

2 Modelos Lineares Multivariados

2.1. Introdução

Neste capítulo, serão apresentados dois modelos lineares, largamente utilizados na estimação de séries temporais multivariadas. O primeiro deles, o modelo *Vector Autoregressive* (VAR), proposto por Sims (1980), foi desenvolvido como um modelo dinâmico, no qual todas as variáveis a serem estudadas são tratadas como endógenas. Sendo assim, o modelo VAR examina as relações lineares existentes entre cada variável endógena e os valores passados das mesmas variáveis (assume-se uma defasagem p), permitindo ainda a inclusão de variáveis exógenas na análise.

Este modelo tem como restrições a escolha do conjunto relevante de variáveis a serem analisadas, que são os valores correntes dos processos, e o número máximo de defasagens envolvidas nas relações entre elas, normalmente escolhido com base em critérios estatísticos, como os de Akaike (1974) ou Schwarz (1978). Por um lado, é desejável incluir o maior número possível de defasagens, de modo a evitar a imposição de restrições falsas sobre a dinâmica do modelo. Por outro lado, quanto maior a ordem de defasagens, maior o número de parâmetros a serem estimados, conseqüentemente, menos graus de liberdade para a estimação. Um cuidado deve ser tomado ao estimar modelos multivariados. Geralmente, eles apresentam elevado número de parâmetros, com reflexo no tamanho de amostra requerido para que se obtenha uma estimação confiável.

Diversos testes estatísticos podem ser utilizados com a finalidade de verificar a adequação de um modelo. Caso alguma das premissas básicas da modelagem não for válida, diz-se que existe um erro de especificação do modelo. Os testes são aplicados em diversos estágios da modelagem, ou seja, a cada momento da elaboração do modelo aplica-se os testes de diagnóstico para verificar se todas as exigências básicas são válidas para o conjunto de dados em questão. A seguir, alguns motivos da realização dos testes:

- 1) Testes com o objetivo de definir a especificação do modelo;
- 2) Testes para inclusão ou não de variáveis defasadas no modelo;
- 3) Testes para as propriedades dos termos de perturbação não-observável (ruído branco);
- 4) Testes para verificação do ajuste do modelo.

Uma condição básica para estimação do modelo VAR é que as séries temporais sob análise sejam estacionárias. Isto indica que as médias e variâncias devem ser constantes ao longo do tempo e o valor das covariâncias entre dois períodos de tempo devem depender apenas da distância ou defasagem entre os dois períodos, e não do período de tempo efetivo em que as covariâncias foram calculadas.

Após a verificação da presença de raízes unitárias nas séries, sendo as mesmas integradas de mesma ordem, ou seja, necessitando as séries do mesmo número de diferenciações para se tornarem estacionárias procedem-se os testes de co-integração. Em existindo relações de co-integração entre as séries, diz-se que as mesmas apresentam uma relação linear estável no longo prazo.

O teste de co-integração visa determinar o número de vetores de co-integração que serão necessários no sistema. Um dos procedimentos para identificar a existência de co-integração é o de Johansen (1988), o qual utiliza Máxima Verossimilhança para estimar os vetores de co-integração. A hipótese nula é de que não há nenhum vetor de co-integração versus a hipótese alternativa de que há pelo menos um vetor de co-integração. Caso o teste detecte a presença de um vetor de co-integração num sistema, então, ao invés de estimar o modelo VAR, deve-se utilizar o modelo VEC.

2.2. Modelo VAR

Algumas propriedades importantes do modelo VAR serão discutidas. As principais utilizações deste modelo são para previsão e análise estrutural. Considere um conjunto de séries temporais observadas no tempo. Um primeiro modelo tem de ser especificado e os parâmetros têm de ser estimados. Em seguida, a adequação do modelo é marcada por diversos instrumentos estatísticos

e o modelo pode ser estimado e utilizado para previsão e análise estrutural ou dinâmico. As principais etapas da modelagem multivariada são:

- 1) Especificação e estimação do modelo;
- 2) Avaliação / verificação do modelo;
 - a. Modelo rejeitado, volta ao passo 1
 - b. Modelo aceito, avança ao passo 3
- 3) Previsão
- 4) Análise estrutural

