

## 经济学与数学模型 (连载—4)

高鸿桢

**编者按：**现在经济学时尚数学模型，不论懂与不懂、必要与不必要的，都要添加些数学模型与运算，有为数学而数学的倾向，于是就难免闹些东施效颦般的笑话，漏洞、错误习以为常。为此，本刊请高鸿桢教授有针对性地写点这方面的文章予以连载，希望有助于初学者能正确运用数学模型。

### 数据分析模型与经济理论

**内容提要：**本文论述建立数据分析模型的一般过程，讨论了关于在建立计量经济模型中经济理论的指导作用，并论及经济数学模型、数据分析与经济理论间关系的不同观点。最后提出在建模时应避免三类建模陷阱。

**关键词：**数据分析模型 经济计量模型 建模陷阱

**作者：**厦门大学经济学院计划统计系教授，博士生导师。

数据分析模型是最常见的数学模型。在经济学中引入数学模型的目的是想将经济理论精确化，分析现实世界中发生的经济现象。经济理论的发展要用实证分析来检验；政策的实施必须有定量信息。经济系统的变化是通过其特征指标来显现的，一个国家国内生产总值的增减、股市指数的变化、法定货币的汇率升降、外汇贮备的多少等无一不是通过数量指标来表示的。因此研究各种经济量之间的关系就成为经济学家的首要任务。统计表明20世纪70年代之后，在核心经济刊物上发表的论文之中有70%以上是关于数据分析模型的作品。诺贝尔经济学奖得主克莱因认为“计量经济学已在经济学科中居于最重要的地位。”

#### 一、建立数据分析模型的一般过程

一般说，建立数据分析模型可能有两种目的：其一是对已有的理论进行检验；其二是想利用已建立的模型对某些变量的将来值进行预测。在经济学中有许多原理、命题并未经过实践的检验，也有一些命题只是在特殊的时间、特殊的地点、在某种特殊的情况下得到的，是否具有一般性，必须用实际数据来检验。至于预测更是大多数经济模型的建立目的。建立适当模型进行准确预测是许多经济学家所追求的目标，也是许多机构不惜重金支持复杂的计量

模型研究的原因。一般说, 计量经济模型的建立要有如下几个步骤: 1. 观察事实; 2. 收集数据; 3. 确定理论模型; 4. 模型具体化; 5. 检验; 6. 模型的改进; 7. 预测。

有的人认为利用已有的理论框架建立一个经济模型很简单, 只需要把实际数据代入模型, 依样画葫芦就可以了, 特别是现代计算技术发达, 有许多可用的现成软件, 只要将数据输入电脑, 结果就出来了! 这种看法是不正确的, 建立一个实际可用的模型, 即使应用现成的理论也要经过艰苦细致的过程。例如, 利用柯布——道格拉斯生产函数理论估计我国的生产函数, 初看似乎是一个很容易的问题, 但真正做起来却并不简单。

柯布——道格拉斯生产函数很简单。

$$Y_t = AK_t^\alpha L_t^\beta \quad (1)$$

其中,  $Y_t$  是第  $t$  期的产出,  $K_t$  是第  $t$  期投入的资本量,  $L_t$  是第  $t$  期投入的劳动力量,  $\alpha$ 、 $\beta$  分别是资本和劳动力的弹性。如果要据此建立我国从1952年到2002年这50年间的生产函数模型, 则可令

