

## 4

### Modelo Poisson bivariado

#### 4.1

##### Introdução

Para a análise de fenômenos discretos no tempo, como por exemplo, o número de vendas de um certo produto num intervalo de tempo, placares de jogos de futebol, ou o número de mortos ou feridos graves em acidentes de trânsito, a distribuição de Poisson é usada com frequência.

Anteriormente foi apresentada a forma de implementação de modelos em espaço de estado lineares para observações provenientes da distribuição de Poisson univariada. Um exemplo de aplicação deste tipo de modelo está descrito em Durbin e Koopman[7], que trabalha com o número de motoristas mortos em acidentes de trânsito.

O uso da distribuição Poisson multivariada é menos comum. Um dos poucos trabalhos nesta área é o de Fernandes[12], que no caso bivariado, considera a modelagem da soma das séries individuais em cada período de tempo. A divisão nas séries individuais é determinada por uma distribuição binomial; este modelo é brevemente apresentado no apêndice, e segue a mesma linha dos modelos introduzidos por Harvey & Fernandes[18], os quais correspondem a uma classe de modelos que utilizam distribuições condicionais na equação das observações. Uma característica importante do modelo proposto por Harvey & Fernandes é que o mesmo admite a existência de correlações negativas, o que não é permitido fazer com o modelo proposto nesta tese.

#### 4.2

##### Modelo Poisson Bivariado

Neste trabalho propõe-se um modelo de espaço de estado para observações com distribuição Poisson bivariada. O modelo proposto pode ser

visto como a versão bivariada do modelo Poisson univariado. Apresentamos também uma metodologia para estimar o modelo considerando a abordagem de DK.

A distribuição Poisson bivariada utilizada nesta tese é a distribuição definida em Kocherlakota & Kocherlakota [23] e Johnson e Kotz [20], que consideram a seguinte distribuição:

$$Y_1 = U + V_1 \quad \text{e} \quad Y_2 = U + V_2,$$

onde  $U$ ,  $V_1$  e  $V_2$  são variáveis aleatórias independentes com distribuição Poisson de médias  $\lambda$ ,  $\delta_1$  e  $\delta_2$ , respectivamente.

A função de probabilidade conjunta é dada por:

$$P(y_1, y_2 | \delta_1, \delta_2) = e^{-(\lambda + \delta_1 + \delta_2)} \sum_{j=0}^{\min(y_1, y_2)} \frac{\lambda^j}{j!} \frac{\delta_1^{y_1-j}}{(y_1-j)!} \frac{\delta_2^{y_2-j}}{(y_2-j)!}, \quad (4-1)$$

se  $y_1$  ou  $y_2$  é negativa  $P(y_1, y_2) = 0$ .

Pode-se provar que as marginais de  $Y_1$  e  $Y_2$  condicionais a  $\delta_1$  e  $\delta_2$  são Poisson, com médias  $\lambda + \delta_1$  e  $\lambda + \delta_2$ , respectivamente.

A principal limitação do modelo é a exclusão de correlação negativa, devido ao fato de que a covariância entre  $Y_1$  e  $Y_2$  é dada por:

$$\text{cov}(Y_1, Y_2 | \delta_1, \delta_2) = \text{cov}(U + V_1, U + V_2 | \delta_1, \delta_2) = \text{var}(U) = \lambda.$$

Assim, a correlação de Pearson entre  $Y_1$  e  $Y_2$  será:

$$\text{corr}(Y_1, Y_2 | \delta_1, \delta_2) = \lambda(\lambda + \delta_1)^{-\frac{1}{2}}(\lambda + \delta_2)^{-\frac{1}{2}}.$$

Esta correlação é sempre positiva, e não pode exceder:

$$\lambda \{\lambda + \min(\delta_1, \delta_2)\}^{-1}, \quad \text{ou} \quad \min \left[ \frac{(\lambda + \delta_1)^{\frac{1}{2}}}{(\lambda + \delta_2)^{\frac{1}{2}}}, \frac{(\lambda + \delta_2)^{\frac{1}{2}}}{(\lambda + \delta_1)^{\frac{1}{2}}} \right].$$

Um conjunto de equações importantes, que são usadas nesta tese na obtenção do modelo gaussiano aproximado, é dado pelas derivadas da função de probabilidade conjunta com respeito aos parâmetros  $\lambda$ ,  $\delta_1$  e  $\delta_2$ :

$$\frac{\partial P(y_1, y_2)}{\partial \lambda} = P(y_1 - 1, y_2 - 1) - P(y_1, y_2) \quad (4-2)$$

$$\frac{\partial P(y_1, y_2)}{\partial \delta_1} = P(y_1 - 1, y_2) - P(y_1, y_2) \quad (4-3)$$

$$\frac{\partial P(y_1, y_2)}{\partial \delta_2} = P(y_1, y_2 - 1) - P(y_1, y_2). \quad (4-4)$$

No caso em que  $\lambda = 0$  as variáveis são independentes, de modo que a distribuição Poisson bivariada se reduz ao produto de duas variáveis aleatórias univariadas independentes com distribuição de Poisson.

