

时间序列分析(初级)

多元时间序列分析

陈垚翰

安徽大学 大数据与统计学院



安徽大学
Anhui University

本章结构

- ARIMAX
- 虚假回归
- 协整
- 误差修正模型
- 向量自回归模型和结构向量自回归模型

- ARIMAX 模型结构

$$\left\{ \begin{array}{l} y_t = \mu + \sum_{i=1}^k \frac{\Theta_i(B)}{\Phi_i(B)} B^{li} x_{it} + \varepsilon_t \\ \varepsilon_t = \frac{\Theta(B)}{\Phi(B)} a_t \quad \text{或者} \quad \varepsilon_t = \frac{\Theta(B)}{\Phi(B)} u_t \end{array} \right. \quad (1)$$

$$\left\{ \begin{array}{l} \varepsilon_t = \frac{\Theta(B)}{\Phi(B)} a_t \quad \text{或者} \quad \varepsilon_t = \frac{\Theta(B)}{\Phi(B)} u_t \end{array} \right. \quad (2)$$

其中, a_t (或者 u_t) 为零均值白噪声序列

互相关系数

- k 阶互相关函数

$$\text{Cov} = \text{Cov}(y_t, x_{t-k}) = E[y_t - E(y_t)][x_{t-k} - E(x_{t-k})]$$

- k 阶互相关系数

$$C_{\rho_k} = \frac{\text{Cov}(y_t, x_{t-k})}{\sqrt{\text{Var}(y_t)}\sqrt{\text{Var}(x_{t-k})}}$$

- $C_{\rho_k} \sim N\left(0, \frac{1}{n-|k|}\right)$

虚假回归：以一元回归模型为例

- 考虑一元回归模型

$$y_t = \beta_0 + \beta_1 x_t + e_t \quad e_t \stackrel{i.i.d.}{\sim} (0, \sigma_e^2)$$

- 假设 x_t 和 y_t 分别由如下相互独立的随机游走序列生成

$$x_t = x_{t-1} + v_t$$

$$y_t = y_{t-1} + w_t$$

其中, $v_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_v^2)$, $w_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_w^2)$; 且 $\text{Cov}(w_t, v_s) = 0, \forall t, s \in [0, T]$

- 原假设和备择假设

$$H_0 : \beta_1 = 0 \leftrightarrow H_1 : \beta_1 \neq 0$$

- OLS 的 t 统计量(自由度为样本容量)

$$t = \frac{\hat{\beta}_1}{\hat{\sigma}_{\beta_1}}, \quad \hat{\sigma}_{\beta_1} = \sqrt{\text{Var}(\hat{\beta}_1)}$$

- 犯第一类错误(拒绝 H_0) 的概率很高

单整的概念

- 如果序列 x_t 平稳, 说明序列不存在单位根, 这时称序列为零阶单整序列, 简记为 $x_t \sim I(0)$
- 假如原序列一阶差分后平稳, 说明序列存在一个单位根, 这时称序列为一阶单整序列, 简记为 $x_t \sim I(1)$
- 假如原序列至少需要进行 d 阶差分才能实现平稳, 说明原序列存在 d 个单位根, 这时称原序列为 d 阶单整序列, 简记为 $x_t \sim I(d)$

协整的概念

- 假定自变量序列为 $\mathbf{x}_1, \dots, \mathbf{x}_k$, 其中, $\mathbf{x}_i = \{x_{it}\}$
- 响应变量序列为 $\{y_t\}$ 构造回归模型

$$y_t = \beta_0 + \sum_{i=1}^k \beta_i x_{it} + \varepsilon_t$$

- 假定回归残差序列 $\{\varepsilon_t\}$ 平稳, 称响应序列 $\{y_t\}$ 与自变量序列 $\mathbf{x}_1, \dots, \mathbf{x}_k$ 之间具有协整关系

协整的概念

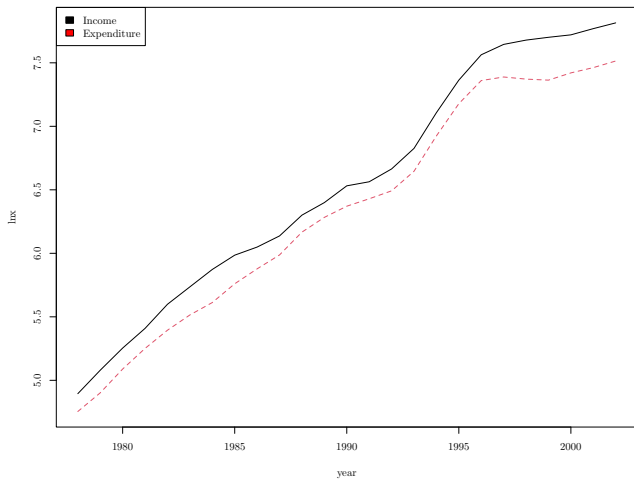


Figure : 中国农村居民家庭人均纯收入与生活消费支出
陈垚翰 (AHU)

