最小二乘法辨识

白噪声情况

邢超

西北工业大学航天学院

SISO 系统的差分方程为

$$y_k + a_1 y_{k-1} + \cdots + a_n y_{k-n} = b_0 u_k + \cdots + b_n u_{k-n} + \xi_k$$

在时刻 $k = n+1, n+2, \cdots, n+N$，有

$$y_{n+1} + a_1 y_n + \cdots + a_n y_1 = b_0 u_{n+1} + \cdots + b_n u_1 + \xi_{n+1}$$
$$y_{n+2} + a_1 y_{n+1} + \cdots + a_n y_2 = b_0 u_{n+2} + \cdots + b_n u_2 + \xi_{n+2}$$
$$\cdots$$
$$y_{n+N} + a_1 y_{n+N-1} + \cdots + a_n y_N = b_0 u_{n+N} + \cdots + b_n u_N + \xi_{n+N}$$

向量形式

$$
\begin{aligned}
Y &= \Phi\theta + \xi \\
Y &= \begin{bmatrix} y_{n+1} & y_{n+2} & \cdots & y_{n+N} \end{bmatrix}^T \\
\Phi &= \begin{bmatrix} -y_n & \cdots & -y_1 & u_{n+1} & \cdots & u_1 \\ -y_{n+1} & \cdots & -y_2 & u_{n+2} & \cdots & u_2 \\ \vdots & & \vdots & \vdots & & \vdots \\ -y_{n+N-1} & \cdots & -y_N & u_{n+N} & \cdots & u_N \end{bmatrix} \\
\theta &= \begin{bmatrix} a_1 & \cdots & a_n & b_0 & \cdots & b_n \end{bmatrix}^T \\
\xi &= \begin{bmatrix} \xi_{n+1} & \xi_{n+2} & \cdots & \xi_{n+N} \end{bmatrix}^T
\end{aligned}
$$

基本的最小二乘法: 辨识准则

辨识准则：残差平方和最小。

$$
\begin{aligned}
J &= \sum_{k=n+1}^{n+N} e^2(k) \\
&= (Y - \Phi\hat{\theta})^T (Y - \Phi\hat{\theta}) \\
\hat{\theta}_{LS} &= \arg\min_{\hat{\theta}} J
\end{aligned}
$$

基本的最小二乘法：求导

$$
\begin{aligned}
\frac{\partial \mathrm{J}}{\partial \hat{\theta}_{\mathrm{k}}} &= \frac{\partial \sum_{\mathrm{i}}(\mathrm{Y}_{\mathrm{i}} - \sum_{\mathrm{m}} \Phi_{\mathrm{i,m}} \hat{\theta}_{\mathrm{m}})^2}{\partial \hat{\theta}_{\mathrm{k}}} \\
&= 2\sum_{\mathrm{i}}(\mathrm{Y}_{\mathrm{i}} - \sum_{\mathrm{m}} \Phi_{\mathrm{i,m}} \hat{\theta}_{\mathrm{m}}) \frac{\partial(\mathrm{Y}_{\mathrm{i}} - \sum_{\mathrm{m}} \Phi_{\mathrm{i,m}} \hat{\theta}_{\mathrm{m}})}{\partial \hat{\theta}_{\mathrm{k}}} \\
&= 2\sum_{\mathrm{i}}(\mathrm{Y}_{\mathrm{i}} - \sum_{\mathrm{m}} \Phi_{\mathrm{i,m}} \hat{\theta}_{\mathrm{m}}) \frac{\partial(-\sum_{\mathrm{m}} \Phi_{\mathrm{i,m}} \hat{\theta}_{\mathrm{m}})}{\partial \hat{\theta}_{\mathrm{k}}} \\
&= -2\sum_{\mathrm{i}}(\mathrm{Y}_{\mathrm{i}} - \sum_{\mathrm{m}} \Phi_{\mathrm{i,m}} \hat{\theta}_{\mathrm{m}}) \Phi_{\mathrm{i,k}} \\
\frac{\partial \mathrm{J}}{\partial \hat{\theta}} &= (-2(\mathrm{Y} - \Phi\hat{\theta})^{\mathrm{T}} \Phi)^{\mathrm{T}} \\
&= -2\Phi^{\mathrm{T}}(\mathrm{Y} - \Phi\hat{\theta})
\end{aligned}
$$

基本的最小二乘法：求解

$$
\begin{aligned}
-2\Phi^{\mathrm{T}}(\mathrm{Y}-\Phi\hat{\theta}_{\mathrm{LS}}) &= 0 \\
\Phi^{\mathrm{T}}\mathrm{Y}-\Phi^{\mathrm{T}}\Phi\hat{\theta}_{\mathrm{LS}} &= 0 \\
\Phi^{\mathrm{T}}\mathrm{Y} &= \Phi^{\mathrm{T}}\Phi\hat{\theta}_{\mathrm{LS}} \\
\hat{\theta}_{\mathrm{LS}} &= (\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}\mathrm{Y}
\end{aligned}
$$