2.2.1. Formulação Matemática

Sejam y_1, y_2, \dots, y_T séries temporais multivariadas, com $y_t = (y_{1t}, y_{2t}, \dots, y_{Kt})'$. Um modelo VAR(p) pode ser expresso matematicamente pela seguinte formulação:

$$y_t = v + A_1 y_{t-1} + A_2 y_{t-2} + \dots + u_t ; t = 0, 1, 2, \dots \quad (2.1)$$

Todos os símbolos utilizados nesta representação possuem significados usuais, isto é,

$y_t = (y_{1t}, y_{2t}, \dots, y_{Kt})'$ é um vetor aleatório ($K \times 1$) de variáveis endógenas;

$v = (v_1, v_2, \dots, v_K)'$ é um vetor fixo de interceptos ($K \times 1$), os quais permitem a possibilidade de média $E(y_t)$ não-nula;

A_i são matrizes fixas de coeficientes, os quais são interpretados como a sensibilidade da de uma variável do modelo com relação a uma defasagem de outra variável;

$u_t = (u_{1t}, u_{2t}, \dots, u_{Kt})'$ é um vetor K -dimensional de ruído branco, ou seja, $E(u) = 0$, $E(u_t u_t') = \Sigma_u$ e $E(u_t u_s') = 0$.

Cada equação do modelo possui um termo específico, o qual pode ser interpretado como o choque correspondente à K -ésima equação. Este termo é o termo de erro u_{it} .

Qualquer modelo VAR(p) pode ser representado por um modelo Kp -dimensional VAR(1), dado por:

$$Y_t = \mathbf{v} + \mathbf{A}y_{t-1} + U_t, \quad (2.2)$$

onde,

$$Y_t := \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}, \quad \mathbf{v} := \begin{bmatrix} \nu \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{A} := \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I_K & 0 & \dots & 0 & 0 \\ 0 & I_K & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_K & 0 \end{bmatrix}, \quad U_t := \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

E as dimensões são: $Y_t \sim (Kp \times 1)$, $\mathbf{v} \sim (Kp \times 1)$, $\mathbf{A} \sim (Kp \times Kp)$ e $U_t \sim (Kp \times 1)$.

Y_t é estável se para $|z| \leq 1$,

$$\det(I_{Kp} - \mathbf{A}z) \neq 0$$

A média do modelo VAR(p) é dada por

$$\boldsymbol{\mu} := E(Y_t) = (I_{Kp} - \mathbf{A})^{-1}\mathbf{v}$$

E as auto-covariâncias são:

$$\Gamma_Y(h) = \sum_{i=0}^{\infty} \mathbf{A}^{h+i} \Sigma_U (\mathbf{A}^i)',$$

onde, $\Sigma_U = E(U_t U_t')$.

Usando a matriz J com dimensão $(K \times Kp)$ dada por:

$$J = [I_K \quad : \quad 0 \quad : \quad \dots \quad : \quad 0].$$

O processo y_t é obtido como $y_t = JY_t$. Pelo fato de Y_t ser um processo bem definido, isto implica que y_t também seja. Então, temos:

$$E(y_t) = J\boldsymbol{\mu}, \text{ constante para todo } t,$$

$$\Gamma_y(h) = J\Gamma_Y(h)J', \text{ também invariantes no tempo.}$$

Seja o polinômio característico do processo VAR(p), dado por:

$$\det(I_{Kp} - \mathbf{A}z) = \det(I_K - A_1 z - \dots - A_p z^p)$$

Dizemos que o processo é estável se, para $|z| \leq 1$,

$$\det(I_K - A_1 z - \dots - A_p z^p) \neq 0$$

Esta é a condição de estabilidade do processo.

2.2.2. Representação Média Móvel (MA) de um processo VAR

A representação de médias móveis é um útil instrumento para examinar a interação entre as variáveis. Chamam-se os coeficientes de A_i de função de resposta ao impulso observados a partir dos choques u_t das variáveis. Estas funções medem o impacto nas variáveis a partir de seus respectivos choques u_t .

Sob a condição de estabilidade, o processo Y_t tem a representação de Média Móvel (MA) definida como:

$$Y_t = \mu + \sum_{i=0}^{\infty} A^i U_{t-i} \quad (2.3)$$

onde, Y_t é expresso em termos do passado e presente do vetor de inovações U_t e do termo média μ .

