

Modelos CAPM com  
observações faltando:  
uma aplicação de regressão com  
parâmetros variando no tempo

Pedro Abreu Pessoa de Mendonça

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

Área de Concentração: Estatística, Finanças  
Orientador: Prof. Pedro Luiz Valls Pereira

São Paulo, Maio de 2003

Modelos CAPM com  
observações faltando:  
uma aplicação de regressão com  
parâmetros variando no tempo

Este exemplar corresponde à redação  
final da dissertação devidamente corrigida  
e defendida por Pedro Abreu Pessoa de Mendonça  
e aprovada pela comissão julgadora.

São Paulo, 28 de maio de 2003

Banca examinadora:

- Prof. Dr. Pedro Luiz Valls Pereira (orientador) (IBMEC-SP)
- Prof. Dr. Luiz Koodi Hotta (IMECC-UNICAMP)
- Prof. Dr. Cicero Augusto Vieira Neto (BM&F)

## **Resumo**

Este trabalho tem por objetivo o estudo do preenchimento de valores ausentes em séries históricas de preços no mercado acionário, utilizando o modelo CAPM com parâmetro variando ao longo do tempo. Para a estimação do parâmetro variante no tempo foi usado o filtro de Kalman, que permite fazer previsões e suavizamento de modelos escritos em espaço de estado. A conclusão geral é que a estimação dos valores ausentes utilizando este modelo com parâmetro variante no tempo é superior ao modelo de regressão com parâmetro fixo.

## **Abstract**

This dissertation targets the study of missing value estimation on time series of stock market prices, using the CAPM model with time varying parameters. The Kalman Filter was used for the estimation of the time varying parameter, which allows predicting and smoothing of space state models. The general conclusion is that the estimation of the missing values using this time varying parameter model is better than the fixed parameter regression.

## **MODELOS CAPM COM OBSERVAÇÕES FALTANDO**

---

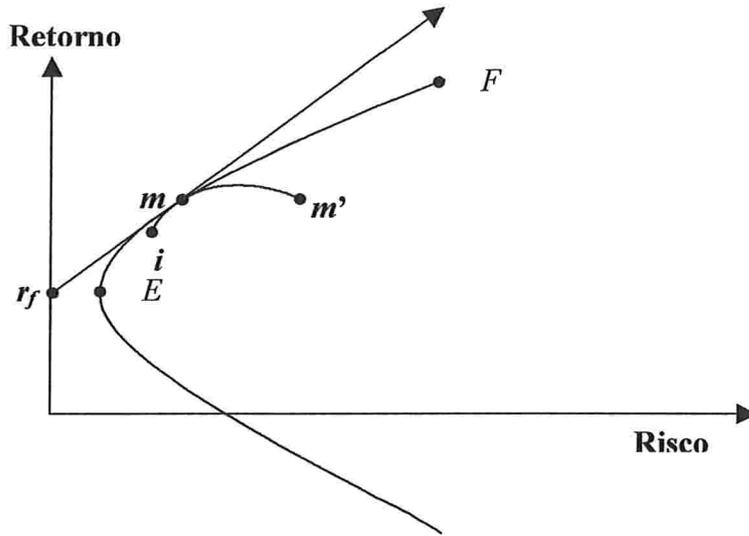
A segunda série foi utilizada na estimação da primeira por ter grande correlação com esta, apresentando assim uma forma de estimar o comportamento da primeira baseado na segunda.

Neste estudo as séries utilizadas são o logaritmo neperiano dos retornos de Petrobrás PN (PETR4) e de Ibovespa (IBOV), bem como os próprios preços de fechamento de 02 de janeiro de 1998 a 14 de novembro de 2002, perfazendo um total de 1.206 dias úteis.

Para a análise, foram retirados valores da série PETR4 aleatoriamente escolhidos e o resultado obtido foi comparado com o resultado real. Os resultados indicam que o modelo com parâmetro variando no tempo é superior ao modelo com parâmetro invariante para a imputação dos dados faltantes.

**Modelo de Precificação de Ativos Financeiros – Capital Asset Pricing Model (CAPM)**

Seja uma carteira  $C$  que consiste de uma proporção  $\omega$  investida diretamente em um ativo qualquer  $i$  e uma proporção  $(1-\omega)$  investida em uma carteira de mercado  $m$ . Esta carteira vai estar em algum ponto da curva  $im$  como ilustrado a seguir, sendo sua localização exata dependente de  $\omega$ .



Desde que  $i$  está contido em  $m$ ,  $C$  realmente possui uma quantia maior que  $\omega$  investida no ativo  $i$  em qualquer ponto da curva  $im$ . Portanto, a carteira de mercado, excluindo o ativo  $i$  denotada por  $m'$  deve estar em uma extensão da curva  $im$  após o ponto  $m$ . Além disto, a curva  $im'$  deve ser tangente à fronteira eficiente  $EF$  no ponto  $m$ , desde que  $m$  é combinação linear de  $i$  e de  $m'$ .

A carteira  $C$  vai ter retorno esperado e desvio padrão iguais a

$$\bar{r}_C = \omega \bar{r}_i + (1 - \omega) \bar{r}_m \quad \text{e} \quad \sigma_C = \left[ \omega^2 \sigma_i^2 + (1 - \omega^2) \sigma_m^2 + 2 \omega (1 - \omega) \text{cov}(i, m) \right]^{1/2}$$

À medida que  $\omega$  é alterado,  $\bar{r}_C$  e  $\sigma_C$  vão se alterar como a seguir.

$$\frac{\partial \bar{r}_C}{\partial \omega} = \bar{r}_i - \bar{r}_m \quad \text{e} \quad \frac{\partial \sigma_C}{\partial \omega} = \frac{(\omega \sigma_i^2 - \sigma_m^2 + \omega \sigma_m^2 + \text{cov}(i, m) - 2 \omega \text{cov}(i, m))}{\sigma_C}$$

Agora a alteração em  $\bar{r}_C$  dada por uma alteração em  $\sigma_C$  é extraída por divisão

$$\frac{\partial \bar{r}_C}{\partial \sigma_C} = \frac{(\bar{r}_i - \bar{r}_m) \sigma_C}{\omega \sigma_i^2 - \sigma_m^2 + \omega \sigma_m^2 + \text{cov}(i, m) - 2 \omega \text{cov}(i, m)}$$

## MODELOS CAPM COM OBSERVAÇÕES FALTANDO

---

Esta última equação pode ser interpretada como o coeficiente angular da reta tangente à curva  $im$ . Pode-se provar que a reta tangente à hipérbole na figura anterior é  $(\bar{r}_m - r_f)/\sigma_m$  onde  $r_f$  é um ativo livre de risco. Igualando a tangente acima à tangente à curva  $im$  no ponto  $m$  resulta em

$$\bar{r}_i = r_f + \left[ (\bar{r}_m - r_f) / \sigma_m^2 \right] \text{cov}(i, m)$$

e esta é a equação do modelo CAPM.

Este modelo pode ser reescrito na forma

$$\bar{r}_i = r_f + (\bar{r}_m - r_f) \beta_i$$

onde  $\beta_i \equiv \frac{\text{cov}(i, m)}{\sigma_m^2}$ .

Trazendo este modelo para o nosso estudo, temos  $m$  sendo o Índice Bovespa e  $i$  como Petrobrás PN. O ativo livre de risco não será considerado para simplificar o estudo.

## Modelos de Espaço de Estado e o Filtro de Kalman

### Representação em Espaço de Estado

A razão de se representar um modelo em espaço de estados é devida ao fato de que os modelos em geral relacionam variáveis observadas e não observadas (parâmetros) e podem ter parâmetros que variam no tempo.

A modelagem em espaço de estado fornece uma metodologia unificada para o tratamento de uma vasta gama de problemas na análise de séries temporais. Nesta abordagem é assumido que o desenvolvimento ao longo do tempo do sistema sob estudo é determinado por uma série não-observável de vetores  $\alpha_1, \alpha_2, \dots, \alpha_n$  os quais são associados a séries de observações  $y_1, \dots, y_n$ ; a relação entre os  $\alpha_t$ 's e os  $y_t$ 's é especificada pelo modelo de espaço de estados. O propósito da análise de espaço de estados é inferir as propriedades relevantes dos  $\alpha_t$ 's a partir do conhecimento das observações  $y_1, \dots, y_n$ .

A representação na forma de espaço de estado fornece um mecanismo recursivo do qual se obtém o erro de previsão um passo à frente e a variância deste erro de previsão. Assim, pode-se obter a função de verossimilhança do modelo e maximizá-la através de métodos numéricos, obtendo estimadores de máxima verossimilhança (ou quase-máxima verossimilhança) para os parâmetros. Os erros de previsão podem ser utilizados para verificar a adequabilidade do modelo.

A *Equação de Medida*, dada a seguir, relaciona o vetor  $N \times 1$  de variáveis observadas no tempo  $t$ ,  $y_t$ , a um vetor  $m \times 1$ ,  $\alpha_t$ , conhecido como o *Vetor de Estado*.

$$y_t = Z_t \alpha_t + d_t + \varepsilon_t, \quad t = 1, \dots, T \quad (1.1)$$

onde  $Z_t$  é uma matriz  $N \times m$ ,  $d_t$  é um vetor  $N \times 1$  e  $\varepsilon_t$  é um vetor  $N \times 1$  de ruídos serialmente não correlacionados com média zero e matriz de covariâncias  $H_t$

Em geral os elementos de  $\alpha_t$  são não-observáveis. Contudo, sabe-se que eles são gerados por um processo de Markov de primeira ordem, e têm transição dada pela equação

$$\alpha_t = T_t \alpha_{t-1} + c_t + R_t \eta_t \quad t = 1, \dots, T \quad (1.2)$$

Onde  $T_t$  é uma matriz  $m \times m$ ,  $c_t$  é um vetor  $m \times 1$ ,  $R_t$  é uma matriz  $m \times g$  e  $\eta_t$  é um vetor  $g \times 1$  de ruídos serialmente não correlacionados com média zero e matriz de covariâncias,  $Q_t$ .

A equação (1.2) é a *Equação de Transição*. A inclusão da matriz  $R_t$  em frente ao ruído  $\eta_t$  é arbitrária. O termo de ruído pode sempre ser redefinido de forma a ter  $R_t Q_t R_t'$  como a matriz de variância-covariância.

A especificação do sistema de espaço de estado é completada por duas outras suposições:

(a) O vetor de estado inicial,  $\alpha_0$ , tem uma média  $a_0$  e matriz de covariâncias  $P_0$ , isto é

$$E(\alpha_0) = a_0 \quad \text{e} \quad \text{Var}(\alpha_0) = P_0 \quad (1.3)$$

(b) Os ruídos  $\varepsilon_t$  e  $\eta_t$  são não-correlacionados entre si em todos os períodos de tempo, e não-correlacionados com o *estado inicial*.