基本的最小二乘法：二阶导数

$$
\begin{aligned}
\frac{\partial^2 J}{\partial \hat{\theta}^2} &= \frac{\partial(-2\Phi^T(Y-\Phi\hat{\theta}))}{\partial\hat{\theta}} \\
\frac{\partial \frac{\partial J}{\partial \hat{\theta}}}{\partial \hat{\theta}_s} &= \frac{\partial(-2\sum_i (Y_i - \sum_m \Phi_{i,m}\hat{\theta}_m)\Phi_{i,k})}{\partial \hat{\theta}_s} \\
&= 2\sum_i \frac{\partial \sum_m \Phi_{i,m}\hat{\theta}_m}{\partial\hat{\theta}_s}\Phi_{i,k} \\
&= 2\sum_i \Phi_{i,s}\Phi_{i,k} \\
\frac{\partial^2 J}{\partial \hat{\theta}^2} &= 2\Phi^T\Phi
\end{aligned}
$$

最小二乘法对输入信号的要求:$\begin{bmatrix} Y_{N\times n} & U_{N\times(n+1)} \end{bmatrix}$

$$
\begin{aligned}
\Phi^T\Phi &= \begin{bmatrix} Y_{N\times n} & U_{N\times(n+1)} \end{bmatrix}^T \begin{bmatrix} Y_{N\times n} & U_{N\times(n+1)} \end{bmatrix} \\
&= \begin{bmatrix} Y^T_{N\times n}Y_{N\times n} & Y^T_{N\times n}U_{N\times(n+1)} \\ U^T_{N\times(n+1)}Y_{N\times n} & U^T_{N\times(n+1)}U_{N\times(n+1)} \end{bmatrix}
\end{aligned}
$$

其中:

$$
Y_{N\times n} = \begin{bmatrix} -y_n & \cdots & -y_1 \\ -y_{n+1} & \cdots & -y_2 \\ \vdots & & \vdots \\ y_{n+N-1} & \cdots & -y_N \end{bmatrix}
$$

$$
U_{N\times(n+1)} = \begin{bmatrix} u_{n+1} & \cdots & u_1 \\ u_{n+2} & \cdots & u_2 \\ \vdots & & \vdots \\ u_{n+N} & \cdots & u_N \end{bmatrix}
$$

最小二乘法对输入信号的要求: $\begin{bmatrix} Y_{N\times n} & U_{N\times(n+1)} \end{bmatrix}$

$$
\begin{aligned}
(Y_{N\times n}^T Y_{N\times n})_{i,j} &= \sum_{k=1}^{N-1+\min\{i,j\}} y_{n-i+k} y_{n-j+k} \\
(Y_{N\times n}^T U_{N\times(n+1)})_{i,j} &= -\sum_{k=1}^{N-1+\min\{i,j-1\}} y_{n-i+k} u_{n+1-j+k} \\
(U_{N\times(n+1)}^T Y_{N\times n}^T)_{i,j} &= -\sum_{k=1}^{N-1+\min\{j,i-1\}} y_{n-j+k} u_{n+1-i+k} \\
(U_{N\times(n+1)}^T U_{N\times(n+1)})_{i,j} &= \sum_{k=1}^{N-2+\min\{i,j\}} u_{n+1-i+k} u_{n+1-j+k}
\end{aligned}
$$

最小二乘法对输入信号的要求: $\begin{bmatrix} R_y & R_{yu} \\ R_{uy} & R_u \end{bmatrix}$

$$\lim_{N\to\infty}\frac{\Phi^T\Phi}{N} = \frac{1}{N}\begin{bmatrix} Y^T_{N\times n}Y_{N\times n} & Y^T_{N\times n}U_{N\times(n+1)} \\ U^T_{N\times(n+1)}Y_{N\times n} & U^T_{N\times(n+1)}U_{N\times(n+1)} \end{bmatrix} = \begin{bmatrix} R_y & R_{yu} \\ R_{uy} & R_u \end{bmatrix}$$

其中:

$$R_y = \begin{bmatrix} R_y(0) & R_y(1) & \cdots & R_y(n-1) \\ R_y(1) & R_y(0) & \cdots & R_y(n-2) \\ \vdots & \vdots & & \vdots \\ R_y(n-1) & R_y(n-2) & \cdots & R_y(0) \end{bmatrix}$$

$$R_{yu} = \begin{bmatrix} -R_{yu}(1) & -R_{yu}(0) & \cdots & -R_{yu}(1-n) \\ -R_{yu}(2) & -R_{yu}(1) & \cdots & -R_{yu}(2-n) \\ \vdots & \vdots & & \vdots \\ -R_{yu}(n) & -R_{yu}(n-1) & \cdots & -R_{yu}(0) \end{bmatrix}$$

最小二乘法对输入信号的要求:$\begin{bmatrix} R_y & R_{yu} \\ R_{uy} & R_u \end{bmatrix}$

$$
\begin{aligned}
R_{uy} &= R_{yu}^T \\
R_{uu} &= \begin{bmatrix} R_u(0) & R_u(1) & \cdots & R_u(n) \\ R_u(1) & R_u(0) & \cdots & R_u(n-1) \\ \vdots & \vdots & & \vdots \\ R_u(n) & R_u(n-1) & \cdots & R_u(0) \end{bmatrix}
\end{aligned}
$$