### 4.3

#### Modelo de Espaço Estado Poisson Bivariado

Cargnoni *et al* [4], afirmam que com freqüência os gráficos das marginais de duas séries revelam padrões de comportamento similares. Isto poderia ser explicado pela existência de algum fator comum não-observável, que afeta ambas as séries de maneira similar. Isto permite supor, por exemplo, a existência de uma tendência comum.

Neste contexto, a formulação em espaço de estado do modelo Poisson bivariado proposto na tese é dada por:

$$y_t/\theta_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} \sim PB(\lambda, \delta_{1t}, \delta_{2t}), \quad (4-5)$$

$$\delta_t = \begin{pmatrix} \delta_{1t} \\ \delta_{2t} \end{pmatrix} = \begin{pmatrix} e^{\theta_{1t}} \\ e^{\theta_{2t}} \end{pmatrix},$$

sendo  $\theta_t$  definido por:

$$\theta_t = \begin{pmatrix} \theta_{1t} \\ \theta_{2t} \end{pmatrix} = Z_t \alpha_t = \begin{pmatrix} Z_{1t} \\ Z_{2t} \end{pmatrix} \alpha_t,$$

portanto

$$\delta_t = \begin{pmatrix} \delta_{1t} \\ \delta_{2t} \end{pmatrix} = \begin{pmatrix} e^{Z_{1t} \alpha_t} \\ e^{Z_{2t} \alpha_t} \end{pmatrix},$$

onde  $\lambda$  é a covariância entre as duas séries observadas, que, neste modelo, se supõe constante.

A equação das observações é dada por (4-5), e a equação do estado dada por

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t, \quad \eta_t \sim N(0, Q_t), \quad (4-6)$$

com todas as matrizes com dimensões apropriadas.

Para entender o modelo anterior é conveniente lembrar o modelo de

espaço de estado linear gaussiano bivariado, dado por:

$$\begin{aligned} y_t &= Z_t \alpha_t + \epsilon_t, & \epsilon_t &\sim N(0, H_t) \\ \alpha_{t+1} &= T_t \alpha_t + R_t \eta_t, & \eta_t &\sim N(0, Q_t) \end{aligned}$$

onde  $\epsilon_t$  e  $\eta_t$  são independentes em todo o período de tempo,  $y_t$  é de dimensão  $2 \times 1$ ,  $\alpha_t$  é  $p \times 1$ ,  $Z_t$  é  $2 \times p$  e as outras matrizes e vetores são das dimensões apropriadas.

Como a densidade do modelo Poisson bivariado descrito anteriormente é construída a partir de uma variável aleatória Poisson comum, presente nos dois processos Poisson observados, é possível pensar num modelo Poisson bivariado com componentes comuns. Neste contexto, o comportamento comum nas duas séries é devido ao fato de ambas terem um vetor de estado com fatores comuns, de modo que as diferenças de comportamento entre as duas séries são dadas apenas pelos elementos da matriz  $Z_t$ , que determinam quais componentes do vetor de estado influenciam em cada série de observações e com que peso. Desta forma o modelo proposto pode ser visto como a versão Poisson do modelo Gaussiano bivariado escrito acima, mas agora considerando a existência de fatores comuns.

Da mesma forma que no caso univariado, a partir de uma análise exploratória dos dados em conjunto com outras técnicas, podemos decidir qual modelo estrutural considerar para cada uma das séries em estudo.

A título de ilustração, considere os exemplos do modelo Poisson bivariado proposto seguintes:

### 1. Modelo de Nível Local

O vetor de estado é unidimensional, com equação dada por:

$$\mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2).$$

A equação que define o sinal é:

$$\theta_t = \begin{pmatrix} \theta_{1t} \\ \theta_{2t} \end{pmatrix} = Z_t \alpha_t = \begin{pmatrix} a \\ b \end{pmatrix} \mu_t. \quad (4-7)$$

Não é difícil mostrar que este modelo é não identificável. Com efeito, a verossimilhança do Modelo Gaussiano Aproximador é não identificável de acordo com a definição apresentada em Harvey [17] (pg: 450-451), e  $\frac{p(y_t|\theta_t)}{g(y_t|\theta_t)}$  assumem o mesmo valor para infinitos valores de  $a$  e  $b$ . Logo,

é preciso que se imponha uma restrição nos parâmetros. Seguindo a estratégia de Harvey [17], propomos a seguinte versão restrita do modelo dado em (4-7), fazendo  $a = 1$ :

$$\theta_t = \begin{pmatrix} \theta_{1t} \\ \theta_{2t} \end{pmatrix} = Z_t \alpha_t = \begin{pmatrix} 1 \\ b \end{pmatrix} \mu_t. \quad (4-8)$$

A figura (4.1) apresenta uma simulação com 100 observações deste modelo para  $\lambda = 1$ ,  $\sigma_\eta^2 = 0.1$ ,  $b = 0.9$ .

