

## 4

### Modelos de Duração Condicional

#### 4.1

##### Introdução

A principal dificuldade enfrentada pelos pesquisadores quando da tentativa de estudar o processo gerador das variáveis financeiras de interesse (preço, volume, intervalo de tempo – frequência, *spreads* de compra e venda, dentre outras), diz respeito à não-uniformidade no processo de chegada de uma informação nova. Em outras palavras, as transações envolvendo ativos financeiros são acordadas em instantes de tempo irregularmente espaçados.

Os modelos de duração condicional, como o próprio nome sugere, buscam identificar a dinâmica que envolve o processo de fechamento de operações, que, em linhas gerais, pode ser entendido como a convergência entre os preços de compra e venda para um mesmo patamar. Dessa forma, o procedimento adotado toma por base não a evolução temporal dos negócios e sim a seqüência de fechamentos, onde a informação relevante passa a ser o intervalo de tempo entre as transações efetivamente realizadas (variáveis aleatórias que seguem um determinado processo estocástico pontual), incorporando, assim, a idéia de frequência e/ou intensidade à modelagem.

Associadas aos intervalos de tempo entre as transações financeiras, podem-se ter, ainda, algumas variáveis estocásticas (volume, *spread* de compra e venda, preço,...) que, ao serem incorporadas ao modelo, adicionam o que se pode denominar de caráter seccional ou explicativo (realizações contemporâneas).

A formulação básica dos modelos de duração condicional estabelece uma relação de dependência entre a variável de interesse e suas realizações passadas, podendo ou não ser incorporadas outras variáveis que expressem a influência do ambiente econômico externo. Além da estrutura de dependência

temporal e cruzada, outra característica passível de ser adicionada, quando da existência de indícios neste sentido, diz respeito ao comportamento sazonal intradiário, donde se originam os grupamentos de negócios ao longo dos diferentes dias de negociação, verificados nas séries financeiras de fechamento de transações.

Cabe observar que existem basicamente dois “estados” distintos para a série de interesse: o primeiro considera cada novo negócio efetivamente transacionado, enquanto o segundo considera como “novos eventos” apenas aqueles advindos de uma alteração significativa (acima de determinado patamar) do nível de preços praticados.

## 4.2

### Histórico

Dentre as principais contribuições ao desenvolvimento dos modelos de duração condicional, tem-se:

- 1948 – Wold – modelos de intensidade condicional tomando por base o intervalo de tempo entre eventos (Modelos ARMA);
- 1955/72 – Cox – incluiu defasamentos como variável explicativa na determinação da intensidade;
- 1980 – Lewis – deu continuidade aos trabalhos de Wold (EARMA);
- 1979 – Lancaster – introduziu os modelos de duração na Econometria;
- 1980/1 – Heckman – influência do desemprego passado no nível atual (séries temporais de curta duração, referentes a um grande número de indivíduos);
- 1998 – Engle e Russell (Econometrica: *Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data*);
- 1999 – Bauwens e Veredas – processo estocástico para duração condicional apoiado sobre um fator estocástico latente, de modo a

- capturar o fluxo aleatório de informações no mercado financeiro (não-observável);
- 2000 – Bauwens e Giot - versão logarítmica dos modelos ACD, evitando a necessidade de adoção das restrições referentes à questão da não-negatividade da variável de interesse, presente na formulação inicial;
  - 2000 – Engle – incorpora a metodologia desenvolvida anteriormente no contexto de volatilidade;
  - 2001 – Zhang, Russell e Tsay (*Econometrica: A nonlinear autoregressive conditional duration model with applications to financial transaction data*) – extensão do modelo original, de modo a capturar a não-linearidade dos dados (TACD) + análise de quebras estruturais;
  - 2001 – Fernandes e Grammig (*CORE: A family of autoregressive conditional duration models*) – ACD + transformação de Box-Cox;
  - 2002 – Manganelli, Simone (*ECB Working Paper Series: Duration, Volume and Volatility Impact of Trades*) – VARM (ACD + ACV + GARCH) e análise da função impulso-resposta.

### 4.3

#### Característica dos dados

Com relação às características dos dados referentes às variáveis de interesse, devem ser destacadas:

- 1) Presença de *clusters* nas séries de duração;

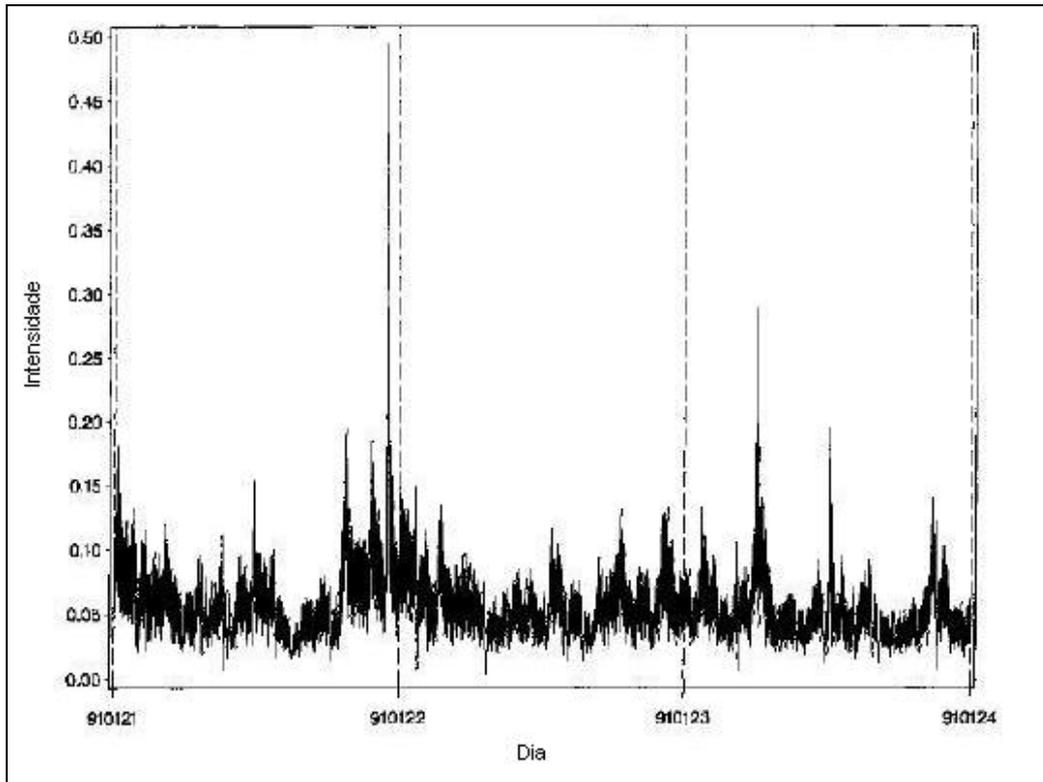


Figura 4.1 – Série de duração – IBM

- 2) Variável de interesse sempre positiva;
- 3) Dependência linear (mais tarde ficou comprovada a existência de não-linearidade nos dados);

Autocorrelação e autocorrelação parcial				
	Duração		Duração s/ sazonalidade	
	acf	pacf	acf	pacf
lag 1	.155	.155	.144	.144
lag 2	.132	.110	.122	.103
lag 3	.126	.094	.113	.085
lag 4	.125	.086	.110	.076
lag 5	.122	.075	.111	.072
lag 6	.108	.056	.098	.053
lag 7	.118	.065	.109	.063
lag 8	.123	.066	.113	.063
lag 9	.116	.054	.102	.047
lag 10	.104	.038	.095	.038
lag 11	.106	.041	.098	.042
lag 12	.106	.039	.095	.036
lag 13	.105	.037	.093	.033
lag 14	.099	.030	.089	.029
lag 15	.097	.028	.088	.028
	Ljung-Box(15) = 9466.43		Ljung-Box(15) = 7807.41	
	Amostra = 46091		Amostra = 46091	

