第14卷第2期
1997年4月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 14, No. 2
Apr., 1997

非线性参数估计中最优解的一个充分条件

郑应文
（福州大学自动化研究所·福州，350002）

摘要：本文研究在一类非线性参数估计问题中，当某个容许参数的输出值与给定值的误差范数为局部极小时，如何判断这个参数是最优解. 文章证明了它为最优参数的一个充分条件是这个误差范数小于某个给定的邻界距离.

关键词：参数估计；非线性映射；曲线曲率；极值问题

1 引 言

考虑下面的参数估计问题：

x 是待估计的参数向量，它的容许集 E 是 Banach 空间 B 中的一个凸集，对于 E 中的不同点 x 将会得到不同的输出 y，其关系为

$$y = f(x).$$

这里 f 是从 B 到 Hilbert 空间 F 上的二次可微映射，f 映射下集合 E 的象集是 D，则 D 是 F 中的连通闭集. 我们的目的，就是要求出 E 中的点 x_0，使得它的象 $y_0 = f(x_0)$ 与 F 空间中的给定点 z 之间的距离最小，即 $y = y_0$ 时

$$J(y) = \| y - z \| = \min, \quad \forall \ y \in D.$$

例如，已知某系统满足以下方程：

$$\frac{\partial u}{\partial t} - \frac{\partial}{\partial y}\left(a^2 \frac{\partial u}{\partial y}\right) - abu = z, \quad y \in [0,1], \quad t \in [0,1],$$

$$u(0,t) = u(1,t) = 0, \quad \forall \ t > 0,$$

$$u(y,0) = u_0(y), \quad \forall \ y \in (0,1).$$

这里 $a_0 \leqslant a \leqslant a_1, b_0 \leqslant b \leqslant b_1, a, b$ 是待定的参数. 我们要从输入 u 与输出 y 中，估计出最优的参数 $\hat{a}, \hat{b}$. E 是由 a, b 构成的凸集，对不同的 a, b，设

$$\eta(a,b) = \frac{\partial u}{\partial t} - \frac{\partial}{\partial y}\left(a^2 \frac{\partial u}{\partial y}\right) - abu.$$

映射 f 是从 (a,b) 到 η 的映射，参数估计就是要从 E 中求出参数 $\hat{a}, \hat{b}$，使

$$\| \eta(\hat{a}, \hat{b}) - z \| = \min.$$

范数的计算一般要用数字仿真，在 $y \in [0,1], t \in [0,T]$ 中取 n 个点，得到 $\eta_1 \cdots \eta_n$ 与 $z_1 \cdots z_n$.

计算
$$J(\eta) = \left[\sum_{i=1}^{n} (\eta_i - z_i)^2\right]^{\frac{1}{2}}$$

作为误差范数，这时的 f 是 2 维空间到 n 维 Hilbert 空间的映射，本文以下的讨论都设 F 空间是 n 维的.

当象集 D 为凸集时，函数 $J(y)$ 有唯一的局部极小值，这一点 y_0 的原象 x_0 即为所求的最

本文于1995年10月21日收到. 1996年12月23日收到修改稿.

佳参数估计，但是当 f 是非线性映射时，凸集 E 的象集 D 可能不是凸集，而只是一个连通闭集. 这时，D 上可能有不止一个点与 z 的距离 $J(y)$ 都是局部极小值，这些点称为局部极值点. 我们的问题在于：判断这时 D 上的某个局部极值点 y_0是否能保证距离 $\|y_0-z\|$ 为 D 到 z 的最小距离.

本文给出闭集 D 的一个邻界距离 Q，证明了 D 上的局部极值点 y_0是 D 到 z 的最小距离点的一个充分条件是

$$\|y_0-z\|<Q.$$

本文还给出用映射 f 的一、二阶导数的界来估计邻界距离 Q 的方法，从而给出在非线性估计问题中确定局部极值点为最优解的一个充分条件. 这个结果改进了 G. Chavent 在[1～2]中关于拟凸集适宜域的结果.