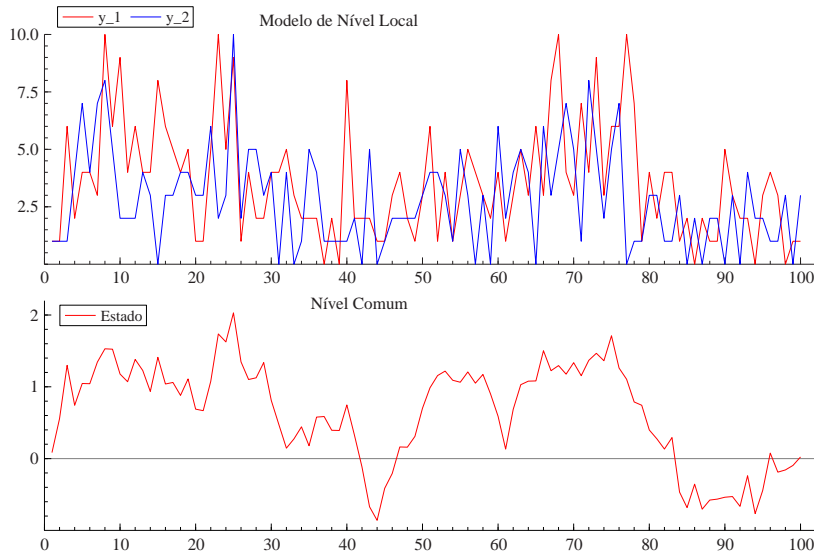


Figura 4.1: Simulação do modelo de nível local Poisson bivariado para  $\lambda = 1$ ,  $\sigma_\eta^2 = 0.1$  e  $b = 0.9$ .

Os gráficos das séries apresentam comportamentos bastante similares e com uma associação positiva entre elas, de acordo com o esperado.

## 2. Modelo Nível e Sazonalidade Locais

Neste caso o vetor de estado é definido por:

$$\begin{aligned} \mu_{t+1} &= \mu_t + \eta_t, & \eta_t &\sim N(0, \sigma_\eta^2) \\ \gamma_{t+1} &= -\sum_{j=1}^{s-1} \gamma_{t+1-j} + \omega_t, & \omega_t &\sim N(0, \sigma_\omega^2), & t &= 1, \dots, n. \end{aligned}$$

Seguindo a mesma idéia da identificação mencionada anteriormente,

é proposta a seguinte equação para o sinal é:

$$\theta_t = \begin{pmatrix} \theta_{1t} \\ \theta_{2t} \end{pmatrix} = Z_t \alpha_t = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ a & b & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \gamma_t \\ \gamma_{t-1} \\ \vdots \\ \gamma_{t-s+2} \end{pmatrix}$$

A figura (4.2) apresenta uma simulação com 100 observações deste modelo para  $\lambda = 1$ ,  $\sigma_\eta^2 = 0.04$ ,  $\sigma_\omega^2 = 0.02$ ,  $s = 4$   $a = 0.8$  e  $b = .9$ . Pode-se ver que os dados apresentam um comportamento de acordo com o esperado.

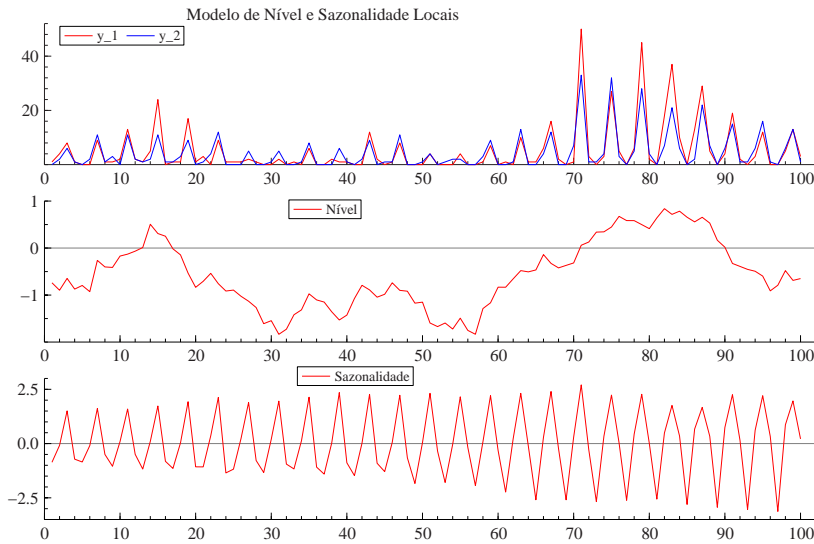


Figura 4.2: Simulação do Modelo Nível e Sazonalidade Locais para  $\lambda = 1$ ,  $\sigma_\eta^2 = 0.04$ ,  $\sigma_\omega^2 = 0.02$ ,  $s = 4$   $a = 0.8$  e  $b = .9$ .