Figura 4.2 – ACF e PACF (duração e duração com ajuste sazonal) – IBM

## 4) Resíduos não-gaussianos;

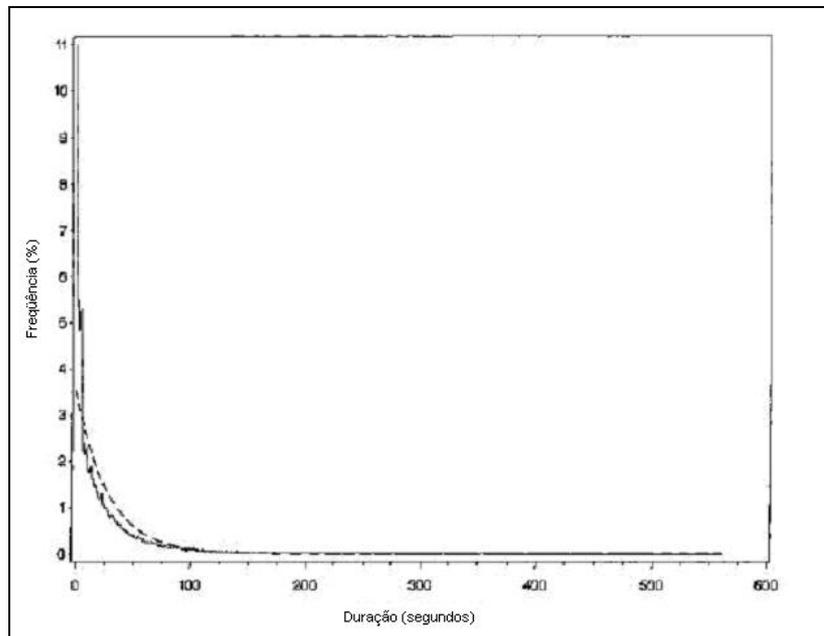


Figura 4.3 – Função de densidade dos resíduos (duração) – IBM

## 5) Existência de um padrão sazonal intradiário;

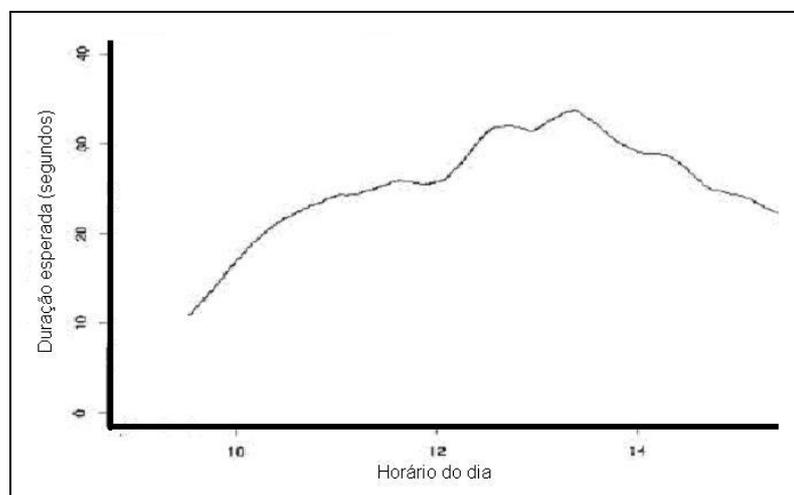


Figura 4.4 – Padrão sazonal (duração) – IBM

Cabe ser observado que, na maioria absoluta dos casos, as variáveis de interesse compartilham características semelhantes às aquelas verificadas quando do emprego dos modelos de volatilidade GARCH. Assim, a primeira proposta sugere o uso de modelos cuja formulação seja análoga aos daquela família.

#### 4.4

##### Processos de intensidade condicional

Esta seção apresenta de maneira sucinta alguns pontos relevantes dos modelos de processos pontuais, quando se consideram eventos correlacionados ao longo do tempo.

Considere um determinado processo estocástico que corresponda simplesmente à seqüência de instantes de tempo  $\{t_0, t_1, \dots, t_n, \dots\}$ , onde  $t_0 < t_1 < \dots < t_n < \dots$ . Para cada instante de tempo, tem-se a função de contagem  $N(t)$  – função degrau (contínua à esquerda com limites à direita), a qual representa o número de eventos associados ao instante de tempo  $t$ . Caso existam características associadas aos diferentes instantes de tempo, tais como preço ou volume, este processo será denominado “*marked point process*”.

Um determinado processo pontual definido entre  $[t_0, \infty)$  “evoluirá sem efeitos passados”, se para qualquer  $t > t_0$ , as realizações associadas aos instantes de tempo  $[t, \infty)$  não apresentar estrutura de dependência com relação às aquelas referentes aos instantes de tempo  $[t_0, t)$ . Em contrapartida, um processo de contagem será dito “condicionalmente ordenável” no instante  $t > t_0$ , se, para um intervalo de tempo suficientemente curto e de maneira condicional a um dado evento  $P$  definido pela realização do processo no intervalo  $[t_0, t)$ , a probabilidade de dois ou mais eventos ocorrerem for infinitesimal, se comparada à probabilidade de ocorrência de um único evento.

O foco do estudo recairá sobre processos pontuais que evoluem ao longo do tempo “considerando efeitos passados” e que se mostrem “condicionalmente ordenáveis”. Assim, a descrição completa deste tipo de processo deve ser formulada em termos da função de intensidade, condicional a toda informação

passada disponível, que deverá conter, pelo menos, os instantes de tempo e a contagem a estes associada. Dessa forma, o processo de intensidade condicional pode ser definido por:

$$\lambda(t|N(t), t_1, \dots, t_{N(t)}) = \lim_{\Delta t \rightarrow 0} \left( \frac{P(N(t + \Delta t) > N(t) | N(t), t_1, \dots, t_{N(t)})}{\Delta t} \right) \quad (4.4.1)$$

Em muitas aplicações, estes processos serão (equivalentemente) denominados de “*hazard function*”, em particular nos casos onde existam vários processos individuais ao invés de um único processo pontual. Conforme apresentado por Lancaster (1990) ou Snyder e Miller (1991), por exemplo, a intensidade condicional das durações ou “tempo de espera” entre eventos corresponde a processos estocásticos condicionalmente ordenáveis. Tomando  $p_i$  como sendo uma família de funções de densidade de probabilidade condicional para o instante de tempo  $t_i$ , o valor da log-verossimilhança, em termos da intensidade condicional, poderá ser expresso por:

$$L = \sum_{i=1}^{N(T)} \log p_i(t_i | t_0, \dots, t_{i-1}) \quad (4.4.2)$$

$$L = \sum_{i=1}^{N(T)} \log \lambda(t_i | i-1, t_0, \dots, t_{i-1}) - \int_{t_0}^T \lambda(t_i | N(u), t_0, \dots, t_{N(u)}) du \quad (4.4.3)$$

A equação (4.4.1) corresponde à formulação geral da função de intensidade de um processo estocástico, no qual a evolução observada no passado influencia a estrutura de probabilidade dos eventos futuros – originalmente proposto por Hawkes (1972) e Rubin (1972). O grau de sucesso oriundo do uso deste tipo de processo depende da parametrização da intensidade condicional.

