状态空间模型与卡门滤器

State Space Models and Kalman Filters

林金龍

March 15, 2014

1 简介

有别于传统回归模型, 状态空间模型(state space model) 探讨影响可观察变量的状态变量行为, 例如在总体经济中, 恒常所得, 自然失业率, 潜在产出均为状态变量。藉由这些状态变量的推估, 我们可以探讨货币政策 (如泰勒法则) 如何建构; 也可以检验恒常所得假说是否成立。因此藉由状态空间模型的发展, 可以使我们对时间序列分析的进展有更深一步的了解。

2 模型

首先以local level model, 介绍状态间空间模型的估算程序。基本上, 状态空间模型的估算是一代迭 (recursive) 的过程。首先我们想要先说明的是, 由我们所观察的时间序列数据转换成状态空间模型, 状态空间模型的表示法不是唯一, 原因在于我们不知时间序列数据的结构为何, 其均为缩减式 (reduce form), 所以我们必须对结构进行猜测, 故时间序列模型转成状态空间模型会有多种。

接着我们开始介绍 local level model, 模型如下所示:[1]

$$y_t = \alpha_t + \epsilon_t, \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

$$\alpha_{t+1} = \alpha_t + \eta_t, \eta_t \sim N(0, \sigma_\eta^2)$$

上两式中,ϵ_t和η_t为mutually independent , 而且与α_1独立。α_1~N(a_1,P_1),a_1,P_1,σ_ϵ^2,σ_ϵ^2已知。α是状态变量, 因此我们无法从数据观察出来, 所以必须对其进行推估。在上述模型中, 有几点需要注意的是此模型所有参数均已知,α的动态方程式类似一 random walk 模型, 但与 random walk 模型不同的是,α并无法观察。

对α的推估有三种方式, 第一种为 forecasting, 即利用过去信息 (Y_{t-1}) 来进行推估。第二种为 filtering, 即利用过去到当期的所有信息 (Y_t) 来进行推估。第三种为 smoothing, 为利用全期的信息 (Y_n) 进行推估。[2]

[1]我们称下列第一个方程式为observation equation; 第二个方程式为 state equation。

[2]Y_{t-1}为过去的信息集合,Y_t为过去到当期的信息集合,Y_n为全期的信息集合。

在介绍估测方法之前, 我们先对y_t的一些统计性质作一介绍,

$$y \sim N(1a_1, \Omega), y = (y_1, \cdots, y_n)', 1 = (1, \cdots, 1)', \Omega = 11'P_1 + \Sigma$$

其中

$$\Sigma_{ij}=\begin{cases} \text{(i-1)}\sigma_\eta^2 \text{ ,i<j} \\ \sigma_\epsilon^2\text{+(i-1)}\sigma_\eta^2\text{, i=j,} \qquad \text{i,j=1,2,}\cdots\text{,n,} \\ \text{(j-1)}\sigma_\eta^2 \text{ ,i>j} \end{cases}$$

我们由上述可以知道,local level model 可以化作下式:

$$y_t = \alpha_1 + \Sigma_{j=1}^{t-1}\eta_j + \epsilon_t, \qquad t = 1, \cdots, n$$

在得知y_t的分配(常态分配) 之后, 我们便可以利用常态分配与条件常态分配的性质来对状态变量(α_t) 进行推估, 在此我们介绍filtering 和 smoothing 的估测过程。

2.1 Filtering

[3] Filtering, 即利用过去到当期的所有信息 (Y_t) 来对α进行推估。我们的目标是当第 t 期y_t的信息进入之后, 求出α的条件期望值与条件变异数为何。

首先我们定义α的条件期望值a_{t+1}=E($\alpha_{t+1}|Y_t$),α的条件变异数P_{t+1}=Var($\alpha_{t+1}|Y_t$)。由上述, 我们所定义的条件期望值

a_{t+1}=E($\alpha_{t+1}|Y_t$)=E($\alpha_t + \eta_t|Y_t$)=E($\alpha_t|Y_t$),

P_{t+1}=Var($\alpha_{t+1}|Y_t$)=Var($\alpha_t + \eta_t|Y_t$)=Var($\alpha_t|Y_t$)+σ_η^2,