As matrizes  $Z_t$ ,  $d_t$  e  $H_t$  na *Equação de Medida* e  $T_t$ ,  $c_t$  e  $R_t$  na *Equação de Transição* serão chamadas de *Matrizes do Sistema*. Será assumido que elas são não-estocásticas. Logo, apesar delas poderem sofrer alterações com o tempo, elas o fazem de uma maneira pré-determinada. Como resultado, o sistema é linear e para qualquer valor de  $t$ ,  $y_t$  pode ser expresso como uma combinação linear de  $\varepsilon_t$ 's e  $\eta_t$ 's passados e presente e do vetor de estado inicial,  $\alpha_0$ .

No caso em que as *Matrizes do Sistema* não sofrerem alterações ao longo do tempo, o modelo é chamado de *Invariante no Tempo* ou *Homogêneo no Tempo*. Modelos estacionários são um caso especial. A *equação de transição* em um *modelo invariante no tempo* é um processo autoregressivo vetorial de primeira ordem.

## Filtragem, Suavizamento e Previsão

O filtro de Kalman é um mecanismo recursivo para se calcular o estimador do vetor de estado dada toda a informação disponível até aquele instante de tempo ou usando toda a informação disponível na amostra.

Uma vez que o fim da série é alcançado, previsões de observações futuras podem ser realizadas. Uma repetição reversa, conhecida como suavizamento habilita estimadores ótimos do vetor de estado a serem calculados em todos os pontos do tempo usando a amostra completa.

### Filtro de Kalman

O objetivo da filtragem é atualizar o conhecimento do sistema a cada vez que uma nova observação  $y_t$  é obtida.

Seja  $a_t$  o estimador ótimo de  $\alpha_t$ , usando as observações até o instante  $t$ . Seja  $P_t$  a matriz de erro quadrado médio (EQM) de  $a_t$ , isto é

$$P_t = [(\alpha_t - a_t)(\alpha_t - a_t)'] \quad (2.1)$$

não pode ser propriamente chamado de matriz de covariâncias de  $a_t$  uma vez que o vetor a ser estimado,  $\alpha_t$ , consiste de variáveis aleatórias em vez de parâmetros fixos.

### Previsão

No instante  $t-1$ ,  $a_{t-1}$  e  $P_{t-1}$  são dados. O melhor previsor de  $\alpha_t$ , usando a informação até  $t-1$  será dado pelas equações de previsão

$$a_{t|t-1} = T_t a_{t-1} + c_t \quad (2.2a)$$

e

$$P_{t|t-1} = T_t P_{t-1} T_t' + R_t Q_t R_t', \quad t = 1, \dots, T \quad (2.2b)$$

enquanto o estimador correspondente de  $y_t$  é

$$\tilde{y}_{t|t-1} = Z_t a_{t|t-1} + d_t, \quad t = 1, \dots, T \quad (2.3a)$$

**Erro de Previsão e sua variância**

O estimador do vetor de erros de previsão, ou de *inovação* é

$$v_t = y_t - \tilde{y}_{t|t-1} = Z_t(\alpha_t - a_{t|t-1}) + \varepsilon_t, \quad t = 1, \dots, T \quad (2.3b)$$

e a variância do erro de previsão é dada por

$$F_t = Z_t P_{t|t-1} Z_t' + H_t \quad (2.3c)$$

**Equações de Atualização**

Uma vez que novas observações se tornam disponíveis, o estimador do vetor de estado pode ser atualizado. As *equações de atualização* são

$$a_t = a_{t|t-1} + P_{t|t-1} Z_t' F_t^{-1} (y_t - Z_t a_{t|t-1} - d_t) \quad (2.4a)$$

e

$$P_t = P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1}, \quad t = 1, \dots, T \quad (2.4b)$$

Como pode-se observar, quanto mais a previsão da observação desviar do seu valor observado, maior será a mudança feita no estimador do estado.

Juntas, as equações (2.2) e (2.4) formam o *Filtro de Kalman*. E dados,  $a_0$  e  $P_0$ , o filtro de Kalman devolve o estimador do estado à medida que cada nova observação se torna disponível. Quando todas as  $T$  observações forem processadas, o estimador  $a_T$  conterá toda a informação necessária para fazer previsões de futuras observações.

***Previsão***

As previsões são dadas adiantando as equações de atualização. O estimador do vetor de estado no tempo  $T + l$ , baseado na informação até o tempo  $T$  é dada por

$$a_{T+l|T} = T_{T+l} a_{T+l-1} + c_{T+l} \quad l = 1, 2, \dots \quad (2.5a)$$

com  $a_{T|T} = a_T$ , enquanto a matriz de erros quadrados médios associada é obtida de

$$P_{T+l|T} = T_{T+l} P_{T+l-1|T} T_{T+l}' + R_{T+l} Q_{T+l} R_{T+l}', \quad l = 1, 2, \dots \quad (2.5b)$$

com  $P_{T|T} = P_T$ . O preditor de  $y_{T+l}$  é

$$\tilde{y}_{T+l|T} = Z_{T+l}a_{T+l|T} + d_{T+l}, \quad l = 1, 2, \dots \quad (2.6a)$$

com estimador de erro quadrado médio dado por

$$EQM(\tilde{y}_{T+l|T}) = Z_{T+l}P_{T+l|T}Z'_{T+l} + H_{T+l} \quad (2.6b)$$

Quando o modelo é Gaussiano, (2.6b) pode ser usada para construir intervalos de previsão.

### **Suavizamento**

O estimador suavizado,  $a_{t|T}$ , é baseado em mais informações que o estimador filtrado, e portanto vai ter uma matriz de EQM,  $P_{t|T}$ , a qual, em geral, é menor que a do estimador filtrado.

Existem basicamente três algoritmos de suavizamento em um modelo linear.

1. Suavizamento em *ponto fixo* (fixed-point) trata de estimadores suavizados computacionais do vetor de estado em algum ponto fixo no tempo. Assim ele fornece  $a_{\tau|t}$  para valores de  $\tau$  em todos os períodos de tempo  $t > \tau$ .
2. Suavizamento em *prazos fixos* (fixed-lag) calcula o estimador suavizado para um prazo fixo, isto é  $a_{t-j|t}$  para  $j = 1, \dots, M$ , onde  $M$  é algum prazo máximo. Ambos algoritmos podem ser utilizados em uma situação *on-line*.
3. Suavizamento em *intervalos fixos* (fixed-interval), por outro lado, trata de utilizar todo o conjunto de estimadores suavizados para um espaço fixo dos dados. Então é uma técnica *off-line* que produz  $a_{t|T}$ ,  $t = 1, \dots, T$ . Este tende a ser o mais usado dos algoritmos para dados econômicos e sociais.

O algoritmo de suavizamento em intervalos fixos consiste de um conjunto de repetições que começam com os valores finais,  $a_T$  e  $P_T$ , dados pelo filtro de Kalman e funciona em reverso.

As equações são

$$a_{t|T} = a_t + P_t^* (a_{t+1|T} - T_{t+1} a_t - c_{t+1}) \quad (2.7a)$$

e

$$P_{t|T} = P_t + P_t^* (P_{t+1|T} - P_{t+1|t}) P_t^{*'} \quad (2.7b)$$

onde

$$P_t^* = P_t T'_{t+1} P_{t+1|t}^{-1}, \quad t = T - 1, \dots, 1 \quad (2.7c)$$

com  $a_{T|T} = a_T$  e  $P_{T|T} = P_T$ . Como está, o algoritmo requer portanto que  $a_t$  e  $P_t$  sejam armazenados para todo  $t$  de forma que eles possam ser combinados com  $a_{t+1|T}$  e  $P_{t+1|T}$ .

Algoritmos de suavização por espaço de estado são mais gerais que os algoritmos clássicos de extração de sinal (signal extraction) desde que eles se aplicam a amostras finitas e a sistemas que não precisam ser invariantes no tempo. A atração das fórmulas clássicas é que estas são fáceis de se interpretar.

### ***Inicialização do Filtro de Kalman***

Se o estado é gerado por um processo estacionário, as condições iniciais do filtro de Kalman são dadas por sua média e variância incondicionais.

Se a equação de transição não é estacionária, e as condições iniciais não são dadas como parte das especificações do modelo, elas podem e devem ser estimadas a partir das próprias observações. Existem duas abordagens:

- A primeira assume que o estado inicial,  $\alpha_0$ , é fixo. Sua distribuição é então degenerada como  $P_0 = 0$ . Contudo, desde que  $\alpha_0$  é desconhecido, seus elementos devem ser estimados tratando-os como parâmetros desconhecidos no modelo.
- A Segunda abordagem assume que  $\alpha_0$  é aleatório e tem uma distribuição *difusa*, isto é, sua matriz de covariâncias é  $P_0 = \kappa I$ , com  $\kappa \rightarrow \infty$ . Isto quer dizer que nada se sabe a

respeito do estado inicial. O resultado é que os valores iniciais são efetivamente construídos das observações iniciais.

Começar o Filtro de Kalman com  $\kappa$  igual a um número grande, porém finito, pode dar uma aproximação do filtro que pode ser obtido com condições iniciais difusas. Contudo, isto não é inteiramente satisfatório, particularmente porque a presença de números grandes no filtro pode levar a instabilidade numérica.

### Modelos Gaussianos e a Função de Verossimilhança

Em um modelo de espaço de estado Gaussiano, os ruídos  $\varepsilon_t$  e  $\eta_t$ , e o vetor de estado inicial,  $\alpha_0$ , são normalmente distribuídos. Sob estas suposições é relativamente direta a obtenção do Filtro de Kalman. A maneira na qual a função de verossimilhança pode ser construída a partir dos erros de previsão segue quase imediatamente.

#### Obtenção do Filtro de Kalman

O estado inicial é normalmente distribuído com média  $a_0$  e matriz de covariâncias  $P_0$ . O vetor de estado no tempo  $t = 1$  é dado por

$$\alpha_1 = T_1 \alpha_0 + c_1 + R_1 \eta_1 \quad (3.1)$$

Então  $\alpha_1$  é uma combinação linear de dois vetores de variáveis aleatórias, ambos tendo distribuição normal multivariada, e um vetor de constantes. Logo também é normal, com média

$$a_{1|0} = T_1 a_0 + c_1 \quad (3.2a)$$

e matriz de covariâncias

$$P_{1|0} = T_1 P_0 T_1' + R_1 Q_1 R_1' \quad (3.2b)$$

A notação  $a_{1|0}$  serve para indicar a média da distribuição de  $\alpha_1$  condicional à informação no tempo  $t = 0$ .

