

噪声的复杂特征建模与数据融合精度评估

易东云,王正明

(国防科技大学数学与系统科学系,湖南长沙 410073)

摘 要: 本文分析噪声的复杂性可归结为模型误差的两个本质特征:混合误差特性和系统误差的不确定性,建立了描述混合误差中次低频部分的自由节点样条参数模型和高频部分的时变自回归模型,提出了结合数据融合过程的系统误差状态模型和精度折扣因子方法,最后,导出了数据融合处理参数估计精度的综合计算过程,实际计算结果说明了方法的合理性。

关键词: 数据融合; 误差; 精度

中图分类号: TN911 **文献标识码:** A **文章编号:** 0372-2112 (2004) 09-1560-03

Complicated Characters of Noise and Precision Evaluation of Data Fusion

YI Dong-yun, WANG Zheng-ming

(Dept. of Mathematics and System Science, National University of Defense Technology, Changsha, Hunan 410073, China)

Abstract: Complicated characters of noise in dynamical measure data are discussed in this paper. The analysis results show that the complicated characters of noise comes from two essential properties of model error. One is complex error, which cannot be described by general statistics model, and another is uncertainty of systematic error. It is built that the mathematic model of including complex error described by optimal knot spline function and time varying AR noise. The model can separate accurately not only the error with secondary low frequency but also the time varying stochastic error from complex error data. Furthermore, data fusion model with systematic error state is presented and precision discount factor method based on the uncertainty of systematic error state is also realized. At last, the precision integrate evaluation process is given. Computing results show the method is adaptive to actual engineering situation well.

Key words: data fusion; error; precision

1 引言

数据融合处理系统中对于跟踪系统多误差源和各种环境干扰造成的复杂特性及其对数据融合结果精度的影响目前尚未开展深入的研究,1998年吴翊等讨论了弹道跟踪数据融合处理系统误差诊断的统计检验问题^[1],2001年宋晋敏等讨论了复杂环境的Markov模型对多传感器目标检测的影响^[2]。吴本忠等人在1999年和2001年分别讨论了混合误差情况下回归模型、非线性回归模型和半参数模型中参数估计的收敛性问题^[3~5],但文献中讨论的混合误差主要是指统计模型中相关噪声的情况或者简单地假设混合误差为三角函数形式,并且结论需要的多项条件在实际中一般很难验证甚至是不成立的,因此上述文献中研究结果在实际处理中难以运用。本文以航天飞行器跟踪系统为背景讨论误差的复杂特性并结合数据融合模型分析其对数据处理结果的影响。

2 误差的复杂特征和数据处理精度

精度是测量数据处理的核心,正确地评价数据处理的精度对于测量设备研制部门、测量设备使用部门和测量数据用户部门具有重要意义,因此一直是数据处理领域的前沿研究课题^[6~9]。

航天飞行器跟踪系统的测量设备复杂,测量体系庞大,测量环境多变;测量数据中误差来源众多,影响因素多达数十乃至上百个,导致测量误差的复杂特性。再者,航天测量数据融合处理模型是一种紧致模型,与实际物理模型存在一定的差异,导致模型误差。模型误差包含不可估计模型化系统误差、非模型化系统误差和非平稳时变相关随机误差,其特点是规律复杂,这三部分的合成表现为混合误差,如图1所示,注意到这里的混合误差与文[3~5]中讨论的混合误差比较,特性更复杂,其特点是无法用通常的统计模型表示;进一步地,测量过程中,哪个测量通道存在何种类型的系统误差难以确切事先知道,虽然可运用系统误差诊断的方法确定系统误差状态^[10],但实际上,并非所有的系统误差状态都是可诊断的,这导致系统误差状态的不确定性,也即是模型误差的不确定性。对于系统误差的不确定性的讨论,Briggs M S等人^[11]在1999年采用Bayes经验模型的方法讨论了天体观测中系统误差的分布,Liu T S等人在2001年^[12]讨论了二值分布情况下大地测量中偏差的分布特性,Grabe M在2001年^[13]则指出国际标准化组织ISO1993年提出的测量不确定度规范(UGM:测量不确定度表达指南,中国JJF1059-1999)中本身既蕴涵了系统误差概率分布的假设。这些文献中的讨论只是限于某个测量通道中系统误差的分布特性,没有多源数据融合系统处理的情

收稿日期:2003-05-28;修回日期:2004-03-05

基金项目:国家自然科学基金(No. 60003013);973计划(No. 5131701.2)

况下系统误差状态的讨论.

实际处理经验表明:正是模型误差的两个本质特征:混合误差特性和系统误差状态的不确定性,影响数据融合处理结果的精度,造成传统精度计算结果严重失真.这个问题长期困扰精度理论研究领域和实际数据使用部门,是数据融合处理理论和应用尚未解决的关键问题.

