向量自回归和自回归条件异方差模型

- 本章介绍时间序列的向量自回归和自回归条件异方差模型。
- **向量自回归模型**是自回归移动平均模型从单个时间序列到多个时间序列的扩展。
- **自回归条件异方差模型**主要考察时间序列数据波动性的变化，在金融领域的风险分析中有重要应用。
- 本章介绍这两种模型的意义和特征、参数估计、检验和应用等。

第一节 向量自回归模型

一、向量自回归模型概述

- ARMA模型分析针对单个时间序列，存在忽略经济变量之间内在联系的缺点。
- 克服这个缺点的方法是把ARMA模型扩展到针对多个时间序列，把ARMA模型中的变量换成向量。
- 因为自回归移动平均模型可相互转换，而且在向量变量的情况下自回归模型比较方便，因此一般主要考虑向量变量的自回归模型，称为"向量自回归模型"（Vector autoregression model，VAR）。

(一)VAR模型的表示形式

- 把*p*阶自回归模型AR(*p*)中的变量和误差项都变成向量，系数变成系数向量或矩阵，就得到一个*p*阶向量自回归模型VAR(*p*)：

$$\mathbf{Y}_t = \boldsymbol{\eta} + \boldsymbol{\Phi}_1 \mathbf{Y}_{t-1} + \cdots + \boldsymbol{\Phi}_p \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t$$

- 引进滞后算子，向量自回归模型可以写成

$$(\mathbf{I}_n - \boldsymbol{\Phi}_1 \mathbf{L} - \boldsymbol{\Phi}_2 \mathbf{L}^2 - \cdots - \boldsymbol{\Phi}_p \mathbf{L}^p)\mathbf{Y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t$$

或者

$$\boldsymbol{\Phi}(\mathbf{L})\mathbf{Y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t$$

引进下列记号：

$$\boldsymbol{\xi}_t = \begin{bmatrix} \mathbf{Y}_t - \boldsymbol{\mu} \\ \mathbf{Y}_{t-1} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{Y}_{t-p+1} - \boldsymbol{\mu} \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 & \cdots & \boldsymbol{\Phi}_{p-1} & \boldsymbol{\Phi}_p \\ \mathbf{I}_n & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ & & \vdots & & \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_n & \mathbf{0} \end{bmatrix}, \quad \mathbf{V}_t = \begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}$$

则p阶向量自回归模型VAR(p)，也可写成一阶自回归形式VAR(1)：

$$\boldsymbol{\xi}_t = \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{V}_t$$

- 其中$\mathbf{V}_t$ 的协方差矩阵为

$$E(\mathbf{V}_t\mathbf{V}_\tau') = \begin{cases} \mathbf{Q} & t = \tau \\ \mathbf{0} & t \neq \tau \end{cases}$$

$$\mathbf{Q} = \begin{bmatrix} \boldsymbol{\Omega} & 0 & \mathrm{L} & 0 \\ 0 & 0 & \mathrm{L} & 0 \\ & & \mathrm{M} & \\ 0 & 0 & \mathrm{L} & 0 \end{bmatrix}$$

- 为了分析的方便，也常常假设 $\boldsymbol{\varepsilon}_t$ 服从多元正态分布，即 $\boldsymbol{\varepsilon}_t \sim iidN[\mathbf{0}, \boldsymbol{\Omega}]$ ，其中 $\boldsymbol{\varepsilon}_t$ 可以包含一定自相关性，即 $\boldsymbol{\Omega}$ 是对称正定矩阵，但不一定是对角线矩阵。
- 当 $\boldsymbol{\varepsilon}_t$ 满足该假设时，上述向量自回归模型也称为"**高斯向量自回归模型**"。

