

基于总体最小二乘的多功能 传感器信号重构方法研究

孙金玮, 刘昕, 孙圣和

(哈尔滨工业大学自动测试及控制系, 黑龙江哈尔滨 150001)

摘要: 基于总体最小二乘法的基本原理和 Moore-Penrose 广义逆矩阵的基本性质, 本文提出了一种新的算法, 用于多功能传感器输入信号的数值重构. 算法既考虑了传感器输出信号的漂移, 同时也考虑了输入信号的误差. 仿真结果表明, 该算法同传统的最小二乘策略相比具有较高的重构准确度和可靠性.

关键词: 总体最小二乘法; Moore-Penrose 广义逆; 信号重构

中图分类号: TN98 **文献标识码:** A **文章编号:** 0372-2112 (2004) 03-0391-04

TL S Algorithm-Based Study on Multi-Functional Sensor Data Reconstruction

SUN Jin-wei, LIU Xin, SUN Sheng-he

(Dept of Automatic Measurement and Control, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

Abstract: A method about multi-functional sensor measurand reconstruction is brought forward, which is based on total-least square algorithm and the general reverse function of Moore-Penrose matrix. It evaluates both the sensor output bias and sensor input error. Emulation results indicate that the proposed algorithm is more accurate and reliable for the measurand reconstruction than conventional LS strategy.

Key words: total-least square algorithm; Moore-Penrose matrix; measurand reconstruction

1 引言

近年来, 多功能传感器的研究逐渐得到重视. 一般而言, 多功能传感器可测量多种非电量信号, 代替普通传感器, 可大大减小测量系统体积和功耗, 在航天、航空和微机械制造业等领域具有重要意义. 有一类多功能传感器事实上是多函数型的, 即假设传感器的输入信号为 x_1, x_2, \dots, x_n , 则其输出为

$$\begin{cases} y_1 = f_1(x_1, x_2, \dots, x_n) \\ y_2 = f_2(x_1, x_2, \dots, x_n) \\ \dots \\ y_m = f_m(x_1, x_2, \dots, x_n) \end{cases}$$

一般情况下, $f_1 \sim f_m$ 为非线性函数, 且传感器的输入信号数量与输出信号数量相等, 即 $m = n$. 多函数型传感器在环境参数测量^[1,2]和一般工业测量^[3,4]设计中已经得到了一定程度的开发应用.

显然, 被测量的获得依赖于对上文测量方程组的求解, 即根据传感器的输出值 $y_1 \sim y_m$ 求出 $x_1 \sim x_n$. 这一过程也可以称作传感器信号重构, 它是多功能传感器一个重要的环节. 信号重构一般有以下几种主要方法^[5]: (1) 第一次序微元积分法; (2) 间隔分析; (3) 统计模型; (4) 模糊变量代数法. 它们可以很好的解决一元情况下的信号重构. 二元函数情况下的多传感

器系统信号重构可以采用分段线性最小二乘法^[6], 但该方法忽略了传感器输入信息的误差, 极易出现矛盾方程组, 使得信息值的分离与重构都十分困难, 甚至无法实现. 基于总体最小二乘法^[7]及 Moore-Penrose 广义逆^[8]的相关原理, 本文提出了一种新的传感器信号重构算法, 在传感器输入、输出信息值都存在误差的情况下对线性条件下的多传感器输入信号进行信号重构, 具有非线性函数关系的传感器同样可以做分段近似线性以应用本算法. 由理论推导可知, 当输入、输出数据被误差严重污染时, 该方法的参数估计精度明显优于最小二乘法^[9].

