

3 Metodologia

3.1 Modelos em Espaço de Estado Lineares Gaussianos

3.1.1 Estrutura Básica

A forma em Espaço de Estado Linear Gaussiana (forma em EE daqui por diante) consiste em duas equações. A primeira delas é chamada *equação das observações*, que descreve a evolução de um processo estocástico p -variado e observável y_t , $t = 1, 2, \dots$ e a outra é chamada *equação do estado*. Especificamente:

$$\begin{aligned} y_t &= Z_t \alpha_t + d_t + \varepsilon_t, & \varepsilon_t &\sim N(0, H_t) \\ \alpha_{t+1} &= T_t \alpha_t + R_t \eta_t, & \eta_t &\sim N(0, Q_t) \\ \alpha_1 &\sim N(a_1, P_1). \end{aligned} \quad (3-1)$$

O processo m -variado α_t é chamado de *estado* e é considerado não-observável. Os erros ε_t e η_t são independentes no tempo, entre si e de α_1 . As matrizes do sistema Z_t , d_t , T_t , R_t , H_t e Q_t são determinísticas. Dada uma série temporal de tamanho n do processo y_t , definam-se $Y_j \equiv (y'_1, \dots, y'_j)'$, $a_{t|j} \equiv E(\alpha_t | Y_j)$ e $P_{t|j} \equiv \text{Var}(\alpha_t | Y_j)$. As equações de predição e de suavização do *filtro de Kalman* fornecem fórmulas recursivas para o cálculo dos momentos condicionais acima quando $j = t - 1$ e para $j = n$, respectivamente. Suas expressões analíticas encontram-se em (3-2) e (3-3). Suas derivações, sob os pressupostos da forma em EE aqui adotada, podem ser obtidas em Durbin & Koopman (2001) e em Harvey (1989).

$$\begin{aligned} v_t &= y_t - Z_t a_{t|t-1} - d_t, & F_t &= Z_t P_{t|t-1} Z_t' + H_t, \\ K_t &= T_t P_{t|t-1} Z_t' F_t^{-1}, & L_t &= T_t - K_t Z_t, & t &= 1, \dots, n, \end{aligned} \quad (3-2)$$

$$a_{t+1|t} = T_t a_{t|t-1} + K_t v_t, \quad P_{t+1|t} = T_t P_{t|t-1} L_t' + R_t Q_t R_t'$$

$$\begin{aligned}
 r_{t-1} &= Z_t' F_t^{-1} v_t + L_t' r_t, & N_{t-1} &= Z_t' F_t^{-1} Z_t + L_t' N_t L_t, \\
 a_{t|n} &= a_{t|t-1} + P_{t|t-1} r_{t-1}, & P_{t|n} &= P_{t|t-1} - P_{t|t-1} N_{t-1} P_{t|t-1}, \\
 r_n &= 0, & N_n &= 0, & t &= 1, \dots, n.
 \end{aligned} \tag{3-3}$$

Há diversos artigos publicados na área de estimação de IBNR mediante o arcabouço da forma em EE. O primeiro deles, que pode ser destacado, é o trabalho de de Jong & Zehnwirth (1983), um precursor do uso do filtro de Kalman na literatura atuarial. Nele, o triângulo é organizado de tal forma que as diagonais formam os vetores das observações. Uma variante deste método também é apresentada em Atherino & Fernandes (2007). Em de Jong (2006) ainda se propõe a forma em EE para se estimar correlações entre valores do triângulo. Também cumpre citar o artigo Verrall (1989), que oferece uma metodologia de estimação dentro do enfoque Bayesiano. Outro trabalho de importância é devido a Wright (1990). Greg Taylor também se utiliza da forma em EE para estimar seu modelo em Taylor (2003), ao qual se emprega o filtro EDF (*Exponential Distribution Filter*). Além de artigos, cita-se também o livro de autoria do próprio Taylor (cf. Taylor, 2000), no qual, em adição a um apanhado de técnicas previamente desenvolvidas na literatura, é oferecida uma abordagem em que cada linha do triângulo *runoff* é vista como um vetor aleatório.