即 a_{t+1}=E($\alpha_t|Y_t$);P_{t+1}=Var($\alpha_t|Y_t$)+σ_η^2。接着我们定义预测误差 (forecasting error),v_t=y_t-a_t和预测变异数(forecasting variance),F_t=$Var(v_t)$。因为v_t=y_t-a_t, 表示在第t期下,v_t的变动代表了y_t的变动, 即v_t所包含的讯息为第 t期y_t所包含的讯息, 据此, 我

[3]以下我们探讨的filtering 称为 Kalman Filter。

们可以得到:[4]

$\mathrm{E}(\alpha_t|Y_t)$=$\mathrm{E}(\alpha_t|Y_{t-1},v_t)$=$\mathrm{E}(\alpha_t|Y_{t-1})$+$\mathrm{Cov}(\alpha_t,v_t)Var(v_t)^{-1}v_t$, 其中

$$Cov(\alpha_t,v_t)=E[\alpha_t(y_t{-}a_t)]=E[\alpha_t(\alpha_t{+}\epsilon_t{-}a_t)]=E[\alpha_t(\alpha_t{-}a_t)]=E[Var(\alpha_t|Y_{t-1})]=P_t;$$

$$Var(v_t)=F_t=Var(\alpha_t+\epsilon_t-a_t)=Var\alpha_t|Y_{t-1}+Var(\epsilon_t)=P_t+\sigma_\epsilon^2$$

因为a_t=$\mathrm{E}(\alpha_t|Y_{t-1})$, 所以我们由上可以得出 $\mathrm{E}(\alpha_t|Y_t)$=a_t+P_t/F_tv_t=a_t+K_tv_t;
$\mathrm{Var}(\alpha_t|Y_t)$=$\mathrm{Var}(\alpha_t|Y_{t-1},v_t)$=$\mathrm{Var}(\alpha_t|Y_{t-1})$-$\mathrm{Cov}(\alpha_t,v_t)^2\mathrm{Var}(v_t)^{-1}$ =P_t-P_t^2/F_t=$P_t(1-K_t)$。

经由上述的计算过程, 我们可以得出变量之间彼此的关系如下:

$$v_t=y_t-a_t,\qquad F_t=P_t+\sigma_\epsilon^2,\qquad K_t=P_t/F_t,$$

$$a_{t+1}=a_t+K_tv_t,\qquad P_{t+1}=P_t(1-K_t)+\sigma_\eta^2$$

我们在模型中假设a_1,P_1,σ_ϵ^2与σ_η^2已知, 因此将这些数值带入, 迭代, 直到其收敛为止, 即可求出我们想要的估测值。

2.1.1 小结

总结以上进行的过程, 我们由观察的数据Y_t=$(y_1,\cdots,y_t)$', 且a_1,P_1,σ_ϵ^2与σ_η^2已知, 可以得出下列的步骤:

step1:

$$F_1=P_1+\sigma_\epsilon^2,\qquad K_1=P_1/F_1,\qquad v_1=y_1-a_1$$

[4]上述的期望值是利用以下的引理得出, 当x,y, 和 z 为随机变量, 且三个变量为联合常态分配时, 根据标准的多变量常态回归理论, 我们可以得出下列条件期望值和条件变异数:

$$E(x|y,z)=E(x|y)+\Sigma_{xz}\Sigma_{zz}^{-1}z,$$

$$Var(x|y,z)=Var(x|y)-\Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{xz}'$$

step2:

$$a_2 = a_1 + K_1 v_1 \Longrightarrow P_2 = P_1(1-K_1) + \sigma_\eta^2 \Longrightarrow F_2 = P_2 + \sigma_\epsilon^2$$

$$\Longrightarrow K_2 = P_2/F_2 \Longrightarrow v_2 = y_2 - a_2$$

step3:

$$\cdots$$

Recursive:

$$\Longrightarrow (a_1, \cdots, a_n), \qquad (P_1, \cdots, P_n), \qquad (F_1, \cdots, F_n)$$

