2022 秋季本科时间序列

第 6 次作业

提交日期：11 月 7 日

注意：本次作业可以参考去年作业 5。

1. 考虑带均值的平稳 AR(1) 过程 $X_t = \mu + \phi X_{t-1} + \varepsilon_t$，$t = 1, \ldots, T$，$\varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$，$|\phi| < 1$。

   (a) 计算随机向量 $\boldsymbol{Z}_t = [1, X_{t-1}]^\intercal$ 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) 按照第 5 讲内容，计算 AR(1) 过程的总体（理论）谱密度函数一般表达式，对 $\phi = 0.5, 0.95$，计算出具体的谱密度函数表达式，并绘图。提示：常数项不影响自协方差函数，进而不影响谱密度函数。

   (c) 固定 $\mu = 1$，考虑 $\phi = 0.5, 0.8, 0.97$，分别记为 $\phi^{(i)}$，$i = 1, 2, 3$。给定 $\phi^{(i)}$，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 $X_0^{(i)} = \mathbb{E}^{(i)}X_t = \mu/(1 - \phi^{(i)})$。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，并与上问计算的样本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。

   (d) 固定 (c) 中 3 个模拟序列不变，对每个序列 $\left\{X_t^{(i)}\right\}$，按照每新增 50 个样本，进行 OLS 估计；即 $k = 1, \ldots, 20$，对样本 $\left\{X_t^{(i)} : t = 1, \ldots, 50k\right\}$ 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。注意该向量有两个元素，应当分开绘图。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。

   (e) 给定 (c) 中 $\phi^{(i)}$ 取值，$i = 1, 2, 3$，并定义 $\boldsymbol{\beta}_0^{(i)} = [\mu, \phi^{(i)}]^\intercal$。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 500 组样本 $\left\{X_t^{(i),j}\right\}$，$j = 1, \ldots, 500$。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 500 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两个系数，应当分开绘图。对 500 组估计值，计算这两个系数的样本标准差 $\hat{\sigma}(\hat{\mu}), \hat{\sigma}(\hat{\phi})$，并在系数估计值样本直方图中，分别添加以 $\mu, \phi^{(i)}$ 为均值、样本标准差 $\hat{\sigma}(\hat{\mu}), \hat{\sigma}(\hat{\phi})$ 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；并讨论当 GDP 中 ϕ 趋近 1 时，两个系数估计值样本分布与理论预测的正态分布间的差别。

   (f) 请根据课件内容，对每组参数取值 i，计算 $T = 100$ 期样本 AR(1) 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，讨论两个系数估计值间的相关性特征。

(g) 重复 (e)–(f) 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 (c) 中的 $1/3$？样本渐近标准误与 (f) 相比呢？