$$y_t = \ln Y_t, \quad C = \ln A, \quad k_t = \ln K_t, \quad l_t = \ln L_t$$

利用

$$y_t = C + \alpha \ln K_t + \beta \ln L_t + \epsilon_t \quad (2)$$

估计参数  $\alpha$ 、 $\beta$  和  $C$ 。 $\epsilon_t$  是随机扰动项。

当我们收集数据将模型具体化时, 困难就出现了, 首先要确定这些变量是什么? 数据如何取得? 例如, 作为产出的  $Y$  是用国内生产总值, 还是用国民收入表示? 如果用国民收入表示, 那么是否能取得连续有效的数据? 中国(大陆)官方统计口径变动很大, 原先是MPS体系后来改为SNA体系, 许多数据无法连续。比如原来统计的社会生产总值是工业、农业、建筑、运输、商业等五个生产部门的纯国民生产总值之和, 其中并不包括诸如银行、保险、教育、不动产等服务业产生的国民生产总值, 因此要进行必要的换算, 才能与以后的SNA体系下的统计数据相连接。例如, 胡祖六<sup>①</sup>等人为探讨中国高速成长的原因, 在建立生产函数时, 利用社会生产总值和国内生产总值的数据, 另行计算了“调整的国民收入”序列作为  $Y_t$ 。再看资本序列  $K_t$ , 这里资本是以存量计, 统计年鉴上的投资额无法直接利用, 数据需要重新计算。例如, 农业部门的固定资产和流动资产的初始值就必须计算, 胡祖六采用了如下方法计算的结果: 1952年农业生产总值为340亿元, 假定资本收益率为18.9%, 在总值中资本金贡献率为25%, 土地贡献率为40%, 则农业资金的初始资本应为:  $(340 \times 25\%) / 18.9\% = 450$  (亿元); 土地的初始资金计为  $(340 \times 40\%) / 18.9\% = 720$  (亿元)。这样的计算实际上是对原问题再添加一些假定, 只要这些假定是合理的, 类似的估算是允许的。劳动力数据可以从年鉴上取得。

在数据如何取得的问题解决后, 我们还必须对数据的质量也就是数据的真实性进行检查。因为经济模型的数据大多数并不是专为建立模型而采集的, 而是政府或其他机构为种种不同的目的而采集的, 因此难免有不真实之处, 有些个别数据虽然是真实的, 但因某种特殊原因与总趋势相差很大, 如何处理这些数据也是建模者应考虑的问题。例如, 我国1958年曾有“大跃进”运动、大刮“浮夸风”粮食亩产万斤, 甚至是数万斤的消息都经常见报, 在这样的

<sup>①</sup> 见参考文献 [1]。

情况下,官方的统计数据自然是不可信的,此后数年政治对统计的干扰严重(还有许多统计资料被列为“机密”,无法利用),统计数据多不真实。建模时可以考虑将这一类不可靠的数据直接删除或者利用虚拟变量处理,如果这类数据较多则可同时采用以上两种方法,最后比较所得结果。文献[1]就是这样处理的。

在用计量经济学方法对参数进行估计之后还必须进行有关的检验,其中包括统计检验和经济检验,统计检验主要是检验数据是否满足计量经济学的基本假定,例如关于方程系数的T检验、关于多重共线性的检验、关于异方差的检验等。当模型在一定显著性水平下能通过有关的统计检验之后,我们还要检查模型所反映的经济关系是否正确,如果出现与经济现象不相符合的情况则必须找出原因,对模型进行修正,直到符合要求为止。得到符合要求的模型后,即可将变量变化的范围向外进行合理的扩大,以便于进行预测。

当我们利用现有理论框架建立计量经济模型时,可以在原有模型的基础上加以推广,使得能适应各种不同的情况。例如在柯布—道格拉斯生产函数(1)中,A设为常数,这就是假定在计算的时间内,该经济系统的技术水平是不变的,这在很多情况下是不符合实际情况的。为反映经济系统的技术改变的情况,我们将(1)中的A改写为时变的,设

$$A_t = e^{\lambda_0 + \lambda_1 t}$$

则通过类似变换,(2)式可写为

$$y_t = \lambda_0 + \lambda_1 + \alpha K_t + \beta l_t + \epsilon_t \quad (3)$$

利用它可以估计参数 $\lambda_0$ 、 $\lambda_1$ 、 $\alpha$ 、 $\beta$ 。这里反映了生产函数中的技术进步因素。

如果要利用它反映多个投入物与产出的关系,则可将(1)式改写为

$$Y_t = A X_1^{\alpha_1} X_2^{\alpha_2} \cdots X_r^{\alpha_r} \quad (4)$$

其中, $X_{it}$ 是第*i*种投入物在第*t*期的投入量, $\alpha_i$ 是相应的弹性。进行类似的取对数变换后,可以估计出参数 $\alpha_1, \alpha_2, \dots, \alpha_r$ 的值。

