第七卷

21世纪
数量经济学

21SHIJI SHULIANG JINGJIXUE

○ 主　编　汪同三　胡祖光

○ 副主编　李富强　李金华　许　冰

方志出版社

21世纪数量经济学

主　编　汪同三　胡祖光

副主编　李富强　李金华　许　冰

第七卷

方志出版社

图书在版编目（CIP）数据

21 世纪数量经济学．第七卷／汪同三，胡祖光主编．
北京：方志出版社，2007.4
ISBN 978-7-80238-040-0

Ⅰ.2... Ⅱ.①汪…②胡… Ⅲ.数量经济学—文集
Ⅳ.F224.0-53

中国版本图书馆 CIP 数据核字（2007）第 057148 号

21 世纪数量经济学（第七卷）

主　　编：汪同三　胡祖光
副 主 编：李富强　李金华　许　冰
责任编辑：颜伦琴

出 版 者：方 志 出 版 社
（北京市建国门内大街 5 号中国社会科学院科研大楼 12 层）
邮编　100732
网址　http://www.fzph.org
发　　行：方志出版社出版发行部
（010）85195814　85196281
经　　销：新华书店总店北京发行所
法律顾问：北京市京诚律师事务所
排　　版：北京开和文化传播中心
印　　刷：北京通州丽源印刷厂

开　　本：787×1092　　1/16
印　　张：30.75
字　　数：558 千
版　　次：2007 年 4 月第 1 版　　2007 年 4 月第 1 次印刷

ISBN 978-7-80238-040-0/F·138　　定价：60.00 元

编审组名单

编审组组长：汪同三

成　　　　员：赵京兴　李金华　李富强　曹曼株

目　录

一

理论与方法研究

多元条件密度函数的投影追踪估计方法*

叶阿忠　　　　　　　　Rob J Hyndman
（福州大学管理学院）（澳大利亚莫纳什大学计量经济和商务统计系）

一　引　言

自从 Huber（1985）与 Friedman，Stuetzle 和 Schroeder（1984）建立多元密度函数的投影追踪估计以来，还没有人应用该方法估计多元条件密度函数。我们知道，多元条件密度函数$f_{Y|X}(y|x)$（$x \in \Re^p$）可以告诉我们许多被解释变量 Y 对解释变量 X 的回归关系的信息，其条件期望就是回归函数，条件方差就是回归误差项的条件方差。因而，对多元条件密度函数进行估计是非常重要的。但令人惊讶的是仅有少量的条件密度函数估计的论文发表。而且，这方面发表的研究几乎都是有关一元条件密度函数的（Bashtannyk 和 Hyndman，2001；Fan，Yao 和 Tong，1996；Hyndman 和 Bashtannyk，1996；Hyndman 和 Yao，2002）。因为高维空间数据的稀松性，对多元条件密度函数进行非参数估计是困难的，除非数据的观察个数很大，这也就是著名的维数诅咒问题（Wand 和 Jones，1995）。为了克服高维空间数据稀松性带来的估计上的困难，投影追踪的降维估计方法已经成功地应用于多元密度函数和多元非参数回归的估计（Friedman，Stuetzle 和 Schroeder，1984；Hardle，1990；Xia，Tong，Li 和 Zhu，2002）。本文提出多元条件密度函数的投影追踪估计方法，通过最小化相对熵，得到了最优初始条件密度函数和每一步的增量函数和方向向量，还给出了估计步骤及其终止法则。

* 国家自然科学基金项目（70371025），教育部人文社会科学研究项目（02JA790014）。

二　投影追踪估计

多元条件密度函数$f_{Y|X}$（$y|x$）（$x\in\Re^p$）的投影追踪估计具有如下形式：

$$g_M(y|x)=g_0(y|x)\prod_{m=1}^{M}h_m(y|\theta_m x) \tag{1}$$

其中g_0是初始条件密度函数，h_m是第m步的待定的函数，θ_m是第m步的待定的方向向量，M是待定的正整数，$\|\theta_m\|=1$，$\int_{-\infty}^{+\infty}g_M(y|x)dy=1$。