2.2 Smoothing

在这一节, 我们探讨 Smoothing, 其为利用全期的信息 (Y_n) 进行对状态变量推估。我们的目标是当第全期 y_t的信息都得知之后之后, 再次回推估α。我们先定义状态估计误差 (state estimation error),x_t: x_t=α_t-a_t, 其变异数为: Var(x_t)=P_t。因为v_t=y_t-a_t =$\alpha_t + \epsilon_t - a_t$=$x_t + \epsilon_t$, 且$x_{t+1} = \alpha_{t+1} - a_{t+1} = \alpha_t + \eta_t - a_t - K_t v_t = x_t + \eta_t - K_t(x_t + \epsilon_t) = L_t x_t + \eta_t - K_t \epsilon_t$。

根据前述的作法, 我们可以得出P_{t+1}=Var(x_{t+1})=Cov(x_{t+1},α_{t+1})=Cov(x_{t+1},α_t+η_t)=L_tCov($x_t, \alpha_t + \eta_t$)+Cov($\eta_t, \alpha_t + \eta_t$)-$K_t$Cov($\epsilon_t, \alpha_t + \eta_t$)=$L_t P_t$+$\sigma_\eta^2$。

2.2.1 Smoothed State

首先我们必须先导出纳入全期 (Y_n) 信息下,α_t的条件期望值, 如前述,v_t的变动代表y_t的变动, 因此我们可以得出全期下的条件期望值, 如下所述:[5]

$$\hat{\alpha_t} = E(\alpha_t|Y_n) = E(\alpha_t|Y_{t-1}, v_t, \cdots, v_n)$$

$$= E(\alpha_t|Y_{t-1}) + Cov[\alpha_t, (v_t, \cdots, v_n)^{'}]Var[(v_t, \cdots, v_n)^{'}]^{-1}(v_t, \cdots, v_n)^{'}$$

[5]导出条件期望值与条件变异数的公式, 参见附注3

$$= a_t + \begin{pmatrix} Cov(\alpha_t, v_t) \\ \vdots \\ Cov(\alpha_t, v_n) \end{pmatrix}' \begin{bmatrix} F_t & & 0 \\ & \ddots & \\ 0 & & F_t \end{bmatrix}^{-1} \begin{pmatrix} v_t \\ \vdots \\ v_n \end{pmatrix}$$

$$= a_t + \sum_{j=t}^{n} Cov(\alpha_t, v_j) F_j^{-1} v_j$$

因为状态估计误差(state estimation error),x_t, 其变动代表α_t的变化, 因此上式的全期条件期望值, 式中的$Cov(\alpha_t, v_j)$=$Cov(x_t, v_j)$,j=t,$\cdots$,n。所以我们可以导出

Cov(x_t, v_t)=E($x_t(x_t + \epsilon_t)$)=Var(x_t)=P_t,

Cov(x_t, v_{t+1})=E($x_t(x_{t+1} + \epsilon_{t+1})$)=E[$x_t(L_t x_t + \eta_t - K_t\epsilon_t)$]=$P_t L_t$,

依据相同的方法, 我们可以得出Cov(x_t, v_{t+2})=$P_t L_t L_{t+1}$, $\cdots$,Cov(x_t, v_{t+n})=$P_t L_t L_{t+1} \ldots L_{n-1}$。

因此我们可以导出:

$$\hat{\alpha_t} = a_t + \sum_{j=t}^{n} Cov(\alpha_t, v_j) F_j^{-1} v_j = a_t + P_t \frac{v_t}{F_t} + P_t L_t \frac{v_{t+1}}{F_{t+1}} + \cdots = a_t + P_t r_{t-1}$$