O processo mais simples ligado a esta classe é o processo Poisson, para o qual o parâmetro  $\lambda$  se mostra constante. Em contrapartida, numa formulação

mais flexível, pode-se propor o uso do chamado processo Poisson não-homogêneo, para o qual a intensidade varia exclusivamente com  $t$ , sendo a “taxa de chegada” (velocidade dos eventos) uma função determinística do tempo. Entretanto, em nenhum caso os eventos passados exercem qualquer tipo de influência na “taxa de chegada”, ou seja, estes processos evoluem “sem considerar efeitos passados”.

Quando a intensidade depende do número de eventos, mas não do instante de tempo no qual tais eventos ocorrem, o processo estocástico é chamado de “*pure birth process*”. Rubin (1972) introduziu os chamados “processos de memória finita” (*m-memory process*), cuja definição advém do fato de que apenas os “ $m$ ” instantes de tempo (correspondentes a realizações de determinado evento) mais recentes se mostram presentes na função de intensidade condicional.

No que diz respeito aos processos de memória “longa”, existem diversas formas de parametrizar a função de intensidade condicional. Dentre estas, são apresentadas duas classes de modelos de processos pontuais. São elas:

- 1) Classe de modelos formulada em termos da evolução temporal dos eventos (*calendar time models*): a representação linear desta classe de modelos pode ser expressa da seguinte forma:

$$\lambda(t|N(t), t_0, \dots, t_{N(t)}) = w + \sum_{i=1}^{N(t)} \pi(t-t_i) \quad (4.4.4)$$

onde cada instante de tempo passado contribui  $\pi(t-t_i)$  para a intensidade no instante  $t$ ;  $\pi$  é chamado medida do grau de infecção (motivado pelo estudo de epidemias). Este tipo de modelo se mostra inapropriado quando do estudo de dados oriundos de transações financeiras, uma vez que o efeito de um determinado evento no passado se mostra independente da história interveniente, ou seja, podem ter sido verificados 0 ou 100 eventos entre este e o atual.

- 2) Classe de modelos formulada em termos da seqüência de eventos, considerando como ponto principal o intervalo de tempo verificado: esta classe de modelos pode ser parametrizada da seguinte forma:

$$\lambda(t|N(t), t_0, \dots, t_{N(t)}) = w + \sum_{i=1}^{N(t)} \pi(t_{N(t)+1-i} - t_{N(t)-i}) \quad (4.4.5)$$

Considerando a segunda classe de modelos apresentada, pode-se perceber que o impacto da duração entre eventos sucessivos dependerá do número de eventos intervenientes. Este tipo de modelo foi introduzido por Wold (1948) e estendido, mais tarde, por Cox (1955). Wold propôs um modelo para intervalos correlacionados mediante o uso de uma estrutura auto-regressiva, similar aos modelos ARMA tradicionais. Tais modelos foram reformulados subsequentelemente por Gaver e Lewis (1980), Lawrence e Lewis (1980) e Jacobs e Lewis (1977) como modelo exponencial autoregressivo com médias móveis – EARMA(p, q). Nesta formulação, assume-se que a duração condicional seja exponencialmente distribuída, sendo que a média do processo segue um processo ARMA. A formulação de um processo aditivo para o ruído com estas propriedades se mostra bastante complexa e os procedimentos de estimação (máxima verossimilhança) virtualmente intratáveis.

Em alguns casos, a intensidade condicional pode ser derivada a partir de pressupostos mais simples. Cox (1955) propôs os chamados “modelos duplamente estocásticos”, onde o pressuposto fundamental dizia respeito à existência de um processo independente, o qual regia o comportamento da taxa de chegada da informação. Supondo  $M(t)$  um determinado processo de contagem denominado “informação” referente a aplicações financeiras; então, a intensidade será condicional tanto a  $M(t)$  quanto ao próprio instante de tempo  $t$ .

Ainda no que diz respeito aos desenvolvimentos propostos por Cox (1955, 1972), tem-se a primeira proposta de análise de dados de duração (generalizado mais tarde – *proportional hazards framework*), onde a intensidade condicional pode ser descrita como:

$$\lambda(t|z_{N(t)}, \dots, z_1) = \lambda(t) \exp\{\beta' z_{N(t)}\} \quad (4.4.6)$$

onde,  $z_i$  corresponde a um vetor de variáveis explicativas associado ao instante de tempo  $i$  (tempo de chegada da informação). Uma das sugestões mencionadas nos trabalhos desenvolvidos propunha a inclusão no vetor de variáveis explicativas de dados passados de duração.

## 4.5

### Principais modelos

Esta seção tem por objetivo apresentar de forma sucinta os principais trabalhos desenvolvidos até o momento, enaltecendo suas características e resultados principais.

#### 4.5.1

##### Modelo ACD

Introduzidos por Engle e Russell (1998), o modelo ACD é mais convenientemente especificado em termos da densidade condicional das durações. Assim, seja  $x_i = t_i - t_{i-1}$  o horizonte de tempo entre duas transações financeiras (duração). Então, a densidade de probabilidade de  $x_i$  condicional às observações passadas pode ser obtida de maneira direta. Tomando  $\psi_i$  como sendo a esperança da  $i$ -ésima duração, tem-se que:

$$E(x_i | x_{i-1}, \dots, x_1) = \psi_i(x_{i-1}, \dots, x_1; \theta) \equiv \psi_i \quad (4.5.1)$$

Assim, a classe de modelos ACD consistirá na parametrização (4.5.1) e no pressuposto de que a variável de interesse irá evoluir segundo o seguinte processo estocástico:

$$x_i = \psi_i \cdot \varepsilon_i \quad (4.5.2)$$

onde,  $\{\varepsilon_i\} \sim \text{iid } p(\varepsilon; \phi)$ , e  $\phi$  e  $\theta$  são determinísticos.

Assim, pode-se propor inúmeras especificações diferentes para os modelos de duração condicional, mediante a adoção de pressupostos distintos quanto à duração esperada ( $\psi$ ) e/ou à distribuição do ruído ( $\varepsilon$ ).

- **ACD (*m*-memory)**: apenas as “*m*” durações mais recentes exercem influência sobre a duração condicional.

$$\psi_i = \omega + \sum_{j=1}^m \alpha_j \cdot x_{i-j} \quad (4.5.3)$$

- **ACD (memória “ilimitada”)**: também chamado de ACD(*m*, *q*), onde “*m*” e “*q*” são os defasamentos; nesta formulação não existe um limite finito para a dependência entre eventos subseqüentes.

$$\psi_i = \omega + \sum_{j=1}^m \alpha_j \cdot x_{i-j} + \sum_{j=1}^q \beta_j \cdot \psi_{i-j} \quad (4.5.4)$$

Cabe ser ressaltado o fato de que os parâmetros que definem o processo estocástico possuem espaço de definição de valores restrito, em função da própria característica dos dados de interesse e de modo a tornar o processo não-explosivo. Assim, tem-se:

$$\left\{ \begin{array}{ll} k > 0 & \\ \alpha_j \geq 0 & \rightarrow j = 1, 2, \dots, m \\ \beta_j \geq 0 & \rightarrow j = 1, 2, \dots, q \\ \sum_{j=1}^m (\alpha_j + \beta_j) < 1 & \end{array} \right.$$

