2024 秋季本科时间序列

第 5 次作业答案

11 月 11 日

1. 给定平稳的 AR(2) 过程 $X_t = \mu+\phi_1 X_{t-1}+\phi_2 X_{t-2}+\epsilon_t$, 我们要推导通项表达式 $\hat{X}_{t+s|t}, \forall s \geq 1$，其中 $Y_t = X_t - \mathbb{E}[X_t]$ 为去均值后的过程。

   X_t 是平稳的 AR(2) 过程，可以表示为：

   $$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

   其中，μ 是常数项，ϵ_t 是白噪声，方差为 σ_ϵ^2。定义去均值后的过程 $Y_t = X_t - \mathbb{E}[X_t] = X_t - \frac{\mu}{1-\phi_1-\phi_2}$。

   给定：

   $$\begin{bmatrix} Y_{t+1} \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} Y_t \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}$$

   这个表示式表明，Y_t 的演化过程由上一期的 Y_t 和 Y_{t-1} 以及当前的白噪声 ϵ_t 决定。

   在给定 t 时，我们希望推导出 $\hat{X}_{t+s|t}$，即在时刻 t 时对时刻 $t+s$ 的预测。

   由于 $Y_t = X_t - \frac{\mu}{1-\phi_1-\phi_2}$，我们可以通过对 Y_t 的预测来推导 X_t 的预测。首先，考虑 $\hat{Y}_{t+s|t}$，这是对 Y_{t+s} 在时刻 t 的条件预测。

   $$\begin{bmatrix} Y_{t+1|t} \\ Y_{t|t} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} Y_{t|t} \\ Y_{t-1|t} \end{bmatrix}$$

对于 $s \geq 1$，递推可得 $Y_{t+s|t}$ 的表达式：

$$\hat{Y}_{t+s|t} = \phi_1^s Y_t + \sum_{i=0}^{s-1} \phi_1^{s-i-1} \phi_2^i Y_{t-i}$$

由 $X_t = Y_t + \frac{\mu}{1-\phi_1-\phi_2}$，我们可以通过预测 Y_t 得到：

$$\hat{X}_{t+s|t} = \hat{Y}_{t+s|t} + \frac{\mu}{1-\phi_1-\phi_2}$$

即为，

$$\hat{X}_{t+s|t} = \phi_1^s X_t + \sum_{i=0}^{s-1} \phi_1^{s-i-1} \phi_2^i X_{t-i} + \frac{\mu}{1-\phi_1-\phi_2}$$

得到：

$$\hat{X}_{t+s|t} = \phi_1^s X_t + \sum_{i=0}^{s-1} \phi_1^{s-i-1} \phi_2^i X_{t-i} + \frac{\mu}{1-\phi_1-\phi_2}$$

这个公式表示了从时刻 t 对 $t+s$ 时刻 X_t 的最优预测，其中 ϕ_1 和 ϕ_2 是 AR(2) 过程的参数，μ 是常数项。

2. 要计算 $L(X_{t+s}|X_t, X_{t-1}) = a_s X_t + b_s X_{t-1}$ 的系数 a_s 和 b_s，我们需要解决以下最小化问题：

$$\min_{a_s, b_s} E\left[(X_{t+s} - a_s X_t - b_s X_{t-1})^2\right]$$

这相当于解以下的正常方程：

$$\begin{cases} E\left[(X_{t+s} - a_s X_t - b_s X_{t-1})X_t\right] = 0 \\ E\left[(X_{t+s} - a_s X_t - b_s X_{t-1})X_{t-1}\right] = 0 \end{cases}$$

首先，计算过程 X_t 的自协方差函数 $\gamma_k = \mathrm{Cov}(X_t, X_{t-k})$：

$$\gamma_0 = E[X_t^2] = (1+\theta^2)\sigma^2, \quad \gamma_1 = E[X_t X_{t-1}] = \theta\sigma^2, \quad \gamma_k = 0 \quad \text{对于} |k| \geq 2$$

计算 $E[X_{t+s}X_t]$ 和 $E[X_{t+s}X_{t-1}]$：

- 对于 $s = 1$：

$$E[X_{t+1}X_t] = \theta\sigma^2, \quad E[X_{t+1}X_{t-1}] = 0$$

- 对于 $s \geq 2$：

$$E[X_{t+s}X_t] = 0, \quad E[X_{t+s}X_{t-1}] = 0$$

对于 $s = 1$，正常方程为：

$$\begin{cases} E[X_{t+1}X_t] = a_1 E[X_t^2] + b_1 E[X_{t-1}X_t] \\ E[X_{t+1}X_{t-1}] = a_1 E[X_t X_{t-1}] + b_1 E[X_{t-1}^2] \end{cases}$$