其中r_{t-1}为预测误差v_t在t-1期之后的加权和。我们由r_{t-1}, 也可得出r_t的估计式, 因为在第n期之后, 我们没有观察值, 所以r_n=0, 将r_t代入r_{t-1}, 可以导出 r_t的递归公式:r_{t-1}=$F_t^{-1} + L_t r_t$。

由$\alpha_t = a_t + P_t r_{t-1}, r_{t-1} = F_t^{-1} + L_t r_t$,t=n,$\cdots$,1。

根据$r_n = 0$, 我们可以采 backward recursion 的方法, 利用观察值的全期信息, 对状态变量进行推估, 此种方法称为 state smoothing recursion。

2.2.2 Smoothed State Variance

根据前面的推估过程, 我们可以依照相同的逻辑来对状态变量的条件变异数进行推估, 如下所述:

$$V_t = Var(\alpha_t|y) = Var(\alpha_t|Y_{t-1}, v_t, \ldots, v_n) = Var(\alpha_t|Y_{t-1}) - Cov[\alpha_t, (v_t, \ldots, v_n)']$$
$$\times Var[(v_t, \ldots, v_n)']^{-1} Cov[\alpha_t, (v_t, \ldots, v_n)']'$$

$$
= P_t + \begin{pmatrix} Cov(\alpha_t, v_t) \\ \vdots \\ Cov(\alpha_t, v_n) \end{pmatrix}' \begin{bmatrix} F_t & & 0 \\ & \ddots & \\ 0 & & F_t \end{bmatrix}^{-1} \begin{pmatrix} Cov(\alpha_t, v_t) \\ \vdots \\ Cov(\alpha_t, v_n) \end{pmatrix}
$$

$$
= P_t - \sum_{j=t}^{n} [Cov(\alpha_t, v_j)]^2 F_j^{-1}
$$

根据前一节的作法, 我们可以求出Cov(α_t, v_j),j=t,. . .,n。并将其代入V_t, 得出V_t如下式:

$$
V_t = P_t - P_t^2 \frac{1}{F_t} - P_t^2 L_t^2 \frac{1}{F_{t+1}} - \cdots - P_t^2 L_t^2 L_{t+1}^2 \cdots L_{n-1}^2 \frac{1}{F_n}
$$

$$
= P_t - P_t^2 N_{t-1}
$$

其中N_{t-1}为预测误差变异数倒数的加权和。根据N_{t-1}, 我们可以导出N_t的估计式, 相同的, 在第n期之后, 因为没有任何的观察值, 因此 N_n=0。将N_t代入N_{t-1}中, 我们得出递归公式:N_{t-1}=$F_t^{-1} + L_t^2 N_t$。

由$V_t = P_t + P_t^2 N_{t-1}, N_{t-1} = F_t^{-1} + L_t^2 N_t$,t=n,· · ·,1。

根据$N_n = 0$, 我们亦采 backward recursion 的方法, 利用观察值的全期信息, 对状态变量进行推估, 此种方法称为 state variance smoothing recursion。

2.2.3 小结

我们由state smoothing recusion 与 state variance smoothing recursion, 给定适当的起始值, 经由代迭的程序, 可以推估出在全期的信息下,α_t的条件期望值与条件变异数。

3 Missing observations

状态空间模型的一个优点是可以处理缺漏值的问题。假设在 local level模型中, 我们观察到的y_j,j=τ,· · ·,$\tau^* - 1$, 为缺漏值, 一个处理的方法是, 定义一组新的时间序

列数据y_t^*, 其中 y_t^*=y_t,t=1,$\cdots$,$\tau-1$,y_t^*=$y_{t+\tau^*-\tau}$,t=τ,$\cdots$,n^*。我们由此一新定义的时间序列数据, 进行如前述的 filtering的推估过程。