Além disso, a representação MA de y_t pode ser encontrada pré-multiplicando Y_t por uma matriz $J = [I_K \quad 0 \quad \dots \quad 0]$ de dimensão $(K \times Kp)$. Aqui, $\mu = J\mu$, $\phi_i = JA^i J'$ e $u_t = JU_t$.

$$\begin{aligned} y_t &= JY_t = J\mu + \sum_{i=0}^{\infty} JA^i J' JU_{t-i} \\ &= \mu + \sum_{i=0}^{\infty} \Phi_i u_{t-i}. \end{aligned} \quad (2.4)$$

O problema de estimar estes parâmetros incorre nas mesmas dificuldades em obter os parâmetros do modelo primitivo a partir do modelo reduzido; a identificação do sistema. Esta metodologia não permite estimação se o sistema é sub-identificado, isto é, tenha um número de equações menores que os números de incógnitas.

Para ser possível identificar o sistema escrito na forma MA, é necessário usar a decomposição de Choleski.

2.2.3. Processo estacionário

Um processo estocástico é estacionário se seus momentos de primeira e segunda ordem são invariantes no tempo. Então, temos o primeiro momento, para todo t ,

$$E(y_t) = \mu$$

Esta condição indica que todos y_t possuem o mesmo vetor finito de média μ .

E o segundo momento, para todo t e $h = 0, 1$,

$$E[(y_t - \mu)(y_{t-h} - \mu)'] = \Gamma_y(h) = \Gamma_y(-h)'$$

E esta condição requer que as auto-covariâncias do processo não dependam do tempo t , mas somente do período h de tempo que separa os vetores y_t e y_{t-h} .

2.2.4. Estimação

Na estimação das equações do modelo VAR, o sistema apresenta uma estrutura fixa, com as mesmas variáveis em todas as equações e com o mesmo número de defasagens, sendo conhecido como “VAR puro”. O estimador para a representação padrão de um processo VAR(p) é definido por mínimos quadrados multivariados (ou mínimos quadrados generalizados) expresso da seguinte forma:

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (2.5)$$

Este resultado é bastante conhecido na literatura e largamente utilizado.

2.2.5. Previsão

Um dos principais objetivos de análise de séries temporais multivariadas é gerar previsão. No contexto dos modelos VAR, interessa-nos o modelo que minimiza o Erro Quadrático Médio (EQM) de previsão, que também é largamente utilizado como a função-perda na fase de estimação. Além disso, o EQM é um

estimador não-viciado da variância dos erros de previsão, útil para a construção de intervalos e regiões de previsão.

A equação de previsão num horizonte h na origem t é dada pelo valor esperado condicional:

$$E_t(y_{t+h}) := E(y_{t+h}|\Omega_t) = E(y_{t+h}|\{y_s|s \leq t\}). \quad (2.6)$$

Supondo normalidade dos termos de erro do modelo, isto é, u_t são normais multivariados com u_t e u_s independentes para $t \neq s$. Sob estas condições, os erros de previsão também são normalmente distribuídos como transformações lineares de vetores normais:

$$y_{t+h} - y_t(h) = \sum_{i=0}^{h-1} \Phi_i u_{t+h-i} \sim \mathcal{N}(0, \Sigma_y(h)) \quad (2.7)$$

Este resultado implica que os erros de previsão das componentes individuais são também normais, dados por:

$$\frac{y_{k,t+h} - y_{k,t}(h)}{\sigma_k(h)} \sim \mathcal{N}(0, 1),$$

onde, $y_{k,t}(h)$ é a k -ésima componente de $y_t(h)$ e $\sigma_k(h)$ é a raiz quadrada do k -ésimo elemento da diagonal principal da matriz $\Sigma_y(h)$.