由式（1）得到如下递推关系

$$g_m(y|x)=g_{m-1}(y|x)h_m(y|\theta_m x)\qquad \int_{-\infty}^{+\infty}g_m(y|x)dy=1,m=1,\cdots,M \tag{2}$$

我们称h_m为增量函数。假定g_{m-1}（$y|x$）给定，我们的目标寻求$f_{Y|X}$（$y|x$）的最佳估计g_m（$y|x$）。我们用g_m（$y|x$）与$f_{Y|X}$（$y|x$）的相对熵来度量g_m（$y|x$）的拟合优度。通过最大化g_m（$y|x$）的拟合优度来选择方向向量θ_m和它对应的增量函数h_m（$y|\theta_m x$）。

三　相对熵

定义1　条件密度函数g（$y|x$）与$f_{Y|X}$（$y|x$）的相对熵为：

$$d_{KL}(f_{Y|X},g)=\int f_{Y|X}(y|x)f_X(x)\log\frac{f_{Y|X}(y|x)}{g(y|x)}dydx \tag{3}$$

引理1　相对熵满足

$$d_{KL}(f_{Y|X},g)\geqslant 0.5\int f_{Y|X}(y|x)f_X(x)p\left(\frac{f_{Y|X}(y|x)}{g(y|x)}-1\right)dydx\geqslant 0 \tag{4}$$

其中当$|z|\leqslant 1$时，p（z）$=0.5z^2$；否则，p（z）$=|z|-0.5$。

引理2

$$\int\left|f_{Y|X}(y|x)-g(y|x)\right|f_X(x)dydx\leqslant[2d_{KL}(f_{Y|X},g)]^{0.5} \tag{5}$$

可见，g_m（$y|x$）与$f_{Y|X}$（$y|x$）的相对熵d_{KL}（$f_{Y|X}$，g_m）越小，g_m（$y|x$）

拟合 $f_{Y|X}(y|x)$ 的程度越好。

四 定理及其证明

定理 1 假定条件密度函数 $f_{Y|X}(y|x)$ 的条件期望为 $\mu_f(x)$，条件方差为 $\sigma_f^2(x)$，则最小化 $d_{KL}(f_{Y|X}, g_0)$ 的最优的高斯条件密度函数 $g_0^{opt}(y|x)$ 的条件期望为 $\mu_f(x)$，条件方差为 $\sigma_f^2(x)$。

证明： 不失一般性，不妨设 $\mu_f(x)=0$，$\sigma_f^2(x)=1$，让 $g_0^{opt}(y|x)$ 是条件期望为零，条件方差为 1 的正态条件密度函数。设 $g(y|x)$ 是条件期望为 $\mu(x)$，条件方差为 $\sigma^2(x)$ 的条件正态密度函数，则

$$d_{KL}(f_{Y|X},g)-d_{KL}(f_{Y|X},g_0^{opt})=\int f_{Y|X}(y|x)f_X(x)\log\frac{g_0^{opt}(y|x)}{g(y|x)}dydx$$

$$=0.5\int\left[\log\sigma^2(x)+\frac{(y-\mu(x))^2}{\sigma^2(x)}-y^2\right]f_{Y|X}(y|x)f_X(x)dydx$$

$$=0.5\int\left[\log\sigma^2(x)+\frac{1+\mu^2(x)}{\sigma^2(x)}-1\right]f_X(x)dx\geqslant 0.5\int\frac{\mu^2(x)}{\sigma^2(x)}f_X(x)dx\geqslant 0$$

可见，仅当 $g=g_0^{opt}$ 时，$d_{KL}(f_{Y|X}, g)$ 达最小。

定理 2 假定方向向量 θ_m 给定，则最小化 $d_{KL}(f_{Y|X}, g_m)$ 的最优增量函数为：

$$h_m^{opt}(y|\theta_m x)=\frac{f_{Y|\theta_m X}(y|\theta_m x)}{g_{m-1}^{(\theta_m)}(y|\theta_m x)} \tag{6}$$

