme Ghysels, Gouriéroux e Jasiak (1995), são mais facilmente medidas pelo espaço de memória que ocupam, do que pelo número de observações propriamente dito.

Tais imensos conjuntos de dados podem ser extremamente valiosos em termos de possibilidade de informação. Entretanto, uma propriedade fundamental deste tipo de dados em alta frequência, também chamados cotação-a-cotação, é que as observações ocorrem a intervalos de tempo variáveis. As negociações não se desenvolvem ao longo do tempo de calendário, igualmente espaçado. Esta natureza esporádica e aleatória da negociação torna a medição da volatilidade problemática para os modelos que utilizam uma abordagem discreta.

Entretanto, nosso modelo de Volatilidade Estocástica (a tempo contínuo) com Deformação Temporal é capaz de lidar com estas dificuldades. Na realidade, sua base de construção é contínua, e a idéia de tempo deformado ou tempo operacional encaixa-se perfeitamente com os tempos obtidos entre negociações para os dados em alta frequência. Uma abordagem da teoria e estimação destes modelos pode ser encontrada em Ghysels, Gouriéroux e Jasiak (1995).

Capítulo 4

Aplicação: Estimação da Volatilidade dos Retornos do IBOVESPA

Vimos no capítulo 1 que os retornos do IBOVESPA são autocorrelacionados. Entretanto, todos os modelos que foram descritos ao longo do trabalho tinham como característica a ausência de autocorrelação para a variável de interesse. Assim, foi necessário que modificássemos de alguma maneira nossa variável de entrada no modelo, os retornos. Baseados em suas funções de autocorrelação e autocorrelação parcial, o procedimento escolhido foi o de ajustar a eles um AR(1). Encontrou-se significativo o seguinte modelo:

$$y_t = 0,1478y_{t-1} + a_t \quad (4.1)$$

A figura 4.1 mostra as funções de autocorrelação dos resíduos estimados, \hat{a}_t , do modelo dado acima, assim como dos seus quadrados. Percebemos claramente a eliminação daquela estrutura de autocorrelação observada na figura 1.2 para os retornos, enquanto os quadrados permanecem altamente correlacionados, satisfazendo-se as características desejadas. Os valores para a estatística de Box-Ljung portmanteau¹ para as autocorrelações dos resíduos, juntamente com os níveis descritivos amostrais são dados na tabela 4.1.

Deste modo, em toda análise posterior, utilizaremos os resíduos do AR(1) para

¹Ver apêndice.

obtermos as estimativas da volatilidade e dos parâmetros associados. As outras variáveis que entrarão nos modelos com Deformação Temporal são os retornos, seus valores absolutos, volumes de negociação em dólares (com escala igual à escala original $\times 10^{-9}$), todos defasados em 1 “lag”, e número de noites entre dois dias consecutivos em que o mercado esteve aberto (para incorporar algum efeito de dias sem negociação, ou seja, efeito do tempo em que o mercado esteve fechado antes do dia em observação; sabe-se, por exemplo, que o mercado reage de maneira diferenciada após um fim-de-semana). A figura 4.2 mostra os resíduos estimados do AR(1), assim como os volumes de negociação diários do IBOVESPA.

Procedeu-se à estimação dos parâmetros e das volatilidades em 7 modelos diferentes: 1 sem deformação temporal, 3 com deformação, utilizando-se a metodologia a tempo discreto, e outros 3 com deformação, utilizando-se a abordagem a tempo contínuo. As tabelas 4.2 e 4.3 nos fornecem um resumo das estimativas encontradas para os parâmetros em cada um dos modelos, assim como seus níveis descritivos amostrais. Os parâmetros c_d , c_v , c_r e c_a são respectivamente os coeficientes das variáveis número de noites entre 2 dias com negociação, volume de negociação reescalado, retornos e valores absolutos dos retornos. Os demais parâmetros são equivalentes aos especificados na parte teórica do trabalho. Apesar dos elevados valores dos níveis críticos (descritivos) amostrais para várias estimativas da abordagem a tempo contínuo, uma análise dos coeficientes obtidos pode nos informar como variações nas variáveis afetam o ritmo de evolução do tempo operacional e a própria volatilidade. Se observarmos o modelo 1 a tempo contínuo, veremos que o coeficiente a é negativo. Olhando a equação (3.24), isto significa dizer que quanto maior for $\Delta g(t)$, a velocidade de variação do tempo operacional, menor será a persistência do processo de volatilidade; ao mesmo tempo, maior será a variância do ruído ν_t . Sendo assim, agora olhando para a forma funcional de $\Delta g(t)$, dada pela equação (3.17), vemos que como c_v é positivo, então, quanto maior o volume de negociação do dia anterior, maior $\Delta g(t)$ e, conseqüentemente, menos persistência e mais ruído para o processo de volatilidade. Do mesmo modo, quanto maior o valor absoluto do retorno do dia anterior, mais volatilidade; entretanto, quedas nos preços (retornos negativos) conduzem a uma volatilidade ainda maior, pois c_r é negativo.

