



改进的非线性连续-离散系统的 极大似然参数估计及其应用¹⁾

李乃宏 吴瑶华 黄文虎

(哈尔滨工业大学航天工程与力学系 哈尔滨 150001)

摘 要

对 Jategaonkar 等人给出的同时计及过程及观测噪声的非线性连续-离散系统的极大似然算法从两个方面进行了改进: 1) 给出了计算灵敏度的最佳摄动有限差分算法, 避免了普通有限差分法计算灵敏度矩阵时需人为选择参数摄动量大小而带来的缺点; 2) 给出了具有快速三角化平方根滤波的极大似然算法, 提高了原算法的数值稳定性。上述改进算法经应用于飞行器系统参数估计证明了方法的有效性。

关键词: 极大似然估计, 非线性系统, 灵敏度, 平方根滤波。

1 引言

工程中的相当一类系统可归结为非线性连续-离散系统模型, 对于此类系统的参数估计, 在 Maine 和 Iliff 线性系统算法的基础上^[1], Jategaonkar 等人已给出了同时计及过程及观测噪声的极大似然算法 (NLMLEKF), 且已在飞行器系统参数估计中得到了应用^[2], 但却存在下列缺点:

1) NLMLEKF 算法中非线性滤波过程的数值稳定性常常难于保证。

2) 采用普通有限差分法计算灵敏度矩阵时, 需人为选取参数摄动量的大小, 由此计算的灵敏度矩阵各元素常常没有一致的精度, 收敛性将受精度最差的影响。

因此, 本文在[1—2]基础上, 给出了改进的非线性连续-离散系统的极大似然算法。

2 非线性连续-离散系统的极大似然参数估计

考虑如下非线性连续-离散系统模型:

1) 国家自然科学基金资助课题, No. 18972024.

本文于1992年2月1日收到

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= f(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\theta}) + F \cdot \boldsymbol{\omega}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \\ \boldsymbol{y}(t) &= h(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\theta}), \\ \boldsymbol{z}_i &= \boldsymbol{y}_i + \boldsymbol{v}_i. \quad (i = 1, 2, \dots, N)\end{aligned}\quad (1)$$

其中 $\boldsymbol{x} \in R^n, \boldsymbol{u} \in R^r, \boldsymbol{y} \in R^m, \boldsymbol{z} \in R^m, \boldsymbol{\theta} \in R^p$ 分别是状态、输入、输出、观测及参数向量, $\boldsymbol{\omega}(t)$ 和 \boldsymbol{v}_i 是相应维的过程及观测噪声向量, 假定为零均值高斯白噪, F 为过程噪声分布阵, N 是数据样本长度, 则极大似然参数估计的负对数似然函数为

$$J(\boldsymbol{\theta}, B_i) = \frac{1}{2} \sum_{i=1}^N (\boldsymbol{z}_i - \tilde{\boldsymbol{y}}_i)^T B_i^{-1} (\boldsymbol{z}_i - \tilde{\boldsymbol{y}}_i) + \frac{1}{2} \sum_{i=1}^N \ln |B_i|, \quad (2)$$

其中 B_i 为新息方差阵, $\tilde{\boldsymbol{y}}$ 是 \boldsymbol{y} 的预报值, 可根据实际应用背景, 采用时变或稳态广义 Kalman 滤波来获取.

以参数向量的某初始值开始, 采用 Gauss-Newton 算法可获得参数的迭代估计为

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + \Delta \boldsymbol{\theta}_{i+1}, \quad (3)$$

$$\Delta \boldsymbol{\theta}_{i+1} = (\nabla_{\boldsymbol{\theta}}^2 J)^{-1} (\nabla_{\boldsymbol{\theta}} J)$$

$$\approx \left\{ \sum_{i=1}^N \left[\frac{\partial \tilde{\boldsymbol{y}}(i)}{\partial \boldsymbol{\theta}} \right]^T B_i^{-1} \frac{\partial \tilde{\boldsymbol{y}}}{\partial \boldsymbol{\theta}}(i) \right\}^{-1} \left\{ \sum_{i=1}^N \left[\frac{\partial \tilde{\boldsymbol{y}}}{\partial \boldsymbol{\theta}}(i) \right]^T B_i^{-1} (\boldsymbol{z}_i - \tilde{\boldsymbol{y}}_i) \right\}. \quad (4)$$