- 向量自回归模型VAR(*p*) 展开，可以写成每个变量对常数项和向量中所有变量的1-*p*阶滞后项回归的，n个方程构成的联立方程组系统

$$Y_{1t}=\eta_1+\phi_{11}^{(1)}Y_{1,t-1}+\cdots+\phi_{1n}^{(1)}Y_{n,t-1}+\cdots+\phi_{11}^{(p)}Y_{1,t-p}+\cdots+\phi_{1n}^{(p)}Y_{n,t-p}+\varepsilon_{1t}$$

$$\cdots\cdots$$

$$Y_{nt}=\eta_n+\phi_{n1}^{(1)}Y_{1,t-1}+\cdots+\phi_{nn}^{(1)}Y_{n,t-1}+\cdots+\phi_{n1}^{(p)}Y_{1,t-p}+\cdots+\phi_{nn}^{(p)}Y_{n,t-p}+\varepsilon_{nt}$$

- 这个展开形式上与一般联立方程组模型相似，但其实有本质差异：

1、VAR模型不强调变量之间关系的理论根据，模型形式、变量、滞后期数等并不以特定经济理论为依据，模型变量也不存在内生、外生之分，每个方程都包含所有的变量；

2、VAR模型的主要作用是进行预测分析而不是经济结构分析；

3、由于模型结构性质的差别，VAR模型的参数估计和检验等与联立方程组模型也有差别。

（二）平稳性分析

1、平稳性定义

如果向量自回归过程 $\mathbf{Y}_t$ 的一阶矩 $E(\mathbf{Y}_t)$ 和二阶矩 $E(\mathbf{Y}_t\mathbf{Y}'_{t-j})$ 关于时期t都是独立的，则称为"协方差平稳的"，或者直接称"平稳的"。平稳意味着向量中包含的各个时间序列都是平稳的。

2、平稳性条件

- 利用一阶自回归形式 $\xi_t = \mathbf{F}\xi_{t-1} + \mathbf{V}_t$ 进行迭代可得

$$\xi_{t+s} = \mathbf{V}_{t+s} + \mathbf{F}\mathbf{V}_{t+s-1} + \mathbf{F}^2\mathbf{V}_{t+s-2} + \cdots + \mathbf{F}^{s-1}\mathbf{V}_{t+1} + \mathbf{F}^s\mathbf{V}_t$$

- 因为平稳过程意味着随着时间的不断增大，扰动的效应必须逐渐消失，因此上述向量自回归过程的平稳性要求随着s不断增大 $\mathbf{F}^s \to \mathbf{0}$

- 因此矩阵**F**的特征方程

$$\left|\mathbf{I}_n\lambda^p-\mathbf{\Phi}_1\lambda^{p-1}-\mathbf{\Phi}_2\lambda^{p-2}-\cdots-\mathbf{\Phi}_p\right|=0$$

的根都必须满足$|\lambda<1|$，也就是在单位圆内。

- 另一种表达方法:

满足方程 $\left|\mathbf{I}_n-\mathbf{\Phi}_1 z-\mathbf{\Phi}_2 z^2-\cdots-\mathbf{\Phi}_p z^p\right|=0$

的所有 z 值（根）都必须在单位圆之外。

(三)性质

1、均值

对于平稳的向量自回归模型$\mathbf{Y}_t$，两边求数学期望得：

$$\boldsymbol{\mu} = E(\mathbf{Y}_t) = \boldsymbol{\eta} + \boldsymbol{\Phi}_1\boldsymbol{\mu} + \cdots + \boldsymbol{\Phi}_p\boldsymbol{\mu}$$

可得到其均值为：

$$\boldsymbol{\mu} = [\mathbf{I}_n - \boldsymbol{\Phi}_1 - \cdots - \boldsymbol{\Phi}_p]^{-1}\boldsymbol{\eta}$$

2、MA(∞)表示方法

- 一个平稳的向量自回归过程可以写成一个白噪声向量的无限移动平均过程**MA(∞)**：

$$\mathbf{Y}_t=\boldsymbol{\mu}+\boldsymbol{\varepsilon}_t+\boldsymbol{\Psi}_1\boldsymbol{\varepsilon}_{t-1}+\boldsymbol{\Psi}_2\boldsymbol{\varepsilon}_{t-2}+\cdots=\boldsymbol{\mu}+\boldsymbol{\Psi}(\mathbf{L})\boldsymbol{\varepsilon}_t$$