As figuras 4.3 e 4.4 mostram as estimativas suavizadas da volatilidade (desvio padrão dos retornos) na abordagem a tempo discreto (baseadas nas observações corrigidas por heteroscedasticidade) e os gráficos de dispersão para os 4 modelos estimados. Percebemos claramente que a volatilidade estimada sem deformação (modelo 1) apresenta uma forma mais suave do que as demais, enquanto o modelo

2 tem a forma mais errática. Já os modelos 3 e 4 fornecem resultados bastante semelhantes. Na figura 4.4, vemos que o modelo 2, além de estimar uma volatilidade mais errática, origina uma amplitude de variação bem maior do que os demais modelos, sendo tal consequência atribuída à presença do módulo dos retornos no modelo.

Quanto à adequação destes modelos a tempo discreto, podemos observar as funções de autocorrelação dos quadrados das observações corrigidas por heteroscedasticidade, dadas na figura 4.5. Comparando com a figura 4.6, que no gráfico de baixo mostra as autocorrelações dos quadrados dos resíduos, vemos que há uma sensível queda nos valores das autocorrelações, sendo que o modelo 2 apresenta a pior performance.

Já a figura 4.7 contém as estimativas suavizadas da volatilidade sem e com deformação a tempo contínuo (modelo 1). Vemos que, ao contrário do que aconteceu na abordagem a tempo discreto, a estrutura de tempo deformado oferece uma estimativa da volatilidade mais suave ao longo do tempo. Tal característica é oriunda do fato do coeficiente do processo autorregressivo do logaritmo da volatilidade ser também um processo estocástico, e não uma constante. Isto faz com que este coeficiente absorva parte do ruído do AR(1). Esta diminuição de ruído pode acarretar vantagens em termos de previsão. Na figura 4.8, mostramos os horizontes de previsão em termos de 1 passo a frente, para um período de aproximadamente 150 dias, baseados nas estimativas filtradas da volatilidade. Também é fornecido o gráfico de dispersão das volatilidades suavizadas estimadas para estes casos, dado na figura 4.9. Percebe-se que o modelo com deformação apresenta uma amplitude de variação um pouco menor.

Em termos de medidas de ajuste do modelo 1 a tempo contínuo, podemos observar as figuras 4.10 e 4.11. Na figura 4.10, a qual mostra as autocorrelações dos quadrados dos resíduos e dos quadrados dos resíduos corrigidos por heteroscedasticidade pelo modelo 1 a tempo contínuo, vemos que há uma sensível diminuição na forte estrutura de autocorrelação apresentada pelos quadrados dos resíduos quando olhamos para os quadrados dos resíduos corrigidos. Isto é um sinal de que o modelo realmente capturou boa parte da variação da volatilidade. Entretanto, percebe-se que ainda ocorre a permanência de autocorrelação. Tal fato é confirmado pelo histograma dos resíduos corrigidos, que apresenta caudas ainda bastante pesadas. Talvez uma abordagem a tempo contínuo que substitui-se o movimento Browniano na equação que modela o preço por algum outro processo que implicasse em uma distribuição a tempo discreto com caudas mais pesadas que a normal oferecesse melhores resultados.

lag	gl	autcor	BoxLjung	prob(gl)
1	0	-0.0053		
2	1	0.0354	3.1097	0.0778
3	2	-0.0025	3.1250	0.2096
4	3	0.0224	4.3472	0.2263
5	4	0.0317	6.7949	0.1471
6	5	-0.0226	8.0406	0.1540
12	11	-0.0023	16.1982	0.1339
18	17	-0.0230	24.3978	0.1090
24	23	-0.0196	27.8603	0.2211
30	29	0.0272	35.6009	0.1855
36	35	0.0126	39.6314	0.2710

Tabela 4.1: Estatística de Box-Ljung para as autocorrelações dos resíduos do AR(1).