(n+1) 阶持续激励信号

- 定义：如果序列 $\{u(k)\}$ 的 (n+1) 阶方阵 R_u 是正定的，则称序列 $\{u(k)\}$ 为 $(n+1)$ 阶持续激励信号。
- 最小二乘法对输入信号的要求为：$\{u(k)\}$ 为 $(n+1)$ 阶持续激励信号
- 若 R_u 为强对角线占优矩阵，则 R_u 正定。以下输入信号均能满足 R_u 正定的要求：
  - 白噪声序列；
  - 伪随机二位式噪声序列；
  - 有色噪声随机信号序列。
- 工程上常用“伪随机二位式噪声序列”、“有色噪声随机信号序列”作为输入信号。

最小二乘估计的概率性质

- 估计的无偏性；
- 估计的一致性；
- 估计的有效性；
- 估计的渐进正态性。

估计的无偏性

若 $\mathrm{E}\{\hat{\theta}\} = \theta$ 则称 $\hat{\theta}$ 是参数 θ 的无偏估计。

$$
\begin{aligned}
\mathrm{Y} &= \Phi\theta + \xi \\
\hat{\theta} &= (\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}\mathrm{Y} \\
\mathrm{E}[\hat{\theta}] &= \mathrm{E}[(\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}\mathrm{Y}] \\
&= \mathrm{E}[(\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}(\Phi\theta + \xi)] \\
&= \mathrm{E}[(\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}\Phi\theta + (\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}\xi] \\
&= \mathrm{E}[\theta + (\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}\xi]
\end{aligned}
$$

LS 无偏估计的充要条件为：

$$
\mathrm{E}[(\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}\xi] = 0
$$

一致性估计

$$\lim_{N\to\infty} P\{|\hat{\theta}-\theta\} = 1$$

设 $\xi\{(k)\}$ 为与 $\{u(k)\}$ 无关的零均值独立同分布随机序列:

$$\begin{aligned}
E(\hat{\theta}-\theta)^2 &= E[(\Phi^T\Phi)^{-1}\Phi^T\xi\xi^T\Phi(\Phi^T\Phi)^{-1}] \\
&= E[\frac{1}{N^2}(\frac{1}{N}\Phi^T\Phi)^{-1}\Phi^T\xi\xi^T\Phi(\frac{1}{N}\Phi^T\Phi)^{-1}] \\
\lim_{N\to\infty} E(\hat{\theta}-\theta)^2 &= \frac{1}{N^2}R^{-1}E[\Phi^T\xi\xi^T\Phi]R^{-1} \\
&= \frac{1}{N^2}R^{-1}\sigma^2E[\Phi^T\Phi]R^{-1} \\
&= \frac{1}{N^2}R^{-1}\sigma^2NRR^{-1} \\
&= \frac{\sigma^2}{N}R^{-1} \\
&= 0
\end{aligned}$$

估计值的有效性

Cramer-Rao 不等式：

$$\mathrm{D}\hat{\theta} = \mathrm{M}^{-1}$$

其中：

$$\mathrm{M} = \mathrm{E}\left[\left(\frac{\partial \ln \mathrm{p}(\mathrm{y}|\theta)}{\partial \theta}\right)^{\mathrm{T}} \left(\frac{\partial \ln \mathrm{p}(\mathrm{y}|\theta)}{\partial \theta}\right)\right]$$

估计值的有效性

$$
\begin{aligned}
y &= \Phi\theta + \xi \\
y &\sim N(\Phi\theta, \sigma^2 I) \\
p(y|\theta) &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left[-\frac{1}{2\sigma^2}(y-\Phi\theta)^T(y-\Phi\theta)\right] \\
\frac{\partial \ln p(y|\theta)}{\partial\theta} &= -\frac{1}{\sigma^2}(y-\Phi\theta)^T\Phi \\
M &= E\left[\frac{1}{\sigma^4}\Phi^T(y-\Phi\theta)(y-\Phi\theta)^T\Phi\right] \\
&= \frac{1}{\sigma^4}E[\Phi^T\xi\xi^T\Phi] \\
\lim_{N\to\infty} M^{-1} &= \sigma^4(\sigma^2 E[\Phi^T\Phi])^{-1} \\
&= \frac{\sigma^2}{N}R^{-1}
\end{aligned}
$$

估计值的渐近正态性

设 $\{\xi(\mathrm{k})\}$ 是零均值且服从正态分布的白噪声序列。则：

$$\begin{aligned}
\mathrm{y} &= \Phi\theta + \xi \\
\mathrm{y} &\sim \mathrm{N}(\Phi\theta, \sigma^2 \mathrm{I}) \\
\hat{\theta} &= (\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}\mathrm{y}
\end{aligned}$$

模型阶次递增算法: 算法特点

- 按模型阶次 n 递推的算法;
- 适合模型阶次 n 未知的情况下应用
- 辨识精度与基本最小二乘相同
- 辨识速度比基本最小二乘有较大提高
- 不需计算高阶矩阵的逆