- 其中 $\boldsymbol{\Psi}(\mathbf{L})$是上述移动平均表达式的滞后算子多项式，$\boldsymbol{\Psi}(\mathbf{L})$与自回归形式的滞后算子多项式 $\boldsymbol{\Phi}(\mathbf{L})$之间的关系为 $\boldsymbol{\Psi}(\mathbf{L})=[\boldsymbol{\Phi}(\mathbf{L})]^{-1}$，意味着

$$(\mathbf{I}_n-\boldsymbol{\Phi}_1\mathbf{L}-\boldsymbol{\Phi}_2\mathbf{L}^2-\cdots-\boldsymbol{\Phi}_p\mathbf{L}^p)(\mathbf{I}_n+\boldsymbol{\Psi}_1\mathbf{L}+\boldsymbol{\Psi}_2\mathbf{L}^2+\cdots)=\mathbf{I}_n$$

二、向量自回归模型参数估计

- 普通最小二乘估计能得到一致估计。
- 因为不同方程的误差之间有相关性，因此ML能得到更有效的估计
- 在模型误差项服从多元正态分布的前提下，模型参数向量的最大似然估计与最小二乘估计是相同的，只是误差方差的估计不同。

- 以误差向量服从多元高斯过程的高斯向量自回归模型为例，说明最大似然估计。
- 一个p阶高斯向量自回归模型即

其中 $$\mathbf{Y}_t = \boldsymbol{\eta} + \boldsymbol{\Phi}_1 \mathbf{Y}_{t-1} + \cdots + \boldsymbol{\Phi}_p \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t$$

$$\boldsymbol{\varepsilon}_t \sim iidN[\mathbf{0}, \boldsymbol{\Omega}]$$

- 如果已经得到向量时间序列 $\mathbf{Y}_t$ 的 $T+p$ 个时期的观测值，那么

$$f_{\mathbf{Y}_T,\cdots,\mathbf{Y}_1|\mathbf{Y}_0,\cdots,\mathbf{Y}_{-p+1}}(\mathbf{Y}_T,\cdots,\mathbf{Y}_1|\mathbf{Y}_0,\cdots,\mathbf{Y}_{-p+1};\boldsymbol{\theta})$$

- 因为 $\boldsymbol{\eta}+\boldsymbol{\Phi}_1\mathbf{Y}_{t-1}+\cdots+\boldsymbol{\Phi}_p\mathbf{Y}_{t-p}$ 在时期t为常数，而 $\boldsymbol{\varepsilon}_t \sim iidN[\mathbf{0},\boldsymbol{\Omega}]$，因此

$$\mathbf{Y}_t|\mathbf{Y}_{t-1},\mathbf{Y}_{t-2},\cdots,\mathbf{Y}_{-p+1} \sim N[\boldsymbol{\eta}+\boldsymbol{\Phi}_1\mathbf{Y}_{t-1}+\cdots+\boldsymbol{\Phi}_p\mathbf{Y}_{t-p},\boldsymbol{\Omega}]$$

- 引进记号：$\mathbf{X}_t = \begin{bmatrix} 1 \\ \mathbf{Y}_{t-1} \\ \vdots \\ \mathbf{Y}_{t-p} \end{bmatrix}$， $\mathbf{\Pi}' = [0 \quad \mathbf{\Phi}_1 \quad \cdots \quad \mathbf{\Phi}_p]$
- 上述 $\mathbf{Y}_t$ 服从的条件分布可以写成下列紧凑形式

$$\mathbf{Y}_t \big| \mathbf{Y}_{t-1}, \mathbf{Y}_{t-2}, \cdots, \mathbf{Y}_{-p+1} \sim N[\mathbf{\Pi}'\mathbf{X}_t, \mathbf{\Omega}]$$