Para obter a distribuição de  $\alpha_1$  condicional a  $y_1$ , escrevemos

$$\alpha_1 = a_{1|0} + (\alpha_1 - a_{1|0}) \quad (3.3a)$$

$$y_1 = Z_1 a_{1|0} + d_1 + Z_1(\alpha_1 - a_{1|0}) + \varepsilon_1 \quad (3.3b)$$

Pode ser visto diretamente de (3.3) que o vetor normalmente distribuído  $(\alpha_1' \ y_1)'$  tem média e matriz de covariâncias dadas por

$$\begin{bmatrix} a_{1|0} \\ Z_1 a_{1|0} + d_1 \end{bmatrix} e \begin{bmatrix} P_{1|0} & P_{1|0} Z_1' \\ Z_1 P_{1|0} & Z_1 P_{1|0} Z_1' + H_1 \end{bmatrix}$$

Assim temos das propriedades da distribuição multivariada que a distribuição de  $\alpha_1$ , condicional a um valor observado de  $y_1$ , é normal multivariada com média

$$a_1 = a_{1|0} + P_{1|0} Z_1' F_1^{-1} (y_1 - Z_1 a_{1|0} - d_1) \quad (3.4a)$$

e matriz de covariâncias

$$P_1 = P_{1|0} - P_{1|0} Z_1' F_1^{-1} Z_1 P_{1|0} \quad (3.4b)$$

onde

$$F_1 = Z_1 P_{1|0} Z_1' + H_1 \quad (3.4c)$$

A repetição dos passos anteriores para  $t = 2, \dots, T$  fornece o filtro de Kalman. Essa obtenção nos permite interpretar  $a_t$  e  $P_t$  como a média e a matriz de covariâncias da distribuição condicional de  $\alpha_t$ . Contudo, essa média condicional é a estimativa de mínimos quadrados médios de  $\alpha_t$ . Além disto, se  $a_t$  for tomado como um estimador ao invés de uma estimativa, ele pode ser usado para minimizar o EQM quando a esperança de todas as observações existentes é tomada, em vez de ser condicionada a somente parte dos valores. Assim, o estimador da média condicional é o estimador de mínimos quadrados de  $\alpha_t$ .

O estimador  $a_t$  é não-viciado no sentido de que a esperança do erro de estimação é zero. Desde que esta esperança pode ser tomada de todas as observações, esta propriedade é chamada de *não-viciosidade incondicional* (unconditional unbiasedness). Além disso, desde que a matriz  $P_t$  é independente das observações, é também a *matriz incondicional de covariâncias dos erros* (unconditional error covariance matrix). Em outras palavras a esperança em (2.1) não precisa ser condicional às observações obtidas até o tempo  $t$ , inclusive.

Quando os ruídos no modelo de espaço de estado não são normalmente distribuídos, não é verdadeiro, em geral, que o filtro de Kalman produz a média condicional do vetor de estado. Contudo, se a atenção for restrita aos estimadores que são

combinações lineares das observações, então  $a_t$  é aquele que minimiza o EQM. Assim  $a_t$  é o estimador linear de mínimos quadrados médios de  $\alpha_t$  baseado em observações até o tempo  $t$  inclusive. Este estimador é incondicionalmente não-viciado e a matriz incondicional de covariâncias do erro de estimação é novamente a matriz  $P_t$  dada pelo filtro de Kalman.

Os pontos acima se aplicam exatamente da mesma forma para  $a_{t|t-1}$  e  $P_{t|t-1}$ . Além disto, a média condicional de  $y_t$  no tempo  $t-1$ , cuja notação é  $\tilde{y}_{t|t-1}$  de (2.3a), pode ser interpretada como a estimador de mínimos quadrados médios de  $y_t$  em um modelo Gaussiano, e como um estimador linear de mínimos quadrados médios em caso contrário.

### **Máxima Verossimilhança**

Em um modelo de espaço de estado as matrizes do sistema vão depender geralmente de um conjunto de parâmetros desconhecidos. Estes serão denotados por  $\psi$ , um vetor  $N \times 1$ , e chamados de hiperparâmetros. A estimação de máxima verossimilhança dos hiperparâmetros pode ser feita usando-se o filtro de Kalman para se construir a função de máxima verossimilhança e então maximizá-la usando um procedimento de otimização numérica.

A função de densidade conjunta de uma gama de  $T$  observações pode ser expressada em termos de distribuições condicionais. Para um modelo univariado

$$L(y, \psi) = \prod_{t=1}^T p(y_t | Y_{t-1}) \quad (3.5)$$

Onde  $p(y_t | Y_{t-1})$  denota a distribuição condicional de  $y_t$  dado o conjunto de informações até o tempo  $t-1$ , isto é,  $Y_{t-1} = \{y_{t-1}, y_{t-2}, \dots, y_1\}$ . A partir do cálculo do filtro de Kalman na seção anterior, será lembrado que, dado  $Y_{t-1}$ ,  $\alpha_t$  é normalmente distribuído com média  $a_{t|t-1}$  e matriz de covariâncias  $P_{t|t-1}$ . Se a equação de medida for escrita da forma

$$y_t = Z_t a_{t|t-1} + Z_t (\alpha_t - a_{t|t-1}) + d_t + \varepsilon_t \quad (3.6)$$

Pode ser visto imediatamente que a distribuição condicional de  $y_t$  é normal com média

$$\tilde{y}_{t|t-1} = Z_t a_{t|t-1} + d_t \quad (3.7)$$

E uma matriz de covariâncias,  $F_t$ , dada por (2.3c). Para um modelo Gaussiano, a função de verossimilhança (3.5) pode ser escrita como

$$\log L(\psi) = -\frac{NT}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log |F_t| - \frac{1}{2} \sum_{t=1}^T v_t' F_t^{-1} v_t \quad (3.8)$$

onde  $v_t$  é o vetor de erros de previsão, (2.3b). A expressão (3.8) é conhecida como a forma de *decomposição do erro de previsão* da verossimilhança.

O filtro de Kalman pode ser inicializado com a média e a matriz de covariâncias da distribuição incondicional de  $\alpha_t$  quando  $\alpha_t$  é estacionário. Para vetores de estado não estacionários, uma função de verossimilhança pode ser formada a partir de todos os  $T$  erros de previsão em (3.8), somente se informações anteriores estiverem disponíveis de forma que  $\alpha_0$  tenha distribuição apropriada com média conhecida,  $a_0$ , e matriz de covariâncias limitada,  $P_0$ .

Se o problema de inicialização é abordado tratando  $\alpha_0$  como um vetor de parâmetros fixos que têm que ser estimados, esses parâmetros podem ser concentrados fora da função de verossimilhança. Contudo, as propriedades dos estimadores obtidas nesta maneira parecem ser menos satisfatórias que estimadores baseados na abordagem *a priori difusa*. Isto é, de fato, a verossimilhança marginal para o modelo com  $\alpha_0$  fixo. O uso de uma verossimilhança marginal é geralmente tido como preferível quando um modelo contém parâmetros de distúrbios, neste caso os elementos de  $\alpha_0$ .

### ***Resíduos e checagem diagnóstica***

Em um modelo Gaussiano, a função densidade conjunta das observações é dada pelo lado direito de (3.8). Esta é também a função densidade conjunta das inovações que são conseqüentemente independentes e normalmente distribuídas, isto é

$$v_t \sim NID(0, F_t) \quad t = 1, \dots, T \quad (3.9)$$

Os resíduos padronizados,  $F_t^{-1/2} v_t$ , são então úteis para checagem diagnóstica, apesar de que se as matrizes do sistema contiverem hiperparâmetros desconhecidos que são substituídos por estimadores, (3.9) não funciona perfeitamente.

Na ausência da suposição de normalidade a média do vetor de inovações ainda é um vetor de zeros, enquanto sua matriz de covariâncias no tempo  $t$  é  $F_t$ . Além disto, pode ser mostrado que as inovações em períodos de tempo diferentes são não correlacionadas.

### ***Observações faltantes e outras irregularidade nos dados***

Observações faltantes não são problema quando se utiliza a abordagem de espaço de estado. Se uma observação está faltando em um ponto no tempo,  $t = \tau$ , as equações de atualização do filtro de Kalman são redundantes. Assim

$$a_\tau = a_{\tau|\tau-1} \quad e \quad P_\tau = P_{\tau|\tau-1}$$

Então as equações de previsão produzem o que é realmente um estimador 2 passos a frente e as equações de atualização são aplicadas da maneira usual quando a observação  $y_{\tau+1}$  se tornar disponível. As equações de suavizamento fornecem um estimador do estado no tempo  $\tau$ . Assim o estimador das observações faltantes, junto com o EQM, pode ser calculado.

A função de verossimilhança é calculada da maneira usual, utilizando as inovações do Filtro de Kalman, exceto, é claro, que não existe inovação correspondente à observação faltante. Logo não existem termos no tempo  $\tau$ , e a inovação no tempo  $\tau + 1$  é um erro de previsão dois passos a frente. Em termos de (3.5), não existe distribuição condicional no tempo  $\tau$ , e a distribuição no tempo  $\tau + 1$  é condicional à informação no tempo  $\tau - 1$ .

O Filtro de Kalman também pode lidar com situações onde observações são agregadas ao longo do tempo, e somente as agregadas são observadas. Isto acontece se uma série de uma variável fluxo, tal como renda nacional, consiste de algumas observações anuais, seguida de observações trimestrais. A solução é estender o vetor de estado para incluir uma variável que é capaz de acumular a série nos pontos onde ela não é observada.

Outros tipos de irregularidade nos dados podem ser tratados pela aproximação de espaço de estado. Por exemplo, um modelo pode ser construído para contabilizar as revisões de dados que são freqüentemente feitas à séries temporais pelas agências governamentais.

### Regressão e parâmetros variantes no tempo

Considere o seguinte modelo de regressão,

$$y_t = x_t' \beta + \varepsilon_t, \quad t = 1, \dots, T \quad (5.1)$$

Onde  $x_t$  é um vetor  $k \times 1$  de variáveis explanatórias exógenas e  $\beta$  é o vetor  $k \times 1$  de parâmetros desconhecidos correspondente.

### Mínimos quadrados recursivos

O estimador de  $\beta_t$  em (5.1), usando a informação até o instante  $t$ , é dado por

$$b_t = \left( \sum_{j=1}^t x_j x_j' \right)^{-1} \sum_{j=1}^t x_j y_j \quad (5.2)$$

Isto é feito pela fórmula de atualização recursiva. Se  $X_t = (x_1, \dots, x_t)'$ , então

$$b_t = b_{t-1} + \left( X_{t-1}' X_{t-1} \right)^{-1} x_t (y_t - x_t' b_{t-1}) / f_t \quad (5.3a)$$

e

$$\left( X_t' X_t \right)^{-1} = \left( X_{t-1}' X_{t-1} \right)^{-1} - \left( X_{t-1}' X_{t-1} \right)^{-1} x_t x_t' \left( X_{t-1}' X_{t-1} \right)^{-1} / f_t \quad (5.3b)$$

onde

$$f_t = 1 + x_t' \left( X_{t-1}' X_{t-1} \right)^{-1} x_t, \quad t = k + 1, \dots, T \quad (5.3c)$$

Um mínimo de  $k$  observações são necessárias para calcular um estimador OLS de  $\beta$ . Se  $X_k$  possui rank completo (full rank), o estimador baseado nas primeiras  $k$  informações é

$$b_k = \left( X_k' X_k \right)^{-1} X_k' y_k = X_k^{-1} y_k \quad (5.4)$$

onde  $y_k = (y_1, \dots, y_k)'$ . Isto dá um valor inicial e os estimadores  $b_{k+1}, \dots, b_T$  podem ser computados sem futuras inversões de matrizes. O estimador final,  $b_t$ , é idêntico ao estimador OLS padrão baseado em todas as  $T$  observações.