系统模型

$$
\begin{aligned}
Y &= \Phi_n \theta_n + \xi \\
\Phi_n &= \begin{bmatrix} u_{n+1} & -y_n & u_n & \cdots & -y_1 & u_1 \\ u_{n+2} & -y_{n+1} & u_{n+1} & \cdots & -y_2 & u_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ u_{n+N} & -y_{n+N-1} & u_{n+N-1} & \cdots & -y_N & u_N \end{bmatrix} \\
&= \begin{bmatrix} X_1 & \cdots & X_{2n+1} \end{bmatrix} \\
\theta_n &= \begin{bmatrix} b_0 & a_1 & b_1 & \cdots & a_n & b_n \end{bmatrix}^T \\
\xi &= \begin{bmatrix} \xi_{n+1} & \cdots & \xi_{n+N} \end{bmatrix}^T \\
Y &= \begin{bmatrix} y_{n+1} & \cdots & y_{n+N} \end{bmatrix}^T
\end{aligned}
$$

从 $n=0$ 开始辨识

$$
\begin{aligned}
\Phi_0 &= X_1 \\
\hat{\theta}_0 &= (\Phi_0^T \Phi_0)^{-1} \Phi_0^T Y \\
&= \frac{\sum\limits_{i=n+1}^{n+N} u_i y_i}{\sum\limits_{i=n+1}^{n+N} u_i^2}
\end{aligned}
$$

从 n 到 n + 1

根据模型阶次为 n 时的参数辨识结果，求模型阶次为 n+1 时的辨识结果。求解时分两步进行，首先求解 $\tilde{P}_n$，其次求解 P_{n+1}。

$$
\begin{aligned}
\Phi_{n+1} &= \begin{bmatrix} \Phi_n & X_{2n+2} & X_{2n+3} \end{bmatrix} \\
&= \begin{bmatrix} \tilde{\Phi}_n & X_{2n+3} \end{bmatrix} \\
\tilde{\Phi}_n &\triangleq \begin{bmatrix} \Phi_n & X_{2n+2} \end{bmatrix} \\
P_n &\triangleq (\Phi_n^T \Phi_n)^{-1} \\
\tilde{P}_n &\triangleq (\tilde{\Phi}_n^T \tilde{\Phi}_n)^{-1} \\
P_{n+1} &= (\Phi_{n+1}^T \Phi_{n+1})^{-1}
\end{aligned}
$$

从 n 到 n + 1：P_{n+1}

$$
\begin{aligned}
P_{n+1} &= \begin{bmatrix} \tilde{\Phi}_n^T\tilde{\Phi}_n & \tilde{\Phi}_n^T X_{2n+3} \\ X_{2n+3}^T\tilde{\Phi}_n & X_{2n+3}^T X_{2n+3} \end{bmatrix}^{-1} \\
&= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\
A_{22} &= (X_{2n+3}^T X_{2n+3} - X_{2n+3}^T\tilde{\Phi}_n\tilde{P}_n\tilde{\Phi}_n^T X_{2n+3})^{-1} \\
A_{12} &= A_{21}^T \\
&= -\tilde{P}_n\tilde{\Phi}_n^T X_{2n+3} A_{22} \\
A_{11} &= \tilde{P}_n - A_{12} X_{2n+3}^T \tilde{\Phi}_n \tilde{P}_n^T
\end{aligned}
$$

从 n 到 n + 1：$\tilde{P}_n$

$$
\begin{aligned}
\tilde{P}_n &= \begin{bmatrix} \Phi_n^T\Phi_n & \Phi_n^T X_{2n+2} \\ X_{2n+2}^T\Phi_n & X_{2n+2}^T X_{2n+3} \end{bmatrix}^{-1} \\
&= \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \\
B_{22} &= (X_{2n+2}^T X_{2n+2} - X_{2n+2}^T \Phi_n P_n \Phi_n^T X_{2n+2})^{-1} \\
B_{12} &= B_{21}^T \\
&= -P_n\Phi_n^T X_{2n+2} B_{22} \\
B_{11} &= P_n - B_{12} X_{2n+2}^T \Phi_n P_n^T
\end{aligned}
$$

计算步聚

- 初始化, 计算 $\mathrm{P}_0 = (\Phi_0^{\mathrm{T}}\Phi_0)^{-1}$
- 计算 $\hat{\theta}_0 = \mathrm{P}_0\Phi_0^{\mathrm{T}}\mathrm{Y}$
- 迭代
  - 根据 P_{n} 计算 $\tilde{\mathrm{P}}_{\mathrm{n}}$
  - 根据 $\tilde{\mathrm{P}}_{\mathrm{n}}$ 计算 $\mathrm{P}_{\mathrm{n+1}}$
  - 计算 $\hat{\theta}_{\mathrm{n+1}} = \mathrm{P}_{\mathrm{n+1}}\Phi_{\mathrm{n+1}}^{\mathrm{T}}\mathrm{Y}$

递推算法推导：模型

假设已获取了数据长度为 N 的输入输出数据，则由最小二乘估计有：

$$
\begin{aligned}
Y_N &= \Phi_N\theta + \xi_N \\
\hat{\theta}_N &= (\Phi_N^T\Phi_N)^{-1}\Phi_N^T Y_N \\
\tilde{\theta}_N &= \theta - \tilde{\theta}_N \\
&= -(\Phi_N^T\Phi_N)^{-1}\Phi_N^T\xi_N
\end{aligned}
$$