2 曲线的邻界距离

定理 1 设 n 维 Hilbert 空间中二次可微曲线 AB 的参数方程为 $g(t)$ $(0\leqslant t\leqslant S)$，长度为 l，曲线上各点的曲率值 $|K(t)|\leqslant\rho=\dfrac{1}{R}$ $(0\leqslant t\leqslant S)$，曲线的全曲率 $\varphi=\int_0^s|K(t)|\mathrm{d}t$ 小于 2π，则曲线两端点距离

$$d=\|g(0)-g(S)\|\geqslant d_0,$$

其中

$$d_0\triangle 2R\sin\frac{\varphi}{2}+(l-R\varphi)\cos\frac{\varphi}{2}.$$

注 这里明显有 $l\geqslant R\varphi$，而且当 $\pi<\varphi<2\pi$ 时，不妨设 $l\leqslant R\varphi-2R\mathrm{tg}\dfrac{\varphi}{2}$，否则 d_0 为负，定理明显成立.

证 分三种情况进行讨论：

1) 平面标准形：AB 为半径 R，弧度 φ 的圆弧 $\overset{\frown}{A_1B_1}$ 和两端引出的切线 A_1A,B_1B 所组成，切线长度 $\eta=A_1A+B_1B=l-R\varphi$，从圆弧 $\overset{\frown}{A_1B_1}$ 中点 C 作切线 MN，不难算出曲线 AB 在 MN 上投影为

$$2R\sin\frac{\varphi}{2}+(l-R\varphi)\cos\frac{\varphi}{2}=d_0,$$

于是有端点距离

$$d=\overline{AB}\geqslant d_0.$$

2) 平面曲线 $A'B'$ 的情形，在 $A'B'$ 上取一点 C'，使 $A'C'$ 与 $B'C'$ 的全曲率相等，都等于 $\varphi/2$，过 C' 作切线 $M'N'$，这时 $A'C'$ 与 $B'C'$ 的长度都不小于 $R\varphi/2$，适当选择平面标准形中 A_1A 与 B_1B 的长度，使 $A'C'$ 与 AC 等长，$B'C'$ 与 BC 等长，则 $A'B'$ 与 AB 上的点一一对应，由于 $A'B'$ 的曲率都不大于 ρ，所以对任意 $t(0\leq t\leq S)$ 都有

$$0\leqslant(V'(t),M'N')\leqslant(V(t),MN)\leqslant\pi.$$

这里 $(V(t),MN)$ 是指曲线 AB 在 $g(t)$ 处的切线 $V(t)$ 与 C 点切线 MN 间的夹角，$(V'(t),M'N')$ 是 $A'B'$ 在 $g'(t)$ 处切线 $V'(t)$ 与 $M'N'$ 的夹角.

我们有

$$\cos(V'(t),M'N')\geqslant\cos(V(t),MN),$$

所以 $A'B'$ 在 $M'N'$ 上的投影

$$d'_0=\int_0^s\cos(V'(t),M'N')\mathrm{d}t\geqslant\int_0^s\cos(V(t),MN)\mathrm{d}t=d_0.$$

故 $\overline{A'B'}\geqslant d'_0\geqslant d_0.$

3) n 维空间曲线情况，对于长度为 l 全曲率为 φ 的 n 维空间曲线 $A''B''$，可以作一条对应

的平面曲线 $A'B'$，使它与 $A''B''$ 的长度相等，对应点的曲率相等，那末由上面分析可知，$\overline{A'B'}$ 的长度 $d'\geqslant d_0$. 而曲线 $A''B''$ 两端点之间的距离 $d''=\overline{A''B''}$，可以由 n 维空间中的 A. Schur 定理得知，有 $d''\geqslant d'$. n 维空间的 A. Schur 定理的证明可参见参考文献[3].　　证毕.