2 算法及原理简介

2.1 总体最小二乘法

考虑超定方程: $Ax = b$ (1)

式中, $A \in R^{m \times n}$, $x \in R^n$, $b \in R^m$, $m > n$. 用最小二乘法解方程 (1) 时, 就相当于找一个解 x_{LS} 使得 $\|b - Ax_{LS}\|_2 = \min_x \|b - Ax\|_2$. 这个问题实际上是在 A 的值域上找一个 b_{LS} 使得 $\|b - b_{LS}\|_2 = \min$; 且 A 的值域定义为 $\text{range}(A) = \{b \in R^m, b = Ax \text{ 对某个 } x \in R^n\}$. 可见, 用最小二乘法 (LS) 解方程 (1) 时, 一般都基于假设仅有观测向量 b 存在着误差 δb , 这时方程 (1) 可表示为:

收稿日期: 2002-08-02; 修回日期: 2003-05-18

基金项目: 国家自然科学基金 (No. 60172071, 60372005)

$$Ax = b + \Delta b \tag{2}$$

令 Δb 的误差平方和为:

$$e = \|b - Ax\|_2^2 = (b - Ax)^T (b - Ax) \tag{3}$$

则 LS 的求解准则就是使 (3) 中的 e 最小. 也可以认为, LS 等同于用一个最小的 Δb 去扰动 b 以便 $b + \Delta b$ 可以用 A 的各列来预测. 即 LS 法只考虑了观测向量 b 的扰动, 而没有考虑系数矩阵 A 的扰动. 但在实际应用中, 矩阵 A 常常由系统的输入、输出数据所构成, 数据中不可避免的含有误差, 所以矩阵 A 通常也含有误差. (由于 A 是由输入、输出信号所构成, 所以信号的误差就是 A 产生扰动的原因.) 因此, 更合理的方法是同时考虑 Δb 和 ΔA 二者的误差. 这就是总体最小二乘法 (TLS) 的基本思想. 现设 A 和 b 同时存在着扰动误差 E 和 e , 则方程 (1) 可改写为:

$$(A + E)x = b + e \tag{4}$$

由于矩阵 A 的扰动误差不为零, 其最小二乘解一般是有偏的. 应用 TLS 方法来考虑方程 (4) 的求解, 式 (4) 可以变换为:

$$(B + D)z = 0 \tag{5}$$

其中, $B = \begin{bmatrix} -\Delta b \\ A \end{bmatrix}$, $D = \begin{bmatrix} -E \\ E \end{bmatrix}$, $z = \begin{bmatrix} 1 \\ x \end{bmatrix}$. 式 (5) 的 TLS 解可以简单表示为求一个解向量 z , 使得 $\|Dz\|_2 = \min_{\|z\|_2=1} \|(b + e) - (A + E)x\|_2$. 下面利用奇异值分解 (SVD) 给出式 (5) 的 TLS 解.

令 $B = USV^T$, 其中 U 、 V 为酉阵, S 为对角阵. 则式 (5) 的 TLS 求解可分以下两种情况:

(1) $S(n, n)$ 明显比 $S(n+1, n+1)$ 大. 此时, 存在唯一解:

$$x_{\text{TLS}} = \frac{1}{v(1, n+1)} \begin{bmatrix} v(2, n+1) \\ \dots \\ v(n+1, n+1) \end{bmatrix} \tag{6}$$

式中 $v(i, n+1)$ 是右奇异阵 V 的第 $n+1$ 列向量中的第 i 个元素.

(2) B 的后面若干个奇异值是重复的 (实际应用中常表现为多个奇异值很接近). 此时, 会有多个 TLS 解. 但在这种情况下, 找出某种意义下的唯一解是可能的. 可能的唯一解有两种: 最小范数解和最佳最小二乘近似解 (本文采用最小范数解). 下面给出最小范数解的算法:

计算 B 的 SVD, 即 $B = USV^T$, 并存储 V ;

利用 $S(p, p) > S(n+1, n+1) + \epsilon$, $S(p+1, p+1)$

... $S(n+1, n+1)$ 决定 p , 其中 $\epsilon > 0$ 是某个很小的数;

令 $V_1 = [v_1, \dots, v_{p+1}]$ 是 V 的一个列分块, 并计算 Householder 矩阵 Q 使得:

$$V_1 Q = \begin{bmatrix} a & 0 & \dots & 0 \\ y & & X & \end{bmatrix}$$

其中 X 代表不感兴趣的块. a 为标量, y 为列向量.