协整检验

- 假设条件

- 原假设: 多元非平稳序列之间不存在协整关系

$$H_0 : \varepsilon_t \sim I(k), k \geq 1$$

- 备择假设: 多元非平稳序列之间存在协整关系

$$H_1 : \varepsilon_t \sim I(0)$$

- 检验步骤

- 建立响应序列与输入序列之间的回归模型
- 对回归残差序列进行平稳性检验

- 案例: 中国农村居民家庭人均纯收入与消费支出

误差修正模型

- 误差修正模型 (Error Correction Model), 简记为 ECM, 用以刻画时间序列的短期影响因素
- 响应序列的当期波动 Δy_t 主要会受到三方面短期波动的影响

$$y_t - y_{t-1} = \beta x_t - \beta x_{t-1} - \varepsilon_{t-1} + \varepsilon_t$$

$$\nabla y_t = \beta \nabla x_t - \text{ECM}_{t-1} + \varepsilon_t$$

- 输入序列的当期波动 Δy_t
 - 上一期的误差 ECM_{t-1}
 - 纯随机波动 ε_t
- 误差修正模型

联立线性方程组

案例 1: 供需均衡

$$Q_t^d = \beta_0 + \beta_1 P_t + \varepsilon_t^d, \beta_1 < 0$$

$$Q_t^s = \alpha_0 + \alpha_1 P_t + \alpha_2 R_t + \varepsilon_t^s, \alpha_1 > 0$$

$$Q_t^d = Q_t^s \quad (= Q_t)$$

案例 2: 国名经济核算

$$C_t = \beta_0 + \beta_1 Y_t + \varepsilon_t^C$$

$$I_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 Y_{t-1} + \varepsilon_t^I$$

$$Y_t = C_t + I_t + G_t$$

- 统计学或者计量经济学的问题
 - 内生性的问题
 - 系统识别性的问题
 - 量测误差相关性问题
 - 经济系统的动态性问题

内生性的问题

- 考虑供给方程和需求方程的联立方程组

$$Q_t^d = \beta_0 + \beta_1 P_t + \varepsilon_t^d, \quad \beta_1 < 0$$

$$Q_t^s = \alpha_0 + \alpha_1 P_t + \alpha_2 R_t + \varepsilon_t^s, \quad \alpha_1 > 0$$

$$Q_t^d = Q_t^s \quad (= Q_t)$$

- OLS

$$Q_t = \delta_0 + \delta_1 P_t + u_t$$

- OLS 估计量 $\hat{\delta}_1$ 不能收敛到 β_1 或者 α_1 . 特别的, 当 $\beta_0 = \alpha_0 = \alpha_2 = 0$ 时,

$$\hat{\delta}_1 \xrightarrow{p} \frac{\sigma_s^2}{\sigma_s^2 + \sigma_d^2} \beta_1 + \frac{\sigma_d^2}{\sigma_s^2 + \sigma_d^2} \alpha_1.$$

系统识别性的问题

- 仍然考虑供给方程和需求方程的联立方程组

$$Q_t^d = \beta_0 + \beta_1 P_t + \varepsilon_t^d, \quad \beta_1 < 0$$

$$Q_t^s = \alpha_0 + \alpha_1 P_t + \alpha_2 R_t + \varepsilon_t^s, \quad \alpha_1 > 0$$

$$Q_t^d = Q_t^s \quad (= Q_t)$$

- 假设 $\text{Cov}(\varepsilon_t^d, \varepsilon_t^s) \neq 0$, $\alpha_2 \neq 0$, $\text{Cov}(R_t, \varepsilon_t^d) = 0$, 则

$$\hat{\beta}_1^{IV} = \frac{\sum_{t=1}^T (R_t - \bar{R})(Q_t - \bar{Q})}{\sum_{t=1}^T (R_t - \bar{R})(P_t - \bar{P})}$$