由定理 1，不难得到

定理 1′　长度为 l，最大曲率小于 $\rho=\frac{1}{R}$ 的空间光滑曲线，如果全曲率 $\varphi\leqslant\pi$，则曲线端点距离

$$d\geqslant 2R\sin\frac{\varphi}{2}.$$

设空间曲线 AB 上的点 $g(t)$ 的切向量 $V(t)$ 的正交补空间为 $T(t)$，则 $T(t)$ 将空间分成两个部分，其中一部分中的点 $h(t)$ 与 $g(t)$ 组成的向量 $h(t)-g(t)$ 与 $V(t)$ 的内积大于零，我们称为点 $g(t)$ 的正向部分，而另一部分称为 $g(t)$ 的反向部分.

对曲线 AB 上的两点 $g(t_1)$ 与 $g(t_2)$，(不妨设 $0\leqslant t_1<t_2\leqslant S$)，取 $g(t_1)$ 的反向部分与 $g(t_2)$ 的正向部分这两部分的交的闭集，记作 $N(t_1,t_2)$. 对于空间中的点 q，设 q 到曲线上点 $g(t)$ 的距离为 $J(t)=\|q-g(t)\|$，如果函数 $J(t)$ 在 t_1,t_2 两点都达到局部极小，也就是 q 到 $g(t_1),g(t_2)$ 的距离为局部极小值，容易看到，若 $g(t_1),g(t_2)$ 不在曲线的端点即 $0<t_1<t_2<S$，则 q 必在它们正交补 $T(t_1)$ 与 $T(t_2)$ 的交集上；若 $t_1=0$，q 必在 $T(t_1)$ 反向部分的闭集上，若 $t_2=S$，q 必在 $T(t_2)$ 正向部分的闭集上，综合各种情况，我们都能得到 $q\in N(t_1,t_2)$.

定义 1　设光滑曲线 $g(t)$ 上的两点为 $g(t_1),g(t_2)$，定义两点 $g(t_1),g(t_2)$ 的邻界距离为

$$Q(t_1,t_2)=\inf_{q\in N(t_1,t_2)}\max[d(q,g(t_1)),d(q,g(t_2))]$$

这里 $d[q,g(t_i)]$ 是指 q 到 $g(t_i)$ 的距离.

定理 2　长度为 l，曲率 $|K(t)|\leqslant\frac{1}{R}$，全曲率 φ 的曲线 AB 两端点 A,B 的邻界距离的一个下界为

$$Q(A,B)\geqslant\begin{cases}R, & \text{当 } \varphi\leqslant\pi \text{ 时,}\\ R\sin\frac{\varphi}{2}+\frac{1}{2}(l-R\varphi)\cos\frac{\varphi}{2}, & \text{当 } \pi<\varphi\leqslant 2\pi \text{ 时.}\end{cases}$$

证　1) 当 $\varphi\leqslant\pi$ 时，由定理 1′ 可知 AB 的距离 $\overline{AB}\geqslant 2R\sin\frac{\varphi}{2}$，又由于曲线的全曲率为 φ，可知 A,B 两点的切线夹角不大于 φ，而对于 $N(A,B)$ 中的任意点 q，可以看到 $\angle AqB\leqslant\varphi$，在 $\triangle AqB$ 中，至少有一角 $\angle qAB$ 或 $\angle qBA$ 是不小于 $(\pi-\varphi)/2$ 的，不失一般性设 $\angle qAB\geqslant\frac{\pi-\varphi}{2}$，根据正弦定理，可以得到：

$$\begin{aligned}\overline{qB}&=\overline{AB}\cdot\sin\angle qAB/\sin\angle AqB\\&\geqslant 2R\sin\frac{\varphi}{2}\cdot\sin\frac{\pi-\varphi}{2}/\sin\varphi=R.\end{aligned}$$

由此可得 $Q(A,B)\geqslant R$.