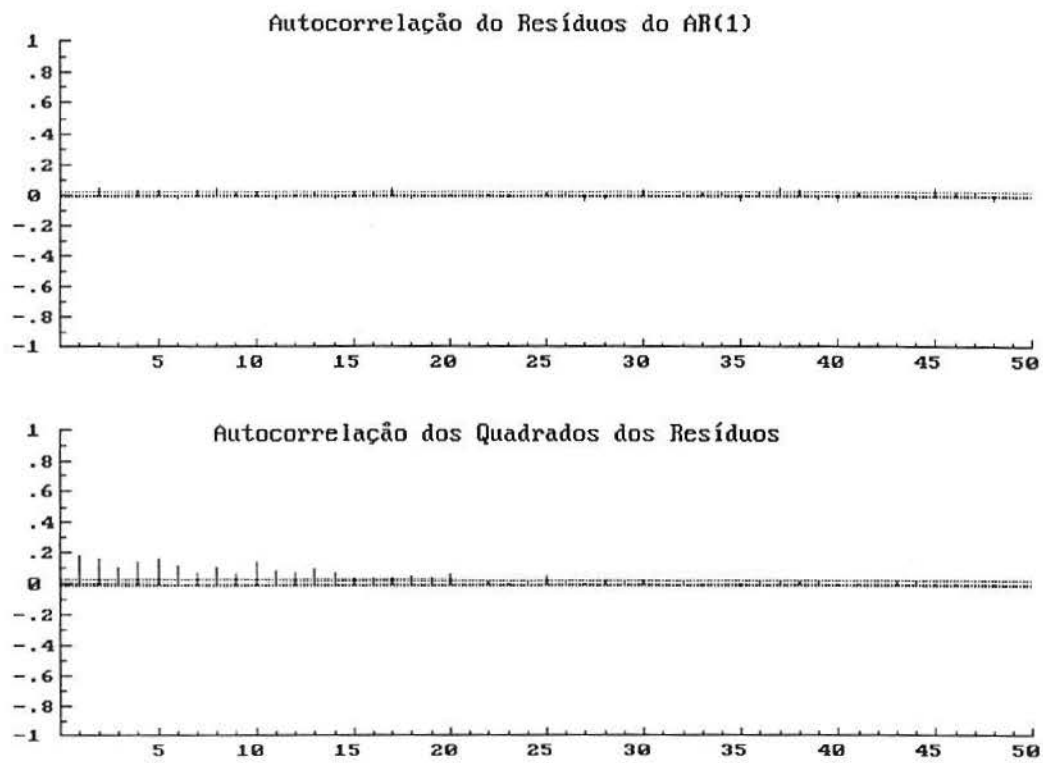


Figura 4.1: Funções de Autocorrelação.

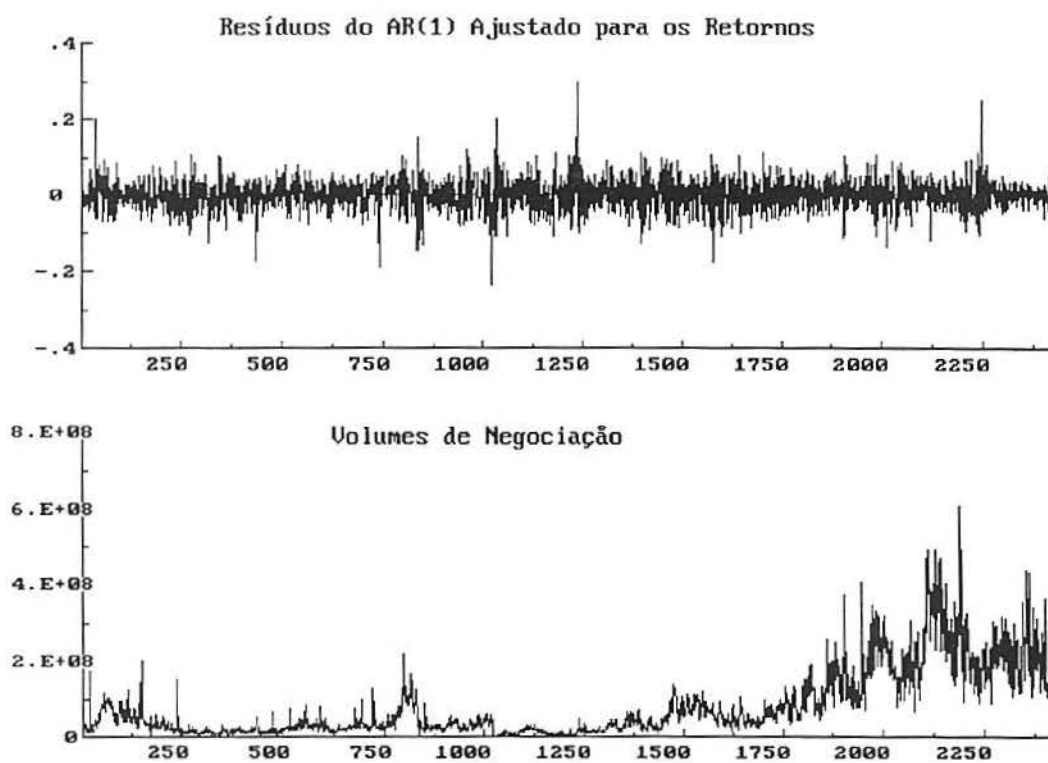


Figura 4.2: Resíduos do AR(1) e volumes de negociação diários em dólares do IBOVESPA.

	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
	Est	Prob	Est	Prob	Est	Prob	Est	Prob
c_d	-	-	0.1043	0.0060	0.0900	0.0124	0.0902	0.0122
c_v	-	-	-0.7770	0.02267	-0.3606	0.6550	-	-
c_r	-	-	-2.8739	0.0005	-2.7738	0.0012	-2.8124	0.0010
c_a	-	-	18.6730	0.0000	-	-	-	-
σ_ξ	1.7054	-	0.4939	-	1.7020	-	1.7019	-
σ_η	0.1370	-	1.6888	-	0.1332	-	0.1324	-
γ_1	0.9716	-	-0.1894	-	0.9721	-	0.9725	-

Tabela 4.2: Modelagem a Tempo Discreto

	Modelo 1		Modelo 2		Modelo 3	
	Est	Prob	Est	Prob	Est	Prob
c_v	7.5695	0.0303	-	-	-	-
c_r	-2.9051	0.3816	-9.5452	0.2682	-	-
c_a	67.0863	0.0000	39.5020	0.0537	2.9767	0.3448
c	3.8200	0.4277	0.3442	0.4477	0.0735	0.0012
a	-29.3738	0.4312	-0.2670	0.4474	-0.0136	0.0010
b	-6.0370	0.0000	-6.0930	0.0000	-6.3573	0.0000

Tabela 4.3: Modelagem a Tempo Contínuo.