A fórmula de atualização de mínimos quadrados recursivos pode ser obtida usando o lema de inversão de matrizes. Contudo, eles são um caso especial do filtro de Kalman. O modelo de regressão em (5.1) pode ser identificado diretamente com uma equação de

medida da forma (3.9) fazendo  $z_t = x_t$ ,  $h_t = 1$  e  $\alpha_t = \beta$ . A equação de transição é simplesmente  $\alpha_t = \alpha_{t-1}$  e então  $T = 1$  enquanto  $Q = 0$ . As equações de previsão são conseqüentemente triviais à medida que elas reduzem as identidades  $a_{t|t-1} = a_{t-1}$  e  $P_{t|t-1} = P_{t-1}$ , enquanto as fórmulas de atualização OLS são obtidas diretamente das equações de atualização (2.4) com

$$P_t = (X_t' X_t)^{-1}, \quad t = k, \dots, T$$

A inicialização em (5.4) é implicada pelo uso de uma priori difusa para inicializar o filtro de Kalman.

Quando o estimador OLS é computado recursivamente, um conjunto de  $T - k$  erros de previsão

$$v_t = y_t - x_t' b_{t-1}, \quad t = k + 1, \dots, T \quad (5.5)$$

são obtidos. Estes têm média zero, variância  $\sigma^2 f_t$  e são não correlacionados; veja (3.16). O erro de previsão padronizado

$$\tilde{v}_t = v_t / \sqrt{f_t}, \quad t = k + 1, \dots, T \quad (5.6)$$

são conhecidos como *resíduos recursivos* e elas figuram em procedimentos para checar a especificação e estabilidade dos modelos de regressão. Lembre-se que se os ruídos em (5.1) são normalmente distribuídos, os resíduos recursivos são também normais, isto é  $\varepsilon_t \sim NID(0, \sigma^2)$  implica  $\tilde{v}_t \sim NID(0, \sigma^2)$ .

A característica dos resíduos recursivos na fórmula de atualização para a soma de quadrado dos resíduos, ou soma de erros quadrados (SSE). Esta repetição é

$$SSE_t = SSE_{t-1} + v_t^2, \quad t = k + 1, \dots, T \quad (5.7)$$

onde

$$SSE_t = (y_t - X_t b_t)' (y_t - X_t b_t), \quad t = k, \dots, T \quad (5.8)$$

Desde que  $SSE_k = 0$  segue que a soma de quadrados residual a partir de todas as  $T$  informações é igual à soma de quadrados dos resíduos recursivos, isto é

$$SSE_T = \sum_{t=k+1}^T v_t^2 \quad (5.9)$$

O uso de (5.7) em conjunto com (5.3) permite que a matriz de covariâncias de  $b_t$  seja estimada como os procedimentos de repetição.

### **Parâmetros de Caminhos Aleatórios (Random Walks)**

O modelo de caminhos aleatórios com parâmetros variantes no tempo é

$$y_t = x_t' \beta_t + \varepsilon_t, \quad t = 1, \dots, T \quad (5.10a)$$

onde  $\varepsilon_t \sim NID(0, \sigma^2)$  e o vetor  $\beta_t$  é gerado pelo processo

$$\beta_t = \beta_{t-1} + \eta_t \quad (5.10b)$$

onde  $\eta_t \sim NID(0, \sigma^2 Q)$ . Por  $\beta_t$  ser não estacionário, ele é capaz de evoluir de forma que o modelo pode acomodar mudanças fundamentais em sua estrutura. A matriz  $k \times k$ ,  $Q$ , PSD (positiva semi-definida), determina a extensão na qual ele pode variar. Se  $Q=0$ , o modelo cai num modelo de regressão, como em (5.1), pois  $\beta_t = \beta_{t-1}$ . Por outro lado, se  $Q$  é PD (positiva definida), todos os  $k$  parâmetros serão variantes no tempo.

O modelo já está na forma de espaço de estados, com  $\beta_t$  sendo o estado. Como em mínimos quadrados recursivos, o uso de uma priori difusa leva as primeiras  $k$  observações a serem usadas para construir valores iniciais. Isto pode ser feito explicitamente expressando os primeiros  $k - 1$  vetores de coeficientes em termos do  $k$ -ésimo. Assim, por exemplo

$$\beta_k = \beta_{k-1} + \eta_t = \beta_{k-2} + \eta_{t-1} + \eta_t \quad (5.11)$$

levando a

$$\beta_{k-2} = \beta_k - \eta_{t-1} - \eta_t$$

e similarmente para os outros vetores de coeficientes. As primeiras  $k$  equações em (5.10a) podem então ser escritas como

$$y_t = x_t' \beta_k + \zeta_t, \quad t = 1, \dots, k \quad (5.12a)$$

onde  $\zeta_k = \varepsilon_k$  e

$$\zeta_t = \varepsilon_t - x_t'(\eta_{t+1} + \dots + \eta_k), \quad t = 1, \dots, k-1 \quad (5.12b)$$

A matriz de covariância do vetor de ruídos,  $\zeta_k = (\zeta_1, \dots, \zeta_k)'$ , é  $\sigma^2 V$ , onde o  $ij$ -ésimo elemento de  $V$  é

$$v_{ij} = \delta_{ij} + [\min(k-i, k-j)] x_i' Q x_j \quad i, j = 1, \dots, k$$

onde  $\delta_{ij}$  é um para  $i = j$  e zero caso contrário.

Escrevendo as equações em (5.12a) na forma matricial, e re-arranjando nos dá

$$\beta_k = X_k^{-1} y_k - \zeta_k \quad (5.13)$$

de onde pode ser visto que, condicionado às  $k$  primeiras observações,  $y_k$ , o vetor de coeficientes  $\beta_k$  é normalmente distribuído com média

$$b_k = X_k^{-1} y_k \quad (5.14a)$$

e matriz de covariâncias  $\sigma^2 P_k$ , onde

$$P_k = (X_k V_k^{-1} X_k')^{-1} \quad (5.14b)$$

**Implementação no software Ox<sup>®</sup> utilizando o pacote SsfPack<sup>®</sup> 2.2**

**Representação em Espaço de Estado**

A representação em espaço de estados gaussiana consiste de uma equação de transição e uma equação de medida formuladas como a seguir

$$\alpha_t = T_t \alpha_{t-1} + c_t + R_t \eta_t, \quad \alpha_1 \sim N(a, P) \quad t = 1, \dots, n \quad (6.1)$$

$$y_t = Z_t \alpha_t + d_t + \varepsilon_t \quad \varepsilon_t, \eta_t \sim \text{NID}(0, I) \quad (6.2)$$

Esta representação no SsfPack<sup>®</sup> do Ox<sup>®</sup> é dada por:

$$\begin{pmatrix} \alpha_{t+1} \\ y_t \end{pmatrix} = \delta_t + \Phi_t \alpha_t + u_t \quad u_t \sim \text{NID}(0, \Omega_t) \quad t = 1, \dots, n$$

$$\delta_t = \begin{pmatrix} c_t \\ d_t \end{pmatrix}, \quad \Phi_t = \begin{pmatrix} T_t \\ Z_t \end{pmatrix}, \quad u_t = \begin{pmatrix} R_t \\ 1 \end{pmatrix} \varepsilon_t, \quad \Omega_t = \begin{pmatrix} H_t H_t' & H_t G_t' \\ G_t H_t' & G_t G_t' \end{pmatrix} \quad (6.3)$$

$$\alpha_1 \sim N(a, P)$$

onde as condições iniciais são

$$\begin{aligned} P &= P_* + \kappa P_\infty, & \kappa &= 10^6 \\ a &= 0, & P_* &= 0, & P_\infty &= I \quad \text{tal que } \alpha_1 \sim N(0, \kappa I) \end{aligned} \quad (6.4)$$

**Representação em Espaço de Estado variante no tempo**

Quando alguns elementos das matrizes do sistema não são constantes, mas mudam ao longo do tempo, são usadas as matrizes índices  $J_\Phi, J_\Omega, J_\delta$  que devem ter as mesmas dimensões de  $\Phi, \Omega$  e  $\delta$  respectivamente. Os elementos das matrizes índices são todos iguais a -1 exceto os elementos cujos elementos correspondentes em  $\Phi, \Omega$  e  $\delta$  são variantes no tempo.

## Valores ausentes

O vetor de observações  $y_t$  com valores ausentes vai ser reduzido para o vetor  $y_t^*$  sem os valores ausentes de forma que a equação de medida seja ajustada de acordo com o último vetor. Por exemplo, a equação de medida  $y_t = c_t + Z_t\alpha_t + G_t\eta_t$  com

$$y_t = \begin{pmatrix} 5 \\ \cdot \\ 3 \\ \cdot \end{pmatrix}, \quad c_t = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \quad Z_t = \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \end{pmatrix}, \quad G_t = \begin{pmatrix} G_{1,t} \\ G_{2,t} \\ G_{3,t} \\ G_{4,t} \end{pmatrix} \quad (6.5)$$

é reduzida para a equação  $y_t^* = c_t^* + Z_t^*\alpha_t + G_t^*\eta_t$  com

$$y_t^* = \begin{pmatrix} 5 \\ 3 \end{pmatrix}, \quad c_t^* = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad Z_t^* = \begin{pmatrix} Z_{1,t} \\ Z_{3,t} \end{pmatrix}, \quad G_t^* = \begin{pmatrix} G_{1,t} \\ G_{3,t} \end{pmatrix} \quad (6.6)$$

## Modelos de Regressão

O modelo de regressão pode ser representado na forma de espaço de estados. O Filtro de Kalman para o modelo de regressão na forma de espaço de estados é equivalente ao algoritmo de “mínimos quadrados recursivos” para o modelo de regressão padrão. A forma de espaço de estados para o modelo de regressão múltipla univariado  $y_t = X_t\beta + \xi_t$  com  $\xi_t \sim \text{NID}(0, \sigma_\xi^2)$  e o vetor de coeficientes  $\beta$  ( $k \times 1$ ) para  $t = 1, \dots, n$  é dado por

$$\alpha_{t+1} = \alpha_t \quad y_t = X_t\alpha_t + G_t\varepsilon_t \quad t = 1, \dots, n \quad (6.7)$$

de forma que as matrizes do sistema serão  $T_t = I_k$ ,  $Z_t = X_t$ ,  $G_t = \sigma_\xi e_1'$ , onde  $e_1' = (1 \ 0 \ \dots)$ , e  $H_t = 0$ .