Com relação ao processo estocástico definido pela equação (4.5.4), pode-se desenvolver facilmente as expressões para a média (condicional e incondicional) como sendo:

$$E(x_i | I_{i-1}) = \psi_i \quad (4.5.5)$$

$$E(x_i) = \mu = \frac{w}{\{1 - \sum(\alpha_j + \beta_j)\}} \quad (4.5.6)$$

## 4.5.2

### Modelo EACD (*Exponential ACD*)

#### 4.5.2.1

##### Características principais

- Distribuição condicional da duração: exponencial;
- Hiperparâmetro  $\lambda$  – estimado juntamente com os parâmetros do modelo ( $\alpha$ 's e  $\beta$ 's), por máxima verossimilhança condicional;
- Função distribuição de probabilidades – exponencial:

$$x \Rightarrow f(x) = \begin{cases} \lambda \cdot \exp[-\lambda \cdot x] \rightarrow x \geq 0 \\ 0 \rightarrow c.c \end{cases}$$

$$\lambda > 0$$

Seja o modelo EACD(1, 1): então, temos que a variância condicional e incondicional podem ser expressas por:

$$\sigma_{I_{i-1}}^2 = \psi_i \quad (4.5.7)$$

$$\sigma^2 = \mu^2 \left( \frac{1 - \beta^2 - 2\alpha\beta}{1 - \beta^2 - 2\alpha\beta - 2\alpha^2} \right) \quad (4.5.8)$$

Pode-se perceber que sempre que  $\alpha > 0$ , o desvio padrão incondicional se mostrará superior à média, caracterizando excesso de dispersão, conforme pode ser comprovado nos resultados apresentados por Engle e Russell (1998).

De forma análoga à desenvolvida no capítulo 2 para os modelos de volatilidade, poder-se-ia reescrever o modelo apresentado em (4.5.4) utilizando-se a estrutura dos modelos ARMA(max(m,q), q) com inovações não-gaussianas. Assim, ter-se-ia:

$$x_i = \omega + \sum_{j=1}^{\max(m,q)} (\alpha_j + \beta_j) \cdot x_{i-j} - \sum_{j=1}^q \beta_j \cdot \eta_{i-j} + \eta_i \quad (4.5.9)$$

onde,  $\eta_i = x_i - \psi_i$ ; seqüência diferenças Martingale (por construção).

A partir da equação (4.5.9), podem-se determinar diretamente as previsões para a variável de interesse (duração). Caso as raízes do polinômio sejam menores do que um (em módulo), então o processo estocástico da duração apresentará reversão à média e o impacto de um dado evento sobre a expectativa com relação às realizações futuras decairá exponencialmente – a velocidade de decaimento dependerá exclusivamente do grau de persistência (raízes próximas do valor unitário).

### 4.5.3

#### Modelo WACD (Weibull ACD)

##### 4.5.3.1 – Distribuição condicional da duração: Weibull

$$x \Rightarrow f(x) = \begin{cases} \frac{\beta}{\delta} \cdot \left(\frac{x-\gamma}{\delta}\right)^{\beta-1} \cdot \exp\left[-\left(\frac{x-\gamma}{\delta}\right)^{\beta}\right] \rightarrow x \geq \gamma \\ 0 \rightarrow c.c \end{cases} \quad (4.5.10)$$

$-\infty < \gamma < \infty \rightarrow$  parâmetro de localização

$\delta \geq 0 \rightarrow$  parâmetro de escala

$\beta > 0 \rightarrow$  parâmetro de forma

$$E(x) = \gamma + \delta \cdot \Gamma\left(1 + \frac{1}{\beta}\right) = \psi \quad (4.5.11)$$

$$\therefore \gamma = 0$$

$$Var(x) = \delta^2 \cdot \left\{ \Gamma\left(1 + \frac{2}{\beta}\right) - \left[ \Gamma\left(1 + \frac{1}{\beta}\right) \right]^2 \right\} = \psi^2 \quad (4.5.12)$$

$$\therefore \delta = \frac{\psi}{\Gamma\left(1 + \frac{1}{\beta}\right)}$$

$$\varepsilon \Rightarrow f(x) = \begin{cases} \frac{\beta}{\delta} \cdot \left(\frac{x}{\delta}\right)^{\beta-1} \cdot \exp\left[-\left(\frac{x}{\delta}\right)^\beta\right] \rightarrow x \geq \gamma \\ 0 \rightarrow c.c. \end{cases} \quad (4.5.13)$$

### 4.5.3.2

#### Verossimilhança condicional

$$L = \prod_{i=1}^N \left[ \frac{\beta}{\delta} \cdot \left(\frac{x_i}{\delta}\right)^{\beta-1} \cdot \exp\left[-\left(\frac{x_i}{\delta}\right)^\beta\right] \right]$$

$$\ln L = \sum_{i=1}^N \ln \left[ \frac{\beta}{\delta} \cdot \left(\frac{x_i}{\delta}\right)^{\beta-1} \cdot \exp\left[-\left(\frac{x_i}{\delta}\right)^\beta\right] \right]$$

$$\ln L = \sum_{i=1}^N \left\{ \ln \left[ \frac{\Gamma\left(1 + \frac{1}{\beta}\right) \cdot \beta}{\psi_i} \right] + (\beta - 1) \cdot \ln \left[ \frac{\Gamma\left(1 + \frac{1}{\beta}\right) \cdot x_i}{\psi_i} \right] - \left[ \frac{\Gamma\left(1 + \frac{1}{\beta}\right) \cdot x_i}{\psi_i} \right]^\beta \right\}$$

$$\ln L = \sum_{i=1}^N \left\{ \ln \left[ \frac{\beta}{x_i} \right] + \beta \cdot \ln \left[ \frac{\Gamma\left(1 + \frac{1}{\beta}\right) \cdot x_i}{\psi_i} \right] - \left[ \frac{\Gamma\left(1 + \frac{1}{\beta}\right) \cdot x_i}{\psi_i} \right]^\beta \right\} \quad (4.5.14)$$

### 4.5.3.3

#### Observações importantes

- Os resultados do modelo exponencial (EACD) são análogos aos modelos com a adoção da distribuição de Weibull –  $\beta = 1$ ;
- Com relação à equivalência entre a distribuição exponencial e Weibull –  $\delta = 1/\lambda$ , ou seja, neste caso,  $\psi = 1/\lambda$ .

Cabe lembrar ainda que a especificação da duração condicional pode ser facilmente generalizada, de modo a incluir, por exemplo, variáveis explicativas. Desse modo ter-se-ia:

$$\psi_i = \psi(x_{i-1}, x_{i-2}, \dots, x_{i-m}, \psi_{i-1}, \psi_{i-2}, \dots, \psi_{i-q}, z_i, z_{i-1}, \dots, z_{i-k}, \theta) \quad (4.5.15)$$

Outro ponto de extrema importância diz respeito à presença de sazonalidade (efeito temporal) intradiário (período sazonal de um dia). Se a duração esperada varia ao longo do tempo dentro do período sazonal, então, o instante de tempo no qual a duração tem início pode ser tratado como uma variável independente. Assim, a duração esperada pode ser entendida como a composição de uma componente estocástica e uma determinística.

Supondo que o efeito determinístico possa ser formulado a partir de uma determinada função multiplicativa, poder-se-ia propor uma correção dos dados de modo a obter uma nova série “sazonalmente ajustada”, a qual seria obtida a partir da seguinte expressão:

$$\tilde{x}_i = x_i / \phi(t_{i-1}; \theta_\phi) \quad (4.5.16)$$

e a duração esperada seria então:

$$E_{i-1}(x_i) = \phi(t_{i-1}; \theta_\phi) \cdot \psi_i(\tilde{x}_{i-1}, \dots, \tilde{x}_1; \theta_\psi) \quad (4.5.17)$$

Os parâmetros inerentes às duas componentes do processo estocástico que governa o comportamento da duração poderiam ser estimados conjuntamente, a partir da maximização da função de verossimilhança, conforme apresentado anteriormente. Dessa forma,  $\psi_i$  pode ser interpretado como sendo a parcela responsável pela duração esperada ser superior ou inferior ao nível médio diário.