2) 当 $\pi<\varphi\leqslant 2\pi$ 时，由定理 1 可知 $\overline{AB}\geqslant 2R\sin\frac{\varphi}{2}+(l-R\varphi)\cos\frac{\varphi}{2}$，故有

$$Q(A,B)\geqslant\max(\overline{qA},\overline{qB})\geqslant\frac{\overline{AB}}{2}\geqslant R\sin\frac{\varphi}{2}+\frac{1}{2}(l-R\varphi)\cos\frac{\varphi}{2}.$$　　证毕.

3　局部极值点为最优解的条件

定义 2　定义曲线 $g(t)$ 邻界距离为

$$Q(g(t))=\inf_{0\leqslant t_1,t_2\leqslant S}(Q(g(t_1),g(t_2))).$$

定理 3 对于 F 空间中的二次可微曲线 $g(t)$ 与 F 中的点 q，如果 q 到 $g(t)$ 的某个局部极小值点 $g(t_0)$ 的距离小于 $Q(g(t))$，那么 $g(t_0)$ 与 q 的距离必是 $g(t)$ 到 q 的最小距离.

证 如果 $g(t_0)$ 是 q 与 $g(t)$ 距离的唯一局部极小值点，那么距离 $d[q,g(t_0)]$ 自然就是 $g(t)$ 到 q 的最小距离.

如果曲线 $g(t)$ 上还有点 $g(t_p)$，它与 q 的距离也是局部极小，那末可知 $q\in N(t_0,t_p)$，由定义 1 可知 $d[q,g(t_0)]$ 与 $d[q,g(t_p)]$ 中至少有一个不小于 $Q(t_0.t_p)$，又由于定义 2 知

$$d[q,g(t_0)]\leqslant Q(g(t))\leqslant Q(t_0,t_p),$$

所以我们有

$$d[q,g(t_p)]\geqslant Q(t_0,t_p)\geqslant d[q,g(t_0)].$$

由此可知 $d[q.g(t_0)]$ 是 q 到 $g(t)$ 的最小距离. 证毕.

定义 3 设 D 是空间 F 的完备闭集，在 D 中作一族二次可微曲线 p，如果对 D 中的任意两点 a,b，都至少有一条曲线 p 经过，我们称这族曲线 p 的集合为 D 的一组线径，记作 P.

定义 4 如果 P 是 F 空间上完备闭集 D 的一组线径，定义 D 的邻界距离为

$$Q(D)=\inf_{p\in P}Q(p).$$

对 D 取不同的线径 P，$Q(D)$ 将会有所不同，但这不影响我们下面的讨论.

定理 4 若 Hilbert 空间 F 中连通闭集 D 上某点 y_0 到点 $z\in F$ 的距离为局部极小，且有 $d(y_0,z)<Q(D)$，则 $d(y_0,z)$ 一定是 D 到 z 的最小距离.

证 对于 D 中异于 y_0 的任意一点 y_p，设过 y_0 与 y_p 的线径为 p，则 $Q(D)\leqslant Q(P)$，y_0 是曲线 p 到 z 的距离极小值点且 $d(y_0,z)<Q(p)$，由定理 3 知 y_0 是到曲线 p 的最小距离点，故 $d(q,y_0)\leqslant d(q,y_p)$，由于 $y_p\in D$ 的任意性，可知 y_0 是 D 到 z 的最小距离点. 证毕.

4 在非线性参数估计中的应用

考虑引言中提出的非线性估计问题，如果 E 中的点 x_0 能使它的象与 F 空间中点 z 的距离 $\|f(x_0)-z\|$ 为函数 $J(x)=\|f(x)-z\|\ (x\in E)$ 的局部极小值，我们来判断 x_0 是否为 E 中的参数最优解.