若 a 不为 0, 则有:

$$x_{\text{TLS}} = \frac{y}{a} \tag{7}$$

若 a 为 0, 就应该减小原定的指数 p , 也就是扩大矩阵 V_1 的列数, 重复第三步, 直到得到一个不为 0 的 a . 此时, 由

就可得到 x_{TLS} 的最小范数解.

观察步骤 可发现: a, y 可通过使 Q 的第一列为 V_1 第一行的复数共轭来获得 (还有其它各种方法, 但这是最简单的一种). 由于确定解 x_{TLS} 只需利用 a 和 y , 所以只要计算 Q 的第一列即可. 于是式 (7) 又可表示为:

$$x_{\text{TLS}} = \frac{V_2 v_1^*}{V_1 v_1^*} = a^{-1} V_2 v_1^* \tag{8}$$

其中, v_1 是矩阵 V_1 的第一行构成的一个行向量, 而 V_2 是 V_1 除 v_1 外其它各行组成的矩阵, 即:

$$V_1 = \begin{bmatrix} v_1 \\ V_2 \end{bmatrix} \tag{9}$$

$a=0$ 的情况对应于 $a = v_1 v_1^* = 0$, 即 v_1 是个零向量, 此时可通过步骤 来求解.

对于 $a=0$ 的情况, 只需有限次的迭代就可得到 x_{TLS} 的最小范数解. 因为若 a 为零, 则需扩大矩阵 V_1 的列数, 直到找到一个不为零的行向量 v_1 . V_1 扩展到极限时为 V , 而 V 的第一行为非零向量 (当 V 是实数矩阵时), 其对应的 a 必不为零. 现证明如下: 因为 V 为实的酉阵, 所以 V 为正交阵. 则由正交阵性质知 $|V_{11}| = 1$ 或 -1 , 所以 V 的第一行为非零向量. 因此对于由步骤 所确定的指数 p , 至多经 p 次迭代就可得到一个不为零的 a .

2.2 Moore-Penrose 广义逆

定义 设矩阵 $A = [a_{ij}]_{m \times n}$, 如果矩阵 $G = [g_{ij}]_{n \times m}$ 满足以下条件:

$$AGA = A; (GA)^T = GA, GAG = G; (AG)^T = AG$$

则矩阵 G 为矩阵 A 的 Moore-Penrose 广义逆, 记为 A^+ .

性质 (1) 对任意矩阵, 其 Moore-Penrose 广义逆存在且唯一; (2) 方程组 $Ax = b$ 的最小二乘最小范数解为 $x = A^+ b$, 其中 A^+ 是 Moore-Penrose 广义逆.

2.3 算法简介

多传感器的传递函数模型可描述为输入变量和输出变量之间的关系. 其模型方程如下:

$$y(t) = f[x(t)] = Fx(t) \tag{10}$$

其中, 输入变量为 $x(t) = [x_1(t), \dots, x_n(t)]$; 输出变量为 $y(t) = [y_1(t), \dots, y_m(t)]$; 传递矩阵为 $F = (F_{mn})$ 当输入、输出为线性关系时, F 转化为实数矩阵. 此时, 信号重构可分两步完成.

(1) 利用总体最小二乘法拟合传递矩阵 F .

