

浙江工业大学

时间序列分析课程总结

2023/2024(2)



学生班级	大数据分析 2101 班
学生姓名	Phlinsia
学生学号	202103150503
成 绩	

理学院

目录

1 时间序列的预处理	2
1.1 概率分布	2
1.2 特征统计量的基本概念	2
1.3 平稳时间序列	3
1.3.1 平稳时间序列的统计量性质	3
1.3.2 平稳性检验方法	3
1.3.3 纯随机性检验（白噪声序列）	4
1.3.4 随机游动序列	5
2 平稳时间序列分析	5
2.1 方法工具	5
2.1.1 差分运算	5
2.1.2 延迟算子 B	5
2.2 AR 和 MA 模型	6
2.2.1 Wold 分解定理	6
2.2.2 一般线性过程	6
2.2.3 AR 模型（自回归过程）	7
2.2.4 $AR(p)$	8
2.2.5 MA 模型（滑动平均过程）	11
2.2.6 MA 的可逆性	12
2.3 ARMA 模型（自回归移动平均）	13
2.3.1 ARMA (p,q)	13
2.3.2 系数多项式	13
2.3.3 统计性质	13
3 非平稳时间序列的分析	14
3.1 差分平稳	14
3.2 ARIMA 模型（自回归差分滑动平均）	14
4 模型拟合	15
4.1 参数估计	15
4.2 残差检验	15
4.2.1 残差的无关性检验	15

4.2.2	残差的正态性检验	15
4.2.3	参数的显著性检验	15
4.2.4	模型优化	16
4.2.5	模型预测	16

插图

1	ARIMA 建模步骤	16
---	----------------------	----

表格

1	不同模型的自相关和偏自相关特性	14
2	差分方式的选择	14
3	非差分的平稳化方法	14

1 时间序列的预处理

1.1 概率分布

任意给定的有限维分布决定了整个时间序列的统计特性。

概率分布的局限性

1. 仅有样本数据，不了解整体特性。
2. 有限维分布通常只能近似，且特殊分布才能写出来。高维则很难写出分布函数。
3. 计算复杂，不易估计。

1.2 特征统计量的基本概念

- 均值函数 $\mu_t = E(X_t)$ ，一般地，不同时刻 μ_t 可能取不同的值
- 方差函数 $\sigma_t^2 = Var(X_t)$
- 自协方差函数 (Cov)

$$\gamma_{t,s} = Cov(X_t, X_s) = E(X_t, X_s) - \mu_t \mu_s$$

t 和 s 均为时刻。同一时间序列不同时间点，度量过去对现在的影响。

- 自相关函数 (ACF)

$$\rho_{t,s} = Corr(X_t, X_s) = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}$$

t 和 s 均为时刻。可利用样本估计总体，且估计较为精确。

- 结论：

$$\begin{aligned} Cov \left[\sum_{i=1}^m c_i Y_{t_i}, \sum_{j=1}^n d_j Y_{s_j} \right] &= \sum_{i=1}^m \sum_{j=1}^n c_i d_j Cov(Y_{t_i}, Y_{s_j}) \\ Var \left[\sum_{i=1}^n c_i Y_{t_i} \right] &= \sum_{i=1}^n c_i^2 Var(Y_{t_i}) + 2 \sum_{i=2}^n \sum_{j=1}^{i-1} c_i c_j Cov(Y_{t_i}, Y_{t_j}) \end{aligned}$$

1.3 平稳时间序列

指的是统计特性随时间平移不变，泛指严平稳和宽平稳，偶尔也特指宽平稳。

严平稳

又称强平稳性。任意有限维分布函数初始时间点无关，只与时间间隔相关。分布函数随时间平移不变，即所有的统计性质均不会随时间变化

严平稳性质

$$\begin{cases} \gamma_0 = \text{Var}(Y_t), & \rho_0 = 1 \\ \gamma_k = \gamma_{-k}, & \rho_k = \rho_{-k} \\ |\gamma_k| \leq \gamma_0, & |\rho_k| \leq 1 \end{cases}$$