O modelo Poisson bivariado proposto tem a mesma estrutura dos modelos de espaço estado não-gaussianos considerados por DK, de modo que é possível estudar o comportamento das séries temporais com distribuição Poisson bivariada considerando a abordagem de DK.

A forma de obtenção da função de verossimilhança do modelo é similar à considerada por DK para o caso univariado, e é descrita a seguir.

## 4.4 Verossimilhança por Monte Carlo

De forma semelhante ao caso univariado, utiliza-se na construção da verossimilhança por Monte Carlo, a técnica de amostragem por importância, o filtro de Kalman suavizado, a simulação suavizada e variáveis antitéticas por locação,

Seja  $g(\theta|y)$  uma densidade de importância gaussiana.

Sabemos que:

$$L_p(\psi) = L_g(\psi) \int \frac{p(y|\theta)}{g(y|\theta)} g(\theta|y) d\theta = L_g(\psi) E_g \left\{ \frac{p(y|\theta)}{g(y|\theta)} \right\}. \quad (4-8)$$

Assim, para achar a verossimilhança no caso bivariado, procuramos em primeiro lugar a verossimilhança do modelo gaussiano aproximado, e obtemos a verossimilhança não-gaussiana utilizando um fator de ajuste desta.

## 4.5 Seleção do Modelo Gaussiano Aproximador

Nesta subseção, deduzimos as expressões das derivadas analíticas, primeiras e segundas, de  $\log p(y_t|\theta_t)$ , sendo esta última a expressão do logaritmo da função de probabilidade do modelo em espaço de estado Poisson Bivariado que vem sendo discutido. As expressões que encontraremos serão úteis tanto para a nossa metodologia quanto para a metodologia original de DK.

Inicialmente, consideremos a metodologia original a qual se concentra em resolver as seguintes fórmulas:

$$H_t = \left( - \frac{\partial^2 \log p(y_t|\theta_t)}{\partial \theta_t \partial \theta_t'} \Big|_{\theta_t = \tilde{\theta}_t} \right)^{-1},$$

$$c_t = y_t - \tilde{\theta}_t - H_t \frac{\partial \log p(y_t|\theta_t)}{\partial \theta_t} \Big|_{\theta_t = \tilde{\theta}_t},$$

$$y_t^* = y_t - c_t = \tilde{\theta}_t + H_t \frac{\partial \log p(y_t|\theta_t)}{\partial \theta_t} \Big|_{\theta_t = \tilde{\theta}_t}.$$

As derivadas incluídas nas equações acima são obtidas utilizando as equações (4-2) e (4-4). Por clareza de exposição, em algumas passagens será

utilizada a notação  $p(y_{1t}, y_{2t})$  no lugar de  $p(y_t|\theta_t)$ . Na expressão acima, as derivadas tomam a forma:

$$\left. \frac{\partial \log p(y_t|\theta_t)}{\partial \theta_t} \right|_{\theta_t = \bar{\theta}_t} = \frac{1}{p(y_t|\theta_t)} \begin{pmatrix} \frac{\partial p(y_t|\theta_t)}{\partial \theta_{1t}} \\ \frac{\partial p(y_t|\theta_t)}{\partial \theta_{2t}} \end{pmatrix}.$$

Como  $\delta_{1t} = e^{\theta_{1t}}$  e  $\delta_{2t} = e^{\theta_{2t}}$ , então:

$$\frac{\partial p(y_t|\theta_t)}{\partial \theta_{1t}} = \frac{\partial p(y_t|\theta_t)}{\partial \delta_{1t}} \frac{\partial \delta_{1t}}{\partial \theta_{1t}} + \frac{\partial p(y_t|\theta_t)}{\partial \delta_{2t}} \frac{\partial \delta_{2t}}{\partial \theta_{1t}}.$$

Usando as equações (4-2)-(4-4) e o fato que  $\frac{\partial \delta_{2t}}{\partial \theta_{1t}} = 0$ , tem-se:

$$\frac{\partial \log p(y_t|\theta_t)}{\partial \theta_{1t}} = \frac{1}{p(y_t|\theta_t)} [p(y_{1t} - 1, y_{2t}) - p(y_{1t}, y_{2t})] e^{\theta_{1t}}.$$

De forma análoga obtém-se:

$$\frac{\partial \log p(y_t|\theta_t)}{\partial \theta_{2t}} = \frac{1}{p(y_t|\theta_t)} [p(y_{1t}, y_{2t} - 1) - p(y_{1t}, y_{2t})] e^{\theta_{2t}}.$$

Assim

$$\frac{\partial^2 \log p(y_t|\theta_t)}{\partial \theta_{1t}^2} = \frac{\partial}{\partial \theta_{1t}} \left[ \frac{1}{p(y_t|\theta_t)} \{p(y_{1t} - 1, y_{2t}) - p(y_{1t}, y_{2t})\} e^{\theta_{1t}} \right].$$

$$\text{Seja } u = \frac{1}{p(y_t|\theta_t)} \{p(y_{1t} - 1, y_{2t}) - p(y_{1t}, y_{2t})\}.$$