其中 $g_{m-1}^{(\theta_m)}(y|\theta_m x)$ 是 $g_{m-1}(y|x)$ 在 $\theta_m x$ 上的一维条件密度函数。此时，相对熵减少

$$d_{KL}(f_{Y|X},g_{m-1})-d_{KL}(f_{Y|X},g_m)=d_{KL}(f_{Y|\theta_m X},g_{m-1}^{(\theta_m)}) \tag{7}$$

证明： 不失一般性，不妨设 $\theta_m=(1, 0, \cdots, 0)$，则 $\theta_m x=x_1$。最小化

$$d_{KL}(f_{Y|X},g_m)=\int f_{Y|X}(y|x)f_X(x)\log\frac{f_{Y|X}(y|x)}{g_{m-1}(y|x)h_m(y|x_1)}dydx$$

$$=\int f_{Y|X}(y|x)f_X(x)\left[\log\frac{f_{YX|YX_1}(y,x|y,x_1)}{f_{X|X_1}(x|x_1)}-\log\frac{g_{m-1}(y,x|y,x_1)}{g_{m-1}(x|x_1)}\right]dydx+$$

$$\int f_{Y|X}(y|x)f_X(x)\left[\log f_{Y|X_1}(y|x_1)-\log g_{m-1}^{(\theta_m)}(y|x_1)h_m(y|x_1)\right]dydx$$

等价于最小化

$$\int f_{Y|X}(y\mid x)f_X(x)[\log f_{Y|X_1}(y\mid x_1)-\log g_{m-1}^{(\theta_m)}(y\mid x_1)h_m(y\mid x_1)]dydx$$

$$=\int f_{Y|X_1}(y\mid x_1)f_{X_1}(x_1)\log\frac{f_{Y|X_1}(y\mid x_1)}{g_{m-1}^{(\theta_m)}(y\mid x_1)h_m(y\mid x_1)}dydx_1=d_{KL}(f_{Y|X_1},g_{m-1}^{(\theta_m)}h_m)$$

也就是最优的增量函数由式（6）给出。此时，相对熵减少

$$d_{KL}(f_{Y|X},g_{m-1})-d_{KL}(f_{Y|X},g_m)=\int f_{Y|X}(y\mid x)f_X(x)\log h_m^{opt}(y\mid x_1)dydx$$

$$=\int f_{Y|X_1}(y\mid x_1)f_{X_1}(x_1)\log h_m^{opt}(y\mid x_1)dydx_1$$

$$=\int f_{Y|X_1}(y\mid x_1)f_{X_1}(x_1)\log\frac{f_{Y|X_1}(y\mid x_1)}{g_{m-1}^{(\theta_m)}(y\mid x_1)}dydx_1$$

$$=d_{KL}(f_{Y|\theta_mX},g_{m-1}^{(\theta_m)})$$

定理3 最小化 d_{KL}（$f_{Y|X}$，$g_{m-1}h_m^{opt}$）的最优方向向量 θ_m^{opt} 最大化

$$W(\theta_m)=\int f_{Y|X}(y\mid x)f_X(x)\log h_m^{opt}(y\mid\theta_mx)dydx \tag{8}$$

证明： 由定理2容易推得。

五 终止法则

从定理2知，m 越大，g_m 就越趋近于 $f_{Y|X}$。显然，M 过大，将增加计算量。所以，在实践中，我们必须确定式（1）中的 M。易见，如果 g_{m-1} 接近于 $f_{Y|X}$，则 h_m^{opt} 接近于1。因而，当 h_m^{opt} 接近于1时，我们可以确定 $M=m-1$，否则，$M\geqslant m$。

六 估计步骤

给定样本数据 $\{X_i, Y_i\}_{i=1}^n$，估计条件密度函数 $f_{Y|X}$（$y\mid x$）的过程如下：

第一步，由定理1，设

$$g_0(y\mid x)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(y-a-\theta_0x)^2}{2\sigma^2}\right) \tag{9}$$

通过最大化对数似然函数

$$ll(a,\theta_0,\sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - a - \theta_0 X_i)^2 \tag{10}$$