Coefficientes de regressão variantes no tempo podem ser introduzidos estabelecendo  $H_t \neq 0$  para  $t = 1, \dots, n$ .

A rotina **GetSsfReg** fornece a estrutura de espaço de estados variante no tempo para o modelo de regressão univariado:

GetSsfReg(mXt, &mPhi, &mOmega, &mSigma, &mJ\_Phi);

**Simulando a partir de modelos de espaço de estados**

Para gerar amostras de distribuições incondicionais implicadas em um modelo estatístico na forma de espaço de estados, se usa a forma apresentada em (4) como um conjunto de equações recursivas. Valores para  $\alpha_{t+1}^{(i)}$  e  $y_t^{(i)}$  na repetição ( $i$ ) podem ser gerados recursivamente a partir de números aleatórios com distribuição normal padrão

$\varepsilon_t^{(i)}$  usando  $\begin{pmatrix} H_t \\ G_t \end{pmatrix} \varepsilon_t^{(i)} = u_t^{(i)}$  e:

$$\begin{cases} \alpha_{t+1}^{(i)} \\ y_t^{(i)} \end{cases} = \delta_t + \Phi_t \alpha_t^{(i)} + u_t^{(i)} \quad t = 1, \dots, n \quad (6.8)$$

com as condições iniciais  $\alpha_1^{(i)} = a + Qu_0^{(i)}$  onde  $u_0^{(i)}$  é um vetor de números aleatórios com distribuição normal padrão e  $Q$  é tal que  $P = QQ'$ .

A rotina **SsfRecursion** implementa a repetição (9) para uma dada amostra de  $u_t^{(i)}$  ( $t = 1, \dots, n$ ).

mD = SsfRecursion(mR, {Ssf});

onde mR representa a matriz de dados com estrutura  $mR = \{u_0^{(i)} \quad u_1^{(i)} \quad \dots \quad u_n^{(i)}\}$  e {Ssf} representa uma das formas de representação a seguir.

mPhi, mOmega

mPhi, mOmega, mSigma

mPhi, mOmega, mSigma, mDelta

mPhi, mOmega, mSigma, mDelta, mJ\_Phi, mJ\_Omega, mJ\_Delta, mXt

A função SsfRecursion retorna a matriz

$$mD = \begin{Bmatrix} \alpha_1^{(i)} & \alpha_2^{(i)} & \dots & \alpha_{n+1}^{(i)} \\ 0 & y_1^{(i)} & \dots & y_n^{(i)} \end{Bmatrix} \quad (6.9)$$

## Filtro de Kalman

O filtro de Kalman é um algoritmo recursivo para a avaliação de momentos da distribuição normal do vetor de estados  $\alpha_{t+1}$  condicional aos dados  $Y_t = (y_1, \dots, y_t)$ , isto é,

$$a_{t+1} = E(\alpha_{t+1} | Y_t), \quad P_{t+1} = \text{cov}(\alpha_{t+1} | Y_t)$$

O filtro de Kalman é dado por

$$\begin{aligned} v_t &= y_t - c_t - Z_t a_t \\ F_t &= Z_t P_t Z_t' + G_t G_t' \\ K_t &= (T_t P_t Z_t' + H_t G_t') F_t^{-1} \\ a_{t+1} &= d_t + T_t a_t + K_t v_t \\ P_{t+1} &= T_t P_t T_t' + H_t H_t' - K_t F_t K_t' \end{aligned} \tag{6.10}$$

onde  $a_1 = a$  e  $P_1 = P_* + \kappa P_\infty$ .

Quando valores ausentes estão presentes, o filtro de Kalman no tempo  $t$  é baseado em  $y_t^*$  em vez de  $y_t$ .

A rotina **KalmanFil** executa o filtro de Kalman e retorna  $v_t$ ,  $F_t$  e  $K_t$

```
mKF = KalmanFil(mYt, {Ssf});
```

## Suavizador de Momentos

O filtro de Kalman é uma repetição progressiva (*forward*) que avalia estimadores um passo a frente. O algoritmo de suavização de momentos associado é uma repetição regressiva (*backward*) que avalia a média e a variância de uma distribuição condicional específica dado o conjunto de dados  $Y_n = (y_1, \dots, y_n)$  usando o resultado do filtro de Kalman. As repetições regressivas são dadas por

$$\begin{aligned}
 e_t &= F_t^{-1}v_t - K_t' r_t \\
 D_t &= F_t^{-1} + K_t' N_t K_t \\
 r_{t-1} &= Z_t' F_t^{-1} v_t + L_t' r_t \\
 N_{t-1} &= Z_t' F_t^{-1} Z_t + L_t' N_t L_t
 \end{aligned}
 \tag{6.11}$$

com  $L_t = T_t - K_t Z_t$  e com a inicialização  $r_n = 0$  e  $N_n = 0$  para  $t = n, \dots, 1$ .

A rotina **KalmanSmo** implementa o suavizador de momentos e retorna  $e_t$ ,  $D_t$ ,  $r_{t-1}$  e  $N_{t-1}$

```
mKS = KalmanSmo(mKF,{Ssf});
```

Para a estimação de máxima verossimilhança é usada a rotina **SsfLikConc** que retorna o estimador de variância e a verossimilhança concentrada:

$$\begin{aligned}
 \tilde{\sigma}^2 &= \frac{1}{Nn-d} \sum_{t=1}^n v_t' (F_t^c)^{-1} v_t \\
 l_c &= -\frac{nN}{2} \log(2\pi) - \frac{nN-d}{2} (\log \tilde{\sigma}^2 + 1) - \frac{1}{2} \sum_{t=1}^n \log |F_t^c|
 \end{aligned}
 \tag{6.12}$$

```
SsfLikConc(&dLogLikConc, &dVar,mYt,{Ssf});
```

### Previsão e Suavizamento

O filtro de Kalman produz a previsão um passo a frente do vetor de estados, isto é a esperança condicional  $E(\alpha_t | Y_{t-1})$ , denotada por  $a_t$ , juntamente com a matriz de variâncias  $P_t$ , para  $t = 1, \dots, n$ .

```
mState = SsfMomentEst(&mPred,mYt,{Ssf});
```

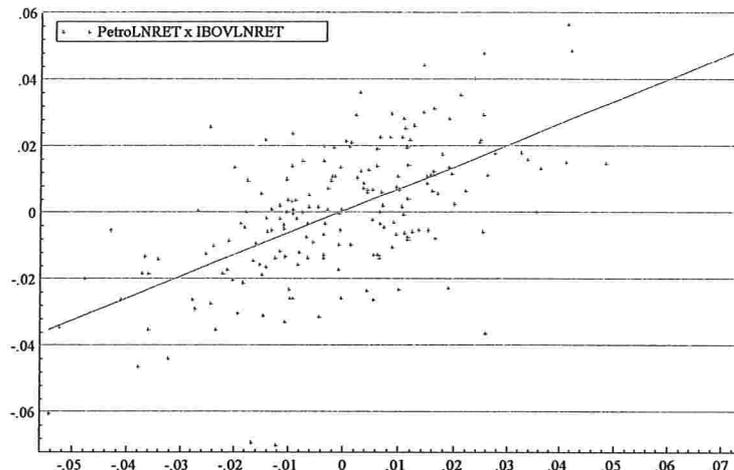
onde a estrutura das matrizes mState e mPred resultantes são dadas por

$$mState = \begin{pmatrix} P_{n+1} \\ a_{n+1}' \end{pmatrix}, \quad mPred = \begin{pmatrix} a_1 & \dots & a_n \\ \hat{y}_1 & \dots & \hat{y}_n \\ \text{diag}(P_1) & \dots & \text{diag}(P_n) \\ \text{diag}(F_1) & \dots & \text{diag}(F_n) \end{pmatrix}
 \tag{6.13}$$

Dadas as funções necessárias foi construído um programa no Ox<sup>®</sup> utilizando-as para a representação em espaço de estados, filtragem, previsão e suavizamento.

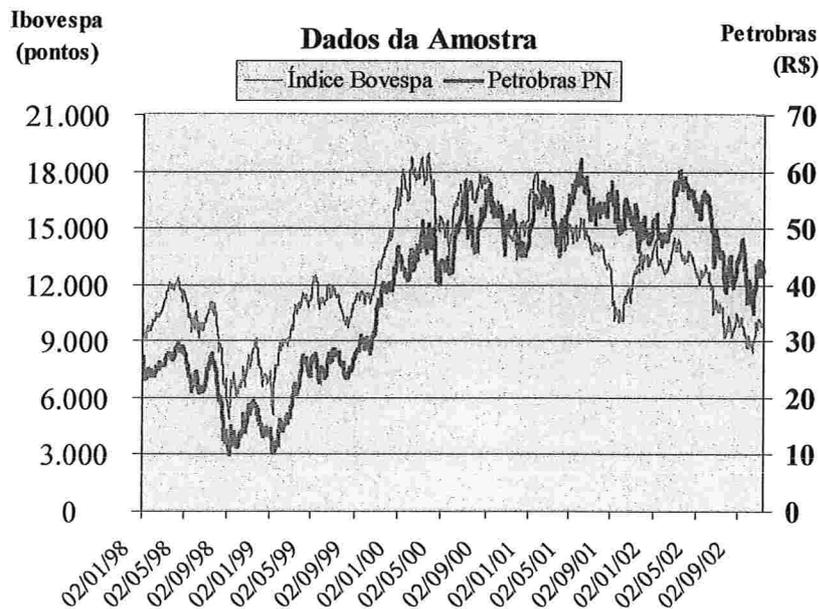
Foram utilizados os dados de Petrobrás PN e do Índice Bovespa para o período entre Janeiro de 1998 a Novembro de 2002.

O gráfico a seguir mostra a relação entre o log do retorno de Ibovespa (eixo x) versus o log do retorno de Petrobrás PN (eixo y).



Para avaliar os dados foram introduzidos valores ausentes aleatórios na série de dados real. Foram rodadas 20 amostras com 60 valores ausentes. Foram estimados tanto os retornos, quanto os preços da Petrobrás PN.

Os valores originais são exibidos abaixo.



## MODELOS CAPM COM OBSERVAÇÕES FALTANDO

A seguir são apresentadas estatísticas descritivas para as séries originais de preços e retornos de Petrobras, bem como para as amostras com missing values.

### Amostras desde Novembro de 2000 com 60 valores ausentes

#### Estatísticas Descritivas

A seguir, apresentamos histogramas contendo as estatísticas descritivas das amostras e seus valores originais.