Então

$$\begin{aligned} \frac{\partial^2 \log p(y_t|\theta_t)}{\partial \theta_{1t}^2} &= \frac{\partial u}{\partial \delta_1} \frac{\partial \delta_1}{\partial \theta_1} e^{\theta_{1t}} + \frac{\partial u}{\partial \delta_2} \frac{\partial \delta_2}{\partial \theta_{1t}} e^{\theta_{1t}} + u \frac{\partial e^{\theta_{1t}}}{\partial \theta_{1t}} \\ &= \frac{\partial u}{\partial \delta_1} e^{2\theta_{1t}} + u e^{\theta_{1t}}. \end{aligned}$$

Por outro lado:

$$\frac{\partial u}{\partial \delta_1} = \frac{\partial}{\partial \delta_1} \left\{ \frac{1}{p(y_t|\theta_t)} [p(y_{1t} - 1, y_{2t}) - p(y_{1t}, y_{2t})] \right\}$$

e

$$\begin{aligned} \frac{\partial}{\partial \delta_1} \{p(y_{1t} - 1, y_{2t}) - p(y_{1t}, y_{2t})\} &= p(y_{1t} - 2, y_{2t}) - p(y_{1t} - 1, y_{2t}) - p(y_{1t} - 1, y_{2t}) + \\ &\quad + p(y_{1t}, y_{2t}) \\ &= p(y_{1t} - 2, y_{2t}) - 2p(y_{1t} - 1, y_{2t}) + p(y_{1t}, y_{2t}). \end{aligned}$$



Assim,

$$\frac{\partial u}{\partial \delta_1} = \frac{\{p(y_{1t} - 2, y_{2t}) - 2p(y_{1t} - 1, y_{2t}) + p(y_{1t}, y_{2t})\} p(y_{1t}, y_{2t})}{p(y_t|\theta_t)^2} - \frac{\{p(y_{1t} - 1, y_{2t}) - p(y_{1t}, y_{2t})\}^2}{p(y_t|\theta_t)^2}$$

Simplificando obtém-se:

$$\frac{\partial u}{\partial \delta_1} = \frac{p(y_{1t} - 2, y_{2t})p(y_{1t}, y_{2t}) - p(y_{1t} - 1, y_{2t})^2}{p(y_t|\theta_t)^2}$$

De modo que:

$$\begin{aligned} \frac{\partial^2 \log p(y_t|\theta_t)}{\partial \theta_{1t}^2} &= \left[ \frac{p(y_{1t} - 2, y_{2t})p(y_{1t}, y_{2t}) - p(y_{1t} - 1, y_{2t})^2}{p(y_t|\theta_t)^2} \right] e^{2\theta_{1t}} \\ &+ \frac{\{p(y_{1t} - 1, y_{2t}) - p(y_{1t}, y_{2t})\} e^{\theta_{1t}}}{p(y_{1t}, y_{2t})}. \end{aligned}$$

Analogamente:

$$\begin{aligned} \frac{\partial^2 \log p(y_t|\theta_t)}{\partial \theta_{2t}^2} &= \left[ \frac{p(y_{1t}, y_{2t} - 2)p(y_{1t}, y_{2t}) - p(y_{1t}, y_{2t} - 1)^2}{p(y_t|\theta_t)^2} \right] e^{2\theta_{2t}} \\ &+ \frac{\{p(y_{1t}, y_{2t} - 1) - p(y_{1t}, y_{2t})\} e^{\theta_{2t}}}{p(y_{1t}, y_{2t})}. \end{aligned}$$

e

$$\frac{\partial^2 \log p(y_t|\theta_t)}{\partial \theta_{1t} \partial \theta_{2t}} = \frac{[p(y_{1t}, y_{2t})p(y_{1t} - 1, y_{2t} - 1) - p(y_{1t}, y_{2t} - 1)p(y_{1t} - 1, y_{2t})] e^{\theta_{1t}} e^{\theta_{2t}}}{p(y_{1t}, y_{2t})^2}$$