3 数学模型

3.1 混合误差的数学建模

混合误差难以用通常的统计模型描述.分析表明,图 1 中的混合误差数据中包含次低频误差和时变随机误差,需要分别建立模型,注意到随机

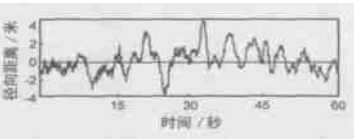


图 1 典型混合误差

误差的相关特性,两者用通常的方法难以分离.设混合误差离散数据为 $\{z(t_i), i=1, 2, \dots, n\}$, $a < t_1 < t_2 \dots < t_n = b$. 采用最优节点的样条函数模型可对混合误差建立数学模型:

$$z(t_i) = c_N^M(t_i, T^M) + e(t_i), i=1, 2, \dots, n \quad (1)$$

其中 $c_N^M(t, T^M) = \sum_{j=1}^M \sum_{j+N}^M (t - T_j)_+^N$ 为 $[a, b]$ 上的具有 M 个内节点的 N 次多项式样条函数. $T^M = (T_1, T_2, \dots, T_M)$, $a < T_1 < T_2 < \dots < T_M < b$ 为 N 次多项式样条函数的 M 个内节点, $t_+^k = t^k I(t > 0)$, $I(t > 0)$ 为示性函数. 通过样条分频算法^[14]中 最优节点个数 M^* 和最优节点 $T^{M^*} = (T_1, T_2, \dots, T_{M^*})$ 的自动调整可将混合误差数据中的次低频误差和时变随机误差分离. 由此分别得到次低频误差的自由节点样条拟合模型和随机误差的时变 AR 描述模型如下:

$$\hat{z}(t) = \hat{\alpha}_0 + \hat{\alpha}_1 t + \dots + \hat{\alpha}_N t^N + \sum_{j=1}^{M^*} \hat{\alpha}_{j+N} (t - T_j)_+^N \quad (2)$$

$$e(t) = \hat{\alpha}_1(k) e(t-1) + \dots + \hat{\alpha}_p(k) e(t-p) + \epsilon(t) \quad (3)$$

时变 AR 模型的具体参数估计过程见参考文献^[15],图 2 表示采用模型(2)和(3)对图 1 中混合误差的分离结果.

3.2 数据融合处理和系统误差状态模型

设 $X(t) = (x(t), y(t), z(t), \dot{x}(t), \dot{y}(t), \dot{z}(t))^T$ 为航天飞行器轨道参数,记号 T 表示矩阵转置,设 $y_k(t)$ 为测元 k 得到的测量数据,

$$y_k(t) = f_k(X(t)) + u_k(t) a_k + e_k(t) \quad (4)$$

其中 $k=1, 2, \dots, N$, $y_k(t)$ 为距离,速度或者角度等的测量信号,共有 N 个测量通道. 已知函数 $f_k(X(t))$ 为真实信号,已知函数向量 $u_k(t) = (u_{k1}(t), \dots, u_{kL}(t))$ 为系统误差项, $a_k = (a_{k1}(t), \dots, a_{kL}(t))^T$ 为系统误差系数向量, L 为可能的系统误差种类, $e_k(k)$ 为零均值随机误差. 采用弹道的节省参数模

型^[11],对于 $t_1 \leq t \leq t_n$ 有弹道的样条函数表示:

$$x(t) = B_x(t) b_x, \dot{x}(t) = \dot{B}_x(t) b_x, y(t) = B_y(t) b_y, \dot{y}(t) = \dot{B}_y(t) b_y, z(t) = B_z(t) b_z, \dot{z}(t) = \dot{B}_z(t) b_z.$$

其中 $B_x(t), b_x$ 分别为 B 样条函数基和样条系数向量^[11],其余类同. 由此, $t_1 \leq t \leq t_n$ 时间段的弹道由样条系数 $b = (b_x^T, b_y^T, b_z^T)^T$ 确定. 引进下面记号:

$$y(t) = (y_1(t), \dots, y_N(t))^T, a = (a_1^T, \dots, a_N^T)^T, e(t) = (e_1(t), \dots, e_N(t))^T, u(t) = \text{diag}(u_1(t), \dots, u_N(t)), f(b, t) = f(X(t)) = (f_1(X(t)), \dots, f_N(X(t)))^T, Y = (y(t_1)^T, \dots, y(t_n)^T)^T, U = (u(t_1)^T, \dots, u(t_n)^T)^T, e = (e(t_1)^T, \dots, e(t_n)^T)^T, F(b) = (f(b, t_1)^T, \dots, f(b, t_n)^T)^T$$