3.1.2 Modelos Estruturais

Um *modelo estrutural* para séries temporais é aquele no qual as componentes não observáveis de nível, inclinação, sazonalidade e ruído são modeladas explicitamente, cf. Harvey (1989). O modelo em (3-4) é o modelo estrutural com as componentes periódica e de nível estocásticas que será considerado nas aplicações do presente trabalho.

$$\begin{aligned}
 y_t &= \mu_t + \gamma_t + x_t' \beta_t + \varepsilon_t, & \varepsilon_t &\sim N(0, \sigma_\varepsilon^2), \\
 \mu_{t+1} &= \mu_t + \zeta_t, & \zeta_t &\sim N(0, \sigma_\zeta^2), \\
 \gamma_{t+1} &= - \sum_{j=1}^{J-1} \gamma_{t+1-j} + \omega_t, & \omega_t &\sim N(0, \sigma_\omega^2).
 \end{aligned} \tag{3-4}$$

O termo de regressão $x_t' \beta_t$ é principalmente motivado pela necessidade de intervenções de *outliers* e de quebras (vide seção 3.1.3). A idéia de se estimar tal modelo estrutural em particular é a de que suas componentes não observáveis

terem capacidade de explicar o comportamento dos sinistros IBNR de forma intuitiva. A componente de nível μ_t explica as informações referentes ao volume de sinistros ocorridos em cada *ano de acidente*. Cabe à componente periódica γ_t captar o padrão da série em cada linha do triângulo, isto é, será responsável por explicar o comportamento do atraso na liquidação dos sinistros.

O modelo estrutural tem a seguinte forma em Espaço de Estado:

$$y_t = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \gamma_t \\ \gamma_{t-1} \\ \vdots \\ \gamma_{t-J+1} \end{pmatrix} + x_t' \beta_t + \varepsilon_t, \quad t = 1, 2, \dots, n$$

$$\begin{pmatrix} \mu_{t+1} \\ \gamma_{t+1} \\ \gamma_t \\ \vdots \\ \gamma_{t-J} \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & -1 & & -1 & -1 \\ 0 & 1 & 0 & & 0 & 0 \\ & & & \vdots & & \\ 0 & 0 & 0 & & 1 & 0 \end{bmatrix} \begin{pmatrix} \mu_t \\ \gamma_t \\ \gamma_{t-1} \\ \vdots \\ \gamma_{t-J+1} \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \begin{pmatrix} \xi_t \\ \omega_t \end{pmatrix}$$

Nas subseções seguintes, serão apresentadas duas propostas que visam resolver o problema da estimação de IBNR sob a ordenação “por linhas” do triângulo de *runoff* proposta na seção 2, e do cálculo do erro médio quadrático associado. Muito essencialmente e, em palavras, o primeiro deles parte para a dedução, bloco a bloco, de uma expressão analítica fechada da matriz de covariância condicional das parcelas do IBNR (cf. expressão (2-6)) empilhados; o segundo, por sua vez, é baseado no aumento do vetor de estado com um *acumulador* das parcelas do IBNR. Ou seja, um método tenta solucionar o problema “construindo blocos de covariâncias”; e o outro, “acumulando a reserva”.

3.1.3

Incorporação de sazonalidade entre o período de origem dos sinistros (linhas do triângulo)

A metodologia descrita neste trabalho independe da unidade de tempo utilizada no triângulo, seja ela mensal, trimestral, semestral etc. Abarca, ainda, naturalmente, a possível existência de comportamentos periódicos entre as colunas do triângulo, após a introdução do modelo (3-4) como previamente comprovado. Entretanto, em determinadas unidades de tempo, pode haver, ao menos teoricamente, certa periodicidade entre as linhas, refletindo, assim, a

possibilidade de existência de um padrão sazonal nas ocorrências de sinistros IBNR.