获得未知参数 a，θ_0，σ^2 的估计。

第二步，假定 $g_{m-1}(y \mid x)$ 和方向向量 θ_m 已知，采用核估计方法估计一元条件密度函数 $f_{Y\mid\theta_m X}(y \mid \theta_m x)$（Hyndman，Bashtannyk，and Grunwald 1996，Bashtannyk and Hyndman 2001）。采用 Monte Carlo 抽样方法估计 $g_{m-1}^{(\theta_m)}(y \mid \theta_m x)$（Friedman，Stuetzle and Schroeder，1984，Zhu，2004），即对于每个 X_i（$1 \leqslant i \leqslant n$），由密度函数 $g_{m-1}(y \mid X_i)$产生 Monte Carlo 随机数据 Y_{i1}，…，Y_{is}（s 是一固定的正整数，为重抽样的数据个数，可取 $s=30$），然后，利用数据 $\{(Y_{ij}, X_i)_{j=1}^{s}\}_{i=1}^{n}$得到 $g_{m-1}^{(\theta_m)}(y \mid \theta_m x)$ 的 Hyndman 核估计。进而得到增量函数 $h_m^{opt}(y \mid \theta_m x) = \dfrac{f_{Y\mid\theta_m X}(y \mid \theta_m x)}{g_{m-1}^{(\theta_m)}(y \mid \theta_m x)}$。

第三步，最大化

$$w(\theta_m) = \frac{1}{n}\sum_{i=1}^{n}\log h_m^{opt}(Y_i \mid \theta_m X_i) \tag{11}$$

得到方向向量 θ_m 的最优估计 θ_m^{opt}。如果 $h_m^{opt}(y \mid \theta_m x)$ 接近于 1，则取 $M = m-1$，否则，让 $g_m(y \mid x) = g_{m-1}(y \mid x)\, h_m^{opt}(y \mid \theta_m^{opt} x)$，重复第二步和第三步，如此反复直到确定 M。

七 主要结论

本文采用投影追踪的降维技术，提出多元条件密度函数的投影追踪估计方法，通过最小化相对熵，得到了最优初始条件密度函数和每一步的增量函数和方向向量，还给出了估计步骤及其终止法则。从而有效地解决了因高维空间数据的稀松性对多元条件密度函数进行非参数估计所带来的“维数诅咒”问题。

参考文献

[1] Bashtannyk, D. M., and Hyndman, R. J., 2001. Bandwidth selection for kernel conditional

density estimation [J] . Computational Statistics and Data Analysis. 36 (3), 279 ~ 298.

[2] Fan, J., Yao, Q., Tong, H., 1996. Estimation of conditional densities and sensitivity measures in nonlinear dynamical system [J] . Biometrika, 83 (1), 189 ~ 206.

[3] Friedman, J. H., Stuetzle, W., Schroeder, A., 1984. Projection pursuit density estimation [J] . J. Amer. Statist. Assoc. 79, 599 ~ 608.

[4] Hardle, W., 1990. Applied Nonparametric Regression [M] . Cambridge University Press.

[5] Huber, P. J., 1985. Projection pursuit [J] . The Annals of Statistics, 13 (2), 435 ~ 475.

[6] Hyndman, R. J., Bashtannyk, D. M., Grunwald, G. K., 1996. Estimating and visualizing conditional densities [J] . J. Comput. Graph. Stat., 5, 315 ~ 336.

[7] Hyndman, R. J., Yao, Q., 2002. Nonparametric estimation and symmetry tests for conditional density functions [J] . Journal of Nonparametric Statistics, 14 (3), 259 ~ 278.

[8] Scott, D. W., 1992. Multivariate density estimation theory practice and visualization [M]. New York, John Wiley.

[9] Wand, M. P., Jone, M. C., 1995. Kernel Smoothing [M] . London, Chapman and Hall.

[10] Xia, Y., Tong, H., Li, W. K., Zhu, L., 2002. An adaptive estimation of dimension reduction space [J] . J. R. Statist. Soc. B, 64, 363 ~ 410.