Um procedimento alternativo sugerido por Engle e Russell (1998) propõe que a série “sazonalmente ajustada” seja obtida (antes do processo de estimação dos parâmetros – máxima verossimilhança) a partir da regressão de  $x$  (dados de duração) sobre uma função *spline* (responsável pela componente determinística) e o posterior cálculo da razão entre a variável de interesse e o valor retornado pela função para o instante de tempo no qual o evento ocorre. Este procedimento conduz a resultados satisfatórios, caso o número de dados seja suficientemente grande (regressão corresponde à expectativa condicional), apesar dos distúrbios apresentarem autocorrelação e indícios de heterocedasticidade.

$$E(x_i | t_{i-1}) = \phi(t_{i-1}; \theta_\phi) + \xi_i \quad (4.5.18)$$

#### 4.5.4

##### Modelo TACD

Introduzidos por Zhang, Russell e Tsay (2001), os modelos TACD tentam preencher algumas das lacunas deixadas pela abordagem proposta por Engle e Russell (1998), tais como: não-linearidade (média condicional) e excesso de dispersão. Os modelos TACD buscam capturar de maneira mais adequada a dinâmica dos dados de duração, mediante adoção de uma formulação na qual a duração condicional esperada seja descrita como uma função não-linear (uso de limitadores – definição de sub-regimes) de variáveis defasadas, quer sejam relacionadas à própria duração, quer estejam ligadas a outros eventos pertinentes. Com relação à questão do excesso de dispersão verificado nos

resultados obtidos por Engle e Russell, o presente modelo adota a distribuição Gama Generalizada dentro da estrutura de múltiplos regimes, permitindo que os parâmetros de forma assumam valores distintos de acordo com o nível da variável de interesse (sub-regimes).

Além do que já fora mencionado, Zhang, Russell e Tsay identificaram uma série de quebras estruturais nos dados referentes ao estudo de caso realizado, o que aumenta a importância de algoritmos de identificação de tais componentes, devendo as mesmas estar presentes na formulação final, pois estas componentes existem na grande maioria das séries financeiras.

Os modelos com limitadores (*thresholds*) podem ser entendidos como uma forma não-linear simples de aproximar sistemas estocásticos complexos, mediante a definição de sub-sistemas a partir do uso de variáveis de limite. Conforme disposto na literatura, os modelos autoregressivos com limitadores (TAR) são capazes de capturar dinâmicas específicas, tais como comportamento periódico assimétrico e saltos (*jumps*).

Seja  $t_i$  o instante de tempo referente ao fechamento da  $i$ -ésima transação financeira, onde  $0 = t_0 < t_1 < \dots < t_N$ , e  $x_i = t_i - t_{i-1}$  a duração desta. Conforme visto, o pressuposto básico dos modelos ACD (4.5.2) estabelece que:

$$\frac{x_i}{\psi_i} \equiv \varepsilon_i \sim iid \quad p(\phi) \tag{4.5.19}$$

onde  $\psi_i = E[x_i | x_{i-1}, \dots, x_1; \theta]$  corresponde à média condicional da duração com vetor de parâmetros  $\theta$ ,  $p(\phi)$  é uma distribuição de probabilidade com média um e vetor de parâmetros  $\phi$ , e cujo espaço de definição pode ser descrito pelo conjunto  $(0, \text{Inf})$ . Os modelos ACD assumem que o termo irregular se mostra multiplicativo e que toda informação passada influencia a duração corrente a partir da média condicional  $\psi_i$ .

O modelo TACD corresponde a uma simples mas poderosa generalização do modelo ACD, no qual a adoção de sub-regimes permite a captura de diferentes comportamentos da variável de interesse, de acordo com o

nível da realização. Desta forma, pode-se observar diferentes graus de persistência, formulação da média condicional e mesmo da função de distribuição de probabilidade do ruído. Estas diferenças combinadas podem reproduzir a dinâmica não-linear apresentada pela série de interesse.

Assim, tomando  $\{x_i\}$  como sendo um processo estocástico positivo e  $\psi_i = E[x_i | I_{t-1}; \theta]$  a média condicional de  $x_i$ , define-se  $R_j = [r_{j-1}, r_j)$ ,  $j = 1, 2, \dots, J$ , para  $J$  igual a um determinado elemento pertencente ao conjunto de números inteiros positivos, onde  $-\text{Inf} = r_0 < r_1 < \dots < r_J = +\text{Inf}$  correspondem aos valores dos limitadores ( $Z_i$ 's). Dessa forma, o processo  $\{x_i\}$  será descrito por um modelo TACD J-regimes (variável  $Z_{i-d} \in R_j$ ), se:

$$\frac{x_i}{\psi_i} \equiv \varepsilon_i \sim \text{independente} \quad p(\phi^{(j)}) \quad (4.5.20)$$

$$\psi_i = \omega^{(j)} + \sum_{l=1}^m \alpha_l^{(j)} \cdot x_{i-l} + \sum_{l=1}^q \beta_l^{(j)} \cdot \psi_{i-l} \quad (4.5.21)$$

onde o parâmetro de defasagem “d” é um inteiro positivo,  $p(\phi^{(j)})$  é uma distribuição de probabilidade com média um e vetor de parâmetros  $\phi^{(j)}$  (assumindo diferentes valores para cada sub-regime) e cujo espaço de definição pode ser descrito pelo conjunto  $(0, \text{Inf})$ . A forma geral da variável de limite  $\{Z_i\}$  é dada por  $Z_i = h(x_i, \dots, x_1; y_i, \dots, y_1)$ , onde  $\{y_i\}$  corresponde ao vetor de variáveis explicativas (estado da economia – caracterizando o mercado financeiro).

Analogamente aos procedimentos já realizados, podem-se reescrever os modelos TACD sob a forma dos modelos TARMA. Assim, para o modelo TACD (J-regimes, m, q) ter-se-ia:

$$x_i = \omega^{(j)} + \sum_{l=1}^{\max(m, q)} (\alpha_l^{(j)} + \beta_l^{(j)}) \cdot x_{i-l} - \sum_{l=1}^q \beta_l^{(j)} \cdot \eta_{i-l} + \eta_i \quad (4.5.22)$$

A equação (4.5.22) corresponde ao modelo TARMA(J-regimes,  $\max(m, q)$ , q), com inovações condicionalmente heterocedásticas.

A seguir são listados algumas definições e resultados importantes:

(1) Uma seqüência positiva  $\{x_i\}$  seguirá um processo TACD (1,1) se

$$\begin{cases} x_i = \psi_i \varepsilon_i^{(j)} \\ \psi_i = w^{(j)} + \alpha^{(j)} x_{i-1} + \beta^{(j)} \psi_{i-1} \end{cases} \quad \text{se } x_i \in R_j \quad (4.5.23)$$

Onde,  $R_j = [r_{j-1}, r_j)$ ,  $j = 1, 2, \dots, J$ , sendo o inteiro positivo  $J$  o número de regimes e  $-\text{Inf} = r_0 < r_1 < \dots < r_J = +\text{Inf}$  correspondem aos valores dos limitadores.  $w^{(j)} > 0$ ,  $\alpha^{(j)} \geq 0$ ,  $\beta^{(j)} \geq 0$  para  $j = 1, 2, \dots, J$ . Para um dado regime  $j$ ,  $\{\varepsilon_i^{(j)}\}$  corresponde a uma seqüência independente e identicamente distribuída com densidade  $f^{(j)}(\cdot)$  positiva  $(0, \text{Inf})$  e  $E[\varepsilon_i^{(j)}] = 1$ .  $\{\varepsilon_i^{(j)}\}$  e  $\{\varepsilon_i^{(k)}\}$  se mostram independentes para  $j \neq k$ .