获得新的数据 u_{n+N+1}, y_{n+N+1} 后，有：

$$
\begin{aligned}
y_{(n+N+1)} &= \Psi^T\theta + \xi_{(n+N+1)} \\
y_{N+1} &= \Psi^T\theta + \xi_{N+1} \\
\Psi_i &= \begin{bmatrix} -y_{(n+i-1)} & \cdots & -y_{(i)} & u_{(n+i)} & \cdots & u_{(i)} \end{bmatrix}^T \\
\begin{bmatrix} Y_N \\ y_{N+1} \end{bmatrix} &= \begin{bmatrix} \Phi_N \\ \Psi_{N+1}^T \end{bmatrix}\theta + \begin{bmatrix} \xi_N \\ \xi_{N+1} \end{bmatrix}
\end{aligned}
$$

递推算法推导:P_{N+1}

$$
\begin{aligned}
\hat{\theta}_{N+1} &= \left(\begin{bmatrix} \Phi_N \\ \Psi_{N+1}^T \end{bmatrix}^T \begin{bmatrix} \Phi_N \\ \Psi_{N+1}^T \end{bmatrix} \right)^{-1} \begin{bmatrix} \Phi_N \\ \Psi_{N+1}^T \end{bmatrix}^T \begin{bmatrix} Y_N \\ y_{N+1} \end{bmatrix} \\
&= (\Phi_N^T \underbrace{\Phi_N}_{N,2n+1} + \underbrace{\Psi_{N+1}}_{2n+1,1} \Psi_{N+1}^T)^{-1} (\Phi_N^T \underbrace{Y_N}_{N,1} + \Psi_{N+1} \underbrace{y_{N+1}}_{1,1}) \\
\hat{\theta}_{N+1} &= P_{N+1}(\Phi_N^T Y_N + \Psi_{N+1} y_{N+1})
\end{aligned}
$$

其中:

$$
\begin{aligned}
P_{N+1} &= (P_N^{-1} + \Psi_{N+1}\Psi_{N+1}^T)^{-1} \\
P_N &= (\Phi_N^T \Phi_N)^{-1}
\end{aligned}
$$

递推算法推导: 矩阵求逆引理

若相应矩阵的逆均存在，则有:

$$(\mathrm{A}+\mathrm{B}\mathrm{C}^{\mathrm{T}})^{-1}=\mathrm{A}^{-1}-\mathrm{A}^{-1}\mathrm{B}(\mathrm{I}+\mathrm{C}^{\mathrm{T}}\mathrm{A}^{-1}\mathrm{B})^{-1}\mathrm{C}^{\mathrm{T}}\mathrm{A}^{-1}$$

所以:

$$\begin{aligned}
\mathrm{P}_{\mathrm{N+1}} &= (\mathrm{P}_{\mathrm{N}}^{-1}+\Psi_{\mathrm{N+1}}\Psi_{\mathrm{N+1}}^{\mathrm{T}})^{-1} \\
&= \mathrm{P}_{\mathrm{N}}-\mathrm{P}_{\mathrm{N}}\Psi_{\mathrm{N+1}}(1+\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P}_{\mathrm{N}}\Psi_{\mathrm{N+1}})^{-1}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P}_{\mathrm{N}} \\
\hat{\theta}_{\mathrm{N+1}} &= \mathrm{A}+\mathrm{B} \\
\mathrm{A} &= \mathrm{P}_{\mathrm{N+1}}\Phi_{\mathrm{N}}^{\mathrm{T}}\mathrm{Y}_{\mathrm{N}} \\
\mathrm{B} &= \mathrm{P}_{\mathrm{N+1}}\Psi_{\mathrm{N+1}}\mathrm{y}_{\mathrm{N+1}} \\
\mathrm{i} &= 1+\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P}_{\mathrm{N}}\Psi_{\mathrm{N+1}}
\end{aligned}$$