Preço/ Retorno	Amostra	Média	Mediana	Desvio Padrão	Assimetria	Curtose
<b>Retornos</b>	1	0.00001	0.00090	0.01895	(0.84591)	4.43586
	2	(0.00270)	(0.00271)	0.02004	(0.16534)	3.79264
	3	0.00038	0.00173	0.02319	(0.29648)	3.74339
	4	(0.00042)	(0.00019)	0.02129	(0.49518)	4.61871
	5	0.00253	0.00282	0.02071	(0.26242)	3.13779
	6	0.00346	0.00189	0.01850	0.05369	3.15746
	7	0.00044	0.00246	0.02200	(0.74385)	4.06463
	8	(0.00445)	(0.00283)	0.01978	(0.64554)	4.36969
	9	0.00308	0.00268	0.01914	(0.05069)	2.71578
	10	(0.00083)	0	0.02043	(0.30012)	4.03700
		<b>Petrobras PN</b>	<b>0.00019</b>	<b>0.00031</b>	<b>0.02070</b>	<b>(0.29575)</b>
<b>Preços</b>	1	52.69	52.99	4.174	(0.20901)	2.354508
	2	53.07	53.50	3.742	(0.22404)	2.632093
	3	52.79	53.27	3.845	(0.15132)	2.453204
	4	52.36	52.43	4.108	0.00261	2.334497
	5	52.61	53.00	3.743	(0.23817)	2.405075
	6	52.98	53.40	3.926	(0.25085)	2.512177
	7	52.49	52.93	4.233	(0.14936)	2.110008
	8	52.61	53.10	3.936	(0.32174)	2.427536
	9	52.76	53.40	4.106	(0.25438)	2.451835
	10	52.63	53.00	4.017	(0.17872)	2.471614
		<b>Petrobras PN</b>	<b>52.75</b>	<b>53.27</b>	<b>3.99900</b>	<b>(0.20976)</b>

Percebe-se claramente as alterações na média, mediana e demais estatísticas quando são introduzidos valores ausentes. Os valores de Curtose para os Retornos são, em sua maioria, superiores aos da normal. Já a assimetria aponta valores negativos e próximos de zero, mostrando uma ligeira concentração abaixo da média.

As autocorrelações exibidas a seguir indicam que ao desconsiderarmos os valores ausentes, há uma alteração na estrutura de autocorrelação da série.

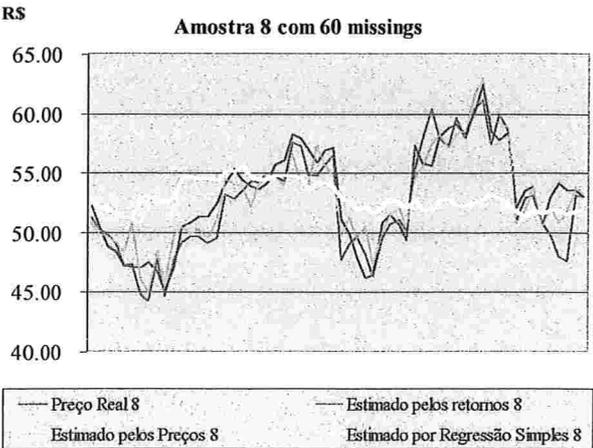
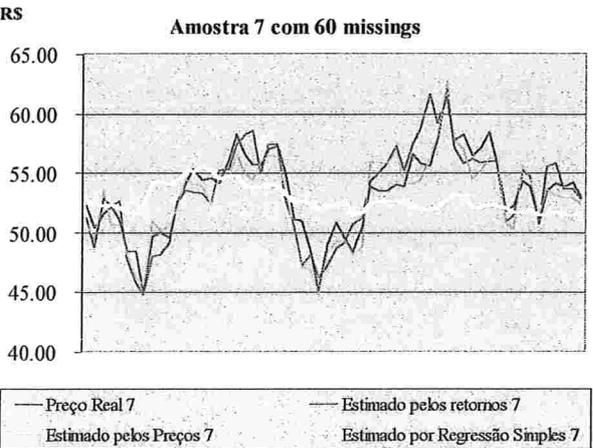
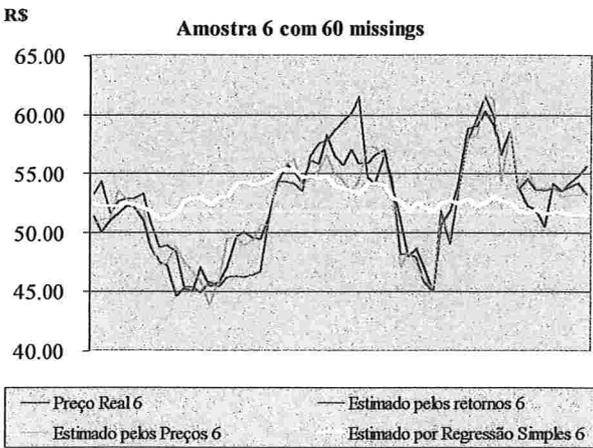
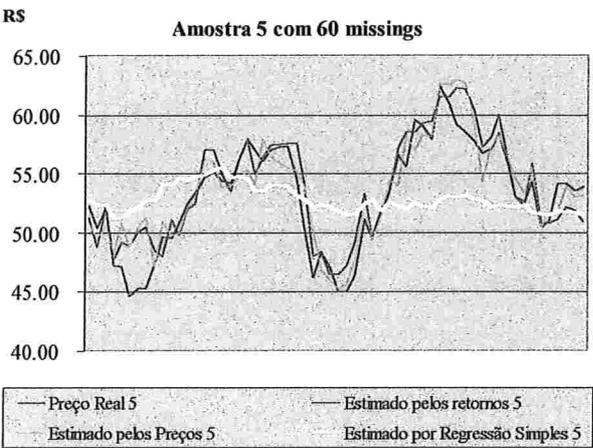
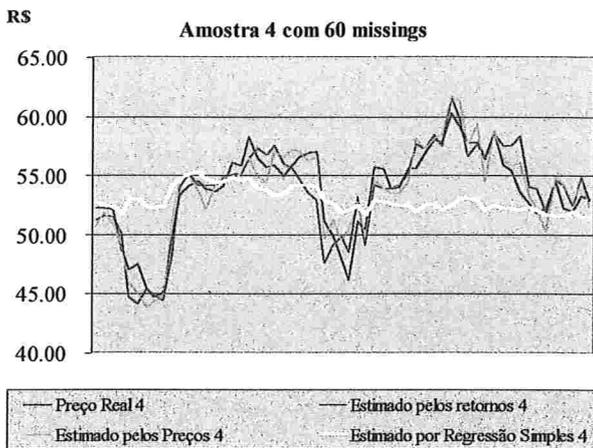
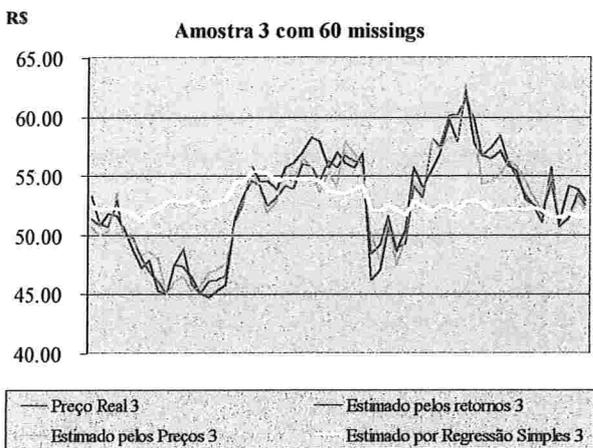
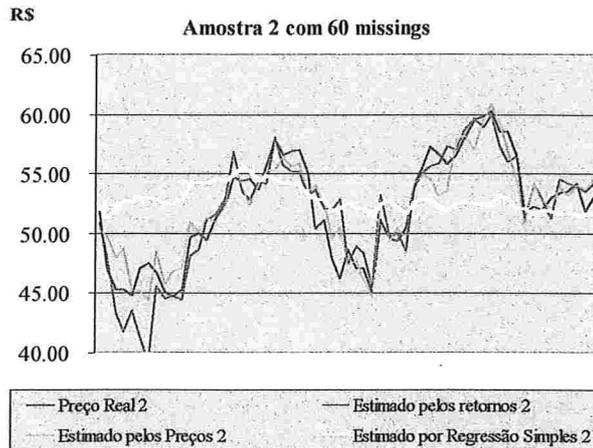
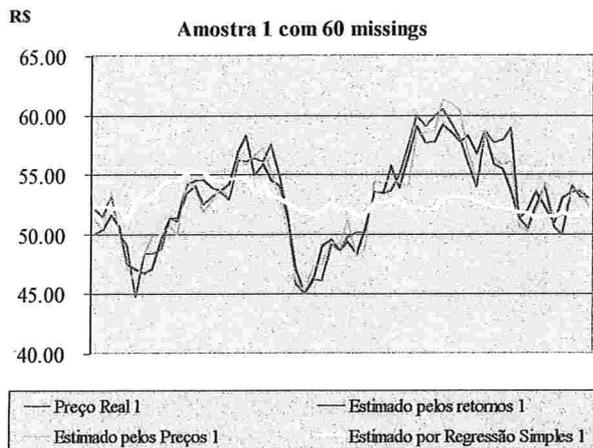
Preço/Retorno	Amostra	Autocorrelação			
		Lag 1	Lag 5	Lag 10	Lag 20
Retornos	1	(0.024)	0.024	0.106	(0.107)
	2	0.173	(0.009)	(0.068)	(0.125)
	3	0.074	0.069	(0.116)	(0.059)
	4	0.131	(0.041)	0.004	0.073
	5	0.039	(0.019)	0.134	(0.012)
	6	0.182	(0.036)	0.017	(0.040)
	7	(0.032)	(0.135)	0.031	(0.067)
	8	0.087	0.043	0.110	(0.074)
	9	0.174	(0.101)	(0.038)	(0.045)
	10	0.169	(0.056)	(0.066)	(0.070)
	<b>Petrobras PN</b>	<b>0.159</b>	<b>(0.048)</b>	<b>0.053</b>	<b>(0.092)</b>
Preços	1	0.756	0.533	0.461	0.100
	2	0.559	0.457	0.405	0.093
	3	0.572	0.459	0.362	0.080
	4	0.679	0.541	0.453	0.109
	5	0.645	0.524	0.455	0.149
	6	0.599	0.525	0.424	0.046
	7	0.714	0.651	0.517	0.088
	8	0.632	0.460	0.418	0.087
	9	0.693	0.525	0.415	0.057
	10	0.607	0.575	0.472	0.089
	<b>Petrobras PN</b>	<b>0.964</b>	<b>0.804</b>	<b>0.631</b>	<b>0.110</b>

Os dados exibidos nos gráficos a seguir, correspondem aos preços reais somente nos momentos onde foram introduzidos os valores ausentes em cada amostra e não são, portanto, seqüenciais. Além dos preços reais, estão exibidos os valores estimados via regressão com parâmetros variantes no tempo a partir da série de retornos e de preços, bem como os estimados por regressão linear com parâmetro fixo (regressão simples).