Na Figura 3.1 está representado um triângulo trimestral. Caso existam indícios empíricos de que sinistros originados em um determinado trimestre de um ano em particular tenham aspectos em comum com os originados no mesmo trimestre de outros anos, esta sazonalidade pode estar incorporada diretamente no modelo. Uma forma de se acrescentar essa componente sazonal seria através de variáveis *dummies*, como no modelo abaixo:

$$y_t = \mu_t + \gamma_t + \sum_{i=2}^J \beta_t^{(i)} d_t^{(i)} + \varepsilon_t, \quad (3-5)$$

no qual

$$d_t^{(i)} = \begin{cases} 1 & y_t \in \{i\text{-ésimo "período"}, \dots, S\text{-ésimo "período"}\}, \quad i = 2, 3, \dots, J \\ 0 & \text{caso contrário.} \end{cases}$$

Origem	Desenvolvimento d					
	w	0	1	2	...	11
2006Q1	$C_{1,0}$	$C_{1,1}$	$C_{1,2}$			$C_{1,11}$
2006Q2	$C_{2,0}$	$C_{2,1}$	$C_{2,2}$...		
2006Q3	$C_{3,0}$	$C_{3,1}$	$C_{3,2}$...		
2006Q4	$C_{4,0}$	$C_{4,1}$	$C_{4,2}$...		
2007Q1	$C_{5,0}$	$C_{5,1}$	$C_{5,2}$			
2007Q2	$C_{6,0}$	$C_{6,1}$	$C_{6,2}$			
2007Q3	$C_{7,0}$	$C_{7,1}$	$C_{7,2}$			
2007Q4	$C_{8,0}$	$C_{8,1}$	$C_{8,2}$			
2008Q1	$C_{9,0}$	$C_{9,1}$	$C_{9,2}$			
2008Q2	$C_{10,0}$	$C_{10,1}$	$C_{10,2}$			
2008Q3	$C_{11,0}$	$C_{11,1}$				
2008Q4	$C_{12,0}$					

Figura 3.1: Triângulo trimestral.

A inclusão dos referidos parâmetros adicionais pode agravar ainda mais a questão da maximização da verossimilhança, já que a quantidade de valores faltantes sempre será cerca de 50% do número de observações do triângulo.

3.2

Primeira Abordagem: o método dos blocos

Considere a forma em EE e toda a notação correspondente a esta e ao filtro de Kalman da subseção 3.1.1. Adicionalmente, defina

$\mathcal{I} \equiv \{t : y_t \text{ é não-ausente}\}$, $\tilde{\mathbf{Y}} \equiv \{y_{i_j} : i_j \in \mathcal{I}, \forall j\}^1$, $\mathbf{Y} \equiv \{y_t : t = 1, \dots, n\}$ e

$$L_t^* \equiv \begin{cases} L_t, & \text{se } t \in \mathcal{I} \\ T_t, & \text{caso contrário} \end{cases} \quad N_t^* \equiv \sum_{k=t+1}^n L_{t+1}^* \dots L_{k-1}^* Z_k^* F_k^{-1} Z_k^* L_{k-1}^* \dots L_{t+1}^*.$$

Observe que $\tilde{\mathbf{Y}}$, no contexto de cálculo de IBNR, consiste na *informação proveniente do triângulo* (cf. Figura 2.3).

O ponto de partida para tudo o que se desenvolverá nesta subseção são expressões recursivas, deduzidas em Durbin & Koopman (2001, seção 4.5), para algumas matrizes de covariâncias condicionais provenientes de manipulações com as equações de suavização do filtro de Kalman. Estas são sumarizadas no seguinte Lema:

Lema 1 *Sejam $t, j = 1, \dots, n$ quaisquer. Então,*

1. $Cov(\alpha_t, \alpha_j | \mathbf{Y}) = P_{t|t-1} L'_t L'_{t+1} \dots L'_{j-1} (I_m - N_{j-1} P_{j|j-1})$, $j \geq t$
sendo que $L'_t L'_{t+1} \dots L'_{j-1} = I_m$ quando $j = t$.
2. $Cov(\varepsilon_t, \varepsilon_j | \mathbf{Y}) = H_t K'_t L'_{t+1} \dots L'_{j-1} W'_j$, $j > t$
sendo que $W_j = H_j (F_j^{-1} Z_j - K'_j N_j L_j)$.
3. $Cov(\varepsilon_t, \alpha_j | \mathbf{Y}) = -H_t K'_t L'_{t+1} \dots L'_{j-1} (I_m - N_{j-1} P_{j|j-1})$, $j > t$.
4. $Cov(\alpha_t, \varepsilon_j | \mathbf{Y}) = -P_{t|t-1} L'_t L'_{t+1} \dots L'_{j-1} W'_j$, $j \geq t$
sendo que $W_j = H_j (F_j^{-1} Z_j - K'_j N_j L_j)$ e $L'_t L'_{t+1} \dots L'_{j-1} = I_m$ quando $j = t$.