[11] Zhu, M., 2004. On the forward and backward algorithms of projection pursuit [J] . The Annals of Statistics, 32 (1), 233 ~ 244.

实验经济学视角下的数据品质分析

王国成

（中国社会科学院数量经济与技术经济研究所）

现代经济学的显著特征之一是基于计量经济学①的实证分析方法的广泛而有效的应用，特别是近年来微观经济计量方法所取得的成功又使当代经济理论的发展进入了一个崭新的阶段（Heckman，2001；Wooldridge，2002），与此同时也使经济理论与现实的联系和定量分析对数据的依赖性等方面的问题更加突出。由于微观主体行为的交互性、经济运行的多变性，导致经济数据的来源与构成愈加复杂，为经济计量学的发展设置了一些难以逾越的障碍，本文的目的在于从实验经济学这一新的角度，试图更深入地认识和探讨数据的品性、质量等特征及具如何提高经济定量分析的有效性。

一　经济计量方法的基本问题

经济计量学，或广义的经济定量分析学科群，是经济学、数学和统计学的交叉学科，是建立在现代经济学理论框架内和（随机）数学理论基础之上，而且要借助计算机等现代科学工具得以实现。一般认为，经济计量模型结构识别、参数估计与假设检验等方法是其核心内容，而将所依赖的经济理论和数据作为前提条件。于是，在进行经济定量分析研究时，始终存在着两方面的基本问题：一是经济理能否论与实际研究对象相吻合；另一是数据品质能否满足定量分析处理基本要求。若模型结构能合理地表现经济运行内在规律，统计学意义上的对数据品质的要求就成为

① 本文不对经济计量学和计量经济学作专门的区别，两者都是 econometrics。

影响经济定量分析有效程度的决定性因素。数据是否真实地记录了现实经济状况，是否满足模型设定时对数据的基本要求，利用这些数据估算参数能否使模型很好地反映经济变量关系和运行变化规律？这就要遵从基本理论中一些隐含的假定，如主体行为特征同质性（理性人及相关假设）、现实经济的结构稳定性和时间一致性（用历史数据和信息建立相应模型能解释现实和预测未来）等都被默认为是事先给定或得到许可的。

（一）基本模型设定

无论是线性或非线性的，还是静态或动态的，可将经济计量模型概括为如下的一般形式：

$$Y = F(X, \beta) + \varepsilon \tag{1}$$

其中 Y 是表示产出结果或经济现象的被解释变量，$X =$（x_1，x_2，…，x_n）是表示投入要素或影响因素的解释变量；$\beta =$（β_0，β_1，…，β_n）是模型参数；ε 是外生干扰项，通常情况下是服从特定分布函数的随机变量，如 $\varepsilon \sim N$（μ，σ^2）；函数关系（模型类别）F（·，·）反映变量关系、经济系统的结构和变化规律，如 C－D 生产函数类 $Y = AK^{\alpha}L^{\beta}e^{u}$。公式（1）最简单的表现形式是常用的线性回归 $Y = \beta\begin{pmatrix} 1 \\ X^T \end{pmatrix} + \varepsilon$，即 $Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n + \varepsilon$（还可扩展考虑多个被解释变量的联立方程组）。若考虑系统结构的动态性或比较动态结构，则有：

$$Y_t = F(X_t, \beta(t)) + \varepsilon \tag{1'}$$

或

$$Y_t = F(X_t, \beta(t), Y_{t-\tau}) + \varepsilon \tag{1''}$$

公式（1″）中的 τ 为一个单位时期。

公式（1）中第一项是依据经济理论构建的数学模型，它能否很好地反映经济实际运行状况取决于经济理论对研究对象的适应性和数学工具所提供的可能性，一旦设定在定量分析过程中就是相对稳定的，即默认了在一定时期内的结构一致性假设；第二项概括了由随机干扰等因素的综合作用与影响，主要运用数学和统计学的分析技术与方法，依据误差最小原理和极大似然原理进行参数估计和显著性检验等。由于是在既定的理论范式、框架和体系内展开定量分析，误差分析是相对于依