为了避免解析推导 $n \times p$ 个灵敏度方程的复杂性, 可由下述有限差分法来计算灵敏度矩阵:

$$\left[\frac{\partial \tilde{\boldsymbol{y}}}{\partial \boldsymbol{\theta}}(i) \right]_{kj} = \frac{\tilde{\boldsymbol{y}}_{pk}(i) - \boldsymbol{y}_k(i)}{\delta \theta_j}. \quad (5)$$

摄动响应 $\tilde{\boldsymbol{y}}_{pk}$ 可由下述摄动系统方程计算:

$$\begin{aligned}\tilde{\boldsymbol{x}}_p(i) &= \tilde{\boldsymbol{x}}_p(i-1) + \int_{t_{i-1}}^{t_i} f[\boldsymbol{x}_p(t), \boldsymbol{u}(i), \boldsymbol{\theta} + \delta \boldsymbol{\theta}] dt, \\ \tilde{\boldsymbol{y}}_p(i) &= h(\tilde{\boldsymbol{x}}_p(i), \boldsymbol{u}(i), \boldsymbol{\theta} + \delta \boldsymbol{\theta}), \\ \hat{\boldsymbol{x}}_p(i) &= \tilde{\boldsymbol{x}}_p(i) + K_p [\boldsymbol{z}(i) - \tilde{\boldsymbol{y}}_p(i)],\end{aligned}\quad (6)$$

其中 K_p 为广义 Kalman 滤波增益阵.

以上即是 Jategaonkar 算法的基本思想, 其同时计及了过程及观测噪声.

3 灵敏度计算的最佳摄动有限差分算法

Jategaonkar 给出的算法采用有限差分法计算灵敏度矩阵, 避免了人为解析推导灵敏度方程的复杂性, 但却需人为选择 p 个摄动量 $\delta \theta_j (j = 1, 2, \dots, p)$. 摄动量的选择直接影响灵敏度计算的精度, $\delta \theta_j$ 太大或太小, 将导致计算的灵敏度受截断误差或舍入误差的影响, 如取 p 个 $\delta \theta_j$ 相同, 则计算的灵敏度矩阵各元素的精度将不同, 参数估计的收敛性将受精度最差的影响, 因此应确定最佳摄动 $\delta \theta_j$, 下面将讨论如何确定最佳 $\delta \theta_j$ 的方法.

首先考虑一阶摄动情形, 对于摄动 $\delta \theta_j$ 有

$$\nabla_{\theta_j} J = J_{\theta_j} = [J(\boldsymbol{\theta}_j + \delta \theta_j) - J(\boldsymbol{\theta}_j)] / \delta \theta_j, \quad (7)$$

其中 $\boldsymbol{\theta}_j + \delta \theta_j = (\theta_1, \theta_2, \dots, \theta_j + \delta \theta_j, \dots, \theta_p)^T, \boldsymbol{\theta}_j = (\theta_1, \theta_2, \dots, \theta_p)^T, \delta \theta_j$ 的初始标准选

择是 10^{-6} .

假定 J_{θ_j} 已知, 则对一阶情形有

$$\delta J = J_{\theta_j} \cdot \delta \theta_j, \quad (8)$$

为了在第 k 个重要位数摄动指标函数 J , 要求

$$\delta J/J = 10^{-k} = \varepsilon. \quad (9)$$

由式(8)—(9)有

$$\delta \theta_j = \varepsilon J / J_{\theta_j}, \quad (10)$$

因此可先由(7)式计算 J_{θ_j} , 再由(7), (10)式改进其计算. 对于迭代参数最优化, 由式(10)计算的 $\delta \theta_j$ 将用于下步迭代计算 J_{θ_j} , 但这一过程不能被继续到极小化, 因为此时 $J \rightarrow \min$, $J_{\theta_j} \rightarrow 0$, 预测的 $\delta \theta_j \rightarrow \infty$.

为此, 进一步给出如下二阶方法. 对于二阶摄动情形

$$\nabla_{\theta_j} J = J_{\theta_j} = [J(\theta_j + \delta \theta_j) - J(\theta_j - \delta \theta_j)] / 2\delta \theta_j, \quad (11)$$

在参数估计过程中, $J(\theta_j + \delta \theta_j)$, $J(\theta_j)$, $J(\theta_j - \delta \theta_j)$ 能够计算, 因此有

$$\nabla_{\theta_j}^2 J = J_{\theta_j \theta_j} = [J(\theta_j + \delta \theta_j) - 2J(\theta_j) + J(\theta_j - \delta \theta_j)] / \delta \theta_j^2. \quad (12)$$