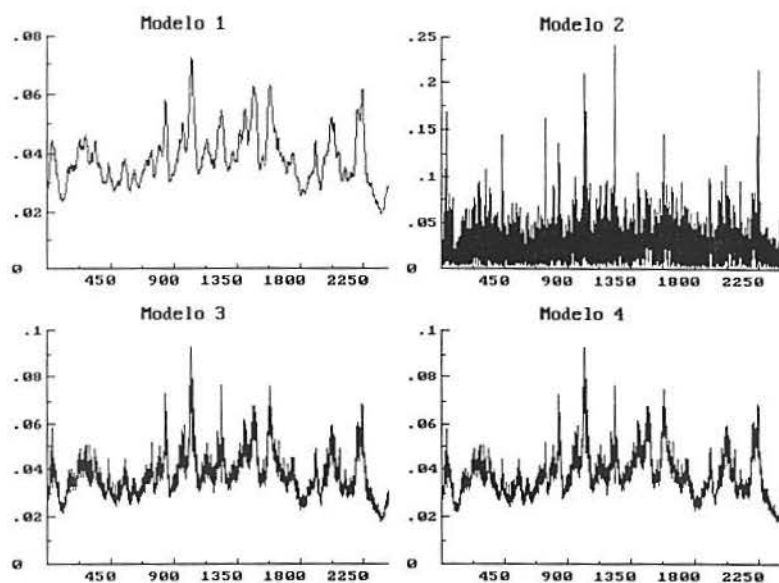


Figura 4.3: Estimativas Suavizadas obtidas dos Modelos a Tempo Discreto.

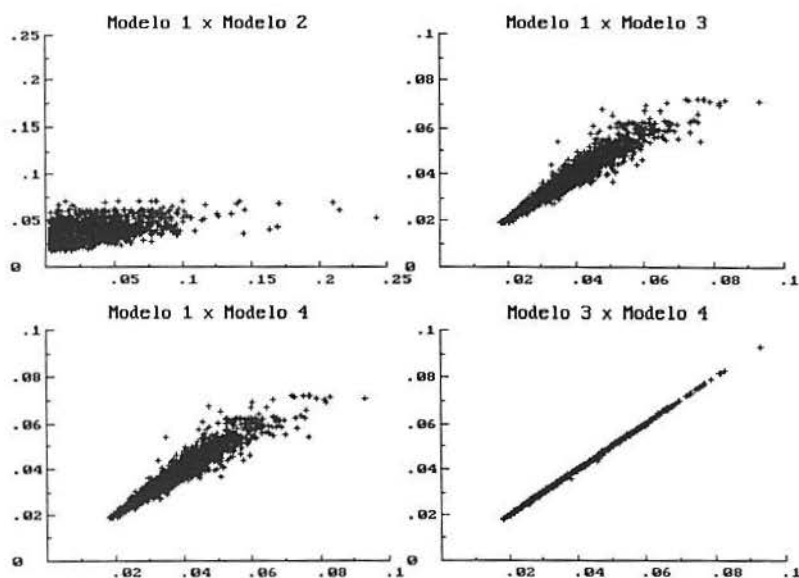


Figura 4.4: Gráficos de Dispersão das Volatilidades obtidas na Modelagem a Tempo Discreto.

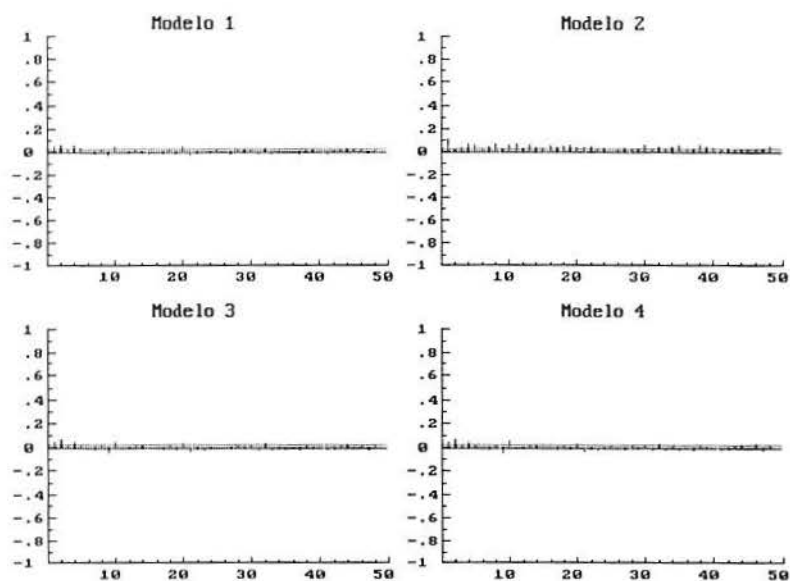


Figura 4.5: Autocorrelações dos quadrados das observações corrigidas por heteroscedasticidade.

