



第6章 极大似然法及其它辨识算法

6.1 极大似然法

作者：李鹏波、胡德文

单位：国防科技大学机电工程与自动化学院

Email: gaoxia73@163.com

中国水利水电出版社版权所有



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page



Page 1 of 24

Go Back

Full Screen

Close

Quit



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)



[Page 2 of 24](#)

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

对参数估计来说，预报误差法、极大似然法适应范围均较广泛，它们不仅适用于线性模型也适用于非线性模型，是处理残差序列相关情况下另一类辨识算法。预报误差法类似于最小二乘法，它并不要求任何关于数据概率分布的统计假设为前提条件，而极大似然估计却属于一种概率性的参数估计法。

随机逼近法是由统计学中，通过连续逼近以获得估计参数发展而来的。它是随机问题的梯度法应用于观测数据被噪声污染，且对此噪声的统计特性不够充分了解的情况。算法十分简便，具有实用价值。

本章介绍极大似然法、预报误差法的原理和递推算法，介绍随机逼近法的原理和随机牛顿法。对几种常用的参数辨识算法优缺点进行比较，分析了它们的应用特点。



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

第6章 极大似然法及其它辨识算法

6.1 极大似然法

1. 极大似然原理

2. 数值解法

3. 递推的极大似然法

4. 极大似然估计的统计性质

Home Page

Title Page

◀ ▶

◀ ▶

Page 3 of 24

Go Back

Full Screen

Close

Quit



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)[Title Page](#)[Page 4 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

6.1 极大似然法

1. 极大似然原理

对极大似然原理描述如下：对于一组观测数据 $\{y_1, y_2, \dots, y_N\}$ ，它所具有的联合概率分布表示了出现该观测结果的可能性。而观测值 $\{y_1, y_2, \dots, y_N\}$ 的联合概率密度 $P(Y, \theta)$ 是与待估参数 θ 的值密切相关的，不同的参数值，将有不同的概率密度函数。当 $\theta = \hat{\theta}_{ML}$ 时，得到该观测值 $\{y_1, y_2, \dots, y_N\}$ 的可能性最大。也就是说，当观测结果为 $\{y_1, y_2, \dots, y_N\}$ 的条件下， $\hat{\theta}_{ML}$ 是接近于参数 θ 真实值的可能性为最大的参数估计值。

极大似然法(Maximum Likelihood Estimation, MLE)需要构造一个以数据和未知参数为自变量的似然函数，并通过极大化这个似然函数，获得模型的参数估计值。模型输出的概率分布将最大可能地逼近实际过程输出的概率分布。极大似然法通常要求具有能够写出输出量的条件概率密度函数的先验知识。

已知参数 θ 的条件下，观测量的概率密度为 $P(Y|\theta)$ ，在 N 次测量 $\{y_1, y_2, \dots, y_N\}$ 后(独立测量)，考虑似然函数

$$L(y_1, y_2, \dots, y_N|\theta) = P(y_1 \cdots y_N|\theta) = \prod_{i=1}^N P(y_i|\theta) \quad (1)$$



如果我们不要求 θ 的分布密度，只要问 θ 的值为多少(最可能的值)，那么，就只要求 θ 使 $L(y_1 y_2 \cdots y_N | \theta) = \max$ 。对于确定了观测值 Y 而言，似然函数仅仅是参数 θ 的函数。由极大似然原理可知， $\hat{\theta}_{ML}$ 满足以下方程。

$$\left. \frac{\partial L}{\partial \hat{\theta}} \right|_{\hat{\theta}=\hat{\theta}_{ML}} = 0 \quad (2)$$

考虑到似然函数一般为指数函数，而指数函数和对数函数都是单调的，为方便求解，上式等价于如下方程。

$$\left. \frac{\partial \ln L}{\partial \hat{\theta}} \right|_{\hat{\theta}=\hat{\theta}_{ML}} = 0 \quad (3)$$

在特殊情况下， $\hat{\theta}_{ML}$ 能够通过求解式(3)得到。但一般情况下，式(3)不容易得到解析解，需要采用数值方法得到其近似解。

下面利用极大似然原理，分析动态系统模型参数的极大似然估计问题。首先分析极大似然估计和最小二乘估计的关系。

1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 5 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

考虑系统模型为线性差分方程

$$y(k) = -a_1 y(k-1) - \cdots - a_n y(k-n) + b_0 u(k) + b_1 u(k-1) + \cdots + b_n u(k-n) + \varepsilon(k) \quad (4)$$

其中 $\varepsilon(k) \sim N(0, \sigma^2)$ 为高斯白噪声。模型的估计问题可写成向量形式

$$Y = \Phi \theta + e \quad (5)$$

其中

$$\begin{aligned} Y &= [y(n+1) \ y(n+2) \ \cdots \ y(n+N)]^T \\ \theta &= [a_1 \ a_2 \ \cdots \ a_n \ b_0 \ b_1 \ \cdots \ b_n]^T \\ e &= [\varepsilon(n+1) \ \varepsilon(n+2) \ \cdots \ \varepsilon(n+N)]^T \\ \Phi &= \begin{bmatrix} -y(n) & \cdots & -y(1) & u(n+1) & \cdots & u(1) \\ -y(n+1) & \cdots & -y(2) & u(n+2) & \cdots & u(2) \\ & \vdots & & & \vdots & \\ -y(n+N-1) & \cdots & -y(N) & u(n+N) & \cdots & u(N) \end{bmatrix} \end{aligned}$$

[Home Page](#)[Title Page](#)[Page 6 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



式(5)也可以写为

$$e = Y - \Phi\theta \quad (6)$$

由于 $\{\varepsilon(k)\}$ 是均值为零的高斯不相关序列，且与 $\{u(k)\}$ 不相关。于是，可得似然函数

$$\begin{aligned} L &= P(Y|\theta, \sigma^2) = P(Y|\theta, \sigma^2) \\ &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2\sigma^2} (Y - \Phi\theta)^T (Y - \Phi\theta) \right\} \end{aligned} \quad (7)$$

