Análise de Séries Temporais
Capítulo VI. Modelos de Espaço de Estados

Fernando Lucambio

Departamento de Estatística
Universidade Federal do Paraná

Outubro, 2023

- VI.6. DLM com erros correlacionados
- VI.6.1 Modelo ARMAX
- VI.9 Modelos Ocultos de Markov e regressão automática alternada
- VI.10 MLD com comutação

Às vezes, é vantajoso escrever o modelo de espaço de estado de uma maneira ligeiramente diferente, como é feito por vários autores; por exemplo, Anderson and Moore (1979) e Hannan and Deistler (1988). Aqui, escrevemos o modelo de espaço de estado como

$$\begin{array}{rcl} X_{t+1} & = & \Phi X_t + \Upsilon U_{t+1} + \Theta W_t, \qquad t = 0, 1, \cdots, n \\ Y_t & = & A_t X_t + \Gamma U_t + \nu_t, \qquad t = 1, 2, \cdots, n, \end{array}$$

onde, na equação de estado, $X_0 \sim N_p(\mu_0, \Sigma_0)$, Φ é $p \times p$ e Υ é $p \times r$, Θ é $p \times m$ e $W_t \sim N_m(0, Q)$ independentes igualmente distribuídas.

Na equação de observações, A_t é $q \times p$, Γ é $q \times r$ e $\nu_t \sim N_q(0, R)$ independentes igualmente distribuídas.

Neste modelo, embora W_t e ν_t ainda sejam séries de ruído branco, ambos independentes de X_0, também permitimos que o ruído de estado e o ruído de observação sejam correlacionados no tempo t; isso é,

$$\text{Cov}(W_s \nu_t) = S\delta_s,$$

onde δ_s é a função delta de Kronecker; observe que S é uma matriz $m \times q$.

A principal diferença entre esta forma do modelo e a especificada na Definição VI.1 é que este modelo inicia o processo de ruído de estado em $t = 0$ para facilitar a notação relacionada à covariância concorrente entre W_t e ν_t. Além disso, a inclusão da matriz Θ nos permite evitar o uso de um processo de ruído de estado singular.

Para obter as inovações, $\epsilon_t = Y_t - A_t X_t^{t-1} - \Gamma U_t$ e a variância da inovação $\Sigma_t = A_t P_t^{t-1} A_t^\top + R$, neste caso, precisamos das previsões de estado um passo à frente. Obviamente, as estimativas filtradas também serão de interesse e serão necessárias para suavização.

O Teorema VI.2, o alisamento, conforme exibido na Seção VI.2, ainda se mantém. A propriedade a seguir gera o preditor X_{t+1}^t a partir do preditor anterior X_t^t quando os termos de ruído são correlacionados e exibe a atualização do filtro.

Teorema VI.5. O filtro de Kalman com ruído correlacionado

Modelo ARMAX

Considere o modelo ARMAX k-dimensinal dado a seguir

$$Y_t = \Upsilon U_t + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \sum_{k=1}^{q} \Theta_k \nu_{t-k} + \nu_t.$$

As observações Y_t são um processo vetorial k-dimensional, os Φs e Θs são matrizes $k \times k$, Υ é $k \times r$, U_t é a entrada $r \times 1$ e ν_t é um processo de ruído branco $k \times 1$; de fato, o modelo acima e aquele apresentado na Definição VI.1 são modelos idênticos, mas aqui escrevemos as observações como Y_t.

Regressão multivariada com erros autocorrelacionados

Na regressão com erros autocorrelacionados, estamos interessados em ajustar o modelo de regressão

$$y_t = \Gamma u_t + \epsilon_t,$$

a um processo vetorial $k \times 1$ y_t, com r regressores u_t onde ϵ_t é o vetor ARMA(p,q) é Γ uma matriz $k \times r$ de parâmetros de regressão.

Notamos que os regressores não precisam variar com o tempo, por exemplo, $u_{t1} = 1$ inclui uma constante na regressão e que o caso $k = 1$ foi tratado na Seção III.8.

Para colocar o modelo na forma de espaço de estados, notamos que $\epsilon_t = y_t - \Gamma u_t$ é um processo ARMA(p,q) k-dimensional. Portanto, se definirmos $H = 0$ e incluímos Γu_t, obtemos

$$
\begin{aligned}
x_{t+1} &= FX_t + G\nu_t, \\
y_t &= \Gamma u_t + Ax_t + \nu_t,
\end{aligned}
$$

onde as matrizes modelo A, F e G são definidas no Teorema VI.6.

O fato de que o modelo acima é uma regressão multivariada com erros autocorrelacionados decorre diretamente do Teorema VI.6, observando que, juntos, $x_{t+1} = Fx_t + G\nu_t$ e $\epsilon_t = Ax_t + \nu_t$ implicam $\epsilon_t = y_t - \Gamma u_t$ é um vetor ARMA(p,q).

Como no caso dos modelos ARMAX, a regressão com erros autocorrelacionados é um caso especial do modelos dinâmicos lineares.

Exemplo VI.12. Mortalidade, Temperatura e Poluição.

VI.9 Modelos Ocultos de Markov e regressão automática alternada

Na introdução mencionamos que o modelo de espaço de estados é caracterizado por dois princípios. Primeiro, há um processo de estado oculto, $\{X_t \; : \; t \; = \; 0, 1, \cdots\}$ assumido como Markoviano, ou seja, X_t é uma Cadeia de Markov. Segundo, as observações, $\{Y_t \, : \, t = 1, 2, \cdots\}$, são independentes condicionalmente, considerando os estados.

