2024 秋季本科时间序列

第 4 次作业

提交日期：10 月 28 日

1. 考虑平稳 AR(2) 过程 $X_t = \mu + \phi X_{t-1} + \psi X_{t-2} + \varepsilon_t$，$t = 1, \ldots, T$，$\varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$，$\phi, \psi$ 取值满足平稳性条件。

   (a) 计算随机向量 $\boldsymbol{Z}_t = [1, X_{t-1}, X_{t-2}]^\intercal$ 的交叉二阶矩矩阵 $\boldsymbol{M} = \mathbb{E}\boldsymbol{Z}_t\boldsymbol{Z}_t^\intercal$，并说明该矩阵满秩。

   (b) 固定 $\mu = 1$，考虑 3 组参数 $(\phi^{(i)}, \psi^{(i)}) = (1.2, -0.36), (1.3, -0.36), (1.8, -0.81)$。给定第 i 组参数，请在 R 或 Python 中编程，生成 $X_t^{(i)}$ 的随机模拟序列，$T = 1000$，起始值为 $X_{-1}^{(i)} = X_0^{(i)} = \mathbb{E}^{(i)}X_t$。绘图展示这 3 个模拟序列，仿照第 5 讲课件使用 R 的 `spec.pgram` 函数（或 Python 对应函数）计算 3 个时间序列的样本谱密度函数的估计，并与其理论谱密度函数进行对比。提示：`spec.pgram` 函数估计样本谱密度，需要对设定 `span` 与 `taper` 两个参数，请自行调整优化，使得样本谱密度估计值较为平滑。

   (c) 固定 (b) 中 3 个模拟序列不变，对每个序列 $\left\{X_t^{(i)}\right\}$，按照每新增 100 个样本，进行 OLS 估计；即 $k = 1, \ldots, 10$，对样本 $\left\{X_t^{(i)} : t = 1, \ldots, 100k\right\}$ 进行 OLS 估计，估计结果记为 $\hat{\boldsymbol{\beta}}^{(i),k}$。对每一个 i，绘制 $\hat{\boldsymbol{\beta}}^{(i),k}$ 关于 k 的序列图。注意该向量有 3 个元素，分开绘图。请讨论随着 ϕ 的增大，OLS 估计随样本量 $100k$ 增加时的收敛性有何变化。

   (d) 给定 (b) 中参数取值，$i = 1, 2, 3$，并定义 $\boldsymbol{\beta}_0^{(i)} = [\mu, \phi^{(i)}, \psi^{(i)}]^\intercal$。按照每次生成 $T = 100$ 个 $\left\{X_t^{(i)}\right\}$ 模拟样本的方式，重复生成 1000 组样本 $\left\{X_t^{(i),j}\right\}$，$j = 1, \ldots, 1000$。对每组样本 $\left\{X_t^{(i),j}\right\}$，进行 OLS 估计得到系数向量 $\hat{\boldsymbol{\beta}}^{(i),j}$。得到 1000 个系数估计向量后，绘制 $\hat{\boldsymbol{\beta}}^{(i),j} - \boldsymbol{\beta}_0^{(i)}$ 关于 j 的样本直方图。注意每个估计向量有 3 个系数，分开绘图。对 1000 组估计值，计算这两个系数的样本标准差 $\hat{\sigma}(\hat{\mu}), \hat{\sigma}(\hat{\phi}), \hat{\sigma}(\hat{\psi})$，并在系数估计值样本直方图中，分别添加以 $\mu, \phi^{(i)}, \psi^{(i)}$ 为均值、样本标准差 $\hat{\sigma}(\hat{\mu}), \hat{\sigma}(\hat{\phi}), \hat{\sigma}(\hat{\psi})$ 为标准差的正态分布密度曲线，对比样本分布与正态分布间的差异；并结合各组 DGP 中特征多项式零点与 1 的距离，讨论 3 个系数估计值样本分布与理论预测的正态分布间的差别。

   (e) 请根据课件内容，对每组参数取值 i，计算 $T = 100$ 期样本 AR(2) 序列系数 OLS 估计值 $\hat{\boldsymbol{\beta}}^{(i)}$ 系数向量的理论与样本渐近协方差矩阵；其中前者用 σ_ε^2 及 $\boldsymbol{M}$ 的总体值计算，后者用样本估计的 $\hat{\sigma}_\varepsilon^2$ 与 $\hat{\boldsymbol{M}}$ 计算。从渐近协方差矩阵中，计算两个系数各自的理论与样本渐近标准误，并对比与 (e) 中所得系数估计值样本标准差的关系。此外，根据理论与样本渐近协方差矩阵，讨论 3 个系数估计值间的相关性特征。

   (f) 重复 (d)–(e) 的内容，但取 $T = 900$。比较此时系数估计值理论与样本渐近标准误与 100 组样本所得估计系数样本标准差的关系。此时估计系数的样本标准差是否近似为 (b) 中的 1/3？样本渐近标准误与 (e) 相比呢？