## 二、经济理论的作用

在建立数据分析模型时,经济理论起什么作用?大多数经济学家认为经济理论是经济数据分析模型的灵魂,经济数据分析模型必须在一定经济理论指导下建立,离开了经济理论数据分析模型就像失去灵魂的躯壳,只是数字游戏,根本无法揭示客观的经济规律。格林指出“理论在经济计量学中的作用是第一位的,有人认为只要有足够的时间分析数据,就一定可以揭示出某些复杂的真理,这种信念对非实验数据来说是毫无道理的乐观。”<sup>①</sup>计量经济学从可能观测的经济数据中推测这些数据发生的机制或理论,计量经济学家面对的是以各种不同目的采集来的数据(暂不考虑实验经济学的数据)。这种数据与实验科学(如物理学、生物学)的数据不同,是事件发生过的记录,而且是无法重复的。物理学家为检验牛顿第二定律 $f=ma$ ,可以通过预先设置改变力*f*的大小来实现,而经济学家要检验柯布—道格拉斯生产函数时却无法自己设置产出的数值,他只能被动地接受数据,在已有的数据的基础上开展自己的研究工作,也就是说他只能“观测”而不能“创造”数据。自然科学中也有只能观测数据的学科如天文学、气象学等,但经济学也不能与之相比。天文学家虽然无法进行实验,却能

<sup>①</sup> 见参考文献[2]第3页。该书曾被誉为“计量经济学的圣经。”

将相关的物理定律应用于遥远的星体，在“其他事物不变”的前提下研究星体的性状。例如天文学家虽然无法取到太阳上的物质的标本，却能依靠光谱分析得知太阳上存在哪些物质，甚至能发现地球上尚未发现的物质。然而经济学家就没有这么幸运，虽然发生的经济事件就在自己的身边，他却从来没有能观察到“其他条件保持不变”下的现象——经济系统错综复杂，牵一发而动全身，很难观测到只有一、两个变量变化而其他变量全都不变的情况。特别是经济现象与人的活动相关，因此理论的抽象更是重要，在这种意义上说，经济理论是经济计量模型的灵魂并不为过。

重视经济理论对计量经济学模型的指导作用是计量经济学的传统。弗里希在给计量经济学所下的定义中强调计量经济学是“统计学、经济理论和数学三者的结合”，在其中经济理论的地位如何，他没明说，但大多数人认为这三者不是平等的，经济理论应处于指导地位，一个计量模型的建立首先应符合相关的经济理论，不可进行“无理论的计量”。20世纪70年代前这种建模思想很流行，但是对于先验经济理论过分依赖的建模思想导致模型本身的缺点。首先是理论模型的选择是先验的，一般与数据集合本身的性质无关，这就具有一定的主观性，而且一经选定，此后的工作就只能在此理论框架下进行。其次，在固定的理论框架下，难以发现新问题发展新理论。事实上，在1974年格兰杰提出“伪回归”（Spurious Regression）问题之前人们对时间序列的数据特征并未特别加以注意，一般都假定经济序列是宽平稳的，直接进行回归分析，而实际上许多经济序列都是非平稳。此后的序列平稳性检验、向量自回归模型（VAR）、误差校正模型（ECM）和协整分析可以说都是以此为契机发展起来的。这些新概念、新方法打破原来的理论框架，使计量经济学的发展进入了一个新阶段，构成了目前所谓“高级计量经济学”的主要内容。这种发展并不是受到某个经济理论指导下发生的，而是计量经济学自身发展的结果。

在这种情况下有人对经济理论在计量经济模型中的作用提出不同的看法。首先，一些经济理论所提出的命题比较含糊，很难依它来指导建模；还有一些经济理论所论问题非常浅显，只不过是简单的心理现象或经济现象的表述。例如，建立消费模型时常见到诸如“作为一条规律，平均来说在收入增加时，人们倾向于增加他们的消费，但消费增加要小于收入的增加”一类的理论依据，实际上这只是简单心理现象的复述，说用它“指导”建模似有言过其实之嫌。