Outro resultado importante é o Lema 2 dado na sequência, o qual se relaciona com distribuições Gaussianas condicionais e, diferentemente de todos os resultados desta seção, não se restringe necessariamente ao contexto de modelos em Espaço de Estado. A prova deste Lema encontra-se no apêndice A.1.

Lema 2 *Sejam x, y e z vetores aleatórios com distribuição conjunta Gaussiana. Se $Cov(y, z) = 0$ e se $Cov(x, z) = 0$, então*

$$E(x|y, z) = E(x|y) \tag{3-6}$$

$$Var(x|y, z) = Var(x|y). \tag{3-7}$$

¹Observe que $\tilde{\mathbf{Y}}$ está sendo definido como um conjunto de vetores aleatórios e não como um vetor aleatório empilhado. Mas, para o que segue, o mesmo pode ser visto como tal.

O próximo resultado, de prova bem direta e apresentada no apêndice A.2, revela uma espécie de *ortogonalidade* entre a parte observada do triângulo e as parcelas não observadas do IBNR.

Lema 3 Para todo $t \notin \mathcal{I}$, ε_t é não-correlacionado com $\tilde{\mathbf{Y}}$.

A linha condutora do desenvolvimento do *método dos blocos* é a obtenção de uma matriz de covariância condicional de todos os y_t , tais que $t \notin \mathcal{I}$, empilhados dado $\tilde{\mathbf{Y}}$. O caminho para isso é estudar, para os mesmos índices t , as covariâncias condicionais dos vetores aleatórios não-observáveis α_t , ε_t e η_t e explorar convenientemente a linearidade da relação entre estes e os y_t . O próximo resultado materializa esta ideia e tem como base de construção os Lemas 1, 2 e 3. Sua prova encontra-se no apêndice A.3.

Lema 4 Para $t, j \notin \mathcal{I}$ arbitrários, tem-se que:

1. $Cov(\varepsilon_t, \varepsilon_j | \tilde{\mathbf{Y}}) = \begin{cases} H_t, & \text{para } t = j \\ 0, & \text{caso contrário.} \end{cases}$
2. $Cov(\varepsilon_t, \alpha_j | \tilde{\mathbf{Y}}) = 0$
3. $Cov(\alpha_t, \alpha_j | \tilde{\mathbf{Y}}) = \begin{cases} P_{t|t-1} L_t^* L_{t+1}^* \dots L_{j-1}^* (I - N_{j-1}^* P_{j|j-1}), & \text{se } t < j \\ P_{t|t-1} - P_{t|t-1} N_{t-1}^* P_{t|t-1}, & \text{se } t = j. \end{cases}$

Estabelecidos os resultados de suporte, agora já existem plenas condições para que se deduzam as expressões computacionais do método dos blocos. Essas se encontram no próximo Teorema, cuja prova está no apêndice A.4.

Teorema 1 Para t, j arbitrários, tem-se que

$$Cov(y_t, y_j | \tilde{\mathbf{Y}}) = \begin{cases} 0 & \text{se } t \in \mathcal{I} \text{ ou } j \in \mathcal{I} \\ Z_t (P_{t|t-1} - P_{t|t-1} N_{t-1}^* P_{t|t-1}) Z_t' + H_t & \text{se } t = j \text{ e } t, j \notin \mathcal{I} \\ Z_t P_{t|t-1} L_t^* L_{t+1}^* \dots L_{j-1}^* (I - N_{j-1}^* P_{j|j-1}) Z_j' & \text{se } t < j \text{ e } t, j \notin \mathcal{I}. \end{cases}$$

Após ter-se chegado ao objetivo desejado, cabem aqui alguns comentários de ordem prática. A implementação computacional da expressão matricial, enunciada no Teorema 1, envolve o armazenamento de algumas matrizes advindas das recursões do filtro de Kalman: $P_{t|t-1}$ e N_{t-1}^* para todo $t \notin \mathcal{I}$,

e ainda as matrizes $L_{\tau}^*, L_{\tau+1}^*, \dots, L_{\tau'-1}^*$, $1 \leq \tau < \tau' \leq n$ nas quais τ e τ' são o primeiro e o último instantes, respectivamente, em que existem valores faltantes. Também, é a mesma fórmula que, quando calculada para todas as combinações possíveis de índices i e j , conduz ao insumo básico - que é a matriz de covariâncias condicionais completa de Y_n dado $\tilde{\mathbf{Y}}$ - para que se calcule uma medida de precisão associada à estimação de *qualquer* combinação linear dos valores faltantes.