Denotando por $z_{(\alpha)}$ o ponto crítico da distribuição normal, temos:

$$\begin{aligned} 1 - \alpha &= \Pr \left\{ -z_{(\alpha/2)} \leq \frac{y_{k,t+h} - y_{k,t}(h)}{\sigma_k(h)} \leq z_{(\alpha/2)} \right\} \\ &= \Pr \left\{ y_{k,t}(h) - z_{(\alpha/2)}\sigma_k(h) \leq y_{k,t+h} \leq y_{k,t}(h) + z_{(\alpha/2)}\sigma_k(h) \right\}. \end{aligned} \quad (2.8)$$

Então, o intervalo de previsão com $(1 - \alpha)\%$, h períodos a frente, para a k -ésima componente de y_t é:

$$y_{k,t}(h) \pm z_{(\alpha/2)}\sigma_k(h)$$

ou

$$[y_{k,t}(h) - z_{(\alpha/2)}\sigma_k(h), y_{k,t}(h) + z_{(\alpha/2)}\sigma_k(h)].$$

Este resultado pode ser estendido para regiões de previsão, para duas ou mais componentes. Defina a matriz $F = [I_N \quad 0]$ com dimensão $(N \times K)$ e note que temos um resultado conhecido de vetores normais multivariados, dado por:

$$[y_{t+h} - y_t(h)]' F' [F \Sigma_y(h) F']^{-1} F [y_{t+h} - y_t(h)] \sim \chi^2(N) \quad (2.9)$$

Assim, a distribuição $\chi^2(N)$ pode ser usada para determinar o elipsóide de previsão com $(1 - \alpha)\%$ para as primeiras N componentes do processo. Para tal, utiliza-se o método de Bonferroni, considerando que, para um intervalo de previsão com $(1 - \frac{\alpha}{N})\%$, a região de previsão resultante tem probabilidade no mínimo $(1 - \alpha)\%$ de conter todas as N variáveis conjuntamente.

2.3. Modelo VEC

Relações de equilíbrio são suspeitas entre muitas variáveis, especialmente econômicas. Suponha que as variáveis de interesse foram coletadas no vetor $y_t = (y_{1t}, y_{2t}, \dots, y_{Kt})$ e a sua relação de equilíbrio de longo prazo seja $\beta' y_t = \beta_1 y_{1t} + \dots + \beta_K y_{Kt} = 0$, onde $\beta = (\beta_1, \dots, \beta_K)'$. Em algum momento particular, essa relação pode não ser satisfeita exatamente, mas talvez tenhamos $\beta' y_t = z_t$, onde z_t é uma variável estocástica que representa os desvios do equilíbrio. Se realmente há um equilíbrio, parece plausível assumir que as variáveis y_t avançam juntas e que z_t é estável. Assim, z_t pode ser conduzido por uma tendência estocástica comum. Em outras palavras, não está excluído que cada variável é integrada, isto é, ainda existe uma combinação linear das variáveis. Variáveis integradas com essas propriedades são chamadas co-integradas.

No Gráfico 2.1, duas séries temporais co-integradas geradas artificialmente são representadas.

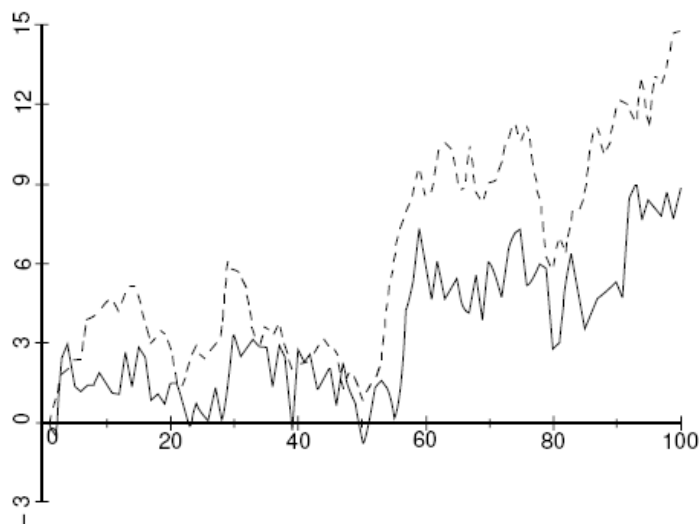


Gráfico 2.1: Séries temporais bivariadas co-integradas

Geralmente, as variáveis em um processo y_t K-dimensional são chamadas co-integradas de ordem (d, b) , brevemente, $y_t \sim CI(d, b)$, se todos os componentes do y_t são $I(d)$ e existe uma combinação linear $z_t = \beta' y_t$, com $\beta = (\beta_1, \dots, \beta_K)' \neq 0$ tal que z_t é $I(d - b)$. Por exemplo, se todos os componentes de z_t $I(1)$ e $\beta' y_t$ é estacionário $I(0)$, então $y_t \sim CI(1, 1)$. O vetor β é chamado de vetor de co-integração.