宽平稳及其条件

1. 均值为常数，不随着 t 的变化而变化。
2. 一二阶矩存在。保证方差有限，即方差存在。
3. 自协方差仅仅依赖于时间间隔，为 T 的函数。

1.3.1 平稳时间序列的统计量性质

- 均值函数和方差函数都是常数。 $\mu_t = \mu, \sigma_t^2 = \sigma^2$
- 自相关和自协方差函数只依赖于时间间隔 T ，与时间起止点无关
- 延迟 k 自协方差函数

$$\text{Cov}(X_t, X_{t+k}) = \gamma(k) = \gamma_k = \gamma(t, t+k)$$

- 延迟 k 自相关函数

$$\rho_k = \frac{\gamma(t, t+k)}{\sqrt{D(X_t)D(X_{t+k})}} = \frac{\gamma_k}{\gamma_0} = \frac{\text{Cov}(X_t, X_{t-k})}{\text{Cov}(X_t, X_t)} = \frac{E(X_t - \mu)(X_{t-k} - \mu)}{E(X_t - \mu)(X_t - \mu)}$$

1.3.2 平稳性检验方法

时序图直观检验

时序图提供了一个直观的方法来初步判断时间序列的平稳性。

- **均值稳定：**若序列值在某常数值附近波动，表明序列的均值 μ 可能是稳定的。
- **方差恒定：**当序列的波动范围有界，且无显著的趋势或周期性特征时，暗示序列的方差 σ^2 可能保持恒定。

自相关图分析

自相关图强调序列值之间的线性依赖关系仅取决于时间间隔：

- 在足够大的时间间隔 k 下，如果自相关系数逐渐衰减至零，这表明序列展现出强烈的短期相关性，而长期依赖性较弱。
- 自相关系数应遵循正态分布。

单位根检验

为了更精确地判断时间序列的平稳性，单位根检验提供了一系列定量工具，通过数学模型检测序列是否存在单位根，从而判断序列是否非平稳。常见的单位根检验方法包括：

- **Dickey-Fuller (DF) 检验：**最早提出的一种单位根检验方法，适用于无趋势和无截距项的简单模型。
- **Augmented Dickey-Fuller (ADF) 检验：**DF 检验的扩展，引入了滞后项以更好地拟合序列的动态特性。
- **Phillips-Perron (PP) 检验：**另一种流行的单位根检验方法，它考虑了序列的异方差性和长记忆效应。

1.3.3 纯随机性检验（白噪声序列）

平稳且非纯随机的时间序列才有研究价值。

白噪声序列的定义和性质

- $\forall t \in T, E(X_t) = \mu$
- $\forall t, s \in T, r(t, s) = \begin{cases} t^2 & t = s \\ 0 & t \neq s \end{cases}$ 即 $\gamma_k = \begin{cases} \sigma_e^2 & k = 0 \\ 0 & k > 0 \end{cases}$
- 自相关函数： $\rho(t, s) = \begin{cases} 1 & t = s \\ 0 & t \neq s \end{cases}$ 即 $\rho_k = \begin{cases} 1 & k = 0 \\ 0 & k > 0 \end{cases}$

二阶矩 $E(X^2)$ 存在（均值、方差都有时，二阶矩存在），是弱平稳序列。

1.3.4 随机游动序列

非平稳，形如 $Y_t = e_1 + e_2 + \cdots + e_t$ 其统计量为

$$\begin{aligned}\mu_t &= t \cdot \mu, & \sigma_t^2 &= t \cdot \sigma_e^2 \\ \gamma_{t,s} &= \min(t, s) \cdot \sigma_e^2 \\ \rho_{t,s} &= \frac{\sqrt{\min(t, s)}}{\sqrt{\max(t, s)}}\end{aligned}$$

2 平稳时间序列分析

2.1 方法工具

2.1.1 差分运算

一步差分

1. 一阶差分

$$\nabla X_t = X_t - X_{t-1}$$

2. 二阶差分

$$\nabla^2 X_t = \nabla X_t - \nabla X_{t-1} = (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2}$$

3. p 阶差分

$$\nabla^p X_t = \nabla^{p-1} X_t - \nabla^{p-1} X_{t-1}$$

k 步差分

$$\nabla_k X_t = X_t - X_{t-k}$$

2.1.2 延迟算子 B

1. 类似于一个时间指针，当序列值乘 k 个延迟算子，即把序列值时间向过去拨了 k 个时刻。

$$X_{t-1} = BX_t$$

$$X_{t-2} = B^2 X_t$$

...

