

姜 微,刘俊昌. 基于半参数混合模型的林业经济增长与政府投资的关系[J]. 江苏农业科学,2017,45(12):270-274.  
doi:10.15889/j.issn.1002-1302.2017.12.067

# 基于半参数混合模型的林业经济增长与政府投资的关系

姜 微<sup>1,2</sup>, 刘俊昌<sup>2</sup>

(1. 中南林业科技大学商学院,湖南长沙 410004; 2. 北京林业大学经济管理学院,北京 100083)

**摘要:**运用中国 1953—2013 年时段数据分阶段分别采用线性模型、广义线性模型、半参数混合模型,以 Ram 提出的两部门理论为框架,研究政府林业投资对区域林业经济增长的作用。结果表明:1953—2013 年,我国政府投资对林业经济增长具有正向的外部作用,而由于两部门之间的要素生产率差异看不出来,经过判断得出政府支出对林业经济增长产生正向总作用,但这个正向的总作用在统计上不显著。在 1953—1980 年,我国政府投资对林业经济增长具有正向的外部作用,而由于两部门之间的要素生产率差异看不出来,使得政府投资对林业经济增长具有一定的正向总作用,而且结合当时的国情看,政府投资对于林业经济增长的正向作用很显著。1981—2013 年,由于非线性的参数有 3 个,政府投资对于林业经济增长的作用就很难判断出来,只能说政府投资对于林业经济增长有一定的正向外部作用。

**关键词:**半参数混合模型;政府投资;林业经济增长;外部作用

**中图分类号:** F326.2    **文献标志码:** A    **文章编号:** 1002-1302(2017)12-0270-04

林业经济发展的基础是投资,而林业由于其特殊性,林业经济投资更多是来自政府的资金投入。林业投资对于当前的林业经济增长产生了直接的推动作用。2000—2013 年《中国林业统计年鉴》统计数据表明,我国林业投资额也从 2000 年的 1 677 712 万元增长至 2013 年的 37 822 690 万元,使得我国林业经济总产值由 2000 年的 35 554 725 万元增长到 2013 年的 473 154 396 万元,增长了 13 倍多。林业不仅提供了公共产品,还提供了包含其他正外部效应的公共服务,如森林游憩、生态旅游、农家乐等,这些都大力鼓励了非国家投资的发展,形成了新的投资环境促进林业经济发展。政府林业投资对于区域林业经济增长的效果如何?究竟有多大影响?是正影响还是负影响?在地区经济发展不平衡且处于供给侧改革的过程中,均衡地区经济发展具有重大意义。但由于理论框架、计量方法、数据来源以及处理方式的不同带来的结论也不尽相同。我国林业经济发展研究是从 1978 年以后通过对各国林业经济发展研究开始进行系统的分析和比较。由于我国国民经济快速发展促进了我国林业经济的快速增长,同时研究方法和理论也有不断的更新。随着十八大会议的召开,生态文明建设被摆放到前所未有的高度,林业对于生态文明的发展起着至关重要的作用。因此,研究林业经济发展对于促进我国生态文明发展有着重要作用。孔凡斌简要阐述了我国林业投资性质、生态性投资机制形成过程以及林业发展面临的机遇与挑战;以 1950—2005 年相关统计数据为主要依据,