对于线性多传感器系统, 由式 (10) 可知, 其传递函数可表示为:

$$y(t) = F_{mn} x(t) \tag{11}$$

其中, F_{mn} 为实数矩阵. 对于此多传感器系统, 经过 N 次实验可得到如下数据对:

$$(x_1(k), \dots, x_n(k), y_1(k), \dots, y_m(k)), \quad k = 1, \dots, N \tag{12}$$

先考虑拟合 F 的第一行元素, 由式 (12) 中的数据对满足式 (11) 可得:

$$\begin{bmatrix} y_1(1) \\ \dots \\ y_1(k) \end{bmatrix} = \begin{bmatrix} x_1(1) & \dots & x_n(1) \\ \dots & \dots & \dots \\ x_1(k) & \dots & x_n(k) \end{bmatrix} \begin{bmatrix} F_{11} \\ \dots \\ F_{1n} \end{bmatrix} \quad (13)$$

通常, k 大于 n . 此时, 对式 (13) 的求解就相当于对式 (1) 的求解. 由于输入信号含通常含有误差, 所以由输入信号所构成的矩阵不可避免会产生扰动, 此时若采用最小二乘法对式 (13) 进行求解, 那么解向量必然是有偏的, 因此采用总体最小二乘法更为合理. 应用总体最小二乘法可得到 (F_{11}, \dots, F_{1n}) 的 TLS 解, 进而可以得到传递矩阵 F .

(2) 利用 Moore-Penrose 广义逆来重构输入信号.
因为传递矩阵已经得到, 所以对于任意测得的输出信号

$$y(i); \text{ 都存在一个输入信号 } x(i) \text{ 使之满足式 (11), 即:} \quad y(i) = F_{mn}x(i) \quad (14)$$

$$\text{则由 Moore-Penrose 广义逆性质可知} \quad x(i) = F_{mn}^+y(i) \quad (15)$$

是其最小二乘最小范数解.

由以上分析可知, 该方法对于线性条件下的多函数型传感器的信号重构是普遍适用的.

3 仿真实验

为了验证算法是否可行, 可利用计算机仿真构造一个线性三输入、三输出多功能传感器系统的物理模型 (见图 1).

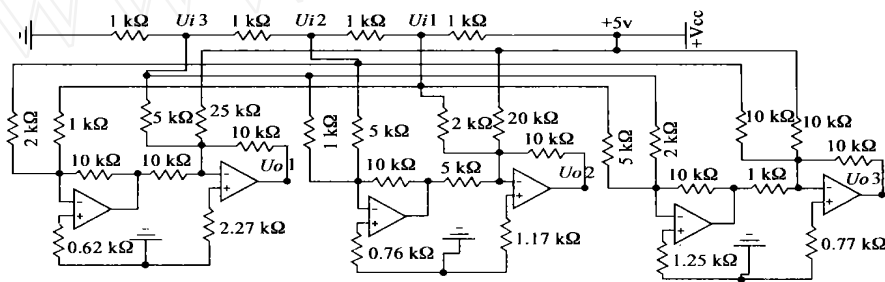


图 1 多功能传感器物理模型

其中, 输入变量为 (U_{i1}, U_{i2}, U_{i3}) ; 输出变量为 (U_{o1}, U_{o2}, U_{o3}) .

由此图可以得到该多功能传感器的实际传递函数为:

$$\begin{bmatrix} U_{o1} \\ U_{o2} \\ U_{o3} \end{bmatrix} = \begin{bmatrix} 10 & 5 & -2 & -2 \\ -5 & 4 & 20 & -2.5 \\ 20 & -1 & 50 & -5 \end{bmatrix} \begin{bmatrix} U_{i1} \\ U_{i2} \\ U_{i3} \\ 1 \end{bmatrix} \quad (16)$$

且可由计算机仿真获得多组数据对 (见表 1). 应用总体最小二乘及广义逆算法, 用表 1 的前六组数据可得到方程 (16) 中实际传递矩阵的拟合矩阵:

表 1 采样数据表 (单位: 伏特)