Assim

$$y_t^* = \tilde{\theta}_t + H_t \frac{1}{p(y_t|\theta_t)} \begin{pmatrix} \{p(y_{1t} - 1, y_{2t}) - p(y_{1t}, y_{2t})\} e^{\theta_{1t}} \\ \{p(y_{1t}, y_{2t} - 1) - p(y_{1t}, y_{2t})\} e^{\theta_{2t}} \end{pmatrix} \quad (4-9)$$

onde

$$H_t = \left( -\frac{1}{p(y_{1t}, y_{2t})^2} \begin{pmatrix} m_{11} + n_{11} & m_{12} \\ m_{12} & m_{22} + n_{22} \end{pmatrix} \right)^{-1}, \quad (4-10)$$

$$m_{11} = \{p(y_{1t} - 2, y_{2t})p(y_{1t}, y_{2t}) - p(y_{1t} - 1, y_{2t})^2\} e^{2\theta_{1t}}$$

$$m_{22} = \{p(y_{1t}, y_{2t} - 2)p(y_{1t}, y_{2t}) - p(y_{1t}, y_{2t} - 1)^2\} e^{2\theta_{2t}}$$

$$m_{12} = \{p(y_{1t}, y_{2t})p(y_{1t} - 1, y_{2t} - 1) - p(y_{1t}, y_{2t} - 1)p(y_{1t} - 1, y_{2t})\} e^{\theta_{1t}} e^{\theta_{2t}}$$

e

$$\begin{aligned} n_{11} &= \frac{1}{p(y_t|\theta_t)} \{p(y_{1t} - 1, y_{2t}) - p(y_{1t}, y_{2t})\} e^{\theta_{1t}} \\ n_{22} &= \frac{1}{p(y_t|\theta_t)} \{p(y_{1t}, y_{2t} - 1) - p(y_{1t}, y_{2t})\} e^{\theta_{2t}} \end{aligned}$$

Para que esta solução seja apropriada, a matriz  $H_t$  deve ser definida-positiva.

Agora passemos para nossa metodologia. Lembrando que o objetivo é a obtenção de matrizes  $H_t$  diagonais (vide expressão (3-15)), o que devemos fazer é  $m_{12} = 0$  na expressão (4-10).

## 4.6 Estimação dos Hiperparâmetros e do Vetor de Estado

Para a estimação dos hiperparâmetros do modelo, denotados por  $\psi$ , se procede da mesma forma que no caso univariado, isto é, maximiza-se a função de verossimilhança com respeito a  $\psi$ . O cálculo da verossimilhança é feito através da equação (4-8). Assim para um  $\psi$  dado, deve-se obter em primeiro lugar o modelo gaussiano aproximado, depois utilizando-se o processo de simulação suavizada calcula-se o ajuste que deve ser feito na verossimilhança gaussiana para se obter a verossimilhança do modelo não-gaussiano, isto é, deve-se obter o valor de  $E_g \left\{ \frac{p(y|\theta)}{g(y|\theta)} \right\}$  na equação (4-8) utilizando-se a simulação suavizada.

Finalmente, considerando alguns critérios de otimização, deve ser decidido se o valor de  $\psi$  é ótimo ou não. Em caso favorável o processo de estimação é finalizado; em caso contrario, é obtido um novo valor para  $\psi$ , utilizando-se algum processo de maximização, o processo é repetido até convergir.

O método de otimização utilizado na implementação do nosso trabalho é baseado no método de Newton que resolve a equação:

$$\partial_1(\psi) = \frac{\partial \log L(y|\psi)}{\partial \psi} = 0, \quad (4-11)$$

usando a série de Taylor de primeiro ordem:

$$\partial_1(\psi) \simeq \tilde{\partial}_1(\psi) + \tilde{\partial}_2(\psi)(\psi - \tilde{\psi}),$$

para algum valor de ensaio  $\tilde{\psi}$ , onde:

$$\tilde{\partial}_1(\psi) = \partial_1(\psi)|_{\psi=\tilde{\psi}}, \quad \tilde{\partial}_2(\psi) = \partial_2(\psi)|_{\psi=\tilde{\psi}},$$

com

$$\partial_2(\psi) = \frac{\partial^2 \log L(y|\psi)}{\partial \psi \partial \psi'}.$$

Fazendo a equação (4-11) igual a zero obtem-se um novo valor para  $\bar{\psi}$  da expressão:

$$\bar{\psi} = \tilde{\psi} - \tilde{\partial}_2(\psi)^{-1} \tilde{\partial}_1(\psi).$$

O processo é repetido até convergir. No caso em que o Hessiano  $\partial_2(\psi)$  é definido-negativo para para todo  $\psi$  a função diz-se côncava e tem um único máximo. O gradiente  $\partial_1(\psi)$  determina a direção do passo ao valor ótimo de  $\psi$  e o Hessiano modifica o tamanho do passo.

Em nosso trabalho o algoritmo de atualização considerado é dado por:

$$\bar{\psi} = \tilde{\psi} + s\tilde{\pi}(\psi),$$

onde

$$\tilde{\pi}(\psi) = -\tilde{\partial}_2(\psi)^{-1} \tilde{\partial}_1(\psi).$$

Existem alguns métodos para achar o valor ótimo de  $s$ , que geralmente está entre 0 e 1. Na prática, não é possível calcular analiticamente  $\partial_1(\psi)$  e  $\partial_2(\psi)$ , de modo que são considerados métodos numéricos na avaliação de  $\partial_1(\psi)$ , e uma aproximação de  $\partial_2(\psi)$ . Por exemplo, o pacote STAMP (*Structural Time Series Analyser, Modeller and Predictor*) de Koopman *et al* [27] e a linguagem Ox de Doornik [5] utilizam o método BFGS (*Broyden-Fletcher-Goldfarb-Shanno*) de aproximação para a matriz do Hessiano. Este método garante que a matriz do Hessiano aproximada é negativa-definida. DK recomendam o livro de Fletcher [13] como uma boa referência do método de otimização de Newton e do BFGS em particular.