1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page

◀ ▶

◀ ▶

Page 7 of 24

Go Back

Full Screen

Close

Quit



式(5)也可以写为

$$e = Y - \Phi\theta \quad (6)$$

由于 $\{\varepsilon(k)\}$ 是均值为零的高斯不相关序列，且与 $\{u(k)\}$ 不相关。于是，可得似然函数

$$\begin{aligned} L &= P(Y|\theta, \sigma^2) = P(Y|\theta, \sigma^2) \\ &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2\sigma^2} (Y - \Phi\theta)^T (Y - \Phi\theta) \right\} \end{aligned} \quad (7)$$

对应的负对数似然函数为

$$-\ln L = \frac{N}{2} \ln \sigma^2 - \frac{N}{2} \ln 2\pi + \frac{1}{2\sigma^2} (Y - \Phi\theta)^T (Y - \Phi\theta) \quad (8)$$

1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 7 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



式(5)也可以写为

$$e = Y - \Phi\theta \quad (6)$$

由于 $\{\varepsilon(k)\}$ 是均值为零的高斯不相关序列，且与 $\{u(k)\}$ 不相关。于是，可得似然函数

$$\begin{aligned} L &= P(Y|\theta, \sigma^2) = P(Y|\theta, \sigma^2) \\ &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2\sigma^2} (Y - \Phi\theta)^T (Y - \Phi\theta) \right\} \end{aligned} \quad (7)$$

对应的负对数似然函数为

$$-\ln L = \frac{N}{2} \ln \sigma^2 - \frac{N}{2} \ln 2\pi + \frac{1}{2\sigma^2} (Y - \Phi\theta)^T (Y - \Phi\theta) \quad (8)$$

根据极大似然原理，求上式对未知参数 $\hat{\theta}$ 和 $\hat{\sigma}^2$ 的偏导数且令其为0，可得

$$\hat{\theta}_{ML} = (\Phi^T \Phi)^{-1} \Phi^T Y \quad (9)$$

上式等同于 $\hat{\theta}_{LS}$ ，这说明当噪声为高斯白噪声时，参数 θ 的极大似然估计和最小二乘估计是等价的。

1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page



Page 7 of 24

Go Back

Full Screen

Close

Quit

由

$$\left. \frac{\partial \ln L}{\partial \hat{\sigma}^2} \right|_{\hat{\sigma}^2 = \hat{\sigma}_{ML}^2} = 0 \quad (10)$$



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page



Page 8 of 24

Go Back

Full Screen

Close

Quit



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

由

$$\left. \frac{\partial \ln L}{\partial \hat{\sigma}^2} \right|_{\hat{\sigma}^2 = \hat{\sigma}_{ML}^2} = 0 \quad (10)$$

得到噪声方差的估计

$$\hat{\sigma}_{ML}^2 = \frac{1}{N} (Y - \Phi \hat{\theta}_{ML})^T (Y - \Phi \hat{\theta}_{ML}) \quad (11)$$

Home Page

Title Page

◀◀ ▶▶

◀ ▶

Page 8 of 24

Go Back

Full Screen

Close

Quit



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

由

$$\left. \frac{\partial \ln L}{\partial \hat{\sigma}^2} \right|_{\hat{\sigma}^2 = \hat{\sigma}_{ML}^2} = 0 \quad (10)$$

得到噪声方差的估计

$$\hat{\sigma}_{ML}^2 = \frac{1}{N} (Y - \Phi \hat{\theta}_{ML})^T (Y - \Phi \hat{\theta}_{ML}) \quad (11)$$

在实际问题中， $\{\varepsilon(k)\}$ 往往不是白噪声序列，而是相关噪声序列。下面讨论残差相关的情况下极大似然估计的求解。

Home Page

Title Page

◀ ▶

◀ ▶

Page 8 of 24

Go Back

Full Screen

Close

Quit

2. 数值解法

考虑模型为

$$A(z^{-1})y(k) = B(z^{-1})u(k) + C(z^{-1})\varepsilon(k) \quad (12)$$

改写为

$$\varepsilon(k) = y(k) + \sum_{i=1}^n a_i y(k-i) - \sum_{i=0}^n b_i u(k-i) - \sum_{i=1}^n c_i \varepsilon(k-i) \quad (13)$$



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page



Page 9 of 24

Go Back

Full Screen

Close

Quit

2. 数值解法

考虑模型为

$$A(z^{-1})y(k) = B(z^{-1})u(k) + C(z^{-1})\varepsilon(k) \quad (12)$$

改写为

$$\varepsilon(k) = y(k) + \sum_{i=1}^n a_i y(k-i) - \sum_{i=0}^n b_i u(k-i) - \sum_{i=1}^n c_i \varepsilon(k-i) \quad (13)$$

在独立观测的前提下，得到输入输出数据 $\{u(k)\}$ 和 $\{y(k)\}$ ，测量N次，得到N值白噪声向量为

$$\varepsilon = \begin{bmatrix} \varepsilon(n+1) \\ \vdots \\ \varepsilon(n+N) \end{bmatrix}, \quad \varepsilon \sim N(0, \sigma^2 I)$$



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page



Page 9 of 24

Go Back

Full Screen

Close

Quit



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

2. 数值解法

考虑模型为

$$A(z^{-1})y(k) = B(z^{-1})u(k) + C(z^{-1})\varepsilon(k) \quad (12)$$

改写为

$$\varepsilon(k) = y(k) + \sum_{i=1}^n a_i y(k-i) - \sum_{i=0}^n b_i u(k-i) - \sum_{i=1}^n c_i \varepsilon(k-i) \quad (13)$$

在独立观测的前提下，得到输入输出数据 $\{u(k)\}$ 和 $\{y(k)\}$ ，测量N次，得到N值白噪声向量为

$$\varepsilon = \begin{bmatrix} \varepsilon(n+1) \\ \vdots \\ \varepsilon(n+N) \end{bmatrix}, \quad \varepsilon \sim N(0, \sigma^2 I)$$

噪声的协方差阵为

$$R = E\{\varepsilon\varepsilon^T\} = \sigma^2 I_{N \times N} \quad (14)$$

[Home Page](#)[Title Page](#)[Page 9 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