其次，从历史上看，许多目前已被广泛承认的“经济理论”，在其发现之时，并没有在什么理论的指导之下，参数也没有明确的经济意义。例如柯布——道格拉斯生产函数，是经济学家道格拉斯在1924年和数学家柯布合作发现的。当时道格拉斯正在研究1899年——1922年美国制造业的生产状况，他发现产出曲线的位置总是在处于两条投入曲线之间，于是他邀请数学家柯布共同研究这个问题结果，发现了如本文（1）式所示的生产函数（当时他们假定了规模报酬不变即 $\alpha + \beta = 1$ ），用最小二乘法估计了 $\alpha$ 、 $\beta$ 的值（ $\alpha = 0.25$ ， $\beta = 0.75$ ）。他们发现这个规律主要是依靠数据分析，而不是理论推导。现在我们可以许多经济书籍中看到：满足一定条件的生产函数必为常数替代弹性的（CES）生产函数：

$$Y = (AC^{-\rho} + BL^{-\rho})^{-\frac{1}{\rho}}$$

其中， $A > 0$ ， $B > 0$ ， $\rho > -1$ ， $C$ 、 $L$ 意义同前。可以证明 $\rho \rightarrow 0$ 时，CES生产函数就成为柯布

——道格拉斯生产函数<sup>①</sup>。这就意味着,我们可以规定一组条件,演绎推导出:满足这一组条件的生产函数就是柯布——道格拉斯生产函数。但是这是事后的推演而不是发现的过程。总之,柯布——道格拉斯生产函数的发现靠的不是理论的演绎推理而是对现实数据特征的归纳总结。

再举一个例子。被作为现代宏观经济学基石之一的菲利普斯曲线,是新西兰经济学家菲利普斯提出的。他分析了英国近百年关于失业和货币工资率变化之间的数据,于1958年发表论文《1861—1957 联合王国的失业与货币工资率的变化之间的关系》。在该文中菲利普斯指出,1861—1957年间,英国名义工资率的变化与失业率呈负相关,若以失业率为横轴货币工资率的变化为纵轴画一直角坐标系,则菲利普斯是一条具有负斜率的从左上方方向右下方倾斜的曲线。菲利普斯论文发表两年后,萨缪尔森和索罗用通货膨胀率代替货币工资

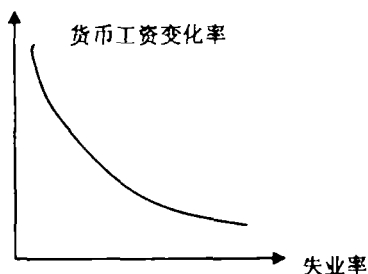


图1 菲利普斯曲线

的变化率,提出了“失业——物价”菲利普斯曲线,表明失业率与通货膨胀率呈反向运动。在经济上升期失业率下降物价上升,在经济回落时失业率上升,物价下降。它说明通货膨胀是由工资成本推动的。菲利普斯曲线完全是依靠统计数据归纳出的原理。在原理形成过程中并没有刻意找哪一个理论作为指导。相反,它的一些结论却与已有的理论相矛盾,例如,它指出了失业与通货膨胀之间的交替关系,这明显与凯恩斯关于失业与通胀不会同时出现的理论相左。

第三,经济理论本身是在不断发展的,理论是否正确,必须在实践中加以检验,用未加检验的理论来指导建模,逻辑上就讲不通。1995年卢卡斯(R. Lucas)荣获诺贝尔经济学奖,瑞典皇家科学院的颁奖词说“卢卡斯提出的理性预期假说,改造了宏观经济分析并深化了我们对经济政策的理解。卢卡斯已经使得直到1970年代为止所发表的大多数经济理论站不住脚。”请注意,这里说的是“大多数”。如果这话是对的,那么以这些“站不住脚”的理论为指导建立起来的模型又有什么意义呢?因此认为计量模型必须有某种理论指导的理由并不充分。

### 三、结束语——避免建模陷阱

“经济学与数学模型”已经连载了四篇文章,编者的意图是“希望有助于初学者”,这个目标不知能否达到。经济数学模型发展迅速,新思想、新概念、新方法不断出现,因此笔者也是个初学者。在结束本文之际笔者介绍三个寓言故事作为初学者和初学者间谈谈避免建模陷阱的体会。