$$X_{t-p} = B^p X_t$$

2. 用 B 表示差分

- p 阶差分：

$$\nabla^p X_t = (1 - B)^p X_t = \sum_{i=0}^p (-1)^i C_p^i X_{t-i}$$

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t$$

$$\nabla^2 X_t = \nabla X_t - \nabla X_{t-1} = (1 - B)X_t - (1 - B)X_{t-1} = (1 - B)^2 X_t$$

...

$$\nabla^P X_t = (1 - B)^P X_t$$

- k 步差分：

$$\nabla_k X_t = (1 - B^k)X_t = X_t - X_{t-k}$$

2.2 AR 和 MA 模型

ARMA 模型为自回归滑动平均混合模型。

2.2.1 Wold 分解定理

奠定了 ARMA 模型的理论基础，即对任意的离散平稳序列 $\{X_t\}$ ，可分解为 2 个不相关的平稳序列之和。这一分解揭示了序列的确定性成分与随机成分之间的基本联系。

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots - \varepsilon_t$$

前面与 X_t 有关的是确定性序列，后面 ε_t 是随机序列。

2.2.2 一般线性过程

可以表示为现在和过去白噪声变量的加权线性组合。

$$Y_t = e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \cdots$$

2.2.3 AR 模型（自回归过程）

AR(p): 称为 p 阶自回归模型

AR(1)

$$X_t = \phi_1 X_{t-1} + \varepsilon_t$$

- X_t : 中心化后的平稳序列 ($\mu = 0$)
- ϕ_1 : X_{t-1} 对 X_t 的影响
- ε_t : 随机扰动

如果 AR(1) 模型的特征根 $|\phi_1| < 1$, 则序列平稳。

$$\mu = \frac{0}{1 - \phi} = 0, \quad \gamma_0 = \frac{\sigma_e^2}{1 - \phi^2} \geq 0, \quad \text{Cov}(\varepsilon_t, X_s) = 0$$

$$\text{Cov}(X_t, X_{t-k}) = \phi^k \gamma_0, \quad \rho_k = \phi^k \quad (\text{指数衰减})$$

AR(2) 二阶自回归模型

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

$$\mu = \frac{0}{1 - \phi_1 - \phi_2} = 0, \quad \gamma_0 = \frac{\sigma_e^2(1 - \phi_2)}{(1 + \phi_2)[(1 + \phi_2)^2 - \phi_1^2]}$$

$$\rho_1 = \frac{\phi_1}{1 - \phi_2}, \quad \gamma_1 = \rho_1 \gamma_0$$

Yule-Walker 方程

描述了 **AR** 模型的自协方差函数如何递归地依赖于模型参数。

$$\text{Cov}(X_t, X_{t-k}) = \text{Cov}(\phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t, X_{t-k})$$

$$\begin{cases} \gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}, & k \geq 2 \\ \rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, & k \geq 2 \end{cases}$$

- 对于 ρ_k 随着 k 增加而指数衰减。
- ACF 有不同模式, 取决于根是复数还是实数。

2.2.4 AR(p)

定义

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t, \quad \phi_p \neq 0$$

性质

- $E(\varepsilon_t) = 0$
- $Var(\varepsilon_t) = \sigma_\varepsilon^2$
- $Cov(\varepsilon_t, \varepsilon_s) = 0 \quad s \neq t$, 随机扰动项是均值为 0 的白噪声
- $Cov(X_s, \varepsilon_t) = 0$
- 当 $\phi_0 = 0$ 时 (常数项为 0), 为中心化的 AR(p) 模型

中心化变换 (去掉常数项)

若 $\phi_0 \neq 0$, 令 $\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$

去中心化后:

$$Y_t = X_t - \mu$$

自回归系数多项式

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t$$

$$\Rightarrow \varepsilon_t = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t = \phi(B) X_t$$

$$\Rightarrow \phi(B) X_t = \varepsilon_t$$

AR(p) 平稳性判别

图示判别直观, 但准确度不高。特征根判别更准确。

特征根判别

1. 写出特征根

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = \varepsilon_t$$

$$\Rightarrow \lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \cdots - \phi_p \lambda^0 = 0$$