将已知值代入：

$$\begin{cases} \theta\sigma^2 = a_1(1+\theta^2)\sigma^2 + b_1\theta\sigma^2 \\ 0 = a_1\theta\sigma^2 + b_1(1+\theta^2)\sigma^2 \end{cases}$$

消去 σ^2，得到：

$$\begin{cases} \theta = a_1(1+\theta^2) + b_1\theta \\ 0 = a_1\theta + b_1(1+\theta^2) \end{cases}$$

解第二个方程，得到 $b_1 = -\frac{a_1\theta}{1+\theta^2}$。将其代入第一个方程，求解 a_1：

$$\theta = a_1(1+\theta^2) - \frac{a_1\theta^2}{1+\theta^2}$$

化简得到：

$$a_1 = \frac{\theta(1+\theta^2)}{1+\theta^2+\theta^4}$$

然后计算 b_1：

$$b_1 = -\frac{a_1\theta}{1+\theta^2} = -\frac{\theta^2}{1+\theta^2+\theta^4}$$

对于 $s \geq 2$，由于相关系数为零，因此 $a_s = 0$，$b_s = 0$。

课件中给出的最优预测器为:

$$\hat{X}_{t+1|t} = \theta \sum_{j=0}^{\infty} (-\theta)^j X_{t-j}$$

展开前两项:

$$\hat{X}_{t+1|t} = \theta X_t - \theta^2 X_{t-1} + \theta^3 X_{t-2} - \cdots$$

使用所有过去数据时，系数为:

- X_t 的系数: θ
- X_{t-1} 的系数: $-\theta^2$
- X_{t-2} 的系数: θ^3
- ...

而使用仅 X_t 和 X_{t-1} 时，系数为:

- X_t 的系数: $a_1 = \frac{\theta(1+\theta^2)}{1+\theta^2+\theta^4}$
- X_{t-1} 的系数: $b_1 = -\frac{\theta^2}{1+\theta^2+\theta^4}$

对于任意 $s \geq 1$:

- 当 $s = 1$ 时，线性预测系数为:

$$a_1 = \frac{\theta(1+\theta^2)}{1+\theta^2+\theta^4}, \quad b_1 = -\frac{\theta^2}{1+\theta^2+\theta^4}$$

- 当 $s \geq 2$ 时，系数均为零:

$$a_s = 0, \quad b_s = 0$$

与课件中使用全部过去数据的最优预测相比，只有当 $s = 1$ 时，系数 a_1 和 b_1 与对应项有所不同。使用全部过去数据时，预测器的系数为 $\theta, -\theta^2, \theta^3$ 等，包含更多的滞后项，因此预测效果更好。

3. 均值向量 $\boldsymbol{\mu} = (\mu_X, \mu_Y)^\intercal$, 协方差矩阵 $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix}$, 设 ρ 为相关系数, 定义为: $\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

由多元正态分布定义有

$$f(x,y) = \frac{1}{2\pi|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

其中 $\boldsymbol{x} = (x,y)^\intercal$, $|\boldsymbol{\Sigma}|$ 表示协方差矩阵的行列式，有 $|\boldsymbol{\Sigma}| = \sigma_X^2\sigma_Y^2 - (\sigma_{XY})^2 = \sigma_X^2\sigma_Y^2(1-\rho^2)$, $\boldsymbol{\Sigma}^{-1} = \frac{1}{|\boldsymbol{\Sigma}|}\begin{bmatrix} \sigma_Y^2 & -\sigma_{XY} \\ -\sigma_{XY} & \sigma_X^2 \end{bmatrix}$，代入 $f(x,y)$ 得到:

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]\right\}$$

接下来计算 X 的边缘密度函数 $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y$，令 $A = \frac{x-\mu_X}{\sigma_X}, \quad B = \frac{y-\mu_Y}{\sigma_Y}$，则指数部分变为:

$$\begin{aligned} &-\frac{1}{2(1-\rho^2)}[A^2 - 2\rho AB + B^2] \\ =&-\frac{1}{2(1-\rho^2)}\left[(B-\rho A)^2 + (1-\rho^2)A^2\right] \end{aligned}$$

所以

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{A^2}{2}\right)\exp\left(-\frac{(B-\rho A)^2}{2(1-\rho^2)}\right)$$