递推算法推导: 化简

$$
\begin{aligned}
\mathrm{A} &= (\mathrm{P_N} - \mathrm{P_N}\Psi_{\mathrm{N+1}}\mathrm{i}^{-1}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N})\Phi_{\mathrm{N}}^{\mathrm{T}}\mathrm{Y_N} \\
&= \mathrm{P_N}\Phi_{\mathrm{N}}^{\mathrm{T}}\mathrm{Y_N} - \mathrm{P_N}\Psi_{\mathrm{N+1}}\mathrm{i}^{-1}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N}\Phi_{\mathrm{N}}^{\mathrm{T}}\mathrm{Y_N} \\
&= \hat{\theta}_{\mathrm{N}} - \mathrm{P_N}\Psi_{\mathrm{N+1}}\mathrm{i}^{-1}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\hat{\theta}_{\mathrm{N}} \\
\mathrm{B} &= (\mathrm{P_N} - \mathrm{P_N}\Psi_{\mathrm{N+1}}\mathrm{i}^{-1}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N})\Psi_{\mathrm{N+1}}\mathrm{y_{N+1}} \\
&= \mathrm{i}^{-1}(\mathrm{P_N}(1 + \Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N}\Psi_{\mathrm{N+1}}) - \mathrm{P_N}\Psi_{\mathrm{N+1}}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N})\Psi_{\mathrm{N+1}}\mathrm{y_{N+1}} \\
&= \mathrm{i}^{-1}(\mathrm{P_N} + \mathrm{P_N}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N}\Psi_{\mathrm{N+1}} - \mathrm{P_N}\Psi_{\mathrm{N+1}}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N})\Psi_{\mathrm{N+1}}\mathrm{y_{N+1}} \\
&= \mathrm{i}^{-1}(\mathrm{P_N}\Psi_{\mathrm{N+1}} + \mathrm{P_N}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N}\Psi_{\mathrm{N+1}}\Psi_{\mathrm{N+1}} \\
&\quad -\mathrm{P_N}\Psi_{\mathrm{N+1}}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N}\Psi_{\mathrm{N+1}})\mathrm{y_{N+1}} \\
&= \mathrm{i}^{-1}(\mathrm{P_N}\Psi_{\mathrm{N+1}} + \mathrm{P_N}\Psi_{\mathrm{N+1}}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N}\Psi_{\mathrm{N+1}} \\
&\quad -\mathrm{P_N}\Psi_{\mathrm{N+1}}\Psi_{\mathrm{N+1}}^{\mathrm{T}}\mathrm{P_N}\Psi_{\mathrm{N+1}})\mathrm{y_{N+1}} \\
&= \mathrm{i}^{-1}\mathrm{P_N}\Psi_{\mathrm{N+1}}\mathrm{y_{N+1}}
\end{aligned}
$$

递推算法推导：结果

$$
\begin{aligned}
\hat{\theta}_{N+1} &= \hat{\theta}_N - P_N\Psi_{N+1} i^{-1}\Psi_{N+1}^T\hat{\theta}_N + i^{-1}P_N\Psi_{N+1}y_{N+1} \\
&= \hat{\theta}_N + i^{-1}P_N\Psi_{N+1}(-\Psi_{N+1}^T\hat{\theta}_N + y_{N+1}) \\
&= \hat{\theta}_N + K_{N+1}(y_{N+1} - \Psi_{N+1}^T\hat{\theta}_N) \\
K_{N+1} &= P_N\Psi_{N+1}(1+\Psi_{N+1}^T P_N\Psi_{N+1})^{-1} \\
P_{N+1} &= P_N - K_{N+1}\Psi_{N+1}^T P_N
\end{aligned}
$$

初值获取方法:

- 基本最小二乘估计
- $\hat{\theta}_0 = 0, P_0 = c^2 I$, 其中 c 为充分大的常数。

收敛性:P_N

$$
\begin{aligned}
P_N &= (P_{N-1}^{-1} + \Psi_N \Psi_N^T)^{-1} \\
P_N^{-1} &= P_{N-1}^{-1} + \Psi_N \Psi_N^T \\
P_{N-1}^{-1} &= P_{N-2}^{-1} + \Psi_{N-1} \Psi_{N-1}^T \\
P_{N-2}^{-1} &= P_{N-3}^{-1} + \Psi_{N-2} \Psi_{N-2}^T \\
P_{N-3}^{-1} &= P_{N-4}^{-1} + \Psi_{N-3} \Psi_{N-3}^T \\
&\vdots \\
P_1^{-1} &= P_0^{-1} + \Psi_1 \Psi_1^T \\
P_N^{-1} &= P_0^{-1} + \sum_{i=1}^{N} \Psi_i \Psi_i^T
\end{aligned}
$$

收敛性

$$
\begin{aligned}
\Phi_N &= \begin{bmatrix} \Psi_1^T \\ \Psi_2^T \\ \vdots \\ \Psi_N^T \end{bmatrix} \\
P_N^{-1} &= \frac{1}{c^2} I + \begin{bmatrix} \Psi_1 & \Psi_2 & \cdots & \Psi_N \end{bmatrix} \begin{bmatrix} \Psi_1^T \\ \Psi_2^T \\ \vdots \\ \Psi_N^T \end{bmatrix} \\
&= \frac{1}{c^2} I + \Phi^T \Phi \\
\lim_{c \to \infty} P_N^{-1} &= \Phi_N^T \Phi_N \\
\hat{\theta}_N &= P_N \Phi_N^T Y_N \\
&= (\Phi_N^T \Phi_N)^{-1} \Phi_N^T Y_N
\end{aligned}
$$

残差与新息的关系

新息 (Innovation)$\tilde{y}_i = y_i - \Psi_i^T \hat{\theta}_{i-1}$ 用来描述 i 时刻的预报误差。残差 $\varepsilon_i = y_i - \Psi_i^T \hat{\theta}_i$ 用来描述 i 时刻的输出偏差。

$$
\begin{aligned}
\varepsilon &= y_i - \Psi_i^T \hat{\theta}_i \\
&= y_i - \Psi_i^T (\hat{\theta}_{i-1} + K_i \tilde{y}_i) \\
&= \tilde{y}_i - \Psi_i^T K_i \tilde{y}_i \\
&= (1 - \Psi_i^T K_i) \tilde{y}_i \\
&= (1 - \Psi_i^T P_{i-1} \Psi_i (\Psi_i^T P_{i-1} \Psi_i + 1)^{-1}) \tilde{y}_i \\
&= \frac{\Psi_i^T P_{i-1} \Psi_i + 1 - \Psi_i^T P_{i-1} \Psi_i}{\Psi_i^T P_{i-1} \Psi_i + 1} \tilde{y}_i \\
&= \frac{\tilde{y}_i}{\Psi_i^T P_{i-1} \Psi_i + 1}
\end{aligned}
$$