U_{i1}	U_{i2}	U_{i3}	U_{o1}	U_{o2}	U_{o3}	组序
0.2664	0.12770	0.01864	1.265	-2.948	1.131	1
0.2882	0.09089	0.01326	1.310	-3.312	1.337	2
0.2869	0.08442	0.01722	1.256	-3.252	1.514	3
0.2866	0.08317	0.01299	1.256	-3.340	1.298	4
0.2993	0.06412	0.01002	1.294	-3.54	1.423	5
0.3008	0.07356	0.006232	1.363	-3.585	1.254	6
0.3201	0.03708	0.004877	1.377	-3.855	1.609	7
0.3203	0.03916	0.003995	1.391	-3.865	1.566	8
0.3204	0.04065	0.003639	1.400	-3.867	1.549	9
0.3205	0.04170	0.003158	1.407	-3.872	1.526	10
0.3206	0.04316	0.002708	1.416	-3.876	1.504	11
0.3223	0.03035	0.005988	1.363	-3.870	1.715	12
0.3227	0.03642	0.003387	1.402	-3.900	1.587	13

$$F = \begin{bmatrix} 9.8928 & 4.9557 & -2.1214 & -1.9638 \\ -5.0185 & 3.9976 & 20.0066 & -2.4943 \\ 20.1276 & -0.9630 & 50.0760 & -5.0412 \end{bmatrix} \quad (17)$$

为了分析方程 (16) 中传递矩阵的拟合效果, 可以将传递矩阵每一行的元素看作是一个行向量, 则传递矩阵各行的拟合相对误差可用下式表示:

$$\frac{x - x_{\text{TLS}}}{x} \quad (18)$$

其中, x 为方程 (16) 中传递矩阵 A 的一个行向量; x_{TLS} 为式 (17) 中和方程 (16) 相对应的拟合行向量. 拟合矩阵 F 的各行向量的拟合相对误差见图 2.

由图 2 可以看出, 各行向量的拟合误差在相同的精度范围内呈离散分布, 这说明应用 TLS 法可对传递矩阵的各个行向量作等精度拟合.

表 2 重构信号比较表 (单位: 伏特)

重构信号			输入信号			组序
U_{i1}	U_{i2}	U_{i3}	U_{i1}	U_{i2}	U_{i3}	
0.3201	0.0372	0.0049	0.3201	0.03708	0.004877	7
0.3203	0.0393	0.0040	0.3203	0.03916	0.003995	8
0.3204	0.0407	0.0036	0.3204	0.04065	0.003639	9
0.3204	0.0419	0.0032	0.3205	0.04170	0.003158	10
0.3205	0.0433	0.0027	0.3206	0.04316	0.002708	11
0.3223	0.0305	0.0060	0.3223	0.03035	0.005988	12
0.3227	0.0365	0.0034	0.3227	0.03642	0.003387	13

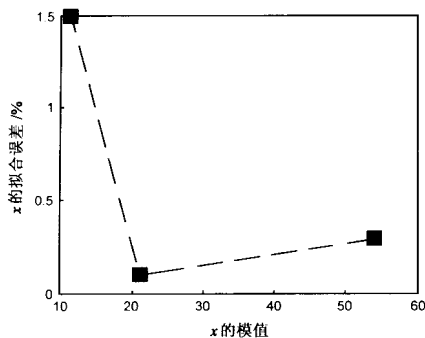


图2 拟合矩阵行向量的和相对误差

利用总体最小二乘及广义逆法对表 1 后七组输出信号重构其输入信号可得到如下重构信号,见表 2.

为了分析信号重构的精度,可以将重构信号看作是三维空间中的一个坐标点,重构信号误差如图 3 所示.由图 3 可知,对于随机给定的几个输入信号点,在相似条件下对比普通最小二乘方法的实验结果^[10],应用总体最小二乘及广义逆法都可在较高的精度下得到其重构信号点.这说明该方法具有较高的重构准确度.