令

$$\theta = [a_1 \cdots a_n \ b_0 \cdots b_n \ c_1 \cdots c_n]^T$$

向量形式的方程组 $Y = \Phi\theta + \varepsilon$ ，也可以写成

$$\varepsilon = Y - \Phi\theta \quad (15)$$

联合概率密度函数

$$P(\varepsilon|\theta) = P(Y|\theta) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{k=n+1}^{n+N} \varepsilon^2(k) \right\} \quad (16)$$

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 10 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

令

$$\theta = [a_1 \cdots a_n \ b_0 \cdots b_n \ c_1 \cdots c_n]^T$$

向量形式的方程组 $Y = \Phi\theta + \varepsilon$, 也可以写成

$$\varepsilon = Y - \Phi\theta \quad (15)$$

联合概率密度函数

$$P(\varepsilon|\theta) = P(Y|\theta) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{k=n+1}^{n+N} \varepsilon^2(k) \right\} \quad (16)$$

当 θ 是某个估计值时, 把 $\varepsilon(k)$ 改写为 $\nu(k)$, 则得到似然函数, 并求对数得

$$\ln L = \frac{N}{2} \ln 2\pi - \frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{k=n+1}^{n+N} \nu^2(k) \quad (17)$$

[Home Page](#)

[Title Page](#)



Page 10 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

由似然函数求极大值，

$$\frac{\partial \ln L}{\partial \hat{\sigma}^2} = 0 \quad (18)$$

得到噪声方差的估计为

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{k=n+1}^{n+N} \nu^2(k) \quad (19)$$

其中

$$\nu(k) = y(k) + \sum_{i=1}^n \hat{a}_i y(k-i) - \sum_{i=0}^n \hat{b}_i u(k-i) - \sum_{i=1}^n \hat{c}_i \nu(k-i) \quad (20)$$



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 11 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

由似然函数求极大值,

$$\frac{\partial \ln L}{\partial \hat{\sigma}^2} = 0 \quad (18)$$

得到噪声方差的估计为

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{k=n+1}^{n+N} \nu^2(k) \quad (19)$$

其中

$$\nu(k) = y(k) + \sum_{i=1}^n \hat{a}_i y(k-i) - \sum_{i=0}^n \hat{b}_i u(k-i) - \sum_{i=1}^n \hat{c}_i \nu(k-i) \quad (20)$$

将式(19)代入式(17), 得

$$\ln L = \text{const} - \frac{N}{2} \ln \frac{1}{N} \sum_{k=n+1}^{n+N} \nu^2(k) \quad (21)$$



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page

◀ ▶

◀ ▶

Page 11 of 24

Go Back

Full Screen

Close

Quit

由似然函数求极大值,

$$\frac{\partial \ln L}{\partial \hat{\sigma}^2} = 0 \quad (18)$$

得到噪声方差的估计为

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{k=n+1}^{n+N} \nu^2(k) \quad (19)$$

其中

$$\nu(k) = y(k) + \sum_{i=1}^n \hat{a}_i y(k-i) - \sum_{i=0}^n \hat{b}_i u(k-i) - \sum_{i=1}^n \hat{c}_i \nu(k-i) \quad (20)$$

将式(19)代入式(17), 得

$$\ln L = \text{const} - \frac{N}{2} \ln \frac{1}{N} \sum_{k=n+1}^{n+N} \nu^2(k) \quad (21)$$

根据极大似然原理, 对数似然函数取极值, 就等价于

$$V(\hat{\theta}_{ML}) = \frac{1}{N} \sum_{k=n+1}^{n+N} \nu^2(k) \Big|_{\hat{\theta}_{ML}} = \min \quad (22)$$

其中 $\nu(k)$ 满足式(20)的约束条件。



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page

◀ ▶

◀ ▶

Page 11 of 24

Go Back

Full Screen

Close

Quit

综合以上分析，极大似然估计 $\hat{\theta}_{ML}$ 就是在式(20)的约束条件下，使得 $V(\hat{\theta}_{ML}) = \min$ 。同时，噪声方差 σ^2 的估计值为 $\hat{\sigma}^2 = \min V(\theta)$ 。因为 $V(\theta)$ 是参数 c_1, c_2, \dots, c_n 的非线性函数，只能通过迭代的方法求解。常用的迭代方法有Lagrangian乘子法和Newton-Raphson法，这里介绍Newton-Raphson法。



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page



Page 12 of 24

Go Back

Full Screen

Close

Quit



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

综合以上分析，极大似然估计 $\hat{\theta}_{ML}$ 就是在式(20)的约束条件下，使得 $V(\hat{\theta}_{ML}) = \min$ 。同时，噪声方差 σ^2 的估计值为 $\hat{\sigma}^2 = \min V(\theta)$ 。因为 $V(\theta)$ 是参数 c_1, c_2, \dots, c_n 的非线性函数，只能通过迭代的方法求解。常用的迭代方法有Lagrangian乘子法和Newton-Raphson法，这里介绍Newton-Raphson法。

(1) 选定初始值 $\hat{\theta}(0)$ ，对于 $\hat{\theta}(0)$ 中的 $a_1, \dots, a_n, b_0, \dots, b_n$ 可按模型

$$\nu(k) = \hat{A}(z^{-1})y(k) - \hat{B}(z^{-1})u(k) \quad (23)$$

用最小二乘法求得，而对于 $\hat{\theta}(0)$ 中的 c_1, c_2, \dots, c_n 可先假定一些值。

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 12 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

综合以上分析，极大似然估计 $\hat{\theta}_{ML}$ 就是在式(20)的约束条件下，使得 $V(\hat{\theta}_{ML}) = \min$ 。同时，噪声方差 σ^2 的估计值为 $\hat{\sigma}^2 = \min V(\theta)$ 。因为 $V(\theta)$ 是参数 c_1, c_2, \dots, c_n 的非线性函数，只能通过迭代的方法求解。常用的迭代方法有Lagrangian乘子法和Newton-Raphson法，这里介绍Newton-Raphson法。