对 y 积分，即对 B 积分，令 $C = \frac{B-\rho A}{\sqrt{1-\rho^2}}$

$$\begin{aligned} f_X(x) &= \exp\left(-\frac{A^2}{2}\right)\cdot\frac{1}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}\exp\left(-\frac{(B-\rho A)^2}{2(1-\rho^2)}\right)\sigma_Y\mathrm{d}B \\ &= \exp\left(-\frac{A^2}{2}\right)\cdot\frac{1}{\sqrt{2\pi}\sigma_X}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}\exp\left(-\frac{(B-\rho A)^2}{2(1-\rho^2)}\right)\sigma_Y\sqrt{1-\rho^2}\,\mathrm{d}C \end{aligned}$$

注意到积分为标准正态分布的积分，结果为 1，故

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi}\sigma_X}\exp\left(-\frac{A^2}{2}\right) \\ &= \frac{1}{\sqrt{2\pi}\sigma_X}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right\} \end{aligned}$$

再计算 Y 的条件密度函数 $f(y|x) = \frac{f(x,y)}{f_X(x)}$，代入已知的 $f(x,y)$ 和 $f_X(x)$，化简可得（注

意其仍然是正态分布）

$$
\begin{aligned}
f(y|x) &= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}\exp\left(-\frac{(B-\rho A)^2}{2(1-\rho^2)}\right) \\
&= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}\exp\left\{-\frac{\left[y-\left(\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right)\right]^2}{2\sigma_Y^2(1-\rho^2)}\right\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_{Y|X}}\exp\left\{-\frac{1}{2}\left(\frac{y-\mathbb{E}(Y|X=x)}{\sigma_{Y|X}}\right)^2\right\}
\end{aligned}
$$

所以 $\mathbb{E}(Y|X=x)=\mu_Y+\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)$

4. (a) i.

$$
\begin{aligned}
\boldsymbol{P}_X\boldsymbol{\xi} &= \boldsymbol{X}(\boldsymbol{X}^\intercal\boldsymbol{X})^{-1}\boldsymbol{X}^\intercal\boldsymbol{X}\boldsymbol{a} \\
&= \boldsymbol{X}\boldsymbol{I}\boldsymbol{a} \\
&= \boldsymbol{X}\boldsymbol{a} \\
&= \boldsymbol{\xi}
\end{aligned}
$$

ii.

$$
\begin{aligned}
(\boldsymbol{X}\boldsymbol{a})^\intercal(\boldsymbol{I}-\boldsymbol{P}_X)\boldsymbol{\zeta} &= \boldsymbol{a}^\intercal\boldsymbol{X}^\intercal(\boldsymbol{\zeta}-\boldsymbol{X}(\boldsymbol{X}^\intercal\boldsymbol{X})^{-1}\boldsymbol{X}^\intercal\boldsymbol{\zeta}) \\
&= \boldsymbol{a}^\intercal\boldsymbol{X}^\intercal\boldsymbol{\zeta}-\boldsymbol{a}^\intercal\boldsymbol{I}\boldsymbol{X}^\intercal\boldsymbol{\zeta} \\
&= 0
\end{aligned}
$$

(b) i. 由 OLS 估计系数表达式可知，$\boldsymbol{Y}$ 对 $\boldsymbol{W}$ 回归的系数为 $(\boldsymbol{W}^\intercal\boldsymbol{W})^{-1}\boldsymbol{W}^\intercal\boldsymbol{Y}$，而 $\boldsymbol{Z}$ 对 $\boldsymbol{W}$ 回归的系数为 $(\boldsymbol{W}^\intercal\boldsymbol{W})^{-1}\boldsymbol{W}^\intercal\boldsymbol{Z}$，故两组回归对应的残差向量分别为

$$
(\boldsymbol{I}-\boldsymbol{W}(\boldsymbol{W}^\intercal\boldsymbol{W})^{-1}\boldsymbol{W}^\intercal)\boldsymbol{Y}=(\boldsymbol{I}-\boldsymbol{P}_W)\boldsymbol{Y}=\tilde{\boldsymbol{Y}}
$$
$$
(\boldsymbol{I}-\boldsymbol{W}(\boldsymbol{W}^\intercal\boldsymbol{W})^{-1}\boldsymbol{W}^\intercal)\boldsymbol{Z}=(\boldsymbol{I}-\boldsymbol{P}_W)\boldsymbol{Z}=\tilde{\boldsymbol{Z}}
$$