2. 判断

- $\lambda(p)$ 的 p 个特征根都在单位圆内, $|\lambda| < 1$
- $\phi(\mu) = 0$ (自回归多项式) 的根都在单位圆外, $\lambda = \frac{1}{B} \quad |B| > 1$

平稳域判别

$\phi_1, \phi_2, \dots, \phi_p$ 特征根在圆内, 主要考察一阶、二阶

- AR (1)

$$\lambda = \phi_1, |\phi_1| < 1,$$

即 $\{|\phi_1| - 1 < \phi_1 < 1\}$

- AR (2)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \Rightarrow \lambda^2 - \phi_1 \lambda - \phi_2 = 0$$

$$\lambda = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}, \quad |\lambda_1| < 1, |\lambda_2| < 1$$

即平稳域为 $\{\phi_1, \phi_2 \mid |\phi_2| < 1 - \phi_1\}$

平稳 AR 模型的统计性质

1. 均值

- 非中心化: $\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$
- 中心化: $\mu = 0$

2. 方差

- 整个的 AR 模型的方差 $AR(1): \frac{\sigma_e^2}{1 - \phi_1^2}$
- 用作预测

$$G_0 = 1$$

$$G_1 = \phi_1$$

$$G_2 = \phi_1^2$$

$$G_3 = \phi_1^3$$

$$G_4 = \phi_1^4$$

...

$$G_0 = 1$$

$$G_1 = \phi_1 G_0$$

$$G_2 = \phi_1 G_1 + \phi_2 G_0$$

$$G_3 = \phi_1 G_2 + \phi_2 G_1$$

$$G_4 = \phi_1 G_3 + \phi_2 G_2$$

$$\dots$$

$$\Rightarrow \text{Var}(X_t) = (G_0^2 + G_1^2 + G_2^2 + \dots)\sigma_e^2$$

$$\text{Var}(X_{t-0}) = G_0^2 \sigma_e^2$$

$$\text{Var}(X_{t-1}) = (G_0^2 + G_1^2) \sigma_e^2$$

$$\text{Var}(X_{t-2}) = (G_0^2 + G_1^2 + G_2^2) \sigma_e^2$$

$$\dots$$

3. 协方差函数

$$\begin{aligned} & \bullet \text{ AR}(1) : r_k = \phi_1^k \frac{\sigma_e^2}{1 - \phi_1^2} \\ & \bullet \text{ AR}(2) : \begin{cases} r_0 = \frac{1 - \phi_2}{(1 + \phi_2)(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)} \sigma_e^2 \\ r_1 = \frac{\phi_1 r_0}{1 - \phi_2} \\ r_k = \phi_1 r_{k-1} + \phi_2 r_{k-2}, \quad k \geq 2 \end{cases} \end{aligned}$$

4. 自相关系数

- 拖尾性、呈指数衰减

$$\rho_k = \frac{r_k}{r_0}$$

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}$$

$$\text{AR}(1) : \quad \rho_k = \frac{r_k}{r_0} = \phi_1^k$$

$$\text{AR}(2) : \quad \rho_k = \frac{r_k}{r_0} = \begin{cases} 1 & k = 0 \\ \frac{\phi_1}{1 - \phi_2} & k = 1 \\ \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} & k \geq 2 \end{cases}$$

5. 偏自相关系数

- 剔除了中间 $k - 1$ 个随机变量的干扰, X_{t-k} 对 X_t 影响的相关度量 (单纯度量两变量间的相关性, 而自相关则有两时刻中间的相关度量)

$$AR(1): \quad \phi_{kk} = \begin{cases} \phi_1 & k = 1 \\ 0 & k \geq 2 \end{cases}$$

$$AR(2): \quad \phi_{kk} = \begin{cases} \frac{\phi_1}{1 - \phi_2} & k = 1 \\ \phi_2 & k = 2 \\ 0 & k \geq 3 \end{cases}$$

$$p \quad \phi_{kk} = 0, \quad k > p$$

对 $AR(1): \phi_{11} \neq 0$, 其他均为 0 对 $AR(2): \phi_{11}, \phi_{22} \neq 0$, 其他均为 0

2.2.5 MA 模型 (滑动平均过程)

定义

$$\begin{cases} X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \\ \theta_q \neq 0 \end{cases}$$

$$\begin{aligned} X_t &= \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \\ &= \varepsilon_t - \theta_1 B \varepsilon_t - \cdots - \theta_q B^q \varepsilon_t \\ &= B \varepsilon_t \end{aligned}$$

$$\text{Hanson } B = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \quad (\text{移动平均系数多项式})$$