另一个方法是, 依照原来的时间序列数据进行 filtering 的推估, 在τ,$\cdots$,τ^*-1这段期间, 对α的条件期望值之估测值均为相同, 唯一的差异在于在这段期间内, 每个估测值的条件变异数会随时间的增加而加大。

4 起始值的设定

这节我们探讨起始值的设定问题, 因为状态空间模型的估算为一代迭过程(recursive procedure), 所以必须设定起始值的大小, 来得出最终的估算结果。当然起始值的设定是任意的, 如之前所设定的a_1与P_1。在这边我们探讨另一设定起始值的方法, 即我们会假设α_1具有diffuse prior density, 其所指的是α_1的期望值a_1为任意设定, 但将其变异数P_1设定为无穷大($P_1 \to \infty$)。经由此设定, 我们得出a_2与P_2如下:

$$a_2 = a_1 + \frac{P_1}{P_1+\sigma_\epsilon^2}(y_1 - a_1)$$

$$P_2 = P_1(1 - \frac{P_1}{P_1+\sigma_\epsilon^2} + \sigma_\eta^2) = \frac{P_1}{P_1+\sigma_\epsilon^2}\sigma_\epsilon^2 + \sigma_\eta^2$$

当$P_1 \to \infty$时, 我们可以得出a_2=y_1,$P_2 = \sigma_\epsilon^2 + \sigma_\eta^2$, 将$a_2$与$P_2$持续代迭, 便可以得出我们要的估算结果, 我们称此一过程为diffuse initialisation of the Kalman filter。明显地, 我们可以看出P_1的设定并不会影响到后续的推估过程。

上面介绍的为filtering 时的起始值设定, 当然, 当我们引入全期的数据进行smoothing 的推估时, 亦可做相同的起始值假设, 不再赘述。

除了探讨a_1与P_1的设定之外,σ_ϵ^2与σ_η^2亦须进行设定, 在我们的模型中, 我们会利用MLE 法将其估计出来, 作为我们的起始值, 而这两个参数的估计在下节会进行说明。

5 估计

前述设定起始值时, 我们说明了a_1 与P_1的设定, 通常利用diffuse prior 的方法进

行设定。我们开始讲述关于σ_ϵ^2与σ_η^2的设定, 基本上此两参数的设定是利用MLE 法所估出来的。首先我们说明概似函数的设定, 考虑y_t的联合密度函数如下:

$$p(y_1,\ldots,y_n)=p(y_1)\prod_{t=2}^{n}p(y_t|Y_{t-1})$$

因为之前说明的预测误差 (v_t), 其变动代表观察值(y_t) 的变动, 是故我们可以将上式转换成v_t的函数, 如下所示:

$$p(v_1,\ldots,v_n)=\prod_{t=1}^{n}p(v_t)$$

其中

$$p(v_1)=p(y_1),\quad p(v_t)=p(y_t|Y_{t-1}),t=2,\ldots,n$$

接着由v_t为a_1与观察值(y_t) 的线性函数与其分配 [6] [7] 我们可以导出下列的概似函

[6]由filtering 的公式中, 我们可以得出v_t为a_1与观察值(y_t) 的线性函数:

$$v_1=y_1-a_1$$

$$v_2=y_2-a_1-K_1(y_1-a_1)$$

$$v_3=y_3-a_1-K_2(y_2-a_1)-K_1(1-K_2)(y_1-a_1)$$

$$\vdots$$

[7]我们亦可导出v_t的分配:

$$v=C(y-1a_1),v=(v_1,\ldots,v_n)'$$

其中

$$C=\begin{bmatrix}1&0&0&&0\\c_{21}&1&0&&0\\c_{31}&c_{32}&1&&0\\&&&\ddots&\vdots\\c_{n1}&c_{n2}&c_{n3}&\cdots&1\end{bmatrix},$$

$$c_{i,i-1}=-K_{i-1},$$

$$c_{ij}=-(1-K_{i-1})(1-K_{i-2})\cdots(1-K_{j+1})K_j,\qquad i=2,\ldots,n;j=1,\ldots,i-2$$