同样可在第 k 个重要位摄动 J , 即 $\delta J/J = \varepsilon$, 由 Taylor 展开有

$$\delta J = J_{\theta_j} \delta \theta_j + J_{\theta_j \theta_j} \delta \theta_j^2 = \varepsilon J, \quad (13)$$

则

$$\delta \theta_j = \{-J_{\theta_j} \pm \sqrt{J_{\theta_j}^2 + 4\varepsilon J J_{\theta_j \theta_j}}\} / 2J_{\theta_j \theta_j} \quad (14)$$

存在 $\delta \theta_j$ 的四个值可供选择 ($\varepsilon > 0, \varepsilon < 0$), 如图 1 所示. 如果 $J_{\theta_j \theta_j} > 0$, 则期望的 $\delta \theta_j$ 是 $\varepsilon > 0$ 的两个解中较小的一个, 相似地, 若 $J_{\theta_j \theta_j} < 0$, 则 $\varepsilon < 0$, 由于 $J_{\theta_j \theta_j}$ 和 ε 同号, 且 ε 与 $J_{\theta_j \theta_j}$ 相乘, 因此可将 $\varepsilon \cdot J_{\theta_j \theta_j}$ 写成 $\varepsilon \cdot |J_{\theta_j \theta_j}|$, 而 ε 始终为正. $\delta \theta_j$ 的最终选择是

$$\delta \theta_j = \min \{ \{-J_{\theta_j} \pm \sqrt{J_{\theta_j}^2 + 4\varepsilon J |J_{\theta_j \theta_j}|}\} / 2J_{\theta_j \theta_j} \}. \quad (15)$$

$\delta \theta_j$ 将用于下步迭代估计时计算 $\partial J / \partial \theta_j$. 在初始迭代时要使用式(11)—(12)去计算 $\partial J / \partial \theta_j$ 及 $\partial^2 J / \partial \theta_j^2$ ($j = 1, 2, \dots, p$).

由以上给出的二阶方法可见, 当 $J \rightarrow \min$ 时, $\partial J / \partial \theta_j \rightarrow 0$, 但 J 和 $J_{\theta_j \theta_j}$ 接近常值, 意味着 $\delta \theta_j$ 接近常值.

由以上方法计算的灵敏度矩阵各元素将有相同的精度, 保证了参数估计能达到最大的收敛度.