3.2.1

Erro médio quadrático da Reserva IBNR estimada: casos total e por ano

Para um vetor $\mathbf{a} = (a_1, a_2, \dots, a_n)'$ qualquer, tem-se que $E(\mathbf{a}'\mathbf{Y}|\tilde{\mathbf{Y}}) = \mathbf{a}'E(\mathbf{Y}|\tilde{\mathbf{Y}})$ e $\text{Cov}(\mathbf{a}'\mathbf{Y}|\tilde{\mathbf{Y}}) = \mathbf{a}'\text{Cov}(\mathbf{Y}|\tilde{\mathbf{Y}})\mathbf{a}$. A reserva IBNR é calculada escolhendo-se um vetor \mathbf{a} , aqui definido por $\mathbf{a}^{(T)}$, de forma que $a_i^{(T)} = 1$ se $i \notin \mathcal{I}$ e $a_i^{(T)} = 0$ caso contrário. Para as parcelas do IBNR relativas a um específico ano de acidente (linhas), utilize o mesmo vetor \mathbf{a} com coordenadas nulas nos lugares apropriados - os que não correspondem ao ano. Esta análise por ano de acidente é interessante porque permite a identificação de fontes de incertezas nas parcelas do IBNR.

Assim, as expressões do método dos blocos para a Reserva IBNR e seu correspondente erro padrão são:

$$\widehat{\text{IBNR}} \equiv \hat{R} = \mathbf{a}'E(\mathbf{Y}|\tilde{\mathbf{Y}}) \quad (3-8)$$

$$\text{ep}(\widehat{\text{IBNR}}) \equiv \text{ep}(\hat{R}) = \sqrt{\mathbf{a}'\text{Cov}(\mathbf{Y}|\tilde{\mathbf{Y}})\mathbf{a}} \quad (3-9)$$

3.2.2

Uso da distribuição log-normal para y_t

Uma alternativa à hipótese de normalidade para y_t do triângulo é a distribuição log-normal, que induz, por primeiros princípios, à modelagem de $z_t \equiv \log y_t$. Essa distribuição foi extensamente explorada na literatura atuarial (cf. Taylor, 2000, capítulo 9), tanto na área de Modelos Lineares Generalizados (MLG) quanto na área de Séries Temporais. Em MLG, podem-se citar os trabalhos de Kremer (1982), Renshaw (1989), Christofides (1990), Verrall (1991) e Doray (1996); na área de Séries Temporais, têm-se de Jong & Zehnwirth (1983), Verrall (1989), de Jong (2004) e de Jong (2006). Apesar de desenvolver seu modelo sem precisar admitir qualquer distribuição (cf. Mack, 1994a), Mack faz um exemplo utilizando log-normalidade para a construção gráfica da reserva (cf. Mack, 1994b).

Sob a adoção do método dos blocos, este pressuposto distribucional alternativo dá origem ao seguinte algoritmo para o cálculo da reserva IBNR estimada e seu correspondente erro médio quadrático:

1. Aplique o filtro de Kalman ao triângulo formado pelos z_t , armazenando todas as matrizes devidas (vide comentários finais da seção 3.2).
2. Utilize o método dos blocos para obter $\hat{z}_t \equiv E(z_t|\tilde{\mathbf{Z}}) = E(\log y_t|\tilde{\mathbf{Y}})$, $\sigma_{\hat{z}_t}^2 \equiv \text{Var}(z_t|\tilde{\mathbf{Z}})$ e $\sigma_{\hat{z}_t, \hat{z}_j} \equiv \text{Cov}(z_t, z_j|\tilde{\mathbf{Z}})$.
3. Calcule²:

$$\hat{y}_t = \exp \left\{ \hat{z}_t + \frac{\sigma_{\hat{z}_t}^2}{2} \right\} \quad (3-10)$$

$$\sigma_{\hat{y}_t}^2 = \exp \left\{ 2\hat{z}_t + \sigma_{\hat{z}_t}^2 \right\} \left(e^{\sigma_{\hat{z}_t}^2} - 1 \right) \quad (3-11)$$

$$\sigma_{\hat{y}_t, \hat{y}_j} = \exp \left\{ \hat{z}_t + \hat{z}_j + \frac{\sigma_{\hat{z}_t}^2}{2} + \frac{\sigma_{\hat{z}_j}^2}{2} \right\} \left(e^{\sigma_{\hat{z}_t, \hat{z}_j}} - 1 \right) \quad (3-12)$$

4. Proceda com os cálculos da subseção anterior.

Essa é uma clara vantagem deste método sobre o do acumulador, que não permite, por construção, a modelagem do logaritmo dos valores do triângulo, como será visto na sequência.

3.3

Segunda Abordagem: o método do acumulador

Este método consiste em adicionar uma componente ao vetor de estado, denotada por δ_t , a qual será responsável pela acumulação das previsões dos valores faltantes:

$$y_t = \begin{bmatrix} Z_t & 0 \end{bmatrix} \begin{bmatrix} \alpha_t \\ \delta_t \end{bmatrix} + d_t + \varepsilon_t$$

$$\begin{bmatrix} \alpha_{t+1} \\ \delta_{t+1} \end{bmatrix} = \begin{bmatrix} T_t & 0 \\ X_t & I \end{bmatrix} \begin{bmatrix} \alpha_t \\ \delta_t \end{bmatrix} + \begin{bmatrix} R_t \\ 0 \end{bmatrix} \eta_t, \quad (3-13)$$

na qual $X_t = 0$ quando $t \in \mathcal{I}$ e $X_t = Z_t$ quando $t \notin \mathcal{I}$ (observação faltante).

Denote por ψ e ψ^\dagger os vetores de parâmetros dos modelos (3-1) e (3-13) respectivamente e denote por \mathcal{L} e \mathcal{L}^\dagger as correspondentes verossimilhanças. Apesar de $\psi = \psi^\dagger$ – com efeito, o modelo em (3-13) é um simples aumento na

²Estas expressões decorrem de direta aplicação de função geradora de momentos e/ou função característica.

equação do estado do modelo em (3-1) cujas matrizes do sistema adicionais não compreendem novos parâmetros –, não é muito direto afirmar o mesmo, ou algo diferente, para os estimadores de máxima verossimilhança provenientes de \mathcal{L} e de \mathcal{L}^\dagger . O próximo resultado, cuja prova encontra-se no apêndice A.5, esclarece essa dúvida e, mais à frente, mostrar-se-á fundamental para a implementação computacional do método do acumulador.

Proposição 1 $\hat{\psi} \equiv \arg \max \mathcal{L}(\psi) = \arg \max \mathcal{L}^\dagger(\psi^\dagger) \equiv \hat{\psi}^\dagger$.

Interpretação: Apesar de o modelo aumentado possuir uma componente a mais de “acumulação”, não há acréscimo de informação à função de verossimilhança, pois, mesmo que a componente adicional seja, recursivamente, função de α_t , este último depende, recursivamente, apenas de si mesmo.

A importância prática da Proposição 1 advém de as implementações se simplificarem muito – sobretudo com a extensão adicional do vetor de estado a ser mostrado na próxima subseção –, pois o vetor de parâmetros poderá ser estimado através do modelo original (com matrizes menores) e, assim, as estimativas obtidas serão utilizadas no filtro de Kalman do modelo aumentado.

