关于隐函数定理与莫尔斯引理结合的评论

Emőke Imre

欧布达大学 *iversity, Hydro*-Bio-Mechanical Systems Research Center，匈牙利布达佩斯 *y*

摘要

本文推广了解析隐函数定理。将定理中的函数 f 关于隐函数的因变量进行积分。利用莫尔斯引理，对积分函数 F 的子几何给出了几何解释。该结果用于分析与非线性参数识别最小化相关的层次化技术。

关键词：隐式、莫尔斯、非线性最小化、层次参数识别

引言

非线性逆问题的层次解

非线性逆问题的解可以通过非线性最小化来确定，但由于噪声的影响，最小二乘优值函数在 M 维地形中具有难以想象的复杂性，这使得求解受到阻碍，其中 M 为参数个数。

极小值可以是全局的，也可以是局部的。这里采用两种启发式方法：（i）从自变量的不同初始值出发寻找局部极值，然后选取最极端的 [1]；或（ii）通过从局部极值处向外迈出一步来扰动该极值，然后观察迭代是否会返回到一个更优的点，或回到同一点 [1]。另一种方法是在局部拟合出一个更平滑的超曲面，并沿其下降 [2]。当全局最小值近似退化时，可能会出现一些额外的几何困难；理论上可通过正则化（[3]）处理，但这需要修改后的算法。

尽管非线性算法中的参数数量可以减少，从而减少由噪声引起的临界点数量并降低数值计算工作量，此外还具有其他优点（由此确定的优值函数的一维截面可用于参数误差估计和隐式正则化），但目前尚无针对非线性情形的层次求解方法报道。

本文的目的

层次方法的概念[4] 在线性方程组中广为人知，例如在线性逆问题的背景下；本文开始将其引入非线性逆问题的求解中。层次解的应用示意如下（图1）。

我们可以假设，$F : R^{M} = R^{n+m} \to \{0U\ R^{+}\}$ 是一个解析最小二乘优值函数，在参数域内有唯一极小值 p_{min}。假设参数空间被分解为直和 p=(x,y)，逆问题的解也相应地被分成两个维数较小的部分（图1中：x^{1}, x^{2}）。

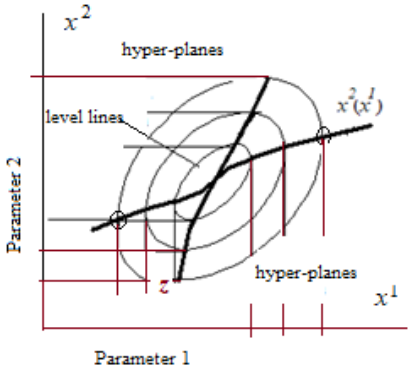

图1. 隐函数 $x^2(x^1)$ 由 $f(x, y)= f(x^1, x^2) =F'_y (x^1, x^2)= 0$ 定义。隐函数 $x^2(x^1)$ 是固定 $x=x^1$ 值时条件极小点集合的原像。如果每个垂直（$y=x^2$ 方向）超平面上的子极小值唯一，那么关于两个参数的 $F (x^1, x^2)$ 极小值与“极小截面” $F [x^1, x^2(x^1)]$ 的极小值相同。

一种最小化是在参数空间中平行于 y方向的平面截面上进行的：通过子最小化（使用最小化方法1），在每个固定的 x 值处搜索 y^2 解的部分。这样确定的子极小值集合位于与条件 $F'_y = f = \mathbf{0}$ 相关的隐函数 y(x) 的图像上。所谓的极小截面 F[x,y(x)] 则使用方法2进行最小化。

本文通过分析优值函数沿 y 平面的几何性质，探讨原逆问题与层次逆问题的解等价的条件。研究发现，要实现等价，F 必须在 y方向上为凸。当优值函数为严格凸目标函数，或优值函数部分严格凸（即在 y 方向上严格凸）且模型对 y 呈线性依赖时，该条件成立。

此外，还提及了一些推论和推断。在严格凸情形下，与各个子空间相关的隐函数图像

隐函数定理的扩展

假设、概念与定理

所用的基本概念如下 [5 至 8]。对于光滑函数 $F : R^M \to R$，一阶导数为零的点称为临界点，其像称为临界值。如果某个临界点处的二阶导数矩阵（海森矩阵）非奇异，则称其为非退化临界点；如果海森矩阵奇异，则称其为退化临界点。文中给出了一些基本定理，这些定理与扩展的隐函数定理相结合。

莫尔斯引理

设 p 是 F 的一个非退化临界点。则在 p 的邻域 U 中存在一个坐标图，使得对于所有 $x_i(p) = 0$，其中 x_i 构成 p 附近的一个局部映射，并且

$$F(x) = F(p) - x_1^2 - \dots - x_\gamma^2 + x_{\gamma+1}^2 \dots - x_n^2 \tag{1}$$