Um processo que consiste em variáveis co-integradas é chamado um processo co-integrado. Estes processos foram introduzidos por Granger (1981) e Engle & Granger (1987). Desde então, tornaram-se populares em trabalhos econométricos teóricos e aplicados.

Lütkepohl (2005) chama um processo K-dimensional y_t integrado de ordem d , brevemente, $y_t \sim I(d)$, se $\Delta^d y_t$ é estável e $\Delta^{d-1} y_t$ não é estável. O processo $I(d)$ é chamado co-integrado se existe uma combinação linear $\beta' y_t$ com $\beta \neq 0$, o qual é integrado de ordem menor que d . Se existir apenas um componente $I(d)$ no vetor y_t e todos os outros componentes forem estáveis $I(0)$, então o vetor y_t será $I(d)$ (Lütkepohl, 2005).

Estritamente falando, as variáveis endógenas de um modelo VAR devem ser todas estacionárias, ou seja, integradas de ordem zero, $I(0)$, portanto nesse contexto os autovalores da matriz A localizam-se todos no interior do círculo unitário. Se não forem, o modelo VEC deve ser estimado ao invés do VAR.

2.3.1. Formulação Matemática

Um processo K-dimensional VAR(p), expresso matematicamente por:

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t \quad (2.10)$$

é chamado co-integrado de *rank* r se

$$\mathbf{\Pi} := -(I_K - A_1 - \cdots - A_p)$$

tem *rank* r e assim $\mathbf{\Pi}$ pode ser escrito como uma matriz-produto $\alpha\beta'$ com α e β sendo de dimensão $(K \times r)$ e de *rank* r . A matriz β é chamada matriz de co-integração e α é chamada *loading matrix*.

Se $r = 0$, Δy_t tem uma representação VAR(p-1) e, para $r = k$,

$$|I_K - A_1 - \cdots - A_p| = |-\mathbf{\Pi}| \neq 0$$

e assim, o VAR não tem raiz unitária e y_t é um processo VAR(p) estável.

Deste modo, y_t tem uma representação VEC dada matematicamente por:

$$\begin{aligned} \Delta y_t &= \mathbf{\Pi} y_{t-1} + \mathbf{\Gamma}_1 \Delta y_{t-1} + \cdots + \mathbf{\Gamma}_{p-1} \Delta y_{t-p+1} + u_t \\ &= \alpha\beta' y_{t-1} + \mathbf{\Gamma}_1 \Delta y_{t-1} + \cdots + \mathbf{\Gamma}_{p-1} \Delta y_{t-p+1} + u_t, \end{aligned} \quad (2.11)$$

onde,

$$\mathbf{\Gamma}_i := -(A_{i+1} + \cdots + A_p), \quad i = 1, \dots, p-1.$$

Se esta representação de correção de erros for dada, é fácil descobrir o VAR correspondente, notando que:

$$\begin{aligned} A_1 &= \mathbf{\Pi} + I_K + \mathbf{\Gamma}_1 \\ A_i &= \mathbf{\Gamma}_i - \mathbf{\Gamma}_{i-1}, \quad i = 2, \dots, p-1, \\ A_p &= -\mathbf{\Gamma}_{p-1}. \end{aligned} \quad (2.12)$$

2.3.2. Estimação

A estimação dos parâmetros de um modelo VEC é feita em dois estágios. No primeiro, estima-se a matriz de co-integração β , por mínimos quadrados ou máxima verossimilhança e, em seguida, substitui-se o verdadeiro valor de β pelo seu estimador $\hat{\beta}$ na equação, ou seja, considerando:

$$\Delta y_t = \alpha \hat{\beta}' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t^* \quad (2.13)$$

E, assim, todos os outros parâmetros são estimados na segunda etapa, tendo os estimadores de dois estágios de α e Γ denotados por:

$$\hat{\alpha}_{2s} = \Delta Y M Y_{-1}' \hat{\beta} (\hat{\beta}' Y_{-1} M Y_{-1}' \hat{\beta})^{-1} \quad (2.14)$$

$$\hat{\Gamma}_{2s} = (\Delta Y - \hat{\alpha}_{2s} \hat{\beta}' Y_{-1}) \Delta X' (\Delta X \Delta X')^{-1} \quad (2.15)$$