Para melhorar a eficiência na estimação dos hiperparâmetros, foram usadas variáveis antitéticas de locação, que são as únicas possíveis de se usar, posto que as antitéticas de escala estão definidas só no caso univariado.

Com respeito à estimação de funções do vetor de estado, note-se em primeiro lugar que esta estimação só pode ser obtida condicional às observações  $y = (y'_1, \dots, y'_n)'$ , a estimação é feita da mesma forma que no

caso univariado, descrito em (3.4), utilizando-se a equação:

$$\hat{x} = \frac{\sum_{t=1}^M x_t w_t}{\sum_{t=1}^M w_t}, \text{ onde } x_t = x(\alpha^{(t)}) \text{ e } w_t = w(\alpha^{(t)}, y). \quad (4-12)$$

Os pesos  $w_t$  são obtidos a partir da equação (3-26):

$$w(\theta, y) = \prod_{t=1}^n \frac{p(y|\theta_t)}{g(\epsilon_t)}.$$

como o estado é linear e gaussiano conforme o modelo proposto na tese.

## 4.7 Simulação

Nesta seção, introduz-se um estudo de simulação com o objetivo de examinar a eficácia do procedimento de estimação. Os programas, tanto os que simulam os processos quanto os que estimam os hiperparâmetros, foram desenvolvidos em Ox 3.0 de Doornik [5], o qual é uma linguagem orientada-objeto com sintaxe similar ao C/C++. A decisão da utilização do Ox foi baseada fundamentalmente na existência dos pacotes SsfPack 2.2 [25] e SsfPack 3.0 beta02 [26], desenvolvidos por Koopman *et al*, que contêm funções para modelagem em forma de espaço de estado para filtragem, alisamento, simulação suavizada, cálculo da verossimilhança, entre outras. No estudo de um problema particular estas funções são usadas dentro de diversos programas em Ox, cuja elaboração geralmente não é um processo simples.

Nos programas construídos para o desenvolvimento desta tese foi utilizado o SsfPack 3.0, fundamentalmente porque considera um tratamento exato para o problema da inicialização, o que não acontece no SsfPack 2.2. Além disto, quase todas as funções do SsfPack 2.0 estão implementadas no SsfPack 3.0, e de acordo com os autores o SsfPack 3.0 possui um tratamento matricial que o faz mais rápido que o SsfPack 2.2.

Lamentavelmente, o processo de estimação dos hiperparâmetros do modelo proposto na tese é um pouco lento, o que torna inviável a execução de muitas simulações. Considerando este problema, foram realizadas simulações para o modelo Poisson bivariado para diferentes modelos estruturais e diferentes valores para os hiperparâmetros. As análises foram replicadas somente para algumas situações particulares que estão descritas adiante. Em todos os casos, as observações seguem um processo  $PB(\lambda, \delta_{1t}, \delta_{2t})$ , dado

$\theta_t$ , com função de probabilidade dada por (4-1). O objetivo das simulações é estudar a habilidade do modelo proposto na estimação dos hiperparâmetros considerados.

A estimação é feita mediante a maximização da verossimilhança do modelo, usando o método de Newton que está implementado no Ox 3.0 no programa BFGS.

O BFGS considera a maximização de funções com respeito aos parâmetros irrestritos, mas como os hiperparâmetros dos diferentes modelos correspondem as variâncias, que são sempre positivas, deve ser feita uma transformação para satisfazer tais restrições ao se utilizar este método para fazer a estimação dos hiperparâmetros.

A transformação considerada foi:

$$\exp(\psi) = \sigma^2,$$

que é sempre positiva.

### Simulações do modelo de Nível Local

Foram simuladas séries de acordo com a seguinte equação de estado:

$$\mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2).$$

A equação que define o sinal é:

$$\theta_t = \begin{pmatrix} \theta_{1t} \\ \theta_{2t} \end{pmatrix} = Z_t \alpha_t = \begin{pmatrix} 1 \\ b \end{pmatrix} \mu_t.$$

Para o modelo acima, foram feitas simulações com diferentes valores de  $\sigma_\eta^2$ ,  $a$ , e  $\lambda$ . A seguir serão apresentados os resultados para  $\sigma_\eta^2 = 0.1$ ,  $b = 0.9$ , e  $\lambda = 1$ , para amostras de diferentes tamanhos, isto é,  $n = 50$  e  $n = 80$ , sendo consideradas 80 replicações em cada caso, considerando o problema de estimar  $\sigma_\eta^2$ .