Esta série consiste no período entre novembro de 2000 a agosto de 2001.

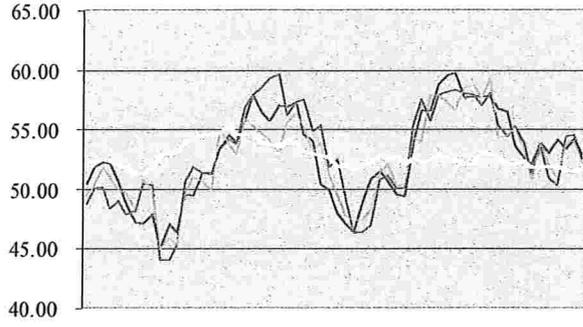
Nota-se claramente a melhor aderência do modelo com parâmetros variando no tempo.

## MODELOS CAPM COM OBSERVAÇÕES FALTANDO



RS

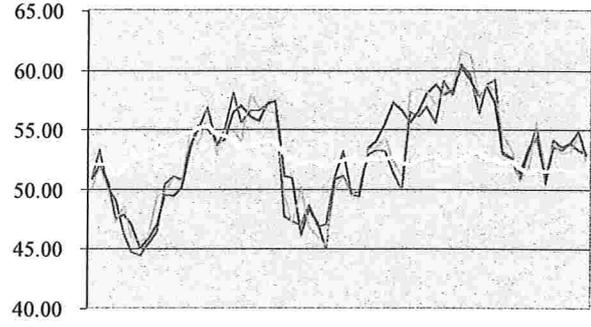
Amostra 9 com 60 missings



— Preço Real 9  
— Estimado pelos retornos 9  
— Estimado pelos Preços 9  
— Estimado por Regressão Simples 9

RS

Amostra 10 com 60 missings



— Preço Real 10  
— Estimado pelos retornos 10  
— Estimado pelos Preços 10  
— Estimado por Regressão Simples 10

Amostra desde Janeiro/1998 com 60 valores ausentes

Estatísticas Descritivas

A seguir, apresentamos histogramas contendo as estatísticas descritivas das amostras e seus valores originais.

Preços/ Retornos	Amostra	Média	Mediana	Desvio Padrão	Assimetria	Curtose
<b>Retornos</b>	1	0.00039	0	0.03054	0.00669	7.53169
	2	0.00033	0	0.03043	0.05092	7.54359
	3	0.00009	0	0.02987	(0.13799)	7.28645
	4	0.00051	0	0.03077	0.03847	7.53787
	5	0.00007	0	0.03064	(0.08246)	7.16302
	6	0.00026	0	0.03044	(0.02030)	7.40547
	7	0.00064	0	0.02914	0.43913	6.08440
	8	0.00006	(0.00005)	0.03057	(0.04369)	7.31639
	9	(0.00017)	(0.00018)	0.03051	0.02083	7.64333
	10	0.00026	0	0.03071	0.05257	7.56368
	<b>Petrobras PN</b>	<b>0.00038</b>	<b>0</b>	<b>0.03037</b>	<b>0.03238</b>	<b>7.38439</b>
<b>Preços</b>	1	39.22	43.90	14.16	(0.427218)	1.78356
	2	38.95	43.61	14.293	(0.392986)	1.74104
	3	39.26	43.93	14.221	(0.434126)	1.79466
	4	39.24	43.90	14.238	(0.426872)	1.78691
	5	39.08	43.75	14.29	(0.408935)	1.76403
	6	39.16	43.80	14.205	(0.421503)	1.77742
	7	39.31	43.93	14.148	(0.422672)	1.77743
	8	39.22	43.90	14.245	(0.428177)	1.78228
	9	39.34	44.10	14.26	(0.437002)	1.78475
	10	39.16	43.95	14.247	(0.421426)	1.77407
	<b>Petrobras PN</b>	<b>39.23</b>	<b>43.90</b>	<b>14.21200</b>	<b>(0.42615)</b>	<b>1.78236</b>

Percebe-se claramente as alterações na média, mediana e demais estatísticas quando são introduzidos valores ausentes. Os valores de Curtose para os Retornos também são superiores aos da normal. Já a assimetria aponta valores negativos e próximos de zero, mostrando uma ligeira concentração abaixo da média. Note que a média desta amostra é inferior à média da primeira amostra, pois esta compreende um período maior que englobou crises econômicas que afetaram as cotações das séries do estudo.

As autocorrelações exibidas a seguir indicam que ao desconsiderarmos os valores ausentes, há uma alteração na estrutura de autocorrelação da série.

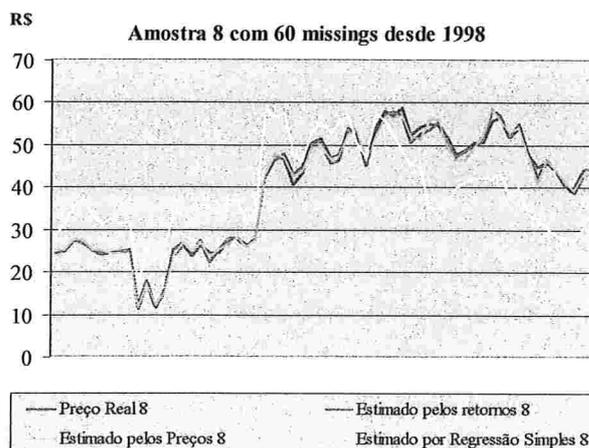
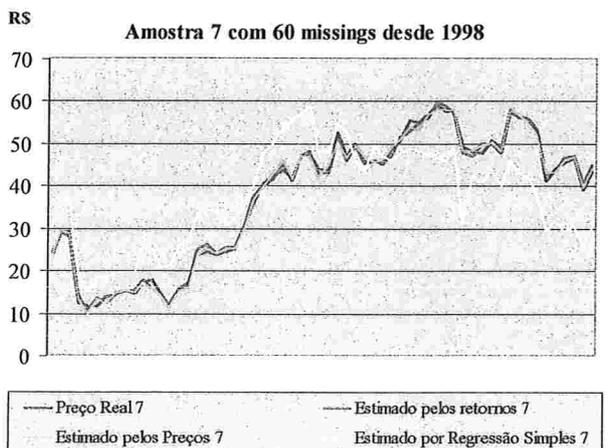
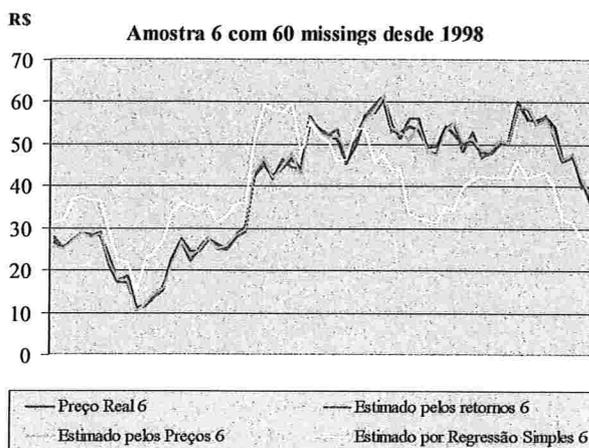
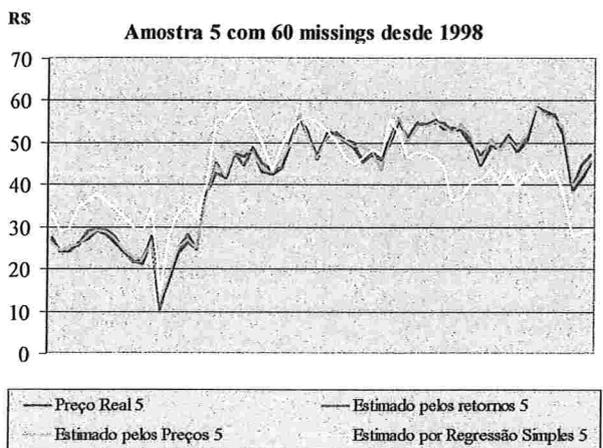
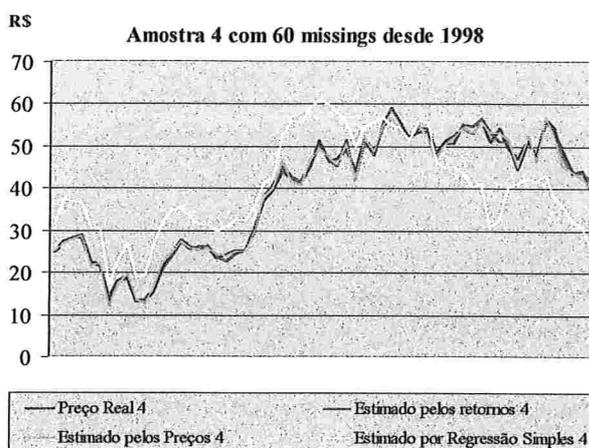
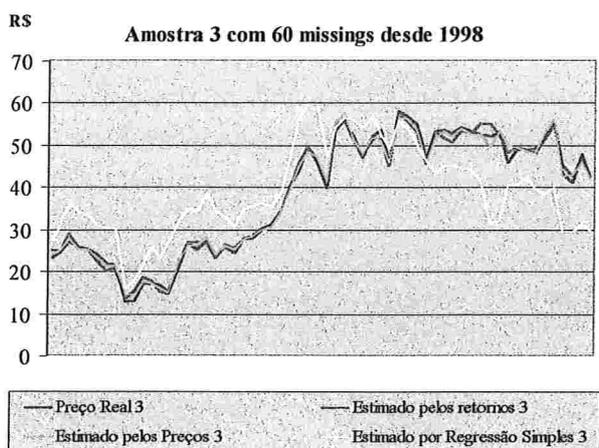
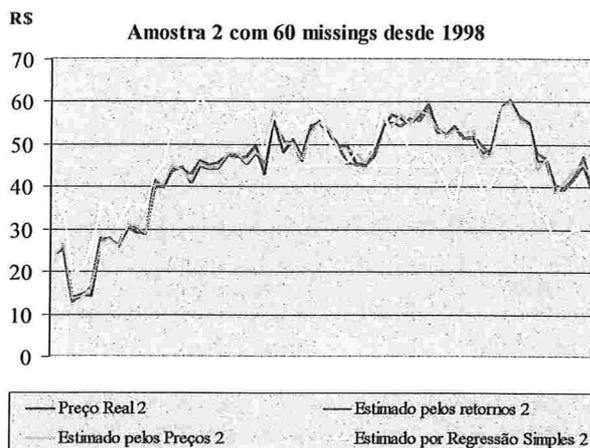
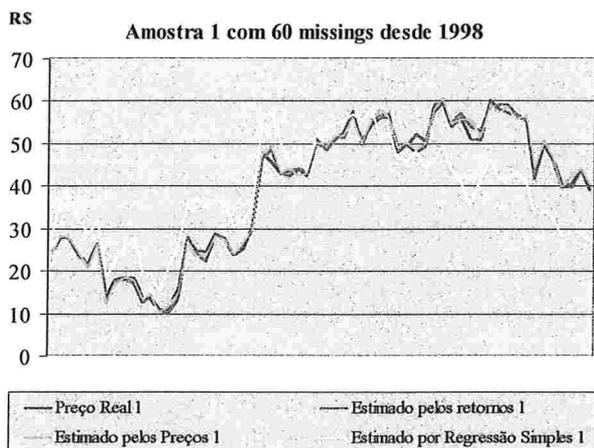
Preços/Retornos	Amostra	Autocorrelação			
		Lag 1	Lag 5	Lag 10	Lag 20
<b>Retornos</b>	1	0.127	0.016	0.039	0.020
	2	0.134	(0.011)	0.039	0.033
	3	0.107	0.019	0.041	0.011
	4	0.114	(0.002)	0.032	(0.005)
	5	0.114	(0.003)	0.066	(0.017)
	6	0.099	(0.006)	0.039	0.013
	7	0.121	(0.054)	0.009	0.012
	8	0.120	(0.007)	0.044	0.007
	9	0.136	0.004	0.036	0.000
	10	0.092	(0.002)	0.027	0.005
	<b>Petrobras PN</b>	<b>0.119</b>	<b>(0.016)</b>	<b>0.042</b>	<b>0.010</b>
<b>Preços</b>	1	0.938	0.926	0.915	0.891
	2	0.958	0.949	0.933	0.904
	3	0.948	0.940	0.926	0.897
	4	0.951	0.937	0.927	0.895
	5	0.958	0.949	0.933	0.903
	6	0.945	0.937	0.920	0.895
	7	0.939	0.930	0.912	0.880
	8	0.953	0.937	0.927	0.896
	9	0.954	0.941	0.928	0.898
	10	0.949	0.938	0.927	0.894
	<b>Petrobras PN</b>	<b>0.997</b>	<b>0.985</b>	<b>0.971</b>	<b>0.939</b>