- 因此第t个观测值的条件密度为

$$f_{\mathbf{Y}_t|\mathbf{Y}_{t-1},\cdots,\mathbf{Y}_{-p+1}}(\mathbf{Y}_t|\mathbf{Y}_{t-1},\cdots,\mathbf{Y}_{-p+1};\boldsymbol{\theta})$$

$$= (2\pi)^{-n/2} \left|\mathbf{\Omega}^{-1}\right|^{1/2} \exp[(-1/2)(\mathbf{Y}_t - \mathbf{\Pi}'\mathbf{X}_t)'\mathbf{\Omega}^{-1}(\mathbf{Y}_t - \mathbf{\Pi}'\mathbf{X}_t)]$$

向量自回归模型的（条件）似然函数为:

$$L(\boldsymbol{\theta}) = f_{\mathbf{Y}_T,\ldots,\mathbf{Y}_1|\mathbf{Y}_0,\ldots,\mathbf{Y}_{-p+1}}(\mathbf{Y}_T,\ldots,\mathbf{Y}_1|\mathbf{Y}_0,\ldots,\mathbf{Y}_{-p+1};\boldsymbol{\theta})$$

$$= \prod_{t=1}^{T} f_{\mathbf{Y}_t|\mathbf{Y}_{t-1},\ldots,\mathbf{Y}_{-p+1}}(\mathbf{Y}_t|\mathbf{Y}_{t-1},\ldots,\mathbf{Y}_{-p+1};\boldsymbol{\theta})$$

$$= (2\pi)^{-Tn/2}\left|\boldsymbol{\Omega}^{-1}\right|^{T/2}\exp\{(-1/2)\sum_{t=1}^{T}[(\mathbf{Y}_t-\boldsymbol{\Pi}'\mathbf{X}_t)'\boldsymbol{\Omega}^{-1}(\mathbf{Y}_t-\boldsymbol{\Pi}'\mathbf{X}_t)]\}$$

- 对数似然函数则为

$$\ln L(\boldsymbol{\theta}) = -(Tn/2)\ln(2\pi) + (T/2)\ln\left|\boldsymbol{\Omega}^{-1}\right|$$

$$-(1/2)\sum_{t=1}^{T}[(\mathbf{Y}_t-\boldsymbol{\Pi}'\mathbf{X}_t)'\boldsymbol{\Omega}^{-1}(\mathbf{Y}_t-\boldsymbol{\Pi}'\mathbf{X}_t)]$$

- 可以证明其中的最大似然估计实际上与最小二乘估计相同，为

$$\hat{\mathbf{\Pi}}' = (\sum_{t=1}^{T} \mathbf{Y}_t \mathbf{X}_t')(\sum_{t=1}^{T} \mathbf{X}_t \mathbf{X}_t')$$

其中$\hat{\mathbf{\Pi}}'$第j行就是$\mathbf{y}_{jt}$对$\mathbf{X}_t$回归得到的回归系数向量。

- $\mathbf{\Omega}$的最大似然估计则是 $\hat{\mathbf{\Omega}} = (1/T)\sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t'$

- 最大似然估计肯定是一致估计，其渐近分布是渐近正态分布

$$\sqrt{T}(\hat{\boldsymbol{\pi}}_T - \boldsymbol{\pi}) \xrightarrow{L} N[\mathbf{0}, (\boldsymbol{\Omega} \otimes \mathbf{Q}^{-1})]$$

其中 $\mathbf{Q} = E(\mathbf{X}_t \mathbf{X}_t')$， $\otimes$ 是克罗内克积。

- 第i个系数估计量 $\hat{\boldsymbol{\pi}}_{iT}$ 渐近分布

$$\sqrt{T}(\hat{\boldsymbol{\pi}}_{iT} - \boldsymbol{\pi}) \xrightarrow{L} N[\mathbf{0}, \sigma_i^2 \mathbf{Q}^{-1})]$$

- 如果 σ_i^2 由其一致估计 $\hat{\sigma}_i^2=(1/T)\sum_{t=1}^{T}\varepsilon_{it}^2$ 代，而 $\mathbf{Q}^{-1}$ 则由一致估计 $[(1/T)\sum_{t=1}^{T}\mathbf{X}_t\mathbf{X}_t']^{-1}$ 代，则可以将近似看作

$$\hat{\boldsymbol{\pi}}_i \sim N(\boldsymbol{\pi}_i,[\sum_{t=1}^{T}\mathbf{X}_t\mathbf{X}_t']^{-1})$$