A tabela [4.1] apresenta os resultados da estimação:

$n$		$\sigma_\eta^2$
80	média	0.14
	d.p	0.09
50	média	0.145
	d.p	0.08

Tabela 4.1: Resultados da Simulação para  $\sigma_\eta^2 = 0.1$ ,  $b = 0.9$  e  $\lambda = 1$ .

O verdadeiro valor do hiperpâmetro é  $\sigma_\eta^2 = 0.1$ , de modo que este é estimado ao parecer com um pequeno sesgo positivo, não existindo uma diferença significativa na estimação do parâmetro de interesse, a medida que mudamos o tamanho da amostra.

É importante destacar aqui a lentidão do processo de estimação, o que parece ser devido as múltiplas contas que requer a implementação do algoritmo de DK.

A título de ilustração do modelo de nível local considere a figura (4.3), a qual corresponde a um gráfico do estado verdadeiro e do estado alisado estimado considerando o modelo com  $N = 50$  observações e com hiperparâmetros dados por  $\sigma_\eta^2 = 0.1$ , com  $b = 0.9$  e  $\lambda = 1$ .

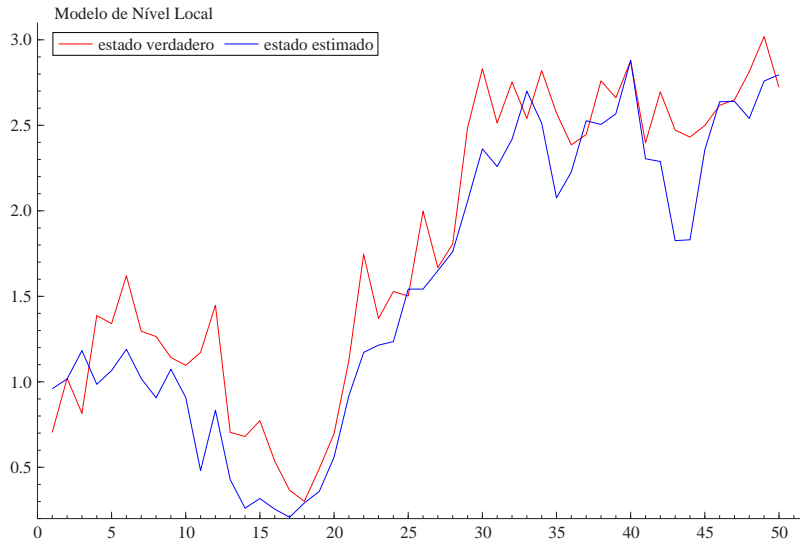


Figura 4.3: Estado estimado para o modelo de nível local Poisson bivariado para  $\lambda = 1$ ,  $\sigma_\eta^2 = 0.1$ ,  $b = 0.9$  e  $N = 50$ .

Da figura (4.3) pode-se observar que o estado estimado acompanha os movimentos do estado verdadeiro.

A figura (4.4) é um gráfico do mesmo tipo mas com  $N = 70$  observações e hiperparâmetro  $\sigma_\eta^2 = 0.05$ , e demais argumentos do modelo sendo  $b = .9$ , e  $\lambda = 1$ . Da figura (4.4) pode-se observar que o estado estimado está bastante próximo do estado verdadeiro.

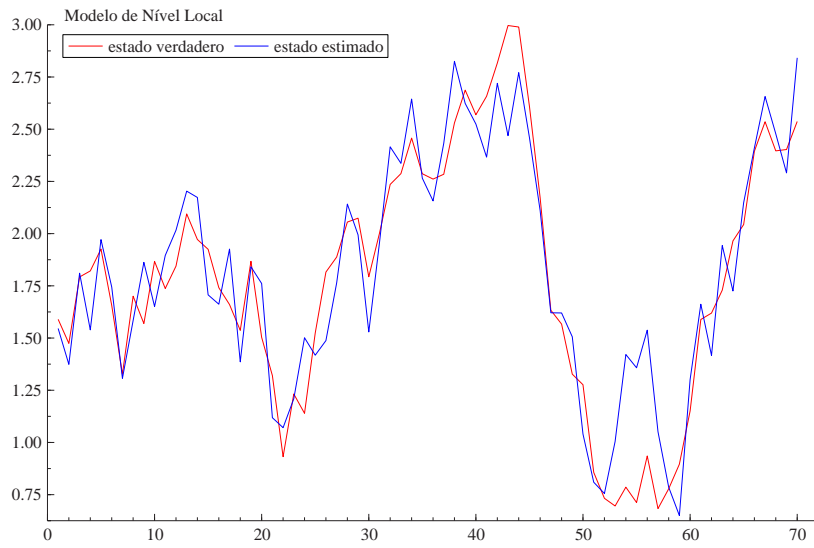


Figura 4.4: Estado Estimado modelo de nivel local Poisson bivariado para  $\lambda = 1$ ,  $\sigma_\eta^2 = 0.05$ ,  $b = 0.9$  e  $N = 70$ .