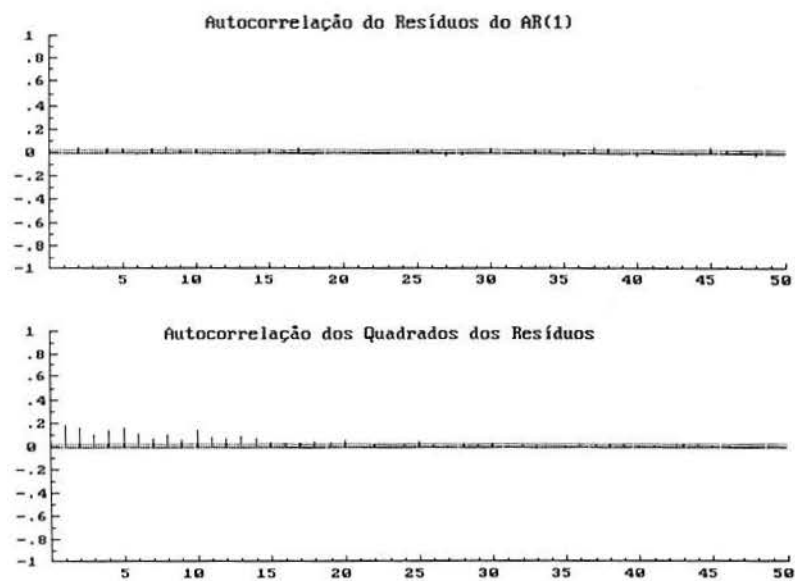


Figura 4.6: Autocorrelações dos resíduos e dos seus quadrados.

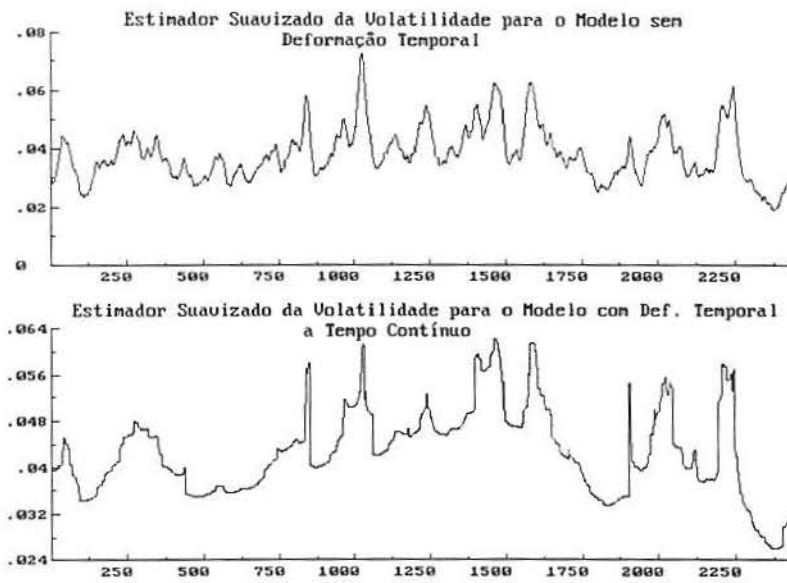


Figura 4.7: Estimativas Suavizadas da Volatilidade. No gráfico de baixo, utilizou-se o Modelo 1 a Tempo Contínuo.

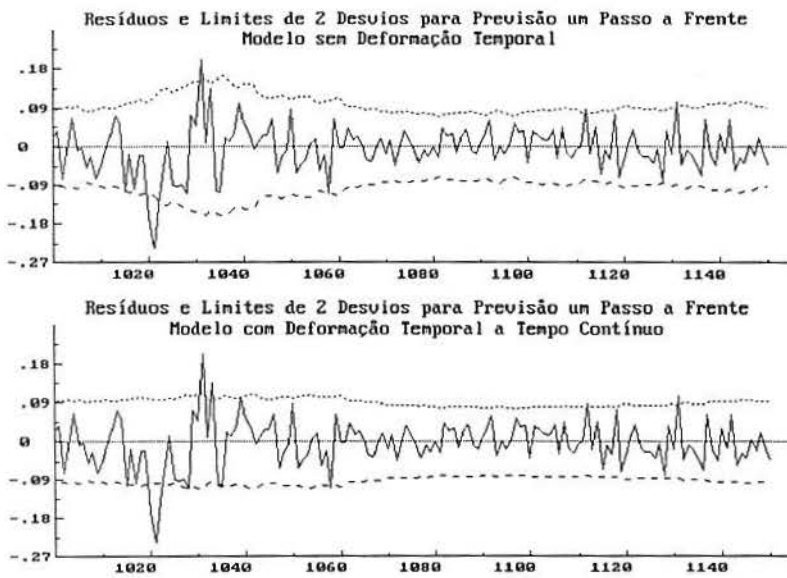


Figura 4.8: Previsões 1 passo a frente para a Volatilidade.