数:

$$logL = logp(y) = -\frac{n}{2}log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}(logF_t + \frac{v_t^2}{F_t})$$

藉由a_1与P_1的设定, 我们藉由numerically 方式得出σ_ϵ^2与σ_η^2的估测值。

6 残差检定

我们的模型中假设残差(ϵ_t, η_t) 是常态分配, 且为mutually indenpendent。因此对于估计所得出的误差进行检定, 检验其是否为常态分配, 是否独立。关于常态的检定方面, 我们可以利用 QQ-plot 法来判断残差是否为常态分配。

在检定残差是否具序列相关上, 我们可以利用 Ljung-Box 统计量,ACF,PACF 等方法来检定其是否存在序列相关。

当然还有许多检定常态分配与序列相关的方法, 我们在此只是提出几种方法来说明, 其余的检定法, 可以参阅上学期所教授的相关课程。

7 延伸

现在进入Durbin and Koopman(2002) 书中第四章, 在本章中模型为一线性 Gaussian 状态空间模型, 设定如下:

$$y_t = Z_t\alpha_t + \epsilon_t, \epsilon_t \sim N(0, H_t),$$

$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t, \eta_t \sim N(0, Q_t), \qquad t = 1, \ldots, n,$$

$$\alpha_1 \sim N(a_1, P_1,)$$

我们并将上设状态空间模型的变量与参数之维度(Dimensions) 列出, 如下表:

由上我们可以得出:

$$v \sim N(0, C\Omega C^{'}), \Omega = Var(y),$$

因为$v \sim N(0, F), F = C\Omega C^{'}$, 我们称$F = C\Omega C^{'}$为对称矩阵的Cholesky 分解。

状态空间模型的变量与参数之维度

Vector		Matrix	
y_t	$p \times 1$	Z_t	$p \times m$
α_t	$m \times 1$	T_t	$m \times m$
ϵ_t	$p \times 1$	H_t	$p \times p$
η_t	$r \times 1$	R_t	$m \times r$
		Q_t	$r \times r$
a_1	$m \times 1$	P_1	$m \times m$

7.1 Filtering

我们可以依照前述 filtering 的做法, 得出代迭过程 (recursive procedure), 推导过程如前, 不再赘述, 在此我们列出 filtering 的变量如下:

$$v_t = y_t - Z_t a_t, \qquad F_t = Z_t P_t Z_t' + H_t,$$

$$K_t = T_t P_t Z_t' F_t^{-1}, \qquad L_t = T_t - K_t Z_t, \qquad t = 1, \dots, n$$

$$a_{t+1} = T_t a_t - K_t v_t, \qquad P_{t+1} = T_t P_t L_t' + R_t Q_t R_t',$$

其中$L_t = T_t - K_t Z_t$, 给定适当的起始值, 则我们便可以进行代迭的过程, 得出我们想要知的估测值。

7.2 Smoothing

相同的, 在给定全期的信息下, 我们亦可依照前述 smoothing 的作法, 导出 state smoothing recursion 如下:

$$r_{t-1} = Z_t' F_t^{-1} + L_t' r_t, \qquad N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t,$$

$$\hat{\alpha}_t = a_t + P_t r_{t-1}, \qquad V_t = P_t - P_t N_{t-1} P_t,$$

$t = n, \ldots, 1$给定$r_n = 0$与$N_n = 0$ 之下, 我们便可依序代迭, 得出估测值。

8 相关软件

关于状态空间模型的估测有很多软件可以使用, 如R,Ox,...等, 我们非别介绍如下:

8.1 R packages about Kalman Filtering

KFM	This provides relatively flexible state space models via the fkf() function: state-space parameters are allowed to be time-varying and intercepts are included in both equations.
KFAS	provides a fast multivariate Kalman filter, smoother, simulation smoother and forecasting.
dlm	A package which also contains tools for converting other multivariate models into state space form.
MARSS	fits constrained and unconstrained multivariate autoregressive state-space models using an EM algorithm.