但是 α_1 无法以 R_t 作为工具变量进行估计

- 1 个外生变量 (R_t), 2 内生变量 (P_t, Q_t)

向量自回归模型

- 将系统变量内生化

$$C_t = \beta_0 + \beta_1 Y_t + \varepsilon_t^C$$

$$I_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 Y_{t-1} + \varepsilon_t^I$$

$$Y_t = C_t + I_t + G_t$$

$$\begin{bmatrix} 1 & 0 & -\beta_1 & 0 \\ 0 & 1 & -\alpha_1 & 0 \\ -1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} C_t \\ I_t \\ Y_t \\ G_t \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \alpha_0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha_2 \\ 0 \\ 0 \end{bmatrix} [Y_{t-1}] + \begin{bmatrix} \varepsilon_t^C \\ \varepsilon_t^I \\ 0 \\ \varepsilon_t^G \end{bmatrix}$$

- 引入内生变量的动态特征

$$\begin{bmatrix} C_t \\ I_t \\ Y_t \\ G_t \end{bmatrix} = \begin{bmatrix} \frac{\beta_0(1-\alpha_1)+\alpha_0\beta_1}{1-\alpha_1-\beta_1} \\ \frac{\beta_1(1-\alpha_0)+\alpha_1\beta_0}{1-\alpha_1-\beta_1} \\ \frac{\beta_0+\alpha_0}{1-\alpha_1-\beta_1} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \frac{\beta_1\alpha_2}{1-\alpha_1-\beta_1} & 0 \\ 0 & 0 & \frac{\alpha_2(1-\beta_1)}{1-\alpha_1-\beta_1} & 0 \\ 0 & 0 & \frac{\alpha_2}{1-\alpha_1-\beta_1} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} C_{t-1} \\ I_{t-1} \\ Y_{t-1} \\ G_{t-1} \end{bmatrix} + \begin{bmatrix} \frac{(1-\alpha_1)\varepsilon_t^C + \beta_1(\varepsilon_t^I + \varepsilon_t^G)}{1-\alpha_1-\beta_1} \\ \frac{(1-\beta_1)\varepsilon_t^I + \alpha_1(\varepsilon_t^C + \varepsilon_t^G)}{1-\alpha_1-\beta_1} \\ \frac{\varepsilon_t^C + \varepsilon_t^I + \varepsilon_t^G}{1-\alpha_1-\beta_1} \\ \varepsilon_t^G \end{bmatrix}$$

向量自回归模型

- 时间序列向量

$$\mathbf{Y}_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{nt} \end{bmatrix}$$

- 向量自回归模型 (Vector Autoregression, VAR)

$$\mathbf{Y}_t = \sum_{j=1}^p \Phi_j \mathbf{Y}_{t-j} + \boldsymbol{\varepsilon}_t \quad (1)$$

$$\mathbb{E}[\boldsymbol{\varepsilon}_t] = \mathbf{0}, \quad \mathbb{E}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s^\top] = \begin{cases} \boldsymbol{\Omega} & \text{if } t = s \\ \mathbf{0} & \text{if otherwise} \end{cases}$$

其中 Φ_j is $n \times n$ 矩阵, 并且 $\boldsymbol{\varepsilon}_t$ 是 $n \times 1$ 向量, 通常也被记为 VAR(p)

向量自回归模型

- 一元 AR(1) 模型

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t \quad \text{平稳如果 } |\beta_1| < 1$$
$$\beta_1 = 1 \quad \text{随机游走 (Random Walk)}$$

- 一元 AR(p) 模型

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \varepsilon_t$$
$$(1 - \lambda_1 B)(1 - \lambda_2 B) \dots (1 - \lambda_p B) Y_t = \beta_0 + \varepsilon_t$$

$$\text{平稳} \quad |\lambda_i| < 1, i = 1, 2, \dots, p$$
$$\text{存在单位根} \quad \exists \lambda_i = 1, \beta_1 + \dots + \beta_p = 1$$

VAR ?