在整个 U 上。此处 γ 值等于 F 在 p 处的指标（即海森矩阵的负特征值个数）。由莫尔斯引理可知，非退化临界点是孤立的。

S^n 上最后一个莫尔斯不等式将指标为 k 的临界点个数与 S^n 的欧拉示性数联系起来；该不等式取等号时如下：

$$\sum_{k=0}^{n}(-1)^k\,(\text{k-index critical points}) = 1 + (-1)^n \tag{2}$$

由此可以推导出 D^n（n 维球）的方程。如果省略一个极小值（指标为 0 的临界点），两边都要减去 1；如果省略一个极大值（指标为 n 的临界点），则要减去 $(-1)^n$。如果函数在球的边缘“向上卷曲”（边缘处的梯度指向外侧）：

$$\sum_{k=0}^{n}(-1)^k\,(\text{k-index critical points}) = 1 \tag{3}$$

在严格凸情形下，边界 D^n（n 维球，即一个半 S^n）上有一个指标为 0 的临界点。

隐函数定理

设 $f : R^{n+m} \to R^m$ 是一个解析函数。R^{n+m} 是 R^n 与 R^m 的直和，其中的一个点记作 $(x, y) = (x_1, \dots, x_n, y_1, \dots, y_m)$。从给定的函数 f 出发，目标是构造一个函数 $g: R^n \to R^m$，使其图像 $(x, g(x))$ 恰好是所有满足 $f(x, y) = \mathbf{0}$ 的 (x, y) 构成的集合。

固定一点 $(a,b) = (a_1, \dots, a_n, b_1, \dots, b_m)$，使得 $f(a, b) = \mathbf{0}$，其中 0 是 R^m 的元素。若矩阵 $[(\partial f_i/\partial y_j)(a, b)]$ 可逆，则存在包含 a 的开集 U、包含 b 的开集 V，以及唯一函数 $g: U \to V$，使得

$$\{\mathbf{x}, \mathbf{g}(\mathbf{x}) \mid \mathbf{x} \in U\} = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in U, \mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}\} \tag{4}$$

若假设 f 在 $U \times V$ 内解析，则隐函数 g 在 U 内也具有相同性质。

扩展

命题 1（隐函数定理的扩展及其几何解释）

假设参数空间分解为直和 p=(x,y)，且解析隐函数定理的条件在 $(a,b) = (a_1, \dots, a_n, b_1, \dots, b_m)$

。

(i) 假设解析隐函数定理中的函数 $f : R^{n+m} \to R^m$ 具有关于 y 的解析偏积分函数，记为 F : R

退化的、沿 y 方向的 F 的“部分”临界点，其局部子几何由 F''_{yy} 的指标决定。（这些点可以是 F 的退化

说明：

考虑解析隐函数定理中的点 (a,b) 及其邻域 U、V。这里引入偏函数 $F^x(y) : R^m \to R^1$, ，用于

Fa(y) : Rm → R 是一个莫尔斯函数，因为其一阶导数为零，且其海森矩阵在 b 处非退化，因此

陈述2（分层最小化的应用）

如果 F 在点 **(*a*,*b*) 处有一个非退化极小值，**那么对于任意直和分解 p=(x,y)，每个可能的隐函数都在局部存在，且关于 y 有一个局部子极小值。然而，对于特殊的直和分解 p=(x,y)，如果 F''_{yy} 在整个参数域上处处正定，则该隐函数全局存在。

解释：

由于 F''_{yy} 处处正定，隐函数可以在局部反复延拓（对任意 x）。Fx(y) 的指标为 0，在 F’

y方向上处处正定，因此 Fx(y) 关于 y方向严格凸。故 Fx(y) 只有一个临界点，即极小值（这一结论也可由最后一个莫尔斯不等式推出）。

陈述3（R^M 上凸函数分层最小化的情形）

对于任意直和分解 p=(x,y)，F 的非退化极小值满足隐函数定理的条件，因此相关隐函数在局部存在。

(i) 对于严格凸函数 F，任意直和分解 p=(x,y) 下，其偏导数 F''_{yy} 全局正定。因此，每个可

(ii) 考虑由 F 的图像生成的欧几里得空间。 对于每条 z 等高线，都有两个坐标超平面，它们垂直于参数 x^i 的坐标方向，并与给定的 z 等高线相切（图 1）。这对超平面构成一系列夹住递减 z 等高线的超平面。切点对是 x^i的一维最小截面上的点；由点对 $[x^i{}_{z-}、x^i{}_{z+}]$ 所定义的区间，是 F 的 z 次水平集在参数 x^i轴上的正交投影。随着 z 增大，这些区间不断缩小；在极小值处，两点重合（图 1）。

推论（参数的一维最小截面的性质）

一维极小截面 $F[[x^i,y(x^i)]$ 是关于参数 x^i 的最小下降线，沿正/负 x^i 方向前进，朝向全局极小点 p_{mino}。一维极小截面 $F[[x^i,y(x^i)]$ 是一个“最深”的敏感性截面（即沿一条经过全局极小值、以 x^i为参数的简单曲线取得的截面）。