8.2 Ox

。关于 Ox 在状态空间模型的使用,Doornik(1998) 发表*Econometrics Journal*中的文章, 其文章中发展 SsfPack 套件, 应用在状态空间模型, 在此我们介绍其文章中关于状态空间模型的设定, 详细内容请参考 Doornik 的文章。[8]

我们说明文章中关于模型的设定, 并简单介绍一些用状态空间模型表示的例子。首先在文章中, 模型的设定如下:

$$\alpha_{t+1} = d_t + T_t \alpha_t + H_t \epsilon_t, \qquad \alpha_1 \sim N(a, P), \qquad t = 1, \ldots, n,$$

[8]同学可以下列网址下载此文章:

http://www.nuff.ox.ac.uk/economics/papers/1998/w6/ssf.pdf。

$$\theta_t = C_t + Z_t \alpha_t,$$

$$y_t = \theta_t + G_t \epsilon_t, \qquad \epsilon_t \sim N(0, I)$$

在 SsfPack 套件中, 上述的模型被表示如下进行估测:

$$\begin{pmatrix} \alpha_{t+1} \\ y_t \end{pmatrix} = \delta_t + \Phi \alpha_t + u_t, u_t \sim N(0, \Omega_t), t = 1, \ldots, n,$$

$$\delta_t = \begin{pmatrix} d_t \\ C_t \end{pmatrix}, \Phi_t = \begin{pmatrix} T_t \\ Z_t \end{pmatrix}, u_t = \begin{pmatrix} H_t \\ G_t \end{pmatrix} \epsilon_t, \Omega_t = \begin{pmatrix} H_t H_t' & H_t G_t' \\ G_t H_t' & G_t G_t' \end{pmatrix},$$

$$\alpha_1 \sim N(a, P)$$

其中δ_t为(m+N)×1,Φ_t为(m+N)× m ,Ω_t为(m+N)×(m+N)。在此模型中, 有 m 个状态变量 (state variable),N 个已观察变量 (observed variable),r 个残差 (distribance)。在 SsfPack 套件中, 给定适当的起始值, 便可进行 filtering 或 smoothing 的估测。

接着我们举一些时间序列相关的例子, 用状态空间模型来表示:

1. Local trend model:

考虑 Local trend model, 如下所示:

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

$$\beta_{t+1} = \beta_t + zeta_t, \zeta_t \sim NID(0, \sigma_\zeta^2)$$

$$y_t = \mu_t + \epsilon_t, \epsilon \sim NID(0, \sigma_\epsilon^2)$$

我们可以得出下式:

$$\begin{pmatrix} \mu_{t+1} \\ \beta_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + \begin{pmatrix} \eta_t \\ \zeta_t \\ \epsilon_t \end{pmatrix}$$

其中

$$\alpha_t = (\mu_t, \beta_t)', \Phi = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \Omega_t = \begin{pmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & \sigma_\zeta^2 & 0 \\ 0 & 0 & \sigma_\epsilon^2 \end{pmatrix}$$

2. Local level model:

考虑 Local level model, 如下所示:

$$y_t = \alpha_t + \epsilon_t, \epsilon_t \sim (0, \sigma_\epsilon^2)$$

$$\alpha_{t+1} = \alpha_t + \eta_t, \eta_t \sim N(0, \sigma_\eta^2)$$

由上可以得出:

$$\begin{pmatrix} \alpha_{t+1} \\ y_t \end{pmatrix} = 0 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \alpha_t + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix}$$

其中,

$$\alpha_t = \alpha_t, \Phi = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \Omega = \begin{pmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_\eta^2 \end{pmatrix}$$

3. ARMA(p,q) Process:

考虑 ARMA(1,1) Process, 如下所示:

$$y_t = \phi_1 y_{t-1} + \ldots + \phi_p y_{t-1} + \xi_t + \theta_1 \xi_{t-1} + \ldots + \theta_q \xi_{t-q}, \xi_t \sim NID(0, \sigma_\xi^2)$$