由此得到弹道联解数据融合模型的向量形式为:

$$Y = F(b) + Ua + e \quad (5)$$

模型(5)可以衍生出两个模型:一是实际数据模型:假设某次测量中系统误差状态处于 $S^j = (S_1^j, \dots, S_N^j)$, $S_k^j = (\hat{l}_{k1}^j, \hat{l}_{k2}^j, \dots, \hat{l}_{kL}^j)$, $\hat{l}_{km}^j \in \{0, 1\}$, $m=1, 2, \dots, L$, \hat{l}_{km}^j 为 1 表示第 k 个测量通道存在第 m 类系统误差,为 0 则没有. 由此有实际数据模型:

$$Y = F(b) + Ua^{S^j} + e \quad (6)$$

其中 $a_k(S_k^j) = (a_{k1} \hat{l}_{k1}^j, \dots, a_{kL} \hat{l}_{kL}^j)^T$, $a^j = (a_1(S_1^j)^T, \dots, a_N(S_N^j)^T)^T$. 二是数据融合处理模型:由于系统误差的复杂性,实际中可能采用的数据融合处理模型为

$$Y = F(b) + Ua^{S^i} + e \quad (7)$$

其中 $S^i = (S_1^i, \dots, S_N^i)$, $S_k^i = (\hat{l}_{k1}^i, \hat{l}_{k2}^i, \dots, \hat{l}_{kL}^i)$, $\hat{l}_{km}^i \in \{0, 1\}$, $m=1, 2, \dots, L$.

模型(6)和(7)的差异即是由系统误差状态的不确定性造成的. 传统数据处理中假设数据融合处理模型总是正确的,此时 $S^i = S^j$,但在实际中由于测量系统和测量环境的复杂性,这个假设并不一定成立. 定义 $d_{ij}(\cdot)$ 表示采用带误差状态 S^i 的数据融合处理模型(7),而实际误差状态为 S^j 时的数据融合模型(6)中某个参数 的估计精度.

假设某次测量采用带误差状态 S^{i_0} 的数据融合处理模型,系统误差状态的不确定性采用分布 $\{ (S^j, p_j), j=1, \dots, L \}$ 来描述,定义待估参数 的精度折合因子为:

$$t = \sum_{j=1}^L d_{i_0 j}(\cdot) p_j / d_{i_0 i_0}(\cdot) \quad (13)$$

折合后,参数 的估计精度修正为: $\hat{\alpha} = \alpha t$,其中 为带误差状态 S^{i_0} 数据融合模型输出的参数 的估计精度.

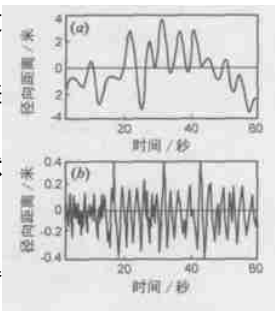
精度折合因子的直观意义在于分子表示数据融合处理系统在各种系统误差状态下参数 估计的平均精度,分母是特定状态下的参数 的估计精度,两者的比例反映了数据融合处理结果的可信度. 估计精度的折合修正反映了系统误差状态的不确定性对数据融合处理结果的影响,传统精度计算方法只是该处理方法的一个特例,即 $t=1$ 的情况.

4 精度的综合计算

有了以上准备,下面给出数据融合处理精度的综合计算步骤:

Step1 由多通道测量数据建立数据融合处理模型;

图 2 混合误差数据的分解 (a) 上图次低频误差; (b) 下图随机误差



Step2 根据数据融合处理的误差分析结果,结合混合误差模型(1),获得各通道混合误差特性的拟和模型(2)和时变随机误差模型(3),并结合实际工程背景经验,给出系统误差状态的分布;

Step3 由数据融合解算的轨道为基础产生一条仿真轨道 $X(t)$;

Step4 根据实测站址产生相应跟踪系统全部的跟踪测元;

Step5 由时变随机误差模型(3)模拟产生全跟踪过程各个测元的随机误差;

Step6 结合拟和模型(2)产生全跟踪过程各个测元的混合误差;

Step7 由 Step2, Step3, Step4 即可产生全跟踪过程的各个测元的仿真测量数据;

Step8 由文献[1]中非线性融合自校准方法求解轨道 $X(t)$;

Step9 给出本次仿真的轨道误差 $X(t) = X(t) - X(t)$;

Step10 重复 Step5 ~ Step9 过程 N 次可得到 N 条轨道误差序列 $\{X_k(t), k=1, 2, \dots, N\}$;

Step11 统计得轨道平均误差均值及标准差

$$\bar{X}(t) = \frac{1}{N} \sum_{k=1}^N X_k(t),$$

$$(\bar{X}(t)) = \sqrt{\frac{1}{N} \sum_{k=1}^N (X_k(t) - \bar{X}_k(t))^2}$$

Step12 结合数据融合模型和系统误差状态的分布计算各轨道参数的精度折合因子;