(1) 选定初始值 $\hat{\theta}(0)$ ，对于 $\hat{\theta}(0)$ 中的 $a_1, \dots, a_n, b_0, \dots, b_n$ 可按模型

$$\nu(k) = \hat{A}(z^{-1})y(k) - \hat{B}(z^{-1})u(k) \quad (23)$$

用最小二乘法求得，而对于 $\hat{\theta}(0)$ 中的 c_1, c_2, \dots, c_n 可先假定一些值。

(2) 计算预测误差

$$\nu(k) = y(k) - \hat{y}(k) \quad (24)$$

给出

$$J = \frac{1}{2} \sum_{k=n+1}^{n+N} \nu^2(k) \quad (25)$$

[Home Page](#)[Title Page](#)[Page 12 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

综合以上分析，极大似然估计 $\hat{\theta}_{ML}$ 就是在式(20)的约束条件下，使得 $V(\hat{\theta}_{ML}) = \min$ 。同时，噪声方差 σ^2 的估计值为 $\hat{\sigma}^2 = \min V(\theta)$ 。因为 $V(\theta)$ 是参数 c_1, c_2, \dots, c_n 的非线性函数，只能通过迭代的方法求解。常用的迭代方法有Lagrangian乘子法和Newton-Raphson法，这里介绍Newton-Raphson法。

(1) 选定初始值 $\hat{\theta}(0)$ ，对于 $\hat{\theta}(0)$ 中的 $a_1, \dots, a_n, b_0, \dots, b_n$ 可按模型

$$\nu(k) = \hat{A}(z^{-1})y(k) - \hat{B}(z^{-1})u(k) \quad (23)$$

用最小二乘法求得，而对于 $\hat{\theta}(0)$ 中的 c_1, c_2, \dots, c_n 可先假定一些值。

(2) 计算预测误差

$$\nu(k) = y(k) - \hat{y}(k) \quad (24)$$

给出

$$J = \frac{1}{2} \sum_{k=n+1}^{n+N} \nu^2(k) \quad (25)$$

并计算

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{k=n+1}^{n+N} \nu^2(k) \quad (26)$$

[Home Page](#)[Title Page](#)[Page 12 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

(3) 计算 J 的梯度 $\frac{\partial J}{\partial \hat{\theta}}$ 和Hessian矩阵 $\frac{\partial^2 J}{\partial \hat{\theta} \partial \hat{\theta}^T}$, 有

$$\frac{\partial J}{\partial \hat{\theta}} = \sum_{k=n+1}^{n+N} \nu(k) \frac{\partial \nu(k)}{\partial \hat{\theta}} \quad (27)$$

其中

$$\frac{\partial \nu(k)}{\partial \hat{a}_i} = y(k-i) - \sum_{j=1}^n \hat{c}_i \frac{\partial \nu(k-j)}{\partial \hat{a}_i} \quad (28)$$

$$\frac{\partial \nu(k)}{\partial \hat{b}_i} = -u(k-i) - \sum_{j=1}^n \hat{c}_i \frac{\partial \nu(k-j)}{\partial \hat{b}_i} \quad (29)$$

$$\frac{\partial \nu(k)}{\partial \hat{c}_i} = -\nu(k-i) - \sum_{j=1}^n \hat{c}_i \frac{\partial \nu(k-j)}{\partial \hat{c}_i} \quad (30)$$

式(28)、式(29)和式(30)均为差分方程, 这些差分方程的初始条件为0, 可通过求解这些差分方程, 分别求出 $\nu(k)$ 关于 $\hat{a}_1, \dots, \hat{a}_n, \hat{b}_0, \dots, \hat{b}_n, \hat{c}_1, \dots, \hat{c}_n$ 的全部偏导数, 而这些偏导数分别为 $\{y(k)\}, \{u(k)\}$ 和 $\{\nu(k)\}$ 的线性函数。

Home Page

Title Page

◀ ▶

◀ ▶

Page 13 of 24

Go Back

Full Screen

Close

Quit



再由向量 $\frac{\partial J}{\partial \hat{\theta}}$ 对参数向量 $\hat{\theta}$ 求偏导数，得二阶偏导数

$$\frac{\partial^2 J}{\partial \hat{\theta} \partial \hat{\theta}^T} = \sum_{k=n+1}^{n+N} \frac{\partial \nu(k)}{\partial \hat{\theta}} \cdot \frac{\partial \nu(k)}{\partial \hat{\theta}^T} + \sum_{k=n+1}^{n+N} \nu(k) \frac{\partial^2 \nu(k)}{\partial \hat{\theta} \partial \hat{\theta}^T} \quad (31)$$

因为 $\nu(k)$ 是个小量， $\sum_{k=n+1}^{n+N} \nu(k) \frac{\partial^2 \nu(k)}{\partial \hat{\theta} \partial \hat{\theta}^T}$ 可忽略。

1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 14 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

再由向量 $\frac{\partial J}{\partial \hat{\theta}}$ 对参数向量 $\hat{\theta}$ 求偏导数，得二阶偏导数

$$\frac{\partial^2 J}{\partial \hat{\theta} \partial \hat{\theta}^T} = \sum_{k=n+1}^{n+N} \frac{\partial \nu(k)}{\partial \hat{\theta}} \cdot \frac{\partial \nu(k)}{\partial \hat{\theta}^T} + \sum_{k=n+1}^{n+N} \nu(k) \frac{\partial^2 \nu(k)}{\partial \hat{\theta} \partial \hat{\theta}^T} \quad (31)$$

因为 $\nu(k)$ 是个小量， $\sum_{k=n+1}^{n+N} \nu(k) \frac{\partial^2 \nu(k)}{\partial \hat{\theta} \partial \hat{\theta}^T}$ 可忽略。

(4) 按Newton-Raphson法计算 $\hat{\theta}$ 的新估计值 $\hat{\theta}(k+1)$ ，有

$$\hat{\theta}(k+1) = \hat{\theta}(k) - \left(\frac{\partial^2 J}{\partial \hat{\theta} \partial \hat{\theta}^T} \right)^{-1} \frac{\partial J}{\partial \hat{\theta}} \bigg|_{\hat{\theta}=\hat{\theta}(k)} \quad (32)$$

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 14 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