Temos focado principalmente nos modelos lineares gaussianos de espaço de estados, mas existe uma área inteira que se desenvolveu em torno do caso em que os estados X_t são uma Cadeia de Markov de valor discreto, e esse será o foco nesta seção. A ideia básica é que o valor do estado no tempo t especifique a distribuição da observação no tempo t.

Na abordagem de Cadeia de Markov, declaramos que a dinâmica do sistema no momento t é gerada por um dos m regimes possíveis que evoluem de acordo com uma Cadeia de Markov ao longo do tempo. O caso em que o regime em particular seja desconhecido para o observador está sob o título de Modelos Ocultos de Markov (HMM).

Embora o modelo satisfaça as condições de ser um modelo de espaço de estados, os HMMs foram desenvolvidos em paralelo. Se o processo de estado é de valor discreto, normalmente se usa o termo Modelo Oculto de Markov e se o processo de estado é de valor contínuo, usa-se o termo Modelo de Espaço de Estados ou uma de suas variantes.

Aqui, assumimos que os estados X_t, são uma Cadeia de Markov que recebe valores em um espaço de estados finito $\{1, 2, \cdots, m\}$, com distribuição estacionária

$$\pi_j = P(X_t = j)$$

e probabilidades de transição estacionárias

$$\pi_{ij} = P(X_{t+1} = j \,|\, X_t = i),$$

para $t = 0, 1, 2, \cdots$ e $i, j \in \{1, 2, \cdots, m\}$.

Como o segundo componente do modelo é que as observações são condicionalmente independentes, precisamos especificar as distribuições e as denotamos por

$$p_j(y_t) = P(Y_t = y_t \,|\, X_t = j)\cdot$$

Nesta seção estendemos o Modelo Oculto de Markov para problemas mais gerais. Conforme indicado anteriormente, o problema de modelar mudanças nos regimes para séries temporais tem sido de interesse em muitos campos diferentes.

As generalizações do modelo de espaço de estado para incluir a possibilidade de mudanças ocorrendo ao longo do tempo foram abordadas permitindo mudanças nas covariâncias de erro (Harrison and Stevens, 1976, Gordon and Smith, 1988, 1990) ou atribuindo distribuições de mistura para os erros de observação ν_t (Peña and Guttman, 1988).

Aproximações à filtragem foram derivadas em todos os artigos mencionados.

Uma aplicação para monitorar transplantes renais foi descrita em Smith and West (1983) e em Gordon and Smith (1990). Gerlach et al. (2000) consideraram uma extensão do modelo de switching AR para permitir mudanças de nível e outliers tanto nas observações quanto nas inovações.

Uma aplicação da ideia de mudança para o rastreamento de múltiplos alvos foi considerada em Bar-Shalom (1978), que obteve aproximações da filtragem de Kalman em termos de médias ponderadas das inovações. Para uma cobertura completa dessas técnicas e outras relacionadas, consulte Cappé, Moulines and Rydén (2009) e Douc, Moulines and Stoffer (2014).

Nesta seção, nos concentraremos no método apresentado em Shumway and Stoffer (1991).

Uma forma de modelar a mudança em uma série de tempo em evolução é assumir a dinâmica de algumas mudanças de modelo subjacentes de forma descontínua em certos pontos indeterminados no tempo.

Nosso ponto de partida é o modelo linear dinâmico (DLM) segundo a Definição VI.1 com as matrizes $\Upsilon = \Gamma = 0$, a saber,

$$X_t = \Phi X_{t-1} + W_t,$$

para descrever a dinâmica dos $p \times 1$ estados e

$$Y_t = A_t X_t + \nu_t,$$

para descrever a dinâmica das $q \times 1$ observações. Lembre-se de que W_t e ν_t são sequências de ruído branco Gaussiano com $\mathrm{Var}(W_t) = Q$, $\mathrm{Var}(\nu_t) = R$ e $\mathrm{Cov}(W_t, \nu_s) = 0$ para todo s e t.

Para incorporar uma estrutura de comutação razoável para a matriz de medição no modelo linear dinâmico (DLM), que seja compatível com ambas as situações práticas descritas anteriormente, assumimos que as m configurações possíveis são estados em um processo independente não estacionário definido pelas probabilidades variáveis no tempo

$$\pi_j(t) \,=\, P(A_t = M_j),$$

para $j = 1, \cdots, m$ e $t = 1, 2, \cdots, n$.

Informações importantes sobre o estado atual do processo de medição são fornecidas pelas probabilidades filtradas de estar no estado j, definidas como as probabilidades condicionais

$$\pi_j(t \,|\, t) \,=\, P(A_t = M_j \,|\, Y_{1:t}),$$

que também variam em função do tempo. Lembre-se de que $Y_{s':s} = \{Y_{s'}, \cdots, Y_s\}$.

As probabilidades filtradas $\pi_j(t \mid t)$, definidas acima, fornecem estimativas variáveis no tempo da probabilidade de estar no estado j dados os dados no tempo t.

Será importante para nós obter estimadores das probabilidades de configuração, $\pi_j(t \mid t)$, os estimadores de estado predito e filtrado X_t^{t-1} e X_t^t e as correspondentes matrizes de covariância de erro P_t^{t-1} e P_t^t. Obviamente, o preditor e os estimadores de filtro dependerão dos parâmetros Θ do DLM. Em muitas situações, os parâmetros serão desconhecidos e teremos que estimá-los. Nosso foco será na estimação por máxima verossimilhança.

Exemplo VI.22. Dados da influenza.