向量自回归模型

- 考虑二元 VAR(1) 模型

$$Y_t = \phi_{11}Y_{t-1} + \phi_{12}X_{t-1} + \varepsilon_{y,t}$$

$$X_t = \phi_{21}Y_{t-1} + \phi_{22}X_{t-1} + \varepsilon_{x,t}$$

Y_t 和 X_t 的动态过程是什么样的？

向量自回归模型

- 考虑二元 VAR(1) 模型

$$Y_t = \phi_{11}Y_{t-1} + \phi_{12}X_{t-1} + \varepsilon_{y,t}$$

$$X_t = \phi_{21}Y_{t-1} + \phi_{22}X_{t-1} + \varepsilon_{x,t}$$

Y_t 和 X_t 的动态过程是什么样的？

- 重新表示 VAR(1)

$$(1 - \phi_{11}B)Y_t - \phi_{12}BX_t = \varepsilon_{y,t}$$

$$-\phi_{21}BY_t + (1 - \phi_{22}B)X_t = \varepsilon_{x,t}$$

向量自回归模型

- 在第一个方程两边同时乘以 $(1 - \phi_{22}B)$, 在第二个方程两边同时乘以 $\phi_{12}B$

$$(1 - \phi_{22}B)(1 - \phi_{11}B)Y_t - (1 - \phi_{22}B)\phi_{12}BX_t = (1 - \phi_{22}B)\varepsilon_{y,t} \\ -\phi_{12}B\phi_{21}BY_t + (1 - \phi_{22}B)\phi_{12}BX_t = \phi_{12}B\varepsilon_{x,t}$$

将第一个方程与第二个方程相加消去 X_t

$$(1 - \phi_{22}B)(1 - \phi_{11}B)Y_t - \phi_{12}\phi_{21}B^2Y_t = (1 - \phi_{22}B)\varepsilon_{y,t} + \phi_{12}B\varepsilon_{x,t}$$

- 整理有

$$[1 - (\phi_{11} + \phi_{22})B + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21})B^2] Y_t = (1 - \phi_{22}B)\varepsilon_{y,t} + \phi_{12}B\varepsilon_{x,t}$$

Y_t 是 ARMA(2, 1)

向量自回归模型

- 如果 $|\lambda_1| < 1$ 且 $|\lambda_2| < 1$, Y_t 平稳

$$\begin{aligned} [1 - (\phi_{11} + \phi_{22})B + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21})B^2] Y_t &= (1 - \phi_{22}B) \varepsilon_{y,t} + \phi_{12}B\varepsilon_{x,t} \\ (1 - \lambda_1B)(1 - \lambda_2B) Y_t &= (1 - \phi_{22}B) \varepsilon_{y,t} + \phi_{12}B\varepsilon_{x,t} \end{aligned}$$

- 类似的,

$$\begin{aligned} [1 - (\phi_{11} + \phi_{22})B + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21})B^2] X_t &= (1 - \phi_{11}B) \varepsilon_{x,t} + \phi_{21}B\varepsilon_{y,t} \\ (1 - \lambda_1B)(1 - \lambda_2B) X_t &= (1 - \phi_{11}B) \varepsilon_{x,t} + \phi_{21}B\varepsilon_{y,t} \end{aligned}$$

如果 $|\lambda_1| < 1$ 且 $|\lambda_2| < 1$, X_t 平稳

假设

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} 0.8 & 0.3 \\ 0.1 & 0.6 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{bmatrix}$$

假设

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} 0.8 & 0.3 \\ 0.1 & 0.6 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{bmatrix}$$

$$[1 - 1.4B + 0.45B^2] Y_t = (1 - 0.6B)\varepsilon_{y,t} + 0.3B\varepsilon_{x,t}$$

$$[1 - 1.4B + 0.45B^2] X_t = (1 - 0.8B)\varepsilon_{x,t} + 0.1B\varepsilon_{y,t}$$

$$[1 - 1.4B + 0.45B^2] = (1 - 0.9B)(1 - 0.5B)$$

假设

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} 0.5 & 0.3 \\ 0.25 & 0.85 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{bmatrix}$$

假设

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} 0.5 & 0.3 \\ 0.25 & 0.85 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{bmatrix}$$

$$[1 - 1.35B + 0.35B^2] Y_t = (1 - 0.85B)\varepsilon_{y,t} + 0.3B\varepsilon_{x,t}$$

$$[1 - 1.35B + 0.35B^2] X_t = (1 - 0.5B)\varepsilon_{x,t} + 0.25B\varepsilon_{y,t}$$

$$1 - 1.35B + 0.35B^2 = (1 - 0.25B)(1 - B)$$

$$(1 - 0.35B)(1 - B)Y_t = (1 - 0.85B)\varepsilon_{y,t} + 0.3B\varepsilon_{x,t}$$

$$(1 - 0.35B)(1 - B)X_t = (1 - 0.5B)\varepsilon_{x,t} + 0.35B\varepsilon_{y,t}$$

向量自回归模型

AR(1) :