ii. $\hat{\boldsymbol{e}}=\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}=\boldsymbol{Y}-\boldsymbol{X}(\boldsymbol{X}^\intercal\boldsymbol{X})^{-1}\boldsymbol{X}^\intercal\boldsymbol{Y}=(\boldsymbol{I}-\boldsymbol{P}_X)\boldsymbol{Y}$，由 (a) 的结论可得 $\hat{\boldsymbol{e}}$ 与 $\boldsymbol{X}$ 的每个列向量互相垂直，故垂直于 $\boldsymbol{Z}$ 与 $\boldsymbol{W}$ (i.e. $\boldsymbol{Z}^\intercal\hat{\boldsymbol{e}}=\boldsymbol{W}^\intercal\hat{\boldsymbol{e}}=0$)

iii. $\boldsymbol{P}_W\hat{\boldsymbol{e}}=\boldsymbol{W}(\boldsymbol{W}^\intercal\boldsymbol{W})^{-1}\boldsymbol{W}^\intercal\hat{\boldsymbol{e}}=0$ (由 ii. 可知 $\boldsymbol{W}^\intercal\hat{\boldsymbol{e}}=0$)

iv. 只需要说明 $\hat{\boldsymbol{\delta}}$ 正好满足回归 $\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{Z}}\boldsymbol{\delta} + \boldsymbol{u}$ 的 OLS 系数估计所满足的一阶条件 $\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{Z}}\boldsymbol{\delta} = 0$.

首先注意到 $\hat{\boldsymbol{e}} = \boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{Y} - \boldsymbol{Z}\hat{\boldsymbol{\delta}} - \boldsymbol{W}\hat{\boldsymbol{\theta}}$，两端同时乘以 $\boldsymbol{I} - \boldsymbol{P}_{\mathrm{W}}$ 有

$$
\begin{aligned}
(\boldsymbol{I} - \boldsymbol{P}_{\mathrm{W}})\hat{\boldsymbol{e}} &= (\boldsymbol{I} - \boldsymbol{P}_{\mathrm{W}})\boldsymbol{Y} - (\boldsymbol{I} - \boldsymbol{P}_{\mathrm{W}})\boldsymbol{Z}\hat{\boldsymbol{\delta}} - (\boldsymbol{I} - \boldsymbol{P}_{\mathrm{W}})\boldsymbol{W}\hat{\boldsymbol{\theta}} \\
\hat{\boldsymbol{e}} &= \tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{Z}}\hat{\boldsymbol{\delta}} \quad (\text{由 i.-iii. 可知})
\end{aligned}
$$

再在两端同乘 $\tilde{\boldsymbol{Z}}^\intercal$ 可得

$$
\begin{aligned}
\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{Z}}\hat{\boldsymbol{\delta}} &= \tilde{\boldsymbol{Z}}^\intercal(\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{Z}}\hat{\boldsymbol{\delta}}) \\
&= \tilde{\boldsymbol{Z}}^\intercal\hat{\boldsymbol{e}} \\
&= 0
\end{aligned}
$$

所以 (1) 中 OLS 估计的结果 $\hat{\boldsymbol{\delta}}$ 同样满足回归 $\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{Z}}\boldsymbol{\delta} + \boldsymbol{u}$ 的一阶条件，得证

(c) 题干中回归可改写为该矩阵形式 $\boldsymbol{Y} = \boldsymbol{1}\alpha + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}$，其中 $\boldsymbol{1}$ 为每个元素均为 1 的 $T \times 1$ 列向量，同上定义回归残差 $\tilde{\boldsymbol{Y}}$ 和 $\tilde{\boldsymbol{X}}$，可推得

$$
\begin{aligned}
\tilde{\boldsymbol{Y}} &= (\boldsymbol{1} - \boldsymbol{P}_1)\boldsymbol{Y} \\
&= \boldsymbol{Y} - \boldsymbol{1}(\boldsymbol{1}^\intercal\boldsymbol{1})^{-1}\boldsymbol{1}^\intercal\boldsymbol{Y} \\
&= \boldsymbol{Y} - \boldsymbol{1}T^{-1}(Y_1 + \cdots + Y_T) \\
&= \boldsymbol{Y} - \boldsymbol{1}\overline{Y} \\
&= \boldsymbol{Y} - \overline{\boldsymbol{Y}} \\
\tilde{\boldsymbol{X}} &= \boldsymbol{X} - \overline{\boldsymbol{X}}
\end{aligned}
$$

因此由 FWL 定理可知，去均值后的回归 $\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{X}}\boldsymbol{\beta} + \boldsymbol{u}$ 得到的 OLS 估计系数 $\hat{\boldsymbol{\beta}}$ 与原回归得到的估计值等价