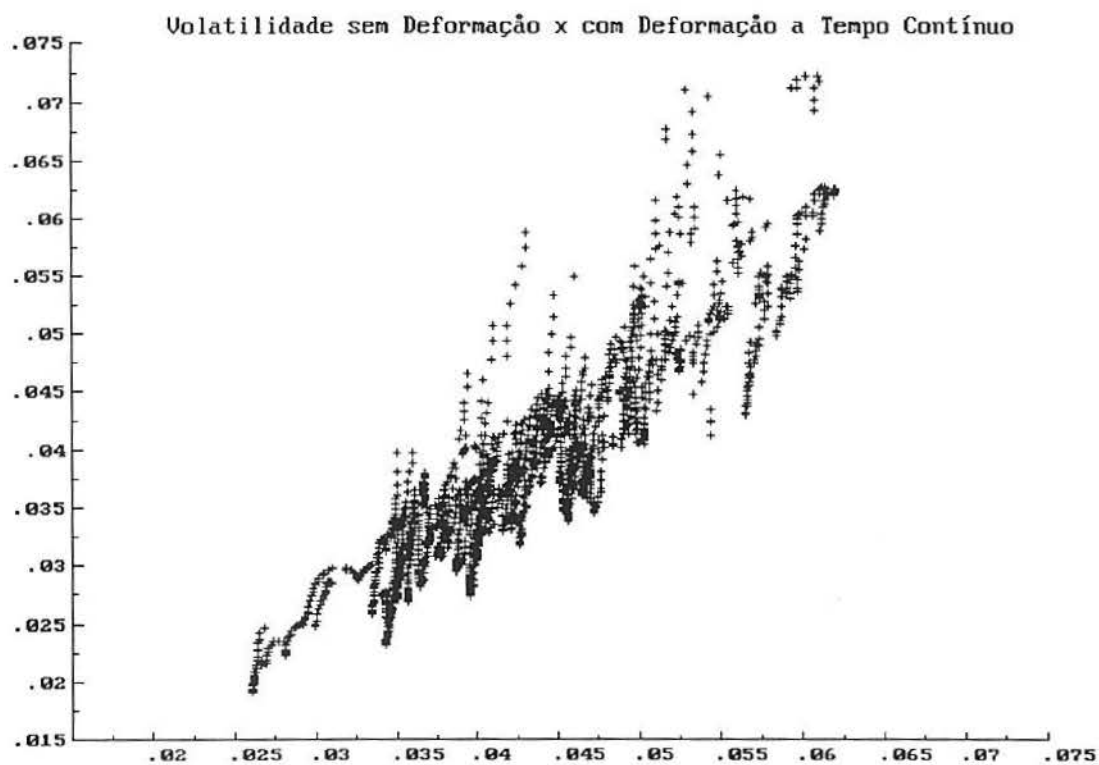


Figura 4.9: Gráfico de dispersão entre as volatilidades suavizadas do modelo sem deformação e do modelo 1 com defromação a tempo contínuo.

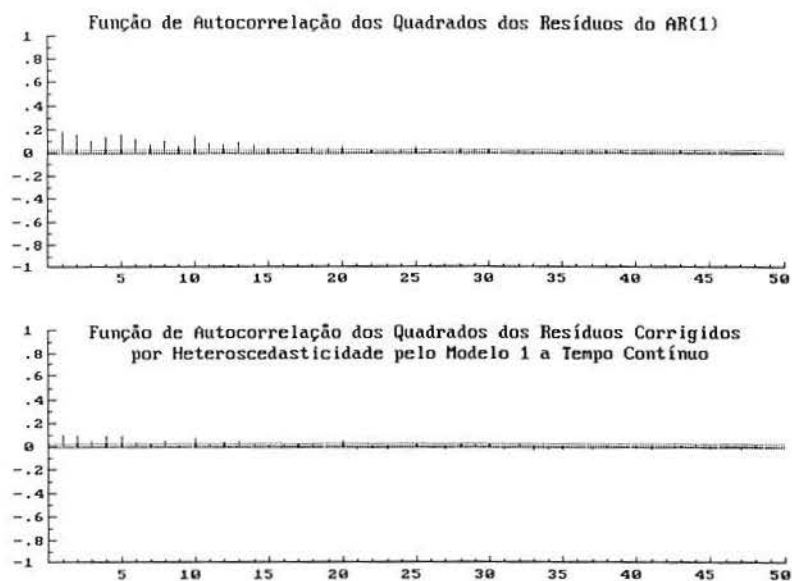


Figura 4.10: Autocorrelações dos Resíduos e Resíduos Corrigidos.

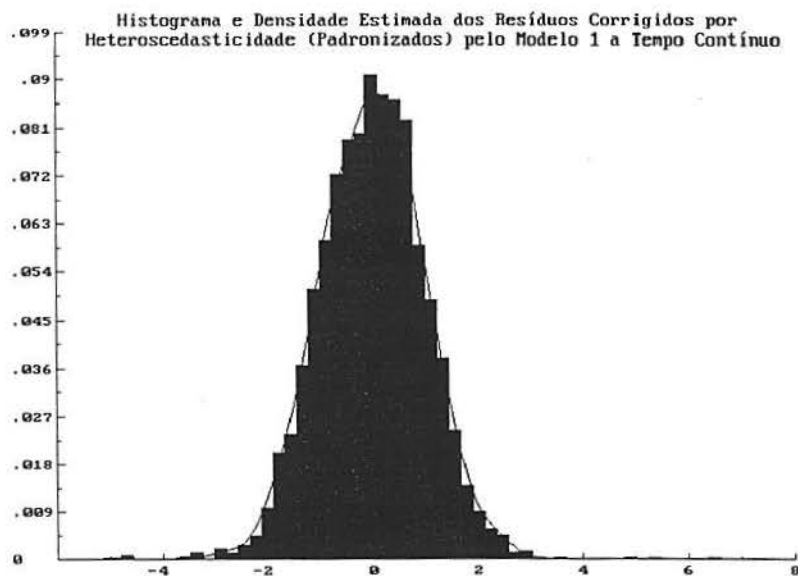


Figura 4.11: Histograma e Densidade Não-Paramétrica Ajustada.