$$\begin{aligned} Y_{t+k} &= \beta_0 + \beta_1 Y_{t+k-1} + \varepsilon_{t+k} \\ &= \beta_0 + \beta_0 \beta_1 + \beta_1^2 Y_{t+k-2} + \varepsilon_{t+k} + \beta_1 \varepsilon_{t+k-1} \\ &\dots \\ &= \beta_0 + \beta_0 \beta_1 + \dots + \beta_0 \beta_1^k + \beta_1^{k+1} Y_{t-1} + \varepsilon_{t+k} + \beta_1 \varepsilon_{t+k-1} + \dots + \beta_1^k \varepsilon_t \end{aligned}$$

- $\partial Y_{t+k} / \partial \varepsilon_t = \beta_1^k$
- 如果 $|\beta_1| < 1$, $\partial Y_{t+k} / \partial \varepsilon_t \rightarrow 0$
- 如果 $\beta_1 = 1$, $\partial Y_{t+k} / \partial \varepsilon_t \rightarrow 0$

向量自回归模型

VAR(1):

$$\begin{aligned} \mathbf{Y}_{t+k} &= \Phi_1 \mathbf{Y}_{t+k-1} + \boldsymbol{\varepsilon}_{t+k} \\ &= \Phi_1^2 \mathbf{Y}_{t+k-2} + \boldsymbol{\varepsilon}_{t+k} + \Phi_1 \boldsymbol{\varepsilon}_{t+k-1} \\ &\dots \\ &= \Phi_1^{k+1} \mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_{t+k} + \Phi_1 \boldsymbol{\varepsilon}_{t+k-1} + \Phi_1^2 \boldsymbol{\varepsilon}_{t+k-2} + \dots + \Phi_1^k \boldsymbol{\varepsilon}_t \end{aligned}$$

- 如果 $\Phi_1^k \rightarrow \mathbf{0}$, 冲击的影响是短期的
- 如果 Φ_1 的特征根的绝对值均小于 1, $\Phi_1^k \rightarrow \mathbf{0}$

$$\begin{vmatrix} \phi_{11} - \lambda & \phi_{12} \\ \phi_{21} & \phi_{22} - \lambda \end{vmatrix} = 0 \Rightarrow (\phi_{11} - \lambda)(\phi_{22} - \lambda) - \phi_{12}\phi_{21} = 0$$

$$\Rightarrow (\phi_{11}\phi_{22} - \phi_{12}\phi_{21}) - (\phi_{11} + \phi_{22})\lambda + \lambda^2 = 0$$

VAR(1) 的 ARMA(2, 1) 表示形式

$$[1 - (\phi_{11} + \phi_{22})B + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21})B^2] Y_t = (1 - \phi_{22}B) \varepsilon_{y,t} + \phi_{12}B\varepsilon_{x,t}$$

$$(1 - \lambda_1B)(1 - \lambda_2B) Y_t = (1 - \phi_{22}B) \varepsilon_{y,t} + \phi_{12}B\varepsilon_{x,t}$$

$$f(x) = 1 - (\phi_{11} + \phi_{22})x + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21})x^2$$

$$= (1 - \lambda_1x)(1 - \lambda_2x)$$

$$f(1/\lambda_i) = 0 \Rightarrow 1 - (\phi_{11} + \phi_{22})\frac{1}{\lambda_i} + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21})\frac{1}{\lambda_i^2} = 0$$

$$\Rightarrow \lambda_i^2 - (\phi_{11} + \phi_{22})\lambda_i + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21}) = 0$$

向量自回归模型

- VAR(p) 可以等价地写成 VAR(1)

$$\begin{bmatrix} \mathbf{Y}_t \\ \mathbf{Y}_{t-1} \\ \mathbf{Y}_{t-2} \\ \vdots \\ \mathbf{Y}_{t-p+1} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & \mathbf{\Phi}_3 & \cdots & \mathbf{\Phi}_{p-1} & \mathbf{\Phi}_p \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_{t-1} \\ \mathbf{Y}_{t-2} \\ \mathbf{Y}_{t-3} \\ \vdots \\ \mathbf{Y}_{t-p} \end{bmatrix}$$

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

其中,

$$\mathbb{E}[\mathbf{v}_t] = \mathbf{0}, \quad \mathbb{E}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s^\top] = \begin{cases} \mathbf{Q} & \text{if } t = s \\ \mathbf{0} & \text{if otherwise} \end{cases}$$