#### 1. 不必邯郸学步。

《庄子·秋水篇》说燕国有一个少年听说赵国首都邯郸人走路的姿态很美,便到邯郸去学习。为了更专心学习赵国人的步法,他决心忘掉以前自己走路方法,但他最终没有学会邯郸人的步法,只好放弃,由于此时他把以前自己的走路方法也忘了,只好爬着回去了。在经济中应用数学方法是主流经济学的趋势,为顺利进入世界主流经济学论坛,学习经济数学方法是必须的。然而我们应该知道,用数学公式表述经济问题只是一种可能的方式,不是唯一的方式。许多诺贝尔经济学奖得主的名成之作就未用数学公式化表述。例如诺斯、科斯等人。

<sup>①</sup> 见参考文献 [3] 第340页。

著名的关于产权的“科斯定理”在其发表之时并未用“定理”的形式叙述。目前我们常见的表述：“只要交易成本为零，财产的法定所有权的分配不影响经济运行的。”是其他人的表述。科斯本人也无法说清楚“科斯定理”的准确含义，只是用简单的数字例子加以说明。其实，在这里重要的是经济学思想而不是数学公式。有人认为马克思主义经济学可以用文字表述，西方经济学不易用文字表述，这也是一种误解。例如张五常的《卖桔者言》就曾倾倒大量青年经济学者<sup>①</sup>。如果你文字表达能力很强，在学习数学模型方法时就大可不必像燕国少年那样忘掉原来的步法。

## 2. 不可南辕北辙。

这是《战国策》上的故事。有人要到楚国去，楚国在南方他却驾车向北走，人家告诉他到楚国应向南走，劝他回头，他回答说：“我的马好！”再劝他，他说：“我的旅费多。”；再劝他，他说：“我的车夫驾车技术高。”其实，这些条件越好，他离楚国就越远了。建立经济计量模型的目的是为验证某一想法或凭借成功的模型进行预测，有点像理科的实验，希望实验所得的结果能支持我们的设想。因此在建模时的指向很重要，如果建立的模型与你要证实或否定的原理关系不大，那么即使你建模所根据的理论正确、设置精巧，你的计算机计算能力超强、计算机软件功能强大，数据收集完整、准确，你的处境是不是和那个要去楚国的人有点相像呢？

## 3. 切勿祠少求多。

《史记》“滑稽列传”中讲到，齐王要淳于髡借兵救齐，所给礼物很少，淳于髡讲了一个故事：有一个人向上天祈祷，希望上天保佑他家的高地打下的粮食装满笼，低田打下的粮食装满大车，五谷丰登粮食满仓。而他的祭品（祠）却只有一只猪蹄和一杯酒。齐王听了以后便多给礼物，借兵成功。计量经济模型的基础是数据，收集数据是建模的基本功，必须有充分多的数据，才有可能建立起有意义的模型。如果实在收集不到足以建模的数据，则宁可放弃。有时我们会看到这样的“模型”：一共只有十几个原始数据却要确定四、五个参数；只有十几年的历史数据却要预测此后二、三十年的情况。为模型做的工作太少，要求模型给出的结论太多。这样建立起来的“模型”当然无法达到目标，就像那位祈祷者一样。

## 参考文献：

1. Hu, Z., Khan, M. S., 1997, "Why China growing so fast," IMF Staff Papers, Vol. 44, pp. 103-131.
2. H. 格林著, 王明舰、王永宏译:《计量经济分析》, 中国社会科学出版社 1998 年版。
3. 冈本哲治等编, 袁镇岳、高鸿桢译:《经济数学》, 辽宁人民出版社 1985 年版。
4. 王少平:《宏观计量的若干前沿理论与应用》, 南开大学出版社 2003 年版。
5. 李子奈:《关于现代计量经济学的研究方法》,《清华大学学报》(哲学社会科学版)1995 年第 3 期。
6. 白雪梅、赵松山:《对动态经济计量建模过程的模型研究》,《中国软科学》2000 年第 5 期。

作者单位：厦门大学计划统计系

邮 编：361005

---

<sup>①</sup> 张五常的作品在大陆风行一时，虽然有其历史原因，但文章行文流畅、说理透彻是受到读者欢迎的主要原因。当然，我们不能像他那样随意批评自己不熟悉的对象，例如他不熟悉马克思主义经济学和博弈论却针对这些领域说一些很偏激的话。