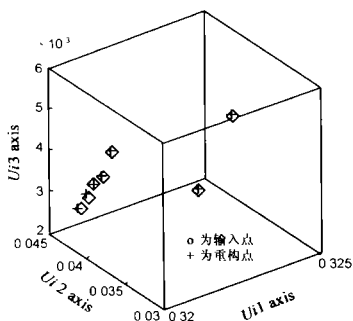


图3 重构信号误差图

4 结论

对于传递矩阵的拟合,可以转化为对超定方程组 $Ax = b$ 的求解.在求解中,一般最小二乘法只考虑了向量 b 的误差,而总体最小二乘法同时考虑了矩阵 A 和向量 b 的误差,因此总体最小二乘法对传递矩阵的参数估计精度更高.此时,可以认为所求传递矩阵为实际传递函数的最佳估计,再通过应用广义逆的性质就可以在较高的精度下重构输入信号.本文的研究成果在理论和技术上是对文献[6]的一个扩展.不仅采用了总体最小二乘算法和广义逆技术使信号重构的准确性和可靠性得到提高,而且适合三元变量情况下的多功能传感器应用.

参考文献:

- [1] K D Schierbaum, U Weimer, W Gopel. Multicomponent gas analysis: An analytical chemistry approach applied to modified SnO_2 sensors[J]. Sensors and Actuators B: Chemical, 1990, 2(4): 71 - 78.
- [2] J Sun, K Shida. Multilayer sensing and aggregation approach to environmental perception with one multifunctional sensor [J]. IEEE Sensors Journal, 2002, 2(2): 62 - 72.
- [3] J Sun, K Shida. A new multifunctional sensor for measuring oil/water two-phase state in pipelines[J]. Japanese Journal of Applied Physics, 2001, 40(3): 1487 - 1492.
- [4] Yuji J, Shida K. New multifunctional tactile sensing technique by selective data processing [J]. IEEE Transactions on Instrumentation and Measurement, 2000, 49(5): 1091 - 1094.
- [5] Tomasz Szafranski, Roman Z Mraweki. Accuracy of measurand reconstruction-comparison of four methods of analysis[A]. Proc of IMTC 98 [C]. St Paul, USA: IEEE, 1998. 32 - 35.
- [6] Alessandra F, Daniele M, Andrea T. Application of an Optimal Look-Up Table to Sensor Data Processing[J]. IEEE Transactions on Instrumentation and Measurement, 1999, 48(4): 813 - 816.
- [7] 张贤达. 现代信号处理[M]. 北京:清华大学出版社, 1993.
- [8] 邢书珍. 工程技术应用数学[M]. 北京:中国铁道出版社, 1990.
- [9] 张洪钺, 黄劲东, 范文雷. 全最小二乘法及其在参数估计中的应用[J]. 自动化学报, 1995, 21(1): 40 - 47.
- [10] Sun J, Zhang Y, Meng L. Reconstruction of signal using LS method for multi-functional sensor [A]. Proc of ISIST 02 [C]. Jinan, China: F-SIST, 2002. 312 - 316.

作者简介:

孙金玮 男, 1964 年 10 月生于哈尔滨市, 分别于 1987 年 7 月和 1990 年 2 月在哈尔滨工业大学电气工程系获得学士和硕士学位, 2000 年 9 月于日本佐贺大学理工学部获得博士学位, 现为哈尔滨工业大学教授、博士生导师, 电子学会高级会员, 主要研究领域为传感技术及其信号处理、测试系统技术, 已发表学术论文 42 篇.

刘昕 男, 1980 年 2 月生于哈尔滨市, 2002 年于哈尔滨工业大学自动测试及控制系获得学士学位, 现为哈工大自动测试及控制系硕士研究生.

孙圣和 男, 1937 年 10 月生于山东, 1961 年毕业于哈尔滨工业大学电磁测量专业, 1962 年毕业于同专业研究生班, 现为哈尔滨工业大学自动化测试与控制研究所所长、教授、博士生导师, 主要研究领域为自动化测试、信号与信息处理和光纤传感技术等, 已发表学术论文 214 篇.