统计性质

1. 常数均值 $E(X_t) = \mu$
2. 常数方差 $\text{var}(X_t) = (1 + \theta_1^2 + \cdots + \theta_q^2) \varepsilon_q^2$
3. 自协方差 q 阶截尾

$$r_k = \begin{cases} (1 + \theta_1^2 + \cdots + \theta_q^2) \sigma_e^2, & k = 0 \\ (-\theta_k + \sum_{i=1}^{q-k} \theta_i \theta_{k+i}) \sigma_e^2, & 1 \leq k \leq q \\ 0, & k > q \end{cases}$$

$$\text{MA}(1) : \begin{cases} (1 + \theta_1^2)\sigma_e^2 & k = 0 \\ -\theta_1\sigma_e^2 & k = 1 \\ 0 & k > 1 \end{cases}$$

$$\text{MA}(2) : \begin{cases} (1 + \theta_1^2 + \theta_2^2)\sigma_e^2 & k = 0 \\ (-\theta_1 + \theta_1\theta_2)\sigma_e^2 & k = 1 \\ -\theta_2\sigma_e^2 & k = 2 \\ 0 & k \geq 3 \end{cases}$$

4. 自相关系数 p 阶截尾

$$\text{MA}(1) : \rho_k = \begin{cases} 1, & \text{if } k = 0 \\ -\frac{\theta}{1 + \theta^2}, & \text{if } k = 1 \\ 0, & \text{if } k \geq 2 \end{cases}$$

$$\text{MA}(2) : \rho_k = \begin{cases} 1, & \text{if } k = 0 \\ \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2}, & \text{if } k = 1 \\ \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}, & \text{if } k = 2 \\ 0, & \text{if } k \geq 3 \end{cases}$$

2.2.6 MA 的可逆性

有限阶 MA 模型一定平稳。

MA(q) 的可逆条件

$$\text{MA}(1) : |\theta_1| < 1$$

$$\text{MA}(2) : |\theta_2| < 1 \quad \theta_2 \pm \theta_1 < 1$$

MA 模型的判别

- 自相关系数 q 阶截尾 $\Rightarrow \text{MA}(q)$
- 偏自相关系数 ∞ 阶截尾（拖尾）

2.3 ARMA 模型（自回归移动平均）

2.3.1 ARMA(p,q)

$$\begin{cases} x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \\ \phi_p \neq 0 \quad \theta_q \neq 0 \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0 \quad s \neq t \\ E(x_t \varepsilon_t) = 0, s < t \end{cases}$$

$\phi_0 = 0$ 时，中心化的 **ARMA**(p,q) 模型

2.3.2 系数多项式

用滞后算子 B 表示，其中 $\Phi(B)$ 和 $\Theta(B)$ 分别是 AR 和 MA 部分的多项式。AR 部分决定平稳，MA 部分决定可逆。

$$\begin{aligned} \Phi(B)Y_t &= \Theta(B)e_t \\ \Phi(B) &= 1 - \phi_1 B - \phi_2 B^2 + \cdots - \phi_p B^p, \\ \Theta(B) &= 1 - \theta_1 B - \theta_2 B^2 + \cdots - \theta_q B^q \end{aligned}$$

2.3.3 统计性质

均值

$$\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$$

协方差

$$r_k = \sigma_e^2 \sum_{i=0}^{\infty} G_i G_{i+k}$$

自相关系数

$$\rho_k = \frac{\sum_{i=0}^{\infty} G_i G_{i+k}}{\sum_{i=0}^{\infty} G_i^2}$$

相关性

模型	自相关系数	偏自相关系数
AR(p)	拖尾	p 阶截尾
MA(q)	q 阶截尾	拖尾
ARMA(p,q)	拖尾	拖尾

表 1: 不同模型的自相关和偏自相关特性

3 非平稳时间序列的分析

3.1 差分平稳

趋势类型	差分策略
线性趋势	一阶差分
曲线趋势	低阶差分
固定周期趋势	步长为周期长度的差分运算

表 2: 差分方式的选择

过差分可能导致有用的信息浪费，降低估计精度；且使方差增加。

方法	描述
对数变换	通过取对数压缩数据范围，减少异方差性。
幂变换	通过幂次变换调整数据分布，适用于多种非线性关系。

表 3: 非差分的平稳化方法

3.2 ARIMA 模型（自回归差分滑动平均）

ARIMA(p,d,q)