4 具有快速三角化平方根滤波的极大似然算法

Jategaonkar 等给出的算法, 采用普通广义卡尔曼滤波进行非线性系统的状态估计,

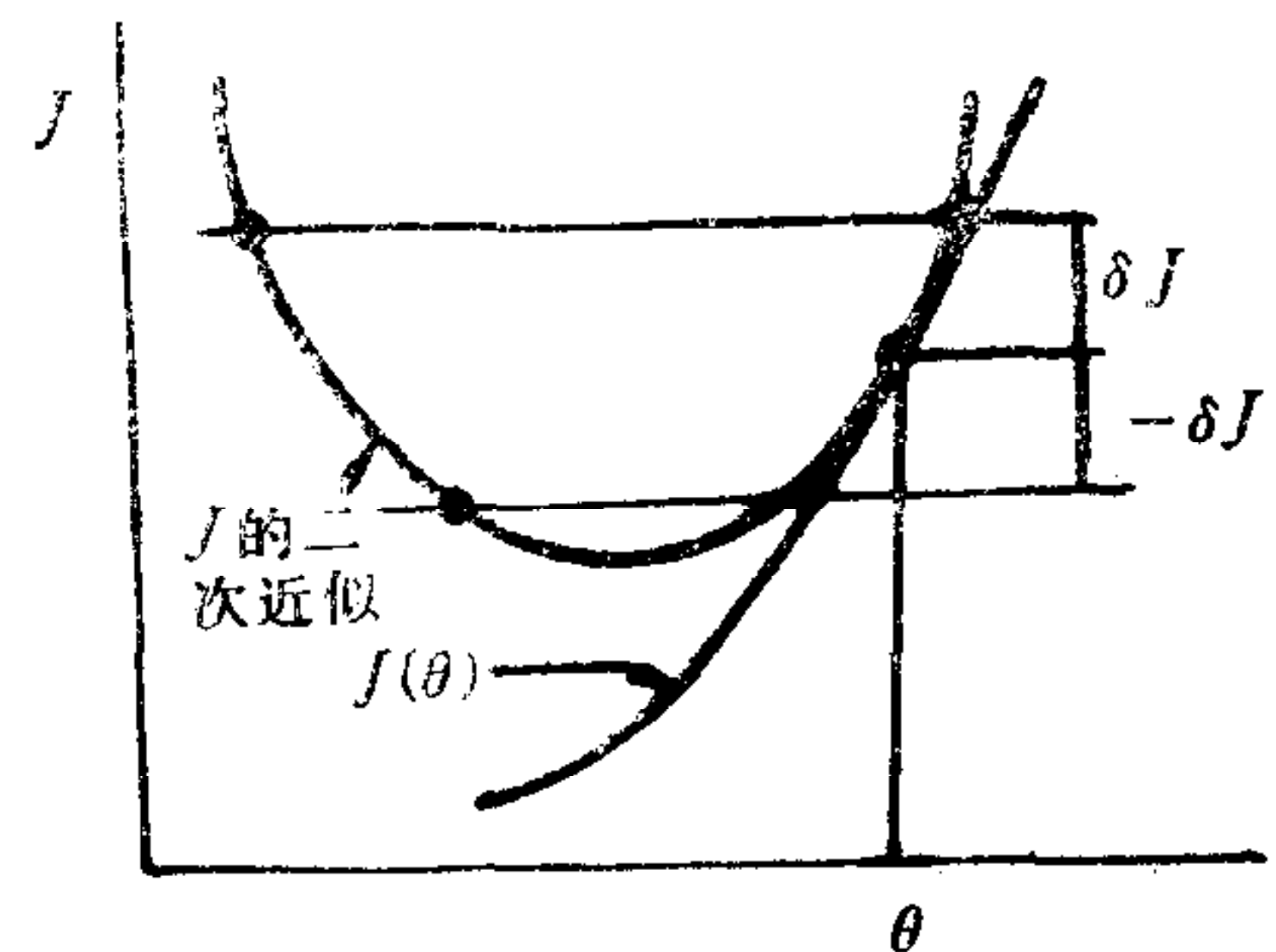


图 1 对于给定 δJ 的摄动

其滤波的数值稳定性常常难于保证,尽管可以采用所谓稳定形式的 Kalman 滤波算法,但正如 B. D. O. Anderson 所指出的,其非负定性不是自动地得到保证^[3]。解决该问题的有效方法是对滤波误差协方差阵 P 进行平方根分解,即 $P = SS^T$, S 为上或下三角阵,得到协方差平方根滤波算法。但该算法在测量修正过程中常会使 S 失去三角性而导致计算量及存贮量加大,为此下面将基于分析的 Cholesky 分解及 Carlson 的工作给出非线性系统的快速三角化平方根滤波算法^[4]。

对于由式(1)所描述的非线性系统有

$$\left. \begin{aligned} A(i) &= \frac{\partial f[\mathbf{x}(t), \mathbf{u}(t), \theta]}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\hat{\mathbf{x}}(i)} \\ B_1(i) &= \frac{\partial f[\mathbf{x}(t), \mathbf{u}(t), \theta]}{\partial \mathbf{u}} \Big|_{\mathbf{x}=\hat{\mathbf{x}}(i), \mathbf{u}=\hat{\mathbf{u}}(i)} \\ \phi(i+1, i) &= I + A(i)\Delta t + A^2(i) \frac{\Delta t^2}{2!} + \dots \\ F(i+1, i) &= \left[I\Delta t + A(i) \frac{\Delta t^2}{2!} + A^2(i) \frac{\Delta t^3}{3!} + \dots \right] \cdot B_1(i). \end{aligned} \right\} \quad (16)$$

时间更新

$$\begin{aligned} \hat{\mathbf{x}}(i+1) &= \hat{\mathbf{x}}(i) + \int_{t_i}^{t_{i+1}} f[\mathbf{x}(t), \bar{\mathbf{u}}(i), \theta] dt, \\ W(i+1/i) &= \phi(i+1, i)S(i/i), \\ S(i+1/i) &= [W(i+1/i)W^T(i+1/i) + \Delta t \cdot F(i+1, i)Q(i)F^T(i+1, i)]^{1/2}, \quad (17) \\ \hat{\mathbf{y}}(i+1) &= h[\hat{\mathbf{x}}(i+1), \mathbf{u}(i), \theta]. \end{aligned}$$

其中 $\bar{\mathbf{u}}(i) = [\mathbf{u}(i+1) + \mathbf{u}(i)]/2$, Q 为过程噪声方差阵。令

$$\begin{aligned} H(i+1) &= \frac{\partial h[\mathbf{x}(t), \mathbf{u}(t), \theta]}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\hat{\mathbf{x}}(i+1)} \\ &= [H_1^T(i+1), H_2^T(i+1), \dots, H_m^T(i+1)]^T, \end{aligned} \quad (18)$$