Step13 由各轨道参数的精度折合因子对轨道估计精度进行修正。

计算实例 某飞行器试验靶场有大型跟踪设备多套,其中主跟踪系统一套,辅助跟踪系统两套,某次实际跟踪测量中共获得 28 个测量通道的测量数据,根据上述精度综合计算的步骤进行计算,弹道六个参数的估计精度计算结果列于表 1。

表 1 结果的说明:表 1 中第二行是传统精度计算的结果,第三行的数值结果表明混合误差的影响为传统精度结果的 3~6 倍,第四行的数值结果表明系统误差的不确定性导致的精度折合因子值在 1.3~2.2 之间,由此,最终精度为原数据处理精度的 4~12 倍左右,这个结果是与设备的实际情况是相符合的,并且获得了实际工程部门的认同,而传统方法计算的精度结果过于乐观,没有反映跟踪的实际情况。根据处理多次实际数据的经验,一般地实际处理结果中误差复杂特性的影响可为随机误差的影响的 3 倍到 15 倍。

表 1 三种情况下参数估计精度的比较

(x 单位为米, \dot{x} 单位为米/秒,其余类同)

参数估计的精度	x	y	z	\dot{x}	\dot{y}	\dot{z}
只考虑随机误差	1.5	1.2	0.3	0.23	0.16	0.02
加入混合误差	5.2	3.8	1.6	0.44	0.41	0.05
加入系统误差状态的不确定性	9.4	5.1	3.5	0.65	0.58	0.08

5 结论

由于误差的复杂特性,传统方法的精度计算结果往往与实际情况相差甚远。本文将航天测量误差的复杂特征归结为模型误差的两个本质特征:混合误差和系统误差状态的不确

定性,通过建立混合误差的自适应样条参数模型和随机误差的时变 AR 模型以及基于系统误差状态分布的精度计算模型,系统地讨论了航天测量误差的复杂特征与数据融合处理精度的内在联系,给出了综合精度计算方法。实测数据处理结果说明了该方法的有效性,客观地考虑了误差的复杂性及其对精度的影响。

参考文献:

- [1] 吴翎,朱炬波.弹道数据处理的融合算法[J].中国科学(E辑),1998,28(6):505-512.
- [2] Sun Qingwei, Wang Jianbin, Song Jinmin et al. Distributed detection with multisensor under complicated disturbing environment [J]. Acta Electronica Sinica, 2001, 29(3): 346-349.
- [3] 吴本忠.混合误差下回归权函数估计的强相合性[J].数学杂志,1999,19(1):93-99.
- [4] 吴本忠,胡舒合.相依误差下非线性回归模型 LS 估计的收敛速度[J].中国科学(A辑),2001,31(7):593-601.
- [5] 朱仲义,韦博成.半参数非线性模型的统计诊断与影响分析[J].应用数学学报,2001,24(4):568-581.
- [6] Frank J Wentz, Matthias Schabel. Precise climate monitoring using complementary satellite data sets[J]. Nature, 2000, (403): 414-416.
- [7] Hammerand D C, Oden J T, Prudhomme S, Kuczma M S. Modeling Error and Adaptivity in Nonlinear Continuum System [R]. No. DE2001-780285/ XAB, USA: NTIS, 2001.
- [8] Pagione M M, Ryan H F, Oaks R D, Summerill J S, Cale M L. Trajectory Prediction Accuracy Report: User Request Evaluation Tool (URET)/Center-TRACON Automation System (CTAS) [R]. No. ADA365862/ XAB, USA: NTIS, 1999.
- [9] Agapov M Y, Bournashev M N. Errors of measurement by laser goniometer[J]. The International Society for Optical Engineering, 2001, (4316): 26-30.
- [10] 朱炬波,易东云.雷达测量系统误差的残差诊断方法[J].系统工程与电子技术,2000,22(7):42-44.
- [11] Briggs M S, Pendleton G N, Kippen R M, et al. Error Distribution of BAISE GRB Location [R]. No. N19990095036/ XAB, USA: NTIS, 1999.
- [12] Liu T S, Finley T D. Estimating Bias Error Distributions [R]. No. N20010000886/ XAB, USA: NTIS, 2001.
- [13] Grabe M. Estimation of measurement uncertainties-An alternative to the ISO Guide [J]. Metrologia, 2001, 38(2): 97-106.
- [14] 易东云,朱炬波,王正明.三频带信号分离的样条函数方法[J].电子学报,1999,27(1):54-58.
- [15] 易东云,王正明.时变混合模型的参数估计[J].控制理论及应用,1999,16(5):732-735.

作者简介:



易东云 男,1965 年生于湖南株洲,现为国防科技大学数学与系统科学系教授、博士生导师,获全军科技进步一等奖两项,合作出版专著两部,发表论文三十余篇,主要研究领域为数据融合和智能计算。