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

其中， W_t 是经过 d 次差分后的序列值， ϕ_i 和 θ_j 分别是自回归和移动平均模型的参数， e_t 表示随机误差项。

4 模型拟合

4.1 参数估计

选择好模型之后，下一步就是估计模型中未知参数的值。对于一个非中心化 ARMA(p, q) 模型，该模型共含有 $p + q + 2$ 个未知参数：

$$\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \mu, \sigma_\epsilon^2$$

其中参数 μ 是序列均值，通常采用矩估计方法，用样本均值估计总体均值即可得到它的估计值。对剩下的 $p + q + 1$ 个未知参数的估计方法有三种：矩估计、极大似然估计和最小二乘估计。

4.2 残差检验

4.2.1 残差的无关性检验

利用 Ljung-Box 对一阶差分序列的残差进行白噪声检验。检验统计量为 LB 检验统计量。原假设和备择假设分别为：

$$\begin{aligned} H_0 : \rho_1 = \rho_2 = \dots = \rho_m = 0, \\ H_1 : \exists k \leq m \quad \text{s.t.} \quad \rho_k \neq 0, \end{aligned}$$

- 若 $p < 0.05$ ，拒绝原假设，残差序列中还残留相关信息，拟合模型不显著，原始序列存在相关性
- 若 $p \geq 0.05$ ，接收原假设，原始序列独立，纯随机；残差是白噪声。

4.2.2 残差的正态性检验

残差的 QQ 图展示了残差分布与理论正态分布之间的对比，如果点大致沿 45° 线分布，则模型的残差满足正态性假设。

4.2.3 参数的显著性检验

参数显著性检验旨在确认模型中的每个参数是否显著非零。通过计算系数估计值与其标准误 (s.e.) 的比值，即 t 统计量，与 t 分布的 5% 临界值 (1.96) 进行比较。若该比值的绝对值超过 1.96，则拒绝零假设，相应的自变量对模型有实质影响。

4.2.4 模型优化

同一个序列可以构造多个有效拟合模型。可引入 AIC 和 BIC 信息准则选出最佳模型。

4.2.5 模型预测

在 R 语言环境下，利用 `forecast` 函数，基于线性最小方差预测原则，未来序列值 x_{t+1} 被建模为历史数据 x_t, x_{t-1}, \dots 的线性组合，以实现最小方差预测。尤其 ARIMA 模型在短期样本外预测中表现优异，但预测期 k 宜控制在模型参数 $p+q$ 之内。预测虽无法消除未来不确定性，却能为决策者提供基于当前趋势的未来场景预览，故对其作用应持理性态度，既不过分低估也不过度依赖。

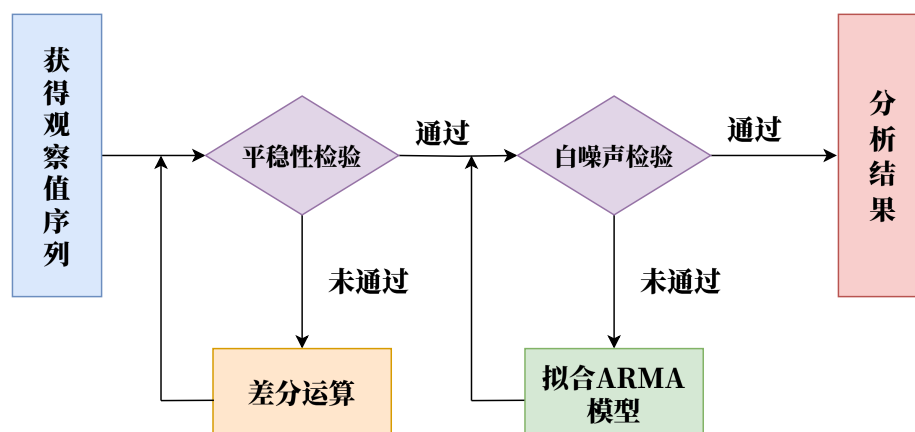


图 1: ARIMA 建模步骤