关于分层最小化的进一步评论

关于等价条件的一些评论

我们发现，如果解析函数 F 在 y方向上部分地、全局地严格凸，那么关于 x 的隐函数全局且唯一地存在，并且在 y方向上使 F 取得子最小值。这两种极小化方式（即 M 维无条件极小化，以及包含 J 维 y 方向子极小化和 M-J 维关于 x 的极小截面极小化的分层极小化）具有相同的解。

从实际意义上说，只有那些会非线性地影响模型的参数，才是非线性最小化所关注的重要参数。模型的线性部分可以被“消去”。如果只有一个参数会非线性地影响模型，那么由于线性子最小化可以自动消去所有其他参数，甚至可以通过一维括定来求解非线性最小化。

值得注意的是，对于具有多个极小值的部分凸函数 F，全局隐函数可能存在。如果两个不同的极小值具有不同的 x 坐标值，那么关于 x 的全局隐函数及其对应的极小截面将穿过这两个极小值。

凸函数坐标网格括定的算法步骤

隐函数点是 F 的一个条件极小点，可按如下方式确定。考虑将 F(x, y) 的图像正交投影到由该图像生成的整个欧几里得空间中的 F - x 坐标平面上。该映射的临界值——即投影发生折叠之处——与条件 $F_y'(x, y) = f(x, y) = \mathbf{0}$ 相关，该条件定义了隐函数。

据此，可以制定一种算法来确定隐函数并进行分层最小化。对于固定的 x，切片函数 $Fx(y) : Rm \to R$ 的唯一子最小值可以通过投影和关于 y 的极小值搜索来确定；由该极小值，则可以通过逆像运算确定隐函数的图像点。

只有当存在三点组 a < b < c（或 c < b < a），且 f(b) 小于 f(a) 和 f(c) 时，才能对一维极小值进行括定。在这种情况下，我们知道函数（若其非奇异）在区间 (a; c) 内有一个极小值。可以对一维极小值进行括定，但在一般的多维情形下，没有类似的过程[1]，除非函数是严格凸的，因为由坐标网格点生成的超平面可以包围递减的凸水平线。

严格凸优值函数的全局极小值可以利用坐标网格，沿每个坐标方向通过三点组进行括定。使用同一坐标网格，在全局极小值搜索之外，将上述算法步骤应用于每个网格点，也可以确定一维极小截面。

正则化

除了噪声之外，当优值函数发生畸变、极小值呈准退化时，还可能出现一些额外的几何困难（例如，在拟合一个偏微分方程的单射解模型时，如果模型取决于时间变量，测量中的测试时间可能会“过短”）。可通过正则化（[1, 3]）处理这一问题。显式正则化是指明确向优化问题中添加某个项的正则化方法。这些项可以是先验、惩罚项或约束。显式正则化通常用于不适

定优化问题。

正则化项，也称惩罚项，会对优化函数施加代价，从而使最优解唯一。隐式正则化指所有其他形式的正则化，例如提前停止、使用鲁棒损失函数以及剔除离群值。

与给定参数相关的一维极小截面具有一些特殊的几何特征。它是一条最短的下降线，可用于确定参数敏感性／误差。由于它是解析的，因此可按如下方式利用它，通过隐式正则化求解准退化极小值问题。

本文提出了一种新的隐式方法，该方法基于一维隐函数的解析特性；这类函数可能具有单射坐标函数（例如，可利用参数向量 p_{min} 中该参数的已知坐标，通过与给定参数相关的一维隐函数的某个单射坐标函数确定 p_{min} ）。即使模型扰动轻微破坏了凸性，并导致出现两个退化极小值，该方法也不会失效。

致谢

谨此致谢 András Stipsicz、András Szűcs、Endre Szabó 和 József Bodnár 的讨论。

参考文献

[1] Press, W.H.; Flannery, B.P.; Teukolsky, S.A.; Wetterling, W.T. (1986): Numerical Recipes。剑桥大学出版社。[2] Meier J, Rudolph, S,. Schanz T：参数反算的有效算法——发表于《Bautechnik》第86卷S1期，第86–97页，岩土工程应用[3] Hegedűs, Cs (2008)：数值分析，ELTE，http://www.inf.elte.hu/Budapest，匈牙利 [4] Zhang, Fuzhen (2005). Zhang, Fuzhen (编)《舒尔补及其应用》。数值方法与算法。第4卷。施普林格。doi:10.1007/b105056。ISBN 0-387-24271-6。[5] Császár, Á. (1983)《实分析》。Tankönyvkiadó，匈牙利语。[6] Milnor, J.W. (1963)《莫尔斯理论》。普林斯顿大学出版社。153页。[7] Milnor, J.W. (1963)《从可微观点看拓扑学》普林斯顿大学出版社，1997年12月14日。157页。[8] Hirsch, M. W.《微分拓扑学》。数学研究生教材，1973年，GTM，第33卷，精装本 ISBN: 978-0-387-90148-0 版权信息：施普林格纽约出版社。