(2) Um processo TACD (1,1)  $\{x_i\}$  será geometricamente ergódico se

$$\text{i. } \alpha^{(j)} > 0, \beta^{(j)} > 0 \rightarrow j = 1, \dots, J \quad (4.5.24)$$

$$\text{ii. } \alpha^{(j)} + \beta^{(j)} < 1 \quad (4.5.25)$$

A condição suficiente (i) do teorema é necessária para provar a irreduzibilidade do processo. Já a condição suficiente (ii) é necessária para que o sistema não tenha um comportamento explosivo, sendo intuitiva sua interpretação, ou seja, uma vez que o processo tem um determinado limite inferior (zero – variável de interesse positiva), o único modo de este apresentar um comportamento explosivo seria através do  $J$ -ésimo sub-regime. A condição de que o coeficiente autoregressivo seja menor do que um (em módulo, porém ambos os termos se mostram positivos – condição (i)) irá “regular” o processo, uma vez que o  $J$ -ésimo regime tenha sido atingido. É importante notar que, assumindo que a condição (ii) seja satisfeita, o processo será estável, mesmo que apresente comportamento explosivo em todos os sub-regimes entre 1 e  $J-1$ .

### 4.5.5

#### Modelo AACD (*Augmented ACD – Transformação Box-Cox*)

Introduzido por Fernandes e Gramming (2002), o modelo abrange a maioria das formulações dispostas na literatura referente aos processos de duração; tem como principal motivação os resultados dos testes empíricos desenvolvidos por Engle e Russell (1998), nos quais os modelos ACD “padrão” superestimam o valor das durações, principalmente após realizações com valores extremos (durações muito longas ou muito curtas). Neste modelo são exploradas as similaridades existentes entre os processos ACD e GARCH, seguindo a mesma linha de pesquisa proposta por Hentschel (1995), a qual apresentava uma família de modelos GARCH com comportamento assimétrico. A grande inovação remete ao uso da transformação de Box e Cox (1964), com parâmetro de forma  $\lambda \geq 0$  para o processo de duração condicional e a introdução de uma função  $g(\cdot)$ , responsável pelo caráter assimétrico das respostas aos choques observados.

Seja o modelo ACD “padrão” (4.5.2), conforme estabelecido por Engle e Russell (1998):

$$x_i = \psi_i \cdot \varepsilon_i \quad (4.5.26)$$

$$\varepsilon_i \sim iid \quad p(\phi)$$

$$\psi_i = E(x_i | I_{i-1}) \quad (4.5.27)$$

Assim, mediante a aplicação da transformação de Box e Cox (parâmetro  $\lambda \geq 0$ ) sobre a duração condicional do processo  $\psi_i$ , tem-se – para ACD(1,1):

$$\frac{\psi_i^\lambda - 1}{\lambda} = w_* + \alpha_* \psi_{i-1}^\lambda \left[ |\varepsilon_{i-1} - b| + c |\varepsilon_{i-1} + b| \right]^\nu + \beta \frac{\psi_{i-1}^\lambda - 1}{\lambda} \quad (4.5.28)$$

onde o parâmetro  $\lambda$  determina a forma da transformação de Box e Cox ( $\lambda < 1$  – côncava e  $\lambda \geq 1$  – convexa).

Desse modo, o modelo autoregressivo aumentado de duração condicional (AACD) será:

$$\psi_i^\lambda = w + \alpha \psi_{i-1}^\lambda \left[ |\varepsilon_{i-1} - b| + c |\varepsilon_{i-1} - b| \right]^\nu + \beta \psi_{i-1}^\lambda \quad (4.5.29)$$

onde,  $w = \lambda w^* - \beta + 1$  e  $\alpha = \lambda \alpha^*$ .

O modelo AACD estabelece uma forma funcional flexível, a qual permite que o processo de duração condicional  $\{\psi_i\}$  apresente respostas diferentes de acordo com o tipo de choque (magnitude). Dentro deste contexto, a função resposta  $g(\varepsilon_i) = [|\varepsilon_i - b| + c |\varepsilon_i - b|]^\nu$  se mostra responsável pela captura do efeito assimétrico observado, através dos parâmetros de locação e rotação (“b” e “c”, respectivamente).

Devido ao fato da variável de interesse, no caso a duração, se mostrar não-negativa, o parâmetro de locação “b” é de fundamental importância na determinação das respostas assimétricas originadas a partir da função  $g(\cdot)$ . Já o parâmetro “c” estará relacionado ao sentido da rotação (sentido horário –  $c < 0$ ).

Ainda no que se refere à função resposta, observa-se que, de fato, o parâmetro de localização afetará principalmente os choques de baixa magnitude, enquanto que o de rotação estará relacionado (dominante) aos de magnitude elevada.

Com relação ao parâmetro de forma  $\nu$ , este desempenhará papel semelhante ao parâmetro  $\lambda$ , estando ligado à forma da função resposta ( $\nu < 1$  – côncava e  $\nu \geq 1$  – convexa).

Ressalte-se o fato de que se pode obter diferentes topologias para os modelos ACD, a partir da simples adoção de restrições nos valores dos parâmetros da equação (4.5.29). A figura 4.5 disposta a seguir apresenta algumas destas formulações.

**Augmented ACD**

$$\psi_i^\lambda = w + \alpha \psi_{i-1}^\lambda \left[ |e_{i-1} - b| + c(e_{i-1} - b) \right]^\nu + \beta \psi_{i-1}^\lambda$$

**Asymmetric Power ACD ( $\lambda = \nu$ )**

$$\psi_i^\lambda = w + \alpha \psi_{i-1}^\lambda \left[ |e_{i-1} - b| + c(e_{i-1} - b) \right]^\lambda + \beta \psi_{i-1}^\lambda$$

**Asymmetric Logarithmic ACD ( $\lambda \rightarrow 0$  and  $\nu = 1$ )**

$$\log \psi_i = w + \alpha \left[ |e_{i-1} - b| + c(e_{i-1} - b) \right] + \beta \log \psi_{i-1}$$

**Asymmetric ACD ( $\lambda = \nu = 1$ )**

$$\psi_i = w + \alpha \psi_{i-1} \left[ |e_{i-1} - b| + c(e_{i-1} - b) \right] + \beta \psi_{i-1}$$

**Power ACD ( $\lambda = \nu$  and  $b = c = 0$ )**

$$\psi_i^\lambda = w + \alpha x_{i-1}^\lambda + \beta \psi_{i-1}^\lambda$$

**Box-Cox ACD ( $\lambda \rightarrow 0$  and  $b = c = 0$ )**

$$\log \psi_i^\lambda = w + \alpha e_{i-1}^\lambda + \beta \psi_{i-1}$$

**Logarithmic ACD Type I ( $\lambda, \nu \rightarrow 0$  and  $b = c = 0$ )**

$$\log \psi_i = w + \alpha \log x_{i-1} + \beta \log \psi_{i-1}$$

**Logarithmic ACD Type II ( $\lambda \rightarrow 0, \nu = 1$  and  $b = c = 0$ )**

$$\log \psi_i = w + \alpha e_{i-1} + \beta \log \psi_{i-1}$$

**Linear ACD ( $\lambda = \nu = 1$  and  $b = c = 0$ )**

$$\psi_i = w + \alpha x_{i-1} + \beta \psi_{i-1}$$

Figura 4.5 – Diferentes topologias AACD

Com relação à distribuição dos erros ou distúrbios aleatórios, os autores propuseram a adoção da função de distribuição de probabilidades Burr. Dessa forma, os procedimentos de estimação dos valores dos parâmetros e hiperparâmetros do modelo tomarão por base a seguinte função:

$$f_B(\varepsilon_i; \theta_B) = \frac{k \cdot \mu_{B,1}^k \cdot \varepsilon_i^{k-1}}{(1 + \gamma \cdot \mu_{B,1}^k \cdot \varepsilon_i^k)^{1+1/\gamma}} \quad (4.5.30)$$

onde,  $k > \lambda > 0$  e

$$\mu_{B,m} = \frac{\Gamma(1 + m/k) \cdot \Gamma(1/\gamma - m/k)}{\gamma^{1+m/k} \cdot \Gamma(1 + 1/\gamma)} \quad (4.5.31)$$

corresponde ao  $m$ -ésimo momento da distribuição, o qual existe para  $m < k/\lambda$ . A família de distribuição Burr abrange tanto as distribuições Weibull ( $\lambda \rightarrow 0$ ) e exponencial ( $\lambda \rightarrow 0$  e  $k = 1$ ) quanto à log-logística ( $\lambda \rightarrow 1$ ).