- 当样本容量较大时，可以利用该渐近分布进行统计推断检验。

三、向量自回归模型检验

（一）需要通过模型检验加以确定的问题包括：

- 滞后期长度的确定；
- 模型的参数是否显著，是否存在特定约束关系等。

（二）似然比检验

$$\ln L(\boldsymbol{\theta}) = \ln L(\hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Pi}})$$

$$= -(Tn/2)\ln(2\pi) + (T/2)\ln\left|\hat{\boldsymbol{\Omega}}^{-1}\right| - (1/2)\sum_{t=1}^{T}(\hat{\boldsymbol{\varepsilon}}_t' \hat{\boldsymbol{\Omega}}^{-1} \hat{\boldsymbol{\varepsilon}}_t)$$

$$= -(Tn/2)\ln(2\pi) + (T/2)\ln\left|\hat{\boldsymbol{\Omega}}^{-1}\right| - (Tn/2)$$

似然比检验实际上就是把不同约束，有约束和无约束的参数估计、最大似然估计分别代入上述似然函数，根据是否有显著差异说明参数约束或者所对应的检验假设是否成立。

- H_0：一组变量数据由 p_0 阶而不是 $p_1 > p_0$ 阶滞后的高斯向量自回归生成。

  H_1：这组变量数据是由 $p_1 > p_0$ 阶滞后的高斯向量自回归生成。
- 分别在两个不同的假设下估计这个系统，也就是分别求每个变量关于常数项和系统中所有变量的 p_0 阶和 p_1 滞后的回归。

- 分别代入上述似然函数公式得到

$$L_0 = -(Tn/2)\ln(2\pi) + (T/2)\ln\left|\hat{\mathbf{\Omega}}_0^{-1}\right| - (Tn/2)$$

$$L_1 = -(Tn/2)\ln(2\pi) + (T/2)\ln\left|\hat{\mathbf{\Omega}}_1^{-1}\right| - (Tn/2)$$

- 得到似然比统计量为

$$2(L_1 - L_0) = T(\ln\left|\hat{\mathbf{\Omega}}_0^{-1}\right| - \ln\left|\hat{\mathbf{\Omega}}_1^{-1}\right|)$$

- Sims（1980）也提出了一种适应小样本情况的修正的似然比检验统计量

$$(T-k)(\ln\left|\hat{\Omega}_0^{-1}\right|-\ln\left|\hat{\Omega}_1^{-1}\right|)$$

- 在确定向量自回归模型的滞后期数等方面，进行参数估计回归时的SIC和AIC信息数标准同样也是重要的参考依据。一般来说，这两个指标较小的选择是较好的。

四、向量自回归模型的应用

（一）预测——向量自回归模型最重要的应用

- 由于向量自回归模型没有当期外生变量，因此在预测方面没有确定外生变量水平的困难，预测比较容易。
- 向量自回归模型的最优预测就是条件期望预测。

- 基于 $\mathbf{Y}_t,\mathbf{Y}_{t-1},\cdots$的 $\mathbf{Y}_{t+1}$的预测如下：

$$\hat{\mathbf{Y}}_{t+1|t}=\boldsymbol{\eta}+\boldsymbol{\Phi}_1\mathbf{Y}_{t-1}+\cdots+\boldsymbol{\Phi}_p\mathbf{Y}_{t-p}$$

- 作为一个动态系统，向量自回归模型很容易得到多步、多期预测。
- 例如作基于 $\mathbf{Y}_t,\mathbf{Y}_{t-1},\cdots$ 的 $\mathbf{Y}_{t+2}$ 、$\mathbf{Y}_{t+3}$ 等预测时，只要把前期预测作为变量值代入模型，就可以得到多步预测$\hat{\mathbf{Y}}_{t+2}$、$\hat{\mathbf{Y}}_{t+3}$等。这种方法的缺点是前期的预测误差会不断积累，导致长期预测的较大偏差。