3.3.1

Extensão adicional do acumulador: acumuladores parciais por ano de origem

Como já visto na seção 3.3, o método do acumulador consiste no acréscimo do acumulador δ_t ao vetor de estado, e no redimensionamento das matrizes do sistema. Entretanto, é conveniente incorporar não só o acumulador da reserva IBNR total, mas também acumuladores parciais para cada ano de origem $w = 2, \dots, J$, pela mesma motivação apresentada no método dos blocos. Para tanto, seja $\boldsymbol{\delta}_t$ um vetor de dimensão $J \times 1$ tal que $\boldsymbol{\delta}_t = (\delta_t^{(2)}, \delta_t^{(3)}, \dots, \delta_t^{(J)}, \delta_t^{(T)})'$, cujo sobrescrito (i) representa a parcela referente a linha i , com $2 \leq i \leq J$, e o total $i = T$. O modelo em (3-13), então, é expandido no modelo em (3-14)

$$y_t = \begin{bmatrix} Z_t & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \alpha_t \\ \delta_t^{(2)} \\ \vdots \\ \delta_t^{(J)} \\ \delta_t^{(T)} \end{bmatrix} + \varepsilon_t,$$

$$\begin{bmatrix} \alpha_{t+1} \\ \delta_{t+1}^{(2)} \\ \vdots \\ \delta_{t+1}^{(J)} \\ \delta_{t+1}^{(T)} \end{bmatrix} = \begin{bmatrix} T_t & 0_{(m \times pJ)} \\ \tilde{X}_t & I_{(pJ \times pJ)} \end{bmatrix} \begin{bmatrix} \alpha_t \\ \delta_t^{(2)} \\ \vdots \\ \delta_t^{(J)} \\ \delta_t^{(T)} \end{bmatrix} + \begin{bmatrix} R_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \eta_t, \quad (3-14)$$

com $\tilde{X}_t = (X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(J)}, X_t^{(T)})'$.

As novas dimensões dos vetores e matrizes deste modelo estão expostas na tabela 3.1. Como visto na seção 3.3, a matriz de transição em (3-14) é formada por três blocos constantes e um variante no tempo, \tilde{X}_t , cujos elementos seguem a regra a seguir.

Para $i = 1, \dots, J$

$$X_t^{(i)} = \begin{cases} \begin{bmatrix} Z_t & 0 & \dots & 0 \end{bmatrix} & t \notin \mathcal{I} \text{ e } t \in \text{linha } i, \\ 0 & \text{caso contrário.} \end{cases}$$

E também

$$X_t^{(T)} = \begin{cases} \begin{bmatrix} Z_t & 0 & \dots & 0 \end{bmatrix} & t \notin \mathcal{I}, \\ 0 & \text{caso contrário.} \end{cases}$$

Por fim, existe uma extensão direta da Proposição 1 aplicável ao modelo (3-14), a qual, como anteriormente discutido na seção 3.3, facilita a estimação dos parâmetros do modelo em situações práticas.

3.3.2

Erro médio quadrático da Reserva IBNR estimada: casos total e por ano

Como já mencionado na seção anterior, uma das vantagens do método do acumulador é obter-se o valor estimado da reserva IBNR e seu erro médio quadrático diretamente do vetor de estado suavizado e sua matriz de variância-covariância, respectivamente.

Vetor ou Matriz	Dimensão
α_t^\dagger	$(m + pJ) \times 1$
Z_t^\dagger	$p \times (m + pJ)$
T_t^\dagger	$(m + pJ) \times (m + pJ)$
R_t^\dagger	$(m + pJ) \times r$

Tabela 3.1: Nova dimensão dos vetores e matrizes para o método do acumulador.

A reserva estimada será dada por $\hat{R} = E(\delta_{n+1}^{(T)} | \tilde{\mathbf{Y}}) + \sum_{t \notin \mathcal{I}} d_t$, exatamente como na equação (2-6). Seu erro médio quadrático será dado pelo último elemento da diagonal principal da matriz $\text{Var}(\alpha_{n+1}^\dagger | \tilde{\mathbf{Y}}) = P_{n+1|n}^\dagger$. Assim, a reserva e seu respectivo erro padrão são dados por:

$$\widehat{\text{IBNR}} \equiv \hat{R} = E(\delta_{n+1}^{(T)} | \tilde{\mathbf{Y}}) + \sum_{t \notin \mathcal{I}} d_t, \quad (3-15)$$

$$\widehat{\text{ep}}(\widehat{\text{IBNR}}) \equiv \widehat{\text{ep}}(\hat{R}) = \sqrt{\text{Var}(\delta_{n+1}^{(T)} | \tilde{\mathbf{Y}}) + \sum_{t \notin \mathcal{I}} H_t}. \quad (3-16)$$