#### 4.5.6

##### Modelo conjunto para duração, volume e volatilidade condicionais

Desenvolvido por Manganeli (2002), o trabalho propõe a utilização das idéias apresentadas por Hasbrouck (1991) num contexto de volatilidade, tornando possível a elaboração de um sistema completo, onde retorno e volatilidade interagem diretamente com as informações referentes à duração e volume das transações financeiras realizadas. A modelagem final envolve duas etapas distintas. Na primeira, as idéias apresentadas por Engle e Russell (1998) são estendidas de modo a tornar possível a adoção de uma formulação explícita para a série de volumes – analogamente aos procedimentos do modelo ACD, a formulação do processo estocástico que governa o comportamento do volume faz uso de uma especificação autoregressiva que multiplica um termo aleatório (ruído), cujo suporte da distribuição compreende os Reais positivos, devido à própria característica da variável de interesse (as realizações são sempre positivas).

Na fase seguinte, duração, volume e retornos são modelados simultaneamente, a partir de um tipo especial de modelo autoregressivo vetorial, onde o valor esperado de cada componente de interesse se mostra definido

como função do valor corrente e defasado de todas aquelas incluídas no sistema. Desta forma, o sistema final passa a incorporar a causalidade existente entre as variáveis de interesse.

Além das questões apresentadas até o presente momento, o trabalho desenvolvido por Manganelli estabelece a função impulso-resposta do sistema (relacionado ao equilíbrio de longo prazo) proposto, fornecendo subsídios para o correto entendimento de pontos relacionados à “resistência” do mercado.

#### 4.5.6.1

##### Formulação de Manganelli

A idéia do modelo econométrico desenvolvido e proposto por Manganelli é apresentar uma extensão do modelo ACD, de modo a tornar possível a análise conjunta de diferentes variáveis de interesse (duração, volume e volatilidade) e o estudo da dinâmica (interação) existente entre estas.

Sendo  $x_i$  e  $z_i$ , respectivamente, a duração e o conjunto de variáveis explicativas (preço, volume e *spread* de compra e venda) correspondente à  $i$ -ésima transação envolvendo ativos financeiros. Então, pode-se dizer que os eventos observados (referentes a  $x_i$  e  $z_i$ ) correspondem à realização de um determinado processo estocástico com a seguinte forma:

$$(x_i, z_i) \sim f(x_i, z_i | \Omega_{i-1}; \theta) \quad (4.5.32)$$

onde,  $\Omega_{i-1}$  corresponde a toda informação disponível até o evento anterior ao  $i$ -ésimo observado e  $\theta$  a um vetor de parâmetros fixos e desconhecidos.

Assim, conforme apresentado por Engle (2000), a densidade conjunta  $(x_i, z_i)$  pode ser reescrita como o produto entre a densidade marginal de  $x_i$  e a condicional de  $x_i$  dado  $z_i$ . Analogamente, sendo  $z$  o vetor formado pelas variáveis correspondentes a volume ( $v$ ) e retornos ( $y$ ), ter-se-ia:

$$(x_i, v_i, y_i) \sim f(x_i, v_i, y_i | \Omega_{i-1}; \theta) \quad (4.5.33)$$

$$(x_i, v_i, y_i) = g(x_i | \Omega_{i-1}; \theta_1) \cdot h(v_i | x_i, \Omega_{i-1}; \theta_2) \cdot k(y_i | x_i, v_i, \Omega_{i-1}; \theta_3) \quad (4.5.34)$$

Dentre os argumentos citados para justificar a relação apresentada em (4.5.34), destaca-se o referente às operações envolvendo investidores com informação privilegiada, nas quais, dependendo do grau de liquidez (intensidade → relacionada à duração) apresentado pelo mercado, os investidores irão decidir pela divisão ou não das transações financeiras (volume) que irão, necessariamente, resultar em alteração do nível de preços praticados (retorno → associado à volatilidade).

Através da modelagem de cada uma das distribuições de probabilidade (marginal e condicional) apresentadas em (4.5.34), é possível o desenvolvimento de uma estrutura completa (sistema) para representar o comportamento das variáveis de interesse, ou seja,  $x_i, v_i, y_i$ . Assim, tem-se:

- ACD: segue a mesma definição proposta por Engle e Russell (1998).

$$x_i = \psi_i \cdot \varepsilon_i \rightarrow \varepsilon_i \sim i.i.d.(1, \sigma_\varepsilon^2) \quad (4.5.35)$$

$$\psi_i = E(x_i | \Omega_i; \theta_x) \quad (4.5.36)$$

- ACV: definido de forma análoga aos modelos de duração, devido às similaridades existentes.

$$v_i = \phi_i \cdot \eta_i \rightarrow \eta_i \sim i.i.d.(1, \sigma_\eta^2) \quad (4.5.37)$$

$$\phi_i = E(v_i | \Omega_i; \theta_v) \quad (4.5.38)$$

- GARCH: modelo escolhido para analisar a dinâmica da volatilidade intradiária.

$$y_i = \sigma_i \cdot \varsigma_i \rightarrow \varsigma_i \sim i.i.d.(0, 1) \quad (4.5.39)$$

$$\sigma_i^2 = E(y_i^2 | \Omega_i; \theta_y) \quad (4.5.40)$$

Já o sistema proposto terá a seguinte formulação:

$$\begin{aligned} x_i &= \psi_i(\theta_x; \Omega_{i-1}) \cdot \varepsilon_i & \varepsilon_i &\sim i.i.d.(1, \sigma_\varepsilon^2) \\ v_i &= \phi_i(\theta_v; x_i, \Omega_{i-1}) \cdot \eta_i & \eta_i &\sim i.i.d.(1, \sigma_\eta^2) \\ y_i &= \sigma_i(\theta_y; x_i, v_i, \Omega_{i-1}) \cdot \zeta_i & \zeta_i &\sim i.i.d.(0, 1) \end{aligned} \quad (4.5.41)$$