首先我们考虑其VAR(1) 的表现型式如下:

$$\alpha_{t+1} = \begin{pmatrix} \phi_1 & 1 & 0 & \cdots & 0 \\ \phi_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \\ \phi_{m-1} & 0 & 0 & \cdots & 1 \\ \phi_m & 0 & 0 & \cdots & 0 \end{pmatrix} \alpha_t + \begin{pmatrix} 1 \\ \theta_1 \\ \vdots \\ \theta_{m-2} \\ \theta_{m-1} \end{pmatrix} \xi_t, \xi_t \sim NID(0, \sigma_\xi^2)$$

$$m = max(p, q+1), y_t = (1, \ldots, 0)\alpha_t$$

上式可以用Doornik 的文章的模型表示, 即为$\alpha_{t+1} = T_\alpha + h\xi_t$, 可用Ssfpack 套件进行估测, 上式为 ARMA 模型表示法的一般化, 以下我们在举几个 ARMA 模型的例子, 以便同学了解。

✓ARMA(1,1) Process:

由$m = max(p, q+1) = max(1, 2) = 2$, 可以将 ARMA(1,1) Process 表示如下:

$$\begin{pmatrix} \alpha_{1,t+1} \\ \alpha_{2,t+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_{1,t} \\ \alpha_{2,t} \end{pmatrix} + \begin{pmatrix} 1 \\ \theta_1 \end{pmatrix} \xi_t$$

$$\alpha_{1,t+1} = \phi_1 \alpha_{1,t} + \alpha_{2,t} + \xi_t$$

$$\alpha_{2,t+1} = \theta_1 \xi_t$$

$$y_t = (1, 0)\alpha_t = \alpha_{1,t}$$

✓ARMA(2,1) Process:

由$m = max(p, q+1) = max(2, 2) = 2$, 可以将 ARMA(2,1) Process 表示如下:

$$\begin{pmatrix} \alpha_{1,t+1} \\ \alpha_{2,t+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{pmatrix} \begin{pmatrix} \alpha_{1,t} \\ \alpha_{2,t} \end{pmatrix} + \begin{pmatrix} 1 \\ \theta_1 \end{pmatrix} \xi_t$$

$$\alpha_{1,t+1} = \phi_1 \alpha_{1,t} + \alpha_{2,t} + \xi_t$$

$$\alpha_{2,t+1} = \phi_2 \alpha_{1,t} + \theta_1 \xi_t$$

$$y_t = \alpha_{1,t}$$

✓ARMA(3,2) Process:

由$m = max(p, q+1) = max(3, 3) = 3$, 可以将 ARMA(3,2) Process 表示如下:

$$\begin{pmatrix} \alpha_{1,t+1} \\ \alpha_{2,t+1} \\ \alpha_{3,t+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & 1 & 0 \\ \phi_2 & 0 & 1 \\ \phi_3 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_{1,t} \\ \alpha_{2,t} \\ \alpha_{3,t} \end{pmatrix} + \begin{pmatrix} 1 \\ \theta_1 \\ \theta_2 \end{pmatrix} \xi_t$$

$$\alpha_{1,t+1} = \phi_1 \alpha_{1,t} + \alpha_{2,t} + \xi_t$$

$$\alpha_{2,t+1} = \phi_2 \alpha_{1,t} + \alpha_{3,t} + \theta_1 \xi_t$$

$$\alpha_{3,t+1} = \phi_3 \alpha_{1,t} + \theta_2 \xi_t$$

$$y_t = \alpha_{1,t}$$

9 参考文献

Durbin, J., and S. J. Koopman (2002), *Time Series Analysis by State Space Methods*, New York:Oxford University Press.

Koopman, S. J., Shephard, N., and J. A. Doornik (1998),“ Statistical algorithms for models in state space using SsfPack 2.2,”*Econometrics Journal*, 1, 1-55.