再由向量 $\frac{\partial J}{\partial \hat{\theta}}$ 对参数向量 $\hat{\theta}$ 求偏导数，得二阶偏导数

$$\frac{\partial^2 J}{\partial \hat{\theta} \partial \hat{\theta}^T} = \sum_{k=n+1}^{n+N} \frac{\partial \nu(k)}{\partial \hat{\theta}} \cdot \frac{\partial \nu(k)}{\partial \hat{\theta}^T} + \sum_{k=n+1}^{n+N} \nu(k) \frac{\partial^2 \nu(k)}{\partial \hat{\theta} \partial \hat{\theta}^T} \quad (31)$$

因为 $\nu(k)$ 是个小量， $\sum_{k=n+1}^{n+N} \nu(k) \frac{\partial^2 \nu(k)}{\partial \hat{\theta} \partial \hat{\theta}^T}$ 可忽略。

(4) 按Newton-Raphson法计算 $\hat{\theta}$ 的新估计值 $\hat{\theta}(k+1)$ ，有

$$\hat{\theta}(k+1) = \hat{\theta}(k) - \left(\frac{\partial^2 J}{\partial \hat{\theta} \partial \hat{\theta}^T} \right)^{-1} \frac{\partial J}{\partial \hat{\theta}} \bigg|_{\hat{\theta}=\hat{\theta}(k)} \quad (32)$$

(5) 重复(2)至(4)的计算步骤，迭代求解新的参数估计值 $\hat{\theta}(k+1)$ ，直至 $\nu(k)$ 方差的相对误差小于某个正小数，所得的参数估计值就是极大似然估计值。

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 14 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

3. 递推的极大似然法 为了进行在线辨识，需要给出递推的极大似然估计算法，即每观测一次数据就递推计算一次参数估计值的算法。
设系统的模型为

$$A(z^{-1})y(k) = B(z^{-1})u(k) + C(z^{-1})\nu(k) \quad (33)$$

式中

$$\begin{cases} A(z^{-1}) = 1 + a_1z^{-1} + \cdots + a_nz^{-n} \\ B(z^{-1}) = b_0 + b_1z^{-1} + \cdots + b_nz^{-n} \\ C(z^{-1}) = 1 + c_1z^{-1} + \cdots + c_nz^{-n} \end{cases} \quad (34)$$

$y(k), u(k)$ 是系统的输入输出量， $\{\nu(k)\}$ 是零均值不相关的随机噪声序列。

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 15 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

3. 递推的极大似然法 为了进行在线辨识，需要给出递推的极大似然估计算法，即每观测一次数据就递推计算一次参数估计值的算法。
设系统的模型为

$$A(z^{-1})y(k) = B(z^{-1})u(k) + C(z^{-1})\nu(k) \quad (33)$$

式中

$$\begin{cases} A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_n z^{-n} \\ B(z^{-1}) = b_0 + b_1 z^{-1} + \cdots + b_n z^{-n} \\ C(z^{-1}) = 1 + c_1 z^{-1} + \cdots + c_n z^{-n} \end{cases} \quad (34)$$

$y(k), u(k)$ 是系统的输入输出量， $\{\nu(k)\}$ 是零均值不相关的随机噪声序列。

令

$$\theta = [a_1 \cdots a_n \ b_0 \cdots b_n \ c_1 \cdots c_n]^T \quad (35)$$

则模型式(33)的参数 θ 的极大似然估计 $\hat{\theta}_{ML}$ 为

$$J(\theta)|_{\hat{\theta}_{ML}} = \frac{1}{2} \sum_{k=n+1}^{n+N} \nu^2(k) \Big|_{\hat{\theta}_{ML}} = \min \quad (36)$$

Home Page

Title Page

◀ ▶

◀ ▶

Page 15 of 24

Go Back

Full Screen

Close

Quit



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

其中 $\nu(k)$ 满足下列关系式

$$\nu(k) = [C(z^{-1})]^{-1}[A(z^{-1})y(k) - B(z^{-1})u(k)] \quad (37)$$

为推导递推公式，记 $\nu(k+n) = \nu_k$ 。如果 ν_k 在 $\hat{\theta}_{ML}$ 点上进行Taylor展开，则可近似表示为

$$\nu_k \approx \nu_k|_{\hat{\theta}_{ML}} + \left. \frac{\partial \nu_k}{\partial \theta} \right|_{\hat{\theta}_{ML}} (\theta - \hat{\theta}_{ML}) \quad (38)$$

[Home Page](#)[Title Page](#)[◀◀](#)[▶▶](#)[◀](#)[▶](#)[Page 16 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

其中 $\nu(k)$ 满足下列关系式

$$\nu(k) = [C(z^{-1})]^{-1}[A(z^{-1})y(k) - B(z^{-1})u(k)] \quad (37)$$

为推导递推公式，记 $\nu(k+n) = \nu_k$ 。如果 ν_k 在 $\hat{\theta}_{ML}$ 点上进行Taylor展开，则可近似表示为

$$\nu_k \approx \nu_k|_{\hat{\theta}_{ML}} + \left. \frac{\partial \nu_k}{\partial \theta} \right|_{\hat{\theta}_{ML}} (\theta - \hat{\theta}_{ML}) \quad (38)$$

设

$$\varphi_k^{(f)} \approx - \left[\left. \frac{\partial \nu_k}{\partial \theta} \right]^T \right|_{\hat{\theta}_{ML}} \quad (39)$$

Home Page

Title Page

◀ ▶

◀ ▶

Page 16 of 24

Go Back

Full Screen

Close

Quit

其中

$$\begin{cases} \frac{\partial \nu_k}{\partial \hat{a}_i} |_{\hat{\theta}_{ML}} = [\hat{C}(z^{-1})]^{-1} y_{k-j} \equiv z^{-j} y_k^{(f)} \\ \frac{\partial \nu_k}{\partial \hat{b}_i} |_{\hat{\theta}_{ML}} = -[\hat{C}(z^{-1})]^{-1} u_{k-i} \equiv -z^{-j} u_k^{(f)} \\ \frac{\partial \nu_k}{\partial \hat{c}_i} |_{\hat{\theta}_{ML}} = -[\hat{C}(z^{-1})]^{-1} \nu_{k-i} \equiv -z^{-j} \hat{\nu}_k^{(f)} \end{cases} \quad (40)$$