准则函数的递推计算

$$
\begin{aligned}
J_i &= (Y_i - \Phi_i\theta_i)^T(Y_i - \Phi_i\theta_i) \\
J_{i-1} &= (Y_{i-1} - \Phi_{i-1}\theta_{i-1})^T(Y_{i-1} - \Phi_{i-1}\theta_{i-1}) \\
Y_i - \Phi_i\theta_i &= Y_i - \Phi_i(\hat{\theta}_{i-1} + K_i\tilde{y}_i) \\
&= \begin{bmatrix} Y_{i-1} \\ y_i \end{bmatrix} - \begin{bmatrix} \Phi_{i-1} \\ \Psi_i^T \end{bmatrix} (\hat{\theta}_{i-1} + K_i\tilde{y}_k) \\
&= \begin{bmatrix} Y_{i-1} - \Phi_{i-1}\hat{\theta}_{i-1} \\ \tilde{y}_i \end{bmatrix} - \begin{bmatrix} \Phi_{i-1} \\ \Psi_i^T \end{bmatrix} K_i\tilde{y}_k
\end{aligned}
$$

准则函数的递推计算

$$
\begin{aligned}
J_i &= J_{i-1} - 2K_i^T\Phi_{i-1}^T(Y_{i-1} - \Phi_{i-1}\hat{\theta}_{i-1})\tilde{y}_i + K_i^T\Phi_{i-1}^T\Phi_{i-1}K_i\tilde{y}_i^2 \\
&\quad +(1 - 2K_i^T\Psi_i + K_i^T\Psi_i\Psi_i^TK_i)\tilde{y}_i^2 \\
&= J_{i-1} - 2K_i^T(\Phi_{i-1}^TY_{i-1} - \Phi_{i-1}^T\Phi_{i-1}\hat{\theta}_{i-1})\tilde{y}_i \\
&\quad +(1 - 2K_i^T\Psi_i + K_i^T\Phi_i\Phi_i^TK_i)\tilde{y}_i^2 \\
&= J_{i-1} + (1 - 2K_i^T\Psi_i + K_i^T\Phi_i\Phi_i^TK_i)\tilde{y}_i^2 \\
&= J_{i-1} + (1 - 2K_i^T\Psi_i + K_i^TP_i^{-1}K_i)\tilde{y}_i^2 \\
&= J_{i-1} + (1 - 2K_i^T\Psi_i + K_i^T\Psi_i)\tilde{y}_i^2 \\
&= J_{i-1} + (1 - K_i^T\Psi_i)\tilde{y}_i^2 \\
&= J_{i-1} + (1 - \Psi_i^TP_{i-1}\Psi_i(\Psi_i^TP_{i-1}\Psi_i + 1)^{-1})\tilde{y}_i^2 \\
&= J_{i-1} + \frac{\Psi_i^TP_{i-1}\Psi_i + 1 - \Psi_i^TP_{i-1}\Psi_i}{\Psi_i^TP_{i-1}\Psi_i + 1}\tilde{y}_i^2 \\
&= J_{i-1} + \frac{\tilde{y}_i^2}{\Psi_i^TP_{i-1}\Psi_i + 1}
\end{aligned}
$$

增益矩阵 $\mathrm{K_i}$ 的计算误差对 $\mathrm{P_i}$ 的影响

当 $\mathrm{K_i}$ 存在误差 $\delta\mathrm{K_i}$ 时：

$$\delta\mathrm{P_i} = \delta\mathrm{K_i}\Psi_i^T\mathrm{P_{i-1}}$$

计算 $\mathrm{P_i}$ 的新形式：

$$\begin{aligned}
\mathrm{P_i} &= (\mathrm{I}-\mathrm{K_i}\Psi_i^T)\mathrm{P_{i-1}} \\
&= (\mathrm{I}-\mathrm{K_i}\Psi_i^T)\mathrm{P_{i-1}} - \mathrm{P_{i-1}}\Psi_i\mathrm{K_i^T} + \mathrm{P_{i-1}}\Psi_i\mathrm{K_i^T} \\
&= (\mathrm{I}-\mathrm{K_i}\Psi_i^T)\mathrm{P_{i-1}} - \mathrm{P_{i-1}}\Psi_i\mathrm{K_i^T} + \mathrm{K_i}(\Psi_i^T\mathrm{P_{i-1}}\Psi_i + 1)\mathrm{K_i^T} \\
&= (\mathrm{I}-\mathrm{K_i}\Psi_i^T)\mathrm{P_{i-1}} - (\mathrm{I}-\mathrm{K_i}\Psi_i^T)\mathrm{P_{i-1}}\Psi_i\mathrm{K_i^T} + \mathrm{K_i}\mathrm{K_i^T} \\
&= (\mathrm{I}-\mathrm{K_i}\Psi_i^T)(\mathrm{P_{i-1}} - \mathrm{P_{i-1}}\Psi_i\mathrm{K_i^T}) + \mathrm{K_i}\mathrm{K_i^T} \\
&= (\mathrm{I}-\mathrm{K_i}\Psi_i^T)\mathrm{P_{i-1}}(\mathrm{I} - \Psi_i\mathrm{K_i^T}) + \mathrm{K_i}\mathrm{K_i^T}
\end{aligned}$$