Conclusão

A principal intenção deste trabalho foi a de modelar a volatilidade dos retornos de ativos financeiros através de Modelos de Volatilidade Estocástica (a tempo contínuo) com Deformação Temporal. Note que no decorrer do trabalho também foi desenvolvida a modelagem a tempo discreto; entretanto o maior interesse concentra-se na abordagem a tempo contínuo, quer por sua maior flexibilidade, quer por seus horizontes de aplicação mais promissores. A intenção foi realizada, tanto a nível teórico, como prático. Nesta parte final do trabalho, salientaremos as vantagens obtidas com a utilização da modelagem com deformação a tempo contínuo.

Sua principal vantagem é o suavizamento das estimativas da volatilidade. A introdução do coeficiente aleatório no processo autorregressivo AR(1) para o logaritmo da variância fez com que ele absorvesse boa parte da erraticidade do processo, a qual, no modelo sem deformação, era atribuída ao termo de perturbação. Sendo assim, o processo duplamente estocástico resultante é preferível por conter menos ruído. Uma segunda vantagem é a possibilidade de incorporação de processos observáveis com um apelo empírico subjacente, ou seja, a idéia de alteração do mercado à medida em que novas informações chegam permite a inclusão destes processos, observados no tempo de calendário, de uma forma natural, incorporados na função de deformação temporal. Uma terceira vantagem é a relação que podemos obter entre variações destes processos e o comportamento da volatilidade. Em nossa aplicação aos dados do IBOVESPA, encontramos que quanto maior o volume negociado no dia anterior, maior a volatilidade do dia em questão. Também encontramos que retornos negativos no dia anterior conduzem a uma volatilidade maior do que retornos positivos (ambos de mesmo módulo), o que significa dizer que, para nosso conjunto de dados, quando os preços decaíam, a volatilidade subsequente era maior do que quando os preços aumentavam. Uma última vantagem a ser citada aqui é a utilização desta abordagem para conjuntos

de dados em alta frequência, os quais surgem aleatória e desigualmente espaçados ao longo do tempo. Lembramos que os modelos a tempo discreto não são capazes de tratar de modo apropriado com este tipo de dados. Para encerrar, uma desvantagem desta modelagem é a impossibilidade de prever-se a volatilidade mais do que 1 passo a frente, embora seja possível fazer previsões condicionais se os processos que influenciam a volatilidade forem modelados através de, por exemplo, modelos autorregressivos.

Apêndice

Estadística de Box-Ljung portmanteau

$$Q(p, q) = T(T + 2) \sum_{\tau=1}^p \frac{r_{\tau}^2}{T - p}$$

Sob a hipótese nula de que todos os coeficientes de autocorrelação até a ordem p são nulos, a estatística Q tem distribuição assintótica Qui-quadrado com q graus de liberdade, onde q é igual a p se nenhum modelo foi ajustado, ou é igual a $p + 1 - n$ se algum modelo foi ajustado (n é igual ao número de parâmetros estimados pelo modelo ajustado).

Teste para Curtose

$$k = \frac{T(b_2 - 3)^2}{24}$$

Teste para Assimetria

$$s = \frac{T^2 b_1}{6}$$

onde

$$b_1 = \frac{m_3^2}{m_2^3};$$

$$b_2 = \frac{m_4}{m_2^2}$$

$$\text{e } m_i = \frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^i.$$

As distribuições assintóticas destas estatísticas, sob as hipóteses, respectivamente, de não haver excesso de curtose e de haver simetria, são cada uma Qui-quadrado com 1 grau de liberdade.

Referências Bibliográficas

- [1] M. Abramowitz and N. C. Stegun. *Handbook of Mathematical functions*. New York: Dover Publications Inc., 1970.
- [2] C. Alexander. *The Handbook of Risk Management and Analysis*. Wiley, 1996.
- [3] R. T. Baille and T. Bollerslev. Prediction in dynamic models with time-dependent conditional variances. *J. Econometrics*, 52:91–113, 1992.
- [4] T. Bollerslev. Generalised autoregressive conditional heteroskedasticity. *J. Econometrics*, 51:307–327, 1986.
- [5] T. Bollerslev, R. Y. Chou, and K. F. Kroner. Arch modeling in finance: a review of the theory and empirical evidence. *Journal of Econometrics*, 52:5–59, 1992.
- [6] T. Bollerslev and R. F. Engle. Modelling the persistence of conditional variances. *Econometric Reviews*, 5:1–50, 1986.
- [7] P. Clark. A subordinated stochastic process model with finite variance for speculative price. *Econometrica*, 41:135–156, 1973.
- [8] W. Dunsmuir. A central limit theorem for parameter estimation in stationary vector time series and its applications to models for a signal observed with noise. *Ann. Statist.*, 7:490–506, 1979.
- [9] R. F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of the united kingdom inflation. *Econometrica*, 50:987–1007, 1982.
- [10] R. F. Engle and G. Gonzales-Rivera. Semiparametric arch models. *J. Economics and Business Statist.*, 9:345–359, 1991.