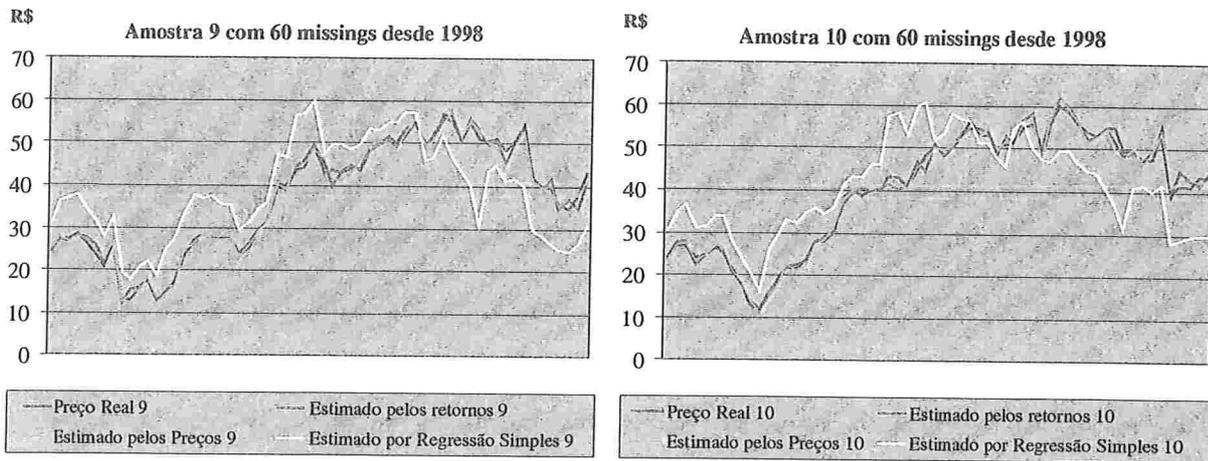
Esta série consiste no período entre janeiro de 1998 a novembro de 2002.

Observe que o eixo vertical aponta uma maior variação do preço de Petrobrás PN. Isto ocorre porque a série analisada compreende um maior período de tempo e englobou crises econômicas que influenciaram fortemente as cotações na bolsa de valores.

Nota-se novamente a melhor aderência do modelo com parâmetros variando no tempo.

## MODELOS CAPM COM OBSERVAÇÕES FALTANDO





Apenas a exibição dos gráficos já aponta uma superioridade do modelo apresentado, mas para uma comparação eficiente, foi avaliado o erro de estimação em todos os casos.

A seguir apresentamos as tabelas com os erros das estimativas, calculados por:

$$\sqrt{\frac{\sum_{i=1}^k (o_i - p_i)^2}{k}}$$

onde:  $k$  é o número de valores ausentes,  $o_i$  é o valor observado  $i$ ,  $p_i$  é o valor estimado  $i$

Tabela I: Erros das estimativas para a amostra com 60 valores ausentes.

Amostra	Estimado por		
	Retorno	Preço	Regressão
1	0.2023	0.1769	0.4437
2	0.2750	0.2194	0.5620
3	0.1682	0.2006	0.5272
4	0.2075	0.1761	0.4556
5	0.2474	0.2611	0.5496
6	0.2486	0.1939	0.5053
7	0.2198	0.1814	0.4269
8	0.2424	0.2086	0.5097
9	0.2446	0.2048	0.4683
10	0.2178	0.2108	0.5027

**Tabela II:** Erros das estimativas para a amostra com 60 valores ausentes desde 1998.

Amostra	Estimado por		
	Retorno	Preço	Regressão
1	0.1622	0.1528	1.1857
2	0.1601	0.1622	1.2011
3	0.1586	0.1420	1.1823
4	0.1531	0.1470	1.2137
5	0.1386	0.1363	1.1176
6	0.1742	0.1550	1.4147
7	0.1366	0.1310	1.1703
8	0.1287	0.1506	1.3007
9	0.1680	0.1615	1.1172
10	0.1797	0.1742	1.1900

Observe que a série de retornos contém mais valores ausentes que a série de preços, uma vez que quando há um valor ausente nos preços no instante  $t$ , este pode gerar dois valores ausentes no retorno nos instantes  $t$  e  $t+1$ . Logo, é de se esperar que os erros estimação a partir da série de preços sejam inferiores ao da série de retornos por esta última conter maior número de valores ausentes, mas o que se observa é que algumas vezes a estimação baseada nas séries de retorno têm erro ligeiramente inferior ao da série de preços, apesar do contrário ocorrer na maioria dos casos.

## Conclusão

Verificando os gráficos e os erros de estimação, nota-se que o modelo com parâmetros variando no tempo é claramente superior ao modelo de regressão com parâmetro fixo.

Este último é ainda menos eficiente quando o período engloba grandes oscilações, como na amostra desde 1998 que engloba crises econômicas que influenciaram a cotação das ações. Isto é comprovado pelo maior erro de estimação (tabela III).

Portanto, é verificado que o ajuste do parâmetro de regressão no tempo permite uma melhor aderência do modelo, mesmo quando existem grandes oscilações na série.

**O programa feito no Ox® é dado a seguir:**

```

#include <oxstd.h> // cabeçalho que inclui a biblioteca padrão Ox
#include <oxdraw.h>
#include <oxfloat.h>
#include <packages/ssfpack/ssfpack.h>
CalcReg(const mx, my,ct,ck);

CalcReg (const mx, my,ct,ck)
{
    decl mphi, momega, msigma, mj_phi, mj_omega, dlik, dvar;
    decl mstate, mpred, vse, mols, mkf, mks, mstab;
    GetSsfReg(mx, &mphi, &momega, &msigma,&mj_phi); // regression in state space
    mj_omega=<0.1,-1,-1;-1,0.5,-1;-1,-1,0.5>;
    SsfLikConc(&dlik, &dvar, my, mphi, momega, msigma, <>,mj_phi, mj_omega, <>, mx);
// regression
    momega *= dvar; //insere a variância na matriz Omega
    mstate = SsfMomentEst(ST_PRED, &mpred, my, mphi, momega, msigma, <>, mj_phi, mj_omega, <>,
mx);
    vse = sqrt(diagonal(mstate[0:ck-1][])); //tira raiz da diagonal da matriz de variancias
    P mols = mstate[ck][ ] | vse | fabs(mstate[ck][ ] ./ vse);
    print ("%r",{"a"}, mols[0][ ]);
// stability tests
    mkf = KalmanFil(my, mphi, momega, msigma, <>, mj_phi, mj_omega, <>, mx);
    mks = KalmanSmo(mkf, mphi, momega, msigma, <>, mj_phi, mj_omega, <>, mx);
    decl msmodist = mks[0:2][0] ~ momega * mks[0:2][1:];
    decl msmostat = SsfRecursion(msmodist, mphi, momega,msigma,<>, mj_phi, mj_omega,<>,<>);
    return(mpred[2][1:]');
}
main()
{
    decl my, my1, my2, my3, mx, i,j, resultado, ct, ck;
    my = loadmat("C:\pvalls\petrodata4.in7"); // load petrobras data
    mx = 1 | my[20][ ]; // X' é o retorno do IBovespa
    ck = rows(mx); // numero de regressores
    println ("Num de Regressores_ck: ", ck);
    my1 = my[43][ ]; // Y' é o retorno da Petrobras com
missings
    ct = columns(my1); // número de observations
    println ("Num de observacoes_ct: ", ct);
    my2 = my1;
    resultado = CalcReg(mx[ ][ ],my2[ ],ct, ck);
    for (i=0;i<=rows(mx')-2;+i)
    {
        if(my2[i]== .NaN)
        {
            my2[i]=resultado[i][ ];
            continue;
        }
    }
    println ("Matriz Yt","%c",{"RealPreco", " MissingPreco ", " EstimadoPreco ",
"CalcRegPreco"},my[22][ ]'~my1'~my2'~resultado);
}

```

## Referências Bibliográficas

- [1] Harvey, Andrew C. (1993), TIME SERIES MODELS, Harvester Wheatsheaf.
- [2] Koopman, S. J., Shephard, N. e Doornik, J.A. (1999), STATISTICAL ALGORITHMS FOR MODELS IN STATE SPACE USING SSFPACK 2.2, Econometrics Journal, volume 2, 113-166.
- [3] Doornik, Jurgen A. (2001), OX WORKSHOP, apresentado na 9ª Escola de Séries Temporais e Econometria, Belo Horizonte, Agosto de 2001.
- [4] Securato, José Roberto (1996), DECISÕES FINANCERIAS EM CONDIÇÕES DE RISCO, Editora Atlas
- [5] Alexander, Gordon J. e Francis, Jack Clark (1986), PORTFOLIO ANALYSIS, Prentice-Hall.
- [6] Durbin, J. e Koopman, S.J. (2001), TIME SERIES ANALYSIS BY STATE SPACE METHODS, Oxford University Press.