其中 $H_j(i+1) (j=1, 2, \dots, m)$ 为 $H(i+1)$ 的第 j 行向量。则快速三角化平方根滤波的测量修正方程为

$$\begin{aligned} \mathbf{f}(i+1) &= S^T(i+1/i)H_j^T(i+1), \\ \alpha_0 &= R_{jj}(i+1), \\ \mathbf{d}_0(i+1) &= 0, \\ \alpha_k &= \alpha_{k-1} + f_k^2(i+1), \\ a_k &= (\alpha_{k-1}/\alpha_k)^{1/2}, \\ c_k &= f_k(i+1)/(\alpha_k\alpha_{k-1})^{1/2}, \\ \mathbf{d}_k(i+1) &= \mathbf{d}_{k-1}(i+1) + S_k(i+1/i)f_k(i+1), \\ S_k(i+1) &= S_k(i+1/i)a_k - \mathbf{d}_{k-1}(i+1)c_k, \\ \hat{\mathbf{x}}(i+1) &= \hat{\mathbf{x}}(i+1) + \mathbf{d}_n(i+1)\{z_j - h_j[\hat{\mathbf{x}}(i+1), \mathbf{u}(i), \theta]\}/\alpha_n. \\ S(i+1/i+1) &= [S_1(i+1/i+1), S_2(i+1/i+1), \dots, S_n(i+1/i+1)]. \end{aligned} \quad (19)$$

其中 $\alpha_0, \alpha_k, a_k, c_k$ 是标量; $\mathbf{d}_0(i+1), \mathbf{d}_k(i+1)$ 是 n 维矢量, $f_k(i+1)$ 及 $S_k(i+1)$

分别是 $f(i+1)$ 及 $S(i+1/i+1)$ 的第 k 个分量; $H_j(i+1)$ 及 z_j 具有同样的含义; R_{jj} 是测量噪声方差阵的第 j 个对角元。

以上的快速三角化平方根滤波算法能够有效地保证滤波误差协方差阵始终是正定的, 计算简单, 且按照 Maybeck 的定义, 其是数值稳定的^[4]。

由以上的平方根滤波算法替换 Jategaonkar 算法中的广义 Kalman 滤波即构成了具有快速三角化平方根滤波的极大似然算法, 其数值稳定性将明显地优于 Jategaonkar 算法。事实上还可由平方根信息滤波或 UDU^T 分解滤波与极大似然法相结合构成具有平方根信息滤波的极大似然算法及具有 UDU^T 分解滤波的极大似然算法, 这里不再讨论。

以上改进的非线性连续-离散系统的极大似然算法经应用于飞行器系统参数估计证明方法正确、有效。限于篇幅, 不再给出具体应用结果。

参 考 文 献

- [1] Maine R E and Iliff K W. Formulation and implementation of a practical algorithm for parameter estimation with process and measurement noise. *SIAM J. Appl. Math.*, 1981, **40**: 558—579.
- [2] Jategaonkar R V and Plaetschke E. Identification of moderately nonlinear flight mechanics systems with additive process and measurement noise. *J. Guidance, Control and Dynamics*, 1990, **13**: 277—285
- [3] BDO 安德森, JB 摩尔著, 卢伯英译, 最佳滤波, 北京: 国防工业出版社, 1983.
- [4] Maybeck P S. Stochastic models, estimation and control. Academic Press, 1979.

IMPROVED MAXIMUM LIKELIHOOD ESTIMATION FOR NONLINEAR CONTINUOUS-DISCRETE SYSTEM AND ITS APPLICATIONS

LI NAIHONG WU YAOSHUA HUANG WENHU

(Dept. of Aerospace Engineering and Mechanics, Harbin Institute of Technology Harbin 150001)

ABSTRACT

In this paper, the maximum likelihood algorithm with process and measurement noise for nonlinear continuous-discrete system given by Jategaonkar and Plaetschke is improved in two aspects: 1) to avoid the disadvantages of sensitivities computation by common finite-difference method, in which the perturbation sizes should be selected optionally, an improved finite-difference method with best perturbations is presented. 2) to improve the numerical stability of Jategaonkar's algorithm, maximum likelihood algorithm with fast triangular square-root decomposition filter is given. Applications to the parameter estimation of flight vehicles demonstrated the advantages of this improved algorithm.

Key words: Maximum likelihood estimation, nonlinear system, sensitivity, square-root filter.