式中, $y_k^{(f)}, u_k^{(f)}, \hat{\nu}_k^{(f)}$ 分别表示 $y_k, u_k, \hat{\nu}_k$ 的滤波值, 满足关系式

$$\begin{cases} y_k^{(f)} = [\hat{C}(z^{-1})]^{-1} y_k \\ u_k^{(f)} = [\hat{C}(z^{-1})]^{-1} u_k \\ \hat{\nu}_k^{(f)} = [\hat{C}(z^{-1})]^{-1} \hat{\nu}_k \end{cases} \quad (41)$$



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)



Page 17 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)

[◀](#) [▶](#)

[◀](#) [▶](#)

Page 17 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

其中

$$\begin{cases} \frac{\partial \nu_k}{\partial \hat{a}_i} |_{\hat{\theta}_{ML}} = [\hat{C}(z^{-1})]^{-1} y_{k-j} \equiv z^{-j} y_k^{(f)} \\ \frac{\partial \nu_k}{\partial \hat{b}_i} |_{\hat{\theta}_{ML}} = -[\hat{C}(z^{-1})]^{-1} u_{k-i} \equiv -z^{-j} u_k^{(f)} \\ \frac{\partial \nu_k}{\partial \hat{c}_i} |_{\hat{\theta}_{ML}} = -[\hat{C}(z^{-1})]^{-1} \nu_{k-i} \equiv -z^{-j} \hat{\nu}_k^{(f)} \end{cases} \quad (40)$$

式中, $y_k^{(f)}, u_k^{(f)}, \hat{\nu}_k^{(f)}$ 分别表示 $y_k, u_k, \hat{\nu}_k$ 的滤波值, 满足关系式

$$\begin{cases} y_k^{(f)} = [\hat{C}(z^{-1})]^{-1} y_k \\ u_k^{(f)} = [\hat{C}(z^{-1})]^{-1} u_k \\ \hat{\nu}_k^{(f)} = [\hat{C}(z^{-1})]^{-1} \hat{\nu}_k \end{cases} \quad (41)$$

即

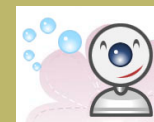
$$\begin{cases} y_k^{(f)} = y_k - \hat{c}_1 y_{k-1}^{(f)} - \cdots - \hat{c}_n y_{k-n}^{(f)} \\ u_k^{(f)} = u_k - \hat{c}_1 u_{k-1}^{(f)} - \cdots - \hat{c}_n u_{k-n}^{(f)} \\ \hat{\nu}_k^{(f)} = \hat{\nu}_k - \hat{c}_1 \hat{\nu}_{k-1}^{(f)} - \cdots - \hat{c}_n \hat{\nu}_{k-n}^{(f)} \end{cases} \quad (42)$$

其中 $\hat{\nu}_k$ 为

$$\hat{\nu}_k = y_k + \sum_{i=1}^n \hat{a}_i y_{k-i} - \sum_{i=0}^n \hat{b}_i u_{k-i} - \sum_{i=1}^n \hat{c}_i \nu_{k-i} \quad (43)$$

那么，向量 $\varphi_k^{(f)}$ 记作

$$\varphi_k^{(f)} = [-y_{k-1}^{(f)} \cdots -y_{k-n}^{(f)} u_{k-1}^{(f)} \cdots u_{k-n}^{(f)} \hat{v}_{k-1}^{(f)} \cdots \hat{v}_{k-n}^{(f)}]^T \quad (44)$$



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)



Page 18 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

那么，向量 $\varphi_k^{(f)}$ 记作

$$\varphi_k^{(f)} = [-y_{k-1}^{(f)} \cdots -y_{k-n}^{(f)} u_{k-1}^{(f)} \cdots u_{k-n}^{(f)} \hat{v}_{k-1}^{(f)} \cdots \hat{v}_{k-n}^{(f)}]^T \quad (44)$$

由于 $J(\theta)$ 是参数 θ 的非线性函数，为了得到极大似然估计的递推形式，先将 $J(\theta)$ 写成递推的形式

$$J_k(\theta) = J_{k-1}(\theta) + \frac{1}{2}\nu_k^2 \quad (45)$$



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)



Page 18 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



那么，向量 $\varphi_k^{(f)}$ 记作

$$\varphi_k^{(f)} = [-y_{k-1}^{(f)} \cdots -y_{k-n}^{(f)} u_{k-1}^{(f)} \cdots u_{k-n}^{(f)} \hat{\nu}_{k-1}^{(f)} \cdots \hat{\nu}_{k-n}^{(f)}]^T \quad (44)$$

由于 $J(\theta)$ 是参数 θ 的非线性函数，为了得到极大似然估计的递推形式，先将 $J(\theta)$ 写成递推的形式

$$J_k(\theta) = J_{k-1}(\theta) + \frac{1}{2}\nu_k^2 \quad (45)$$

设 $\hat{\theta}_{k-1}$ 是 $k-1$ 时刻的极大似然估计值，那么 $J_{k-1}(\theta)$ 在 $\hat{\theta}_{k-1}$ 点上进行Taylor展开，并考虑到 $J_{k-1}(\theta)$ 在 $\hat{\theta}_{k-1}$ 点上关于 θ 的一阶导数近似为零，则有

$$J_k(\theta) = \frac{1}{2}[\theta - \hat{\theta}_{k-1}]^T P_{k-1}^{-1}[\theta - \hat{\theta}_{k-1}] + \frac{1}{2}\eta_k + \frac{1}{2}\nu_k^2 \quad (46)$$

其中 $P_{k-1}^{-1} = \frac{\partial^2 J_k(\theta)}{\partial \theta \partial \theta^T} |_{\hat{\theta}_{k-1}}$ 是正定对称阵， η_k 是 $J_{k-1}(\theta)$ Taylor展开时的残差项。令

$$J_k^*(\theta) = 2J_k(\theta) \quad (47)$$

1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)[Title Page](#)

Page 18 of 24

[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)