De acordo com a forma funcional segundo a qual são especificadas a duração, o volume e a volatilidade condicional, tem-se diferentes modelos econométricos para avaliar o comportamento conjunto das variáveis de interesse e testar as hipóteses cabíveis com relação às microestruturas de mercado. Dentre as diferentes especificações possíveis, tem-se os modelos autoregressivos médias móveis vetoriais (VARMA), onde a dependência entre as variáveis, relacionadas ao evento mais recente, é formulada de maneira explícita. Assim, tem-se:

$$\begin{aligned} \psi_i &= a_0 + \sum_{l=1}^q (a_{1,l} \psi_{i-l} + a_{2,l} \phi_{i-l} + a_{3,l} \sigma_{i-l}^2) + \sum_{m=1}^p (a_{4,m} x_{i-m} + a_{5,m} v_{i-m} + a_{6,m} y_{i-m}^2) \\ \phi_i &= b_0 + \sum_{l=1}^q (b_{1,l} \psi_{i-l} + b_{2,l} \phi_{i-l} + b_{3,l} \sigma_{i-l}^2) + b_4 x_i + \sum_{m=1}^p (b_{5,m} x_{i-m} + b_{6,m} v_{i-m} + b_{7,m} y_{i-m}^2) \\ \sigma_i^2 &= c_0 + \sum_{l=1}^q (c_{1,l} \psi_{i-l} + c_{2,l} \phi_{i-l} + c_{3,l} \sigma_{i-l}^2) + c_4 x_i + c_5 v_i + \sum_{m=1}^p (c_{6,m} x_{i-m} + c_{7,m} v_{i-m} + c_{8,m} y_{i-m}^2) \end{aligned}$$

ou, utilizando notação matricial,

$$\mu_i = \gamma + A_1 \mu_{i-1} + \dots + A_q \mu_{i-q} + B_0 \tau_i + B_1 \tau_{i-1} + \dots + B_p \tau_{i-p} \quad (4.5.42)$$

onde,  $\mu_i' = (\psi_i, \phi_i, \sigma_i^2)$ ,  $\tau_i' = (x_i, v_i, y_i^2)$ ,  $\gamma$  é um vetor de coeficientes e  $A_1, \dots, A_q$  e  $B_1, \dots, B_p$  são matrizes de coeficientes.

De modo a facilitar os cálculos e demonstrações subseqüentes, será feito uso apenas do modelo de primeira ordem:

$$\mu_i = \gamma + A\mu_{i-1} + B\tau_i + C\tau_{i-1} \quad (4.5.43)$$

ou

$$\begin{bmatrix} \psi_i \\ \phi_i \\ \sigma_i^2 \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \\ c_0 \end{bmatrix} + \begin{bmatrix} a_1^{(1)} & a_2^{(1)} & a_3^{(1)} \\ b_1^{(1)} & b_2^{(1)} & b_3^{(1)} \\ c_1^{(1)} & c_2^{(1)} & c_3^{(1)} \end{bmatrix} \begin{bmatrix} \psi_{i-1} \\ \phi_{i-1} \\ \sigma_{i-1}^2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ b_4 & 0 & 0 \\ c_4 & c_5 & 0 \end{bmatrix} \begin{bmatrix} x_i \\ v_i \\ y_i^2 \end{bmatrix} + \begin{bmatrix} a_4^{(1)} & a_5^{(1)} & a_6^{(1)} \\ b_5^{(1)} & b_6^{(1)} & b_7^{(1)} \\ c_6^{(1)} & c_7^{(1)} & c_8^{(1)} \end{bmatrix} \begin{bmatrix} x_{i-1} \\ v_{i-1} \\ y_{i-1}^2 \end{bmatrix}$$

onde a parametrização (4.5.34) impõe que B seja uma matriz triangular inferior, com todos os elementos da diagonal principal iguais a zero. Além desta, podem ser impostas restrições adicionais. Por exemplo, ao impor que A e C sejam diagonais e que B seja nula, tem-se os modelos ACD, ACV e GARCH independentes. É importante notar que a restrição de a matriz A ser diagonal é equivalente a assumir que as médias condicionais das componentes do sistema não influenciam umas às outras. Considerando tal restrição, a verossimilhança (4.5.34) pode ser reescrita como:

$$f(x_i, v_i, y_i | \Omega_i; \theta) = g(x_i | \Omega_i; \theta_x) \cdot h(v_i | x_i, \Omega_i; \theta_v) \cdot k(y_i | x_i, v_i, \Omega_i; \theta_y) \quad (4.5.44)$$

onde,  $\theta' = (\theta_x' \theta_v' \theta_y')$ . Tal procedimento torna possível que a estimação de cada componente do sistema seja feita de forma independente (separada).

A seguir serão apresentadas algumas das proposições<sup>1</sup> constantes do trabalho de Manganeli. São elas:

- 1) O processo (4.5.42) será estacionário de segunda ordem (covariância) se e somente se  $|\lambda| < 1$  para todos os valores de  $\lambda$  que satisfazem:

$$|I_n \lambda^p - H_1 \lambda^{p-1} - H_2 \lambda^{p-2} - \dots - H_p| = 0 \quad (4.5.45)$$

Onde,  $H_i = (B_0 - I)^{-1} (B_i + A)$ ,  $i=1, \dots, p$ .

<sup>1</sup>As provas e demonstrações são apresentadas no apêndice I.

2) A função impulso-resposta do modelo (4.5.43), considerando  $i > 0$ , será:

$$\frac{\partial E(\mu_i | \Omega_i)}{\partial \tau_0} \equiv \Phi_i = D^{i-1} (I - B)^{-1} (AB + C) \quad (4.5.46)$$

Onde,  $D = (I - B)^{-1} (A + C)$ .

3) Considerando o modelo formado por (4.5.41) e (4.5.43). Sendo  $\theta_{(p \times 1)} = [\theta_x' \ \theta_v' \ \theta_y']$  e  $k_i = \text{vec}(\Phi_i(\theta))$ . Se  $\sqrt{N}(\hat{\theta} - \theta) \xrightarrow{a} N(0, Q)$ , então:

$$\sqrt{N}(\hat{k}_i - k_i) \xrightarrow{a} N(0, G_i Q G_i')$$

onde,  $G_i = \frac{\partial k_i}{\partial \theta}$ .

A  $j$ -ésima coluna da matriz de derivadas  $G_i$  pode ser estimada numericamente:

$$G_i^{(j)} = \frac{k_i(\hat{\theta} + e_j \Delta) - k_i(\hat{\theta})}{\Delta}$$

onde,  $e_j$  corresponde à  $j$ -ésima coluna da matriz identidade  $(p \times p)$  e  $\Delta$  a um número pequeno.

4) Função de previsão do sistema:

sendo  $\mu_i = \gamma + A\mu_{i-1} + B\tau_i + C\tau_{i-1}$ , então por substituição recursiva, ter-se-á:

$$\mu_i = \gamma + A(\gamma + A\mu_{i-2} + B\tau_{i-1} + C\tau_{i-2}) + B\tau_i + C\tau_{i-1}$$

$$\mu_i = \gamma + A\gamma + A^2\mu_{i-2} + B\tau_i + (AB + C)\tau_{i-1} + AC\tau_{i-2}$$

⋮

$$\mu_i = \sum_{k=0}^{i-1} A^k \gamma + B\tau_i + \sum_{k=1}^{i-1} A^{k+1} (AB + C)\tau_{i-k} + A^{i-1} C\tau_0 + A^i \mu_0 \quad (4.5.47)$$

Como  $E(\mu_i | \Omega_0) = E(\tau_i | \Omega_0)$ , pela própria definição, então, calculando a expectativa condicional, tem-se:

$$E(\mu_i | \Omega_i) = \sum_{k=0}^{i-1} A^k \gamma + BE(\tau_i | \Omega_i) + \sum_{k=1}^{i-1} A^{k+1} (AB + C) \tau_{i-k} + A^{i-1} C \tau_0 + A^i \mu_0$$

$$E(\mu_i | \Omega_i) = (I - B)^{-1} \left\{ (I - A)^{-1} \gamma + \sum_{k=1}^{i-1} A^{k+1} (AB + C) \tau_{i-k} + A^{i-1} C \tau_0 + A^i \mu_0 \right\} \quad (4.5.48)$$