- [11] R. F. Engle, D. M. Lilien, and R. P. Robins. Estimating time-varying risk premium in the term structure: the arch-m model. *Econometrica*, 55:391–407, 1987.
- [12] E. Ghysels, C. Gouriéroux, and J. Jasiak. Market time and asset price movements theory and estimation. *Discussion Paper CIRANO and CREST*, 1995.
- [13] E. Ghysels and J. Jasiak. Stochastic volatility and time deformation: an application to trading volume and leverage effects. *Paper presented at the 1994 Western Finance Association Meetings, Santa Fe*, 1994.
- [14] I. I. Gihman and A. V. Skorohod. *Stochastic Differential Equations*. Springer-Verlag, 1972.
- [15] J. Hamilton. *Time Series Analysis*. Princeton: Princeton University Press, 1994.
- [16] L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.
- [17] A. C. Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press, 1989.
- [18] A. C. Harvey and N. Shephard. Estimation and testing of stochastic variance models. *STICERD Econometrics Discussion Paper, LSE*, 1993a.
- [19] A. C. Harvey and N. Shephard. Estimation of an asymmetric stochastic volatility model for asset returns. *STICERD Econometrics Discussion Paper, LSE*, 1993b.
- [20] P. Y. Hong. The autocorrelation structure for the garch-m process. *Economic Letters*, 37:129–132, 1991.
- [21] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300, 1987.
- [22] S. W. Lee and B. E. Hansen. Asymptotic theory for the garch(1,1) quasi-maximum likelihood estimator. *Econometric Theory*, 10:29–52, 1994.
- [23] R. L. Lumsdaine. Consistency and asymptotic normality of the quasi-maximum likelihood estimator in igarch(1,1) and covariance stationary garch(1,1) models. *Econometrica*, 64:575–596, 1996.

- [24] B. Mandelbrot and H. Taylor. On the distribution of stock prices differences. *Operations Research*, 15:1057–1062, 1967.
- [25] B. J. Murteira. *Probabilidades e Estatística*. McGraw-Hill, 1990.
- [26] D. B. Nelson. Conditional heteroskedasticity in asset pricing: a new approach. *Econometrica*, 59:347–370, 1991.
- [27] A. Pagan. The econometrics of financial markets. *Lectures, The Australian National University and The University of Rochester*, 1993.
- [28] M. B. Priestley. *Spectral Analysis and Time Series*. London: Academic Press, 1981.
- [29] N. Shephard. Statistical aspects of arch and stochastic volatility. *Paper prepared for the 1994 Séminaire Européen de Statistique at Nuffield College, Oxford*, 1994.
- [30] J. H. Stock. Estimating continuous time processes subject to time deformation. *Journal of the American Statistical Association*, 83(401):77–84, 1988.
- [31] S. Taylor. *Modelling Financial Time Series*. John Wiley, 1986.
- [32] A. A. Weiss. Asymptotic theory for arch models: estimation and testing. *Econometric Theory*, 2:107–131, 1986.
- [33] P. Whittle. *Prediction and Regulation*. Oxford: Blackwell, 1983.
- [34] R. Wu. *Stochastic Differential Equations*. Pitman Publishing, 1985.

Índice

- AGARCH, modelo, 21
- ARCH, modelos
 - características gerais, 14
 - definição, 13
 - estimação, 15
 - nomenclatura, 5
 - previsão, 16
- ARCH-M, modelo, 23
- ARIMA, processo, 32
- ARMA, processo, 7
- autorregressivo, processo
 - tempo contínuo, 41
 - tempo discreto, 7
- Browniano, movimento, 40
- deformação temporal, 1, 37
 - função de, 43
 - modelos de V.E. com, 38
- diferença martingala, processo, 5
- EGARCH, modelo, 22
- erros de predição, 29
- estacionariedade
 - forte, 6
 - fraca, 6
- estimação
 - GMM, 27
 - QML, 28
 - QML irrestrita, 31
 - variáveis explanatórias, 35, 39
- fatos estilizados, 4
- GARCH, modelos
 - características gerais, 18
 - definição, 17
 - estimação, 19
 - nomenclatura, 5
 - previsão, 20
- IBOVESPA, 1
- IGARCH, modelo, 22
- Kalman, filtro de
 - definição, 28
 - estimador filtrado, 29
 - estimador suavizado, 29
- limite estocástico
 - continuidade, 40
 - convergência, 39
 - diferenciabilidade, 40
 - equação diferencial, 41
- médias móveis, processo de, 6
- observações corrigidas por heteroscedasticidade, 31
- quase-verossimilhança, função de, 16
- retorno, 3
- ruído branco, processo, 5
- tempo operacional, 38, 42
- UCARIMA, processo, 32

variabilidade condicional estocástica,
12

volatilidade, 1, 3

Volatilidade Estocástica, modelos de
com deformação temporal, 38

a tempo contínuo, 41

a tempo discreto, 38

definição, 23

estimação, 26

GMM, 27

quase-máxima verossimilhança,
28

propriedades, 24

Wiener, processo de, 40