增益矩阵 K_i 的计算误差对 P_i 的影响

当 K_i 存在误差 δK_i 时：

$$
\begin{aligned}
\delta P_i \;=\;& (I-(K_i+\delta K_i)\Psi_i^T)P_{i-1}(I-\Psi_i(K_i+\delta K_i)^T) \\
& +(K_i+\delta K_i)(K_i+\delta K_i)^T - P_i \\
=\;& -\delta K_i\Psi_i^T P_{i-1}(I-\Psi_i K_i^T)+K_i\delta K_i^T \\
& -(I-K_i\Psi_i^T)P_{i-1}\Psi_i\delta K_i^T+\delta K_i K_i^T \\
& +\delta K_i\Psi_i^T P_{i-1}\Psi_i\delta K_i^T+\delta K_i\delta K_i^T \\
& +(I-K_i\Psi_i^T)P_{i-1}(I-\Psi_i K_i^T)+K_iK_i^T-P_i \\
=\;& -\delta K_i\Psi_i^T P_{i-1}(I-\Psi_i K_i^T)+K_i\delta K_i^T \\
& -(I-K_i\Psi_i^T)P_{i-1}\Psi_i\delta K_i^T+\delta K_i K_i^T+O(\delta K_i) \\
=\;& -\delta K_i\Psi_i^T P_i^T+\delta K_i K_i^T-P_i\Psi_i\delta K_i^T+K_i\delta K_i^T+O(\delta K_i) \\
=\;& -\delta K_i K_i^T+\delta K_i K_i^T-K_i\delta K_i^T+K_i\delta K_i^T+O(\delta K_i) \\
=\;& O(\delta K_i)
\end{aligned}
$$

递推算法的稳定性: 差分方程

$$
\begin{aligned}
y_i &= \Psi_i^T\theta + \xi_i \\
\tilde{\theta}_i &\stackrel{\text{def}}{=} \theta - \hat{\theta}_i \\
&= \theta - [\hat{\theta}_{i-1} + K_i(y_i - \Psi_i^T\hat{\theta}_{i-1})] \\
&= \tilde{\theta}_{i-1} - K_i(y_i - \Psi_i^T\hat{\theta}_{i-1}) \\
&= \tilde{\theta}_{i-1} - K_i(\Psi_i^T\theta + \xi_i - \Psi_i^T\hat{\theta}_{i-1}) \\
&= \tilde{\theta}_{i-1} - K_i(\Psi_i^T\tilde{\theta}_{i-1} + \xi_i) \\
&= (I - K_i\Psi_i^T)\tilde{\theta}_{i-1} - K_i\xi_i \\
&= P_iP_{i-1}^{-1}\tilde{\theta}_{i-1} - K_i\xi_i \\
&= A_i\tilde{\theta}_{i-1} - K_i\xi_i \\
A_i &= P_iP_{i-1}^{-1}
\end{aligned}
$$

递推算法的稳定性：特征值

$$
\begin{aligned}
A_i x &= \lambda x \\
(P_{i-1}^{-1} + \Psi_i \Psi_i^T)^{-1} P_{i-1}^{-1} x &= \lambda x \\
P_{i-1}^{-1} x &= [P_{i-1}^{-1} + \Psi_i \Psi_i^T] \lambda x \\
(1-\lambda) P_{i-1}^{-1} x &= \lambda \Psi_i \Psi_i^T x \\
(1-\lambda) x^T P_{i-1}^{-1} x &= \lambda x^T \Psi_i \Psi_i^T x
\end{aligned}
$$

其中：P_{i-1}^{-1} 正定，与 $\Psi_i \Psi_i^T$ 非负定，所以 $0 < \lambda \leq 1$。即：$\tilde{\theta}_i \leq \tilde{\theta}_0$。

状态模型：

$$
\begin{aligned}
\theta_{i+1} &= \theta_i \\
y_i &= \Psi_i^T \theta_i + \xi_i
\end{aligned}
$$

Kalman 滤波器：

$$
\begin{aligned}
\hat{\theta}_i &= \hat{\theta}_{i-1} + K_i (y_i - \Psi_i^T \hat{\theta}_{i-1}) \\
K_i &= S_i \Psi_i (\Psi_i^T S_i \Psi_i + \sigma^2)^{-1} \\
S_i &= P_{i-1} \\
P_i &= (I - K_i \Psi_i^T) P_{i-1} \\
\hat{\theta}_0 &= 0
\end{aligned}
$$