由于 E 是凸集，E 中的任意两点 x_1,x_2 可用直线连接，这些直线的象 $f(x_1,x_2)$ 组成 E 的象集 D 中的一组线径. 那么，可以得到，这组线径的最小曲率半径 R，最大长度 L 和最大全曲率 φ 分别为[2]：

$$R=\inf_{x_1,x_2\in E}\ \inf_{t\in(0,1)}\frac{\|f'(x_t)(x_2-x_1)\|^2}{\|f''(x_t)(x_2-x_1,x_2-x_1)\|},$$

$$L=\sup_{x_1,x_2\in E}\int_0^1\|f'(x_t)(x_2-x_1)\|\,\mathrm{d}t,$$

$$\varphi=\sup_{x_1,x_2\in E}\int_0^1\frac{\|f''(x_t)(x_2-x_1,x_2-x_1)\|}{\|f'(x_t)(x_2-x_1)\|}\mathrm{d}t.$$

其中 $x_t=(1-t)x_1+tx_2$.

如果映射 f 的一二阶微分是有界的，即对任意 $x_1,x_2\in E$，及 E 的任意内点 u 有

$$\|f''(u)(x_2-x_1,x_2-x_1)\|\leqslant K\|x_2-x_1\|^2,$$

$$m\|x_2-x_1\|\leqslant\|f'(u)(x_2-x_1)\|\leqslant M\|x_2-x_1\|.$$

设 E 中最长直线段 $\max\|x_2-x_1\|=C$，那么我们有

$$R\geqslant m^2/K,\quad L\leqslant MC,\quad \varphi\leqslant KC/m.$$

于是，我们可以确定 E 上的这个非线性映射 f 的邻界距离为：

当 $KC/m \leqslant \pi$ 时，取 $Q(f) = m^2/K$；

当 $\pi < KC/m \leqslant 2\pi$ 时，$Q(f) = R\sin\frac{\varphi}{2} + \frac{1}{2}(l - \varphi R)\cos\frac{\varphi}{2}$，注意到这里 $Q(f)$ 是 R 的增函数，L,φ 的减函数，所以可分别用它们的下界与上界来估计，取

$$Q(f) = \frac{m^2}{K}\sin\frac{KC}{2m} + \frac{1}{2}(M - m)\cos\frac{KC}{2m}.$$

这样，当求出的局部最佳参数 x_0 满足条件 $\| f(x_0) - z \| < Q(f)$ 时，可以断定 x_0 就是所求的最小距离意义下的最优参数.

在进行参数估计时，如果某个容许的参数向量 x 符合条件 $\| f(x) - z \| < Q(f)$，那么，以这点 x 为初始点开始搜索寻优，当达到其局部极值点 x_0 时，就能够断定 x_0 就是容许参数集上的最优解.

参 考 文 献

1　Chavent, G.. On the uniqueness of local minima for general abstract nonlinear least squares problems. Inverse Problems, (4):1988, 417—433

2　Chavent, G.. A new sufficient condition for the well-posedness of non-linear least square problem arising in identification and control. Proc. Problems Inverse, INRIA France, 1990, 1—12

3　陈省身，陈维恒. 微分几何讲义. 北京：北京大学出版社，1983，273—313

A Sufficient Condition of Optimum Solution in Non-Linear Parameter Estimation

ZHENG Yingwen

(Institute of Automation, Fuzhou University · Fuzhou, 350002, PRC)

Abstract: In a kind of non-linear parameter estimation, a sufficient condition for an admissible parameters being optimum is provided. When the error norm between the given data and the output of a local-minima parameter is smaller than a given adjacent distance, the parameter is optimun solution.

Key words: parameter estimation; non-linear mapping; norm distance; curvature of curves, extreme value

本文作者简介

郑应文　1947 年生. 1969 年北京大学数学力学系毕业，1981 年厦门大学控制理论研究生毕业，获硕士学位，1989 年至 1990 年在法国 INRIA 进修. 现为福州大学自动化研究所副研究员. 目前承担的课题有关于 CIMS 设计自动化，离散事件对策，无穷维系统理论等方面内容.