Page 19 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

由式(38)、式(39)和式(46)推导得到

$$\begin{aligned} J_k^*(\theta) &\approx [\theta - \hat{\theta}_{k-1}]^T P_{k-1}^{-1} [\theta - \hat{\theta}_{k-1}] + \left\{ \nu_k |_{\hat{\theta}_{k-1}} + \frac{\partial \nu_k}{\partial \theta} |_{\hat{\theta}_{k-1}} (\theta - \hat{\theta}_{k-1}) \right\}^2 + \eta_{k-1} \\ &= [\theta - \hat{\theta}_{k-1}]^T [P_{k-1}^{-1} + \varphi_k^{(f)} \varphi_k^{(f)T}] [\theta - \hat{\theta}_{k-1}] - 2\nu_k |_{\hat{\theta}_{k-1}} \varphi_k^{(f)T} (\theta - \hat{\theta}_{k-1}) \\ &\quad + \nu_k^2 |_{\hat{\theta}_{k-1}} + \eta_k \end{aligned} \quad (48)$$

记 $\tilde{\theta}_{k-1} = \theta - \hat{\theta}_{k-1}$ ，并将式(48)配成二次型

$$J_k^*(\theta) = [\tilde{\theta}_{k-1} - r_k]^T P_k^{-1} [\tilde{\theta}_{k-1} - r_k] + \eta_k^* \quad (49)$$

其中

$$\begin{cases} P_k^{-1} = P_{k-1}^{-1} + \varphi_k^{(f)} \varphi_k^{(f)T} \\ r_k = P_k \varphi_k^{(f)} \nu_k |_{\hat{\theta}_{k-1}} \equiv G_k \hat{\nu}_k \\ \eta_k^* = r_k^T P_k^{-1} r_k + \nu_k^2 |_{\hat{\theta}_{k-1}} + \eta_k \end{cases} \quad (50)$$



显然， η_k^* 是大于零的，由式(49)可知，若 k 时刻的估计值 $\hat{\theta}_k$ 使得

$$\tilde{\theta}_{k-1}|\hat{\theta}_{k-1} = r_k = G_k \hat{\nu}_k \quad (51)$$

1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 20 of 24

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)



显然， η_k^* 是大于零的，由式(49)可知，若 k 时刻的估计值 $\hat{\theta}_k$ 使得

$$\tilde{\theta}_{k-1}|\hat{\theta}_{k-1} = r_k = G_k \hat{\nu}_k \quad (51)$$

$J_k^*(\theta)$ 必然取得极小值。同时，类似于最小二乘法的推导，对式(50)利用矩阵求逆引理，可得

$$P_k = P_{k-1} - \frac{P_{k-1} \varphi_k^{(f)} \varphi_k^{(f)T} P_{k-1}}{1 + \varphi_k^{(f)T} P_{k-1} \varphi_k^{(f)}} \quad (52)$$

1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 20 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

显然， η_k^* 是大于零的，由式(49)可知，若 k 时刻的估计值 $\hat{\theta}_k$ 使得

$$\tilde{\theta}_{k-1}|\hat{\theta}_{k-1} = r_k = G_k \hat{\nu}_k \quad (51)$$

$J_k^*(\theta)$ 必然取得极小值。同时，类似于最小二乘法的推导，对式(50)利用矩阵求逆引理，可得

$$P_k = P_{k-1} - \frac{P_{k-1} \varphi_k^{(f)} \varphi_k^{(f)T} P_{k-1}}{1 + \varphi_k^{(f)T} P_{k-1} \varphi_k^{(f)}} \quad (52)$$

增益矩阵的递推公式为

$$G_k = P_{k-1} \varphi_k^{(f)} [1 + \varphi_k^{(f)T} P_{k-1} \varphi_k^{(f)}]^{-1} \quad (53)$$

[Home Page](#)[Title Page](#)[Page 20 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



于是，将递推的极大似然估计算法(RML)归纳为

$$\begin{aligned}
 \hat{\theta}_k &= \hat{\theta}_{k-1} + G_k \hat{\nu}_k \\
 G_k &= P_{k-1} \varphi_k^{(f)} [1 + \varphi_k^{(f)T} P_{k-1} \varphi_k^{(f)}]^{-1} \\
 P_k &= [I - G_k \varphi_k^{(f)T}] P_{k-1} \\
 \hat{\nu}_k &= y_k - \varphi_k^T \hat{\theta}_{k-1} \\
 \varphi_k &= [-y_{k-1} \cdots -y_{k-n} \ u_{k-1} \cdots u_{k-n} \ \hat{\nu}_{k-1} \ \hat{\nu}_{k-n}]^T \\
 \varphi_k^{(f)} &= [-y_{k-1}^{(f)} \cdots -y_{k-n}^{(f)} \ u_{k-1}^{(f)} \cdots u_{k-n}^{(f)} \ \hat{\nu}_{k-1}^{(f)} \ \hat{\nu}_{k-n}^{(f)}]^T \\
 y_k^{(f)} &= y_k - \hat{c}_1 y_{k-1}^{(f)} - \cdots - \hat{c}_n y_{k-n}^{(f)} \\
 u_k^{(f)} &= u_k - \hat{c}_1 u_{k-1}^{(f)} - \cdots - \hat{c}_n u_{k-n}^{(f)} \\
 \hat{\nu}_k^{(f)} &= \hat{\nu}_k - \hat{c}_1 \hat{\nu}_{k-1}^{(f)} - \cdots - \hat{c}_n \hat{\nu}_{k-n}^{(f)}
 \end{aligned} \tag{54}$$

1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page



Page 21 of 24

Go Back

Full Screen

Close

Quit

式中，向量 $\varphi_k^{(f)}$ 还可以写成如下的递推形式

$$\varphi_{k+1}^{(f)} = \begin{bmatrix} -\hat{c}_1 & \cdots & -\hat{c}_n & & & \\ 1 & & \mathbf{0} & & & \mathbf{0} \\ & \ddots & & & & \\ \mathbf{0} & & 1 & 0 & & \\ & \mathbf{0} & & -\hat{c}_1 & \cdots & -\hat{c}_n \\ & & & 1 & & \mathbf{0} \\ & & & & \ddots & \\ & & \mathbf{0} & & 1 & 0 \\ & & & & & -\hat{c}_1 & \cdots & -\hat{c}_n \\ & \mathbf{0} & & & & 1 & & \mathbf{0} \\ & & & & & & \ddots & \\ & & & & & \mathbf{0} & & 1 & 0 \end{bmatrix} \varphi_k^{(f)} + \begin{bmatrix} -y_k \\ \mathbf{0} \\ u_k \\ \mathbf{0} \\ \hat{v}_k \\ \mathbf{0} \end{bmatrix} \quad (55)$$



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page



Page 22 of 24

Go Back

Full Screen

Close

Quit

式中，向量 $\varphi_k^{(f)}$ 还可以写成如下的递推形式

$$\varphi_{k+1}^{(f)} = \begin{bmatrix} -\hat{c}_1 & \cdots & -\hat{c}_n & & & \\ 1 & & \mathbf{0} & & & \mathbf{0} \\ & \ddots & & & & \\ \mathbf{0} & & 1 & 0 & & \\ & \mathbf{0} & & -\hat{c}_1 & \cdots & -\hat{c}_n \\ & & & 1 & & \mathbf{0} \\ & & & & \ddots & \\ & & \mathbf{0} & & 1 & 0 \\ & & & & & -\hat{c}_1 & \cdots & -\hat{c}_n \\ & \mathbf{0} & & & & 1 & & \mathbf{0} \\ & & & & & & \ddots & \\ & & & & & & & 1 & 0 \end{bmatrix} \varphi_k^{(f)} + \begin{bmatrix} -y_k \\ \mathbf{0} \\ u_k \\ \mathbf{0} \\ \hat{v}_k \\ \mathbf{0} \end{bmatrix} \quad (55)$$

极大似然估计递推算法类似于增广最小二乘法，所不同的只是向量 $\varphi_k^{(f)}$ 的构造不一样。初始值一般取 $\hat{\theta}_0 = \varepsilon$ (ε 为充分小的向量)， $P_0 = I$ ， $\hat{v}_k = 0$ ($k = 0, 1, \dots, n-1$)；并利用输入输出数据构造 φ_1 ，使之不为全零向量，同时置 $\varphi_1^{(f)} = \varphi_1$ 。在这些初始状态下，利用式(54)，便可逐步递推计算 G_k ， P_k 和 $\hat{\theta}_k$ 。

RML算法的辨识效果好于RLS，对噪声模型的参数辨识精度也好于RELS。



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

Home Page

Title Page

◀ ▶

◀ ▶

Page 22 of 24

Go Back

Full Screen

Close

Quit



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

4. 极大似然估计的统计性质

当输入信号 $\{u(k)\}$ 是 $2n$ 阶持续激励的，则极大似然估计具有良好的统计性质。下面直接给出结论。

(1) 渐近无偏性

极大似然参数估计值 $\hat{\theta}_{ML}$ 是 θ 的无偏估计，即

$$\lim_{N \rightarrow \infty} E\{\hat{\theta}_{ML}\} = \theta \quad (56)$$

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 23 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

4. 极大似然估计的统计性质

当输入信号 $\{u(k)\}$ 是 $2n$ 阶持续激励的，则极大似然估计具有良好的统计性质。下面直接给出结论。

(1) 渐近无偏性

极大似然参数估计值 $\hat{\theta}_{ML}$ 是 θ 的无偏估计，即

$$\lim_{N \rightarrow \infty} E\{\hat{\theta}_{ML}\} = \theta \quad (56)$$

(2) 渐近一致性

极大似然估计 $\hat{\theta}_{ML}$ 几乎必然收敛于参数真值 θ ，即

$$\hat{\theta}_{ML} \xrightarrow{N \rightarrow \infty} \theta, \quad a.s. \quad (57)$$

其中a.s.表示几乎肯定(almost surely)的意思。

[Home Page](#)[Title Page](#)

Page 23 of 24

[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

(3) 渐近正态性

设 $\hat{\theta}_{ML}$ 是由 N 个独立同分布随机变量 y 的样本 $y(1), y(2), \dots, y(N)$ 得到的参数 θ 的极大似然估计量，那么当 $N \rightarrow \infty$ 时， $\hat{\theta}_{ML}$ 的分布收敛于正态分布，即

$$\lim_{N \rightarrow \infty} \sqrt{N}(\hat{\theta}_{ML} - \theta) \rightarrow \beta, \quad \beta \sim N(0, \overline{M}_{\theta}^{-1}) \quad (58)$$

其中 \overline{M}_{θ} 表示在参数 θ 条件下的平均Fisher信息矩阵。

[Home Page](#)[Title Page](#)

Page 24 of 24

[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)



1. 极大似然原理
2. 数值解法
3. 递推的极大似然法
4. 极大似然估计的统...

(3) 渐近正态性

设 $\hat{\theta}_{ML}$ 是由 N 个独立同分布随机变量 y 的样本 $y(1), y(2), \dots, y(N)$ 得到的参数 θ 的极大似然估计量, 那么当 $N \rightarrow \infty$ 时, $\hat{\theta}_{ML}$ 的分布收敛于正态分布, 即

$$\lim_{N \rightarrow \infty} \sqrt{N}(\hat{\theta}_{ML} - \theta) \rightarrow \beta, \quad \beta \sim N(0, \overline{M}_{\theta}^{-1}) \quad (58)$$

其中 \overline{M}_{θ} 表示在参数 θ 条件下的平均Fisher信息矩阵。

(4) 渐近有效性

极大似然估计量 $\hat{\theta}_{ML}$ 与参数真值 θ 之间的偏差 $\tilde{\theta} = \theta - \hat{\theta}_{ML}$ 的协方差阵达到Cramer-Rao不等式的下界。

$$\text{Cov} \hat{\theta}_{ML} \longrightarrow \overline{M}_{\theta}^{-1} \quad (59)$$

而由Cramer-Rao不等式, $\text{Cov} \hat{\theta}_{ML} \geq \overline{M}_{\theta}^{-1}$ 。

[Home Page](#)[Title Page](#)[Page 24 of 24](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)