对中国林业投资总量、结构及森林资源发展绩效进行评价,分析了经济体制转型期林业投资机制和投资方向的变化路径,并就林业投融资机制问题给出了政策建议<sup>[1]</sup>。刘珉运用经济学投资理论,通过历史比较的方法,对林业投资规模、林业投资结构、林业投资资金来源、林业投资经济效益、林业投资阶段几方面展开分析<sup>[2]</sup>。于江龙等根据我国 1981—2009 年林业投资与林业产值数据,运用协整检验、误差修正模型、Granger 因果检验等分析林业投资与产业经济增长之间的关系,发现我国林业投资与产业经济增长之间存在长期稳定的关系,产业经济增长对林业投资存在很强的依赖性<sup>[3]</sup>。田淑英运用 DEA 评价模型,对 1993—2010 年我国林业投入产出效率进行测算,并结合我国林业发展的现实背景分析,发现林业投资资金的利用效率不高、人力资源投入不足和林业经济发展水平偏低导致近几年林业投入产出效率降低<sup>[4]</sup>。周家春等利用 2000—2010 年的相关数据进行林业固定资产投资投入的双对数模型回归分析发现,林业固定资产投入的总量增加对农业经济增长具有明显促进作用;从结构上看,营林固定资产投资、森工固定资产投资对林业经济增长的影响程度不同<sup>[5]</sup>。孔凡斌等从影响林业经济增长的因素、林业经济增长方式的转变以及林业产业结构调整 3 个方面入手,对当前林业经济增幅方面的研究加以梳理综述,并在此基础上总结当前林业经济增长研究存在的问题并提出参考性建议<sup>[6]</sup>。廖文梅以南方 13 个省(区)的林业数据利用动态偏离-份额分析法进行分析,发现南方集体林区林业第一产业不具有竞争力优势,而林业第二、第三产业高于全国平均水平,具有竞争力优势<sup>[7]</sup>。才琪等采用 Stata 方法利用科布-道格拉斯生产函数分析,利用 VAR 模型测算中央林业投资的滞后期,使用协整关系检验得出林业经济增长和中央林业投资存在长期协整关系;结合格兰杰因果检验,判断中央林业投资和林业经济增长之间短期的因果关系;利用脉冲响应函数,更直观地判断

收稿日期:2016-11-16

基金项目:国家社会科学基金(编号:16CJY051);湖南省教育厅项目(编号:14C1171)。

作者简介:姜 微(1982—),女,湖南岳阳人,博士研究生,讲师,主要从事林业经济理论与政策、生态经济研究。E-mail:zhongyang1q2w@163.com。

二者变动在长期内对自身和相互之间的影响<sup>[8]</sup>。纵观已有研究,针对我国林业经济增长与林业投资的研究方法,此前的相关研究多采用偏离-份额法(static shift-share method)、DEA、VAR 模型以及 VAR 修正模型,对我国林业经济与投资之间的关系比较多,如中央投资对于林业经济的增长作用、固定资产投资分析等。但是对于政府投资与林业经济增长的关系实证分析很少,因此,本研究尝试利用半参数混合模型,分析 1953—2013 年我国林业产业经济增长的变化趋势,探寻影响我国政府投资对于林业经济增长的影响,以期为我国林业产业经济发展理论研究提供补充。

## 1 理论分析及模型介绍

### 1.1 生产函数理论

1986 年, Ram 提出改进生产函数模型,即政府部门和非政府部门两部门在其部门间资源分配对于经济增长的影响<sup>[9-20]</sup>。假设经济总产出( $Y$ )包括政府部门产出( $G$ )和非政府部门产出( $N$ ),各个部门产出由资金投入( $K$ )和劳动投入( $L$ ),且政府部门产出( $G$ )对非政府部门产出( $N$ )会产生外部的作用。公式如下:

$$Y = G + N; \quad (1)$$

$$N = N(L_N, K_N, G); \quad (2)$$

$$G = G(L_G, K_G)。 \quad (3)$$

满足约束条件——各个投入总要素等于两部门的投入要素之和,即

$$K = K_G + K_N; \quad (4)$$

$$L = L_G + L_N。 \quad (5)$$

式中: $L_G$  代表政府的资金投入; $L_N$  代表非政府的资金投入; $K_G$  代表政府的劳动投入; $K_N$  代表非政府的资金投入。令两部门要素的边际产出相等,即

$$\frac{G_L}{N_L} = \frac{G_K}{N_K} = 1 + \theta。 \quad (6)$$

式中: $G_L$  表示政府部门对劳动投入的边际产出; $G_K$  表示政府部门对资金投入的边际产出; $N_L$  表示非政府部门对劳动投入的边际产出; $N_K$  表示非政府部门对资金投入的边际产出; $\theta$  表示两部门的边际产出大小差异。当  $\theta > 0$ , 代表政府部门投入要素的边际产出更高;反之,非政府部门投入要素的边际产出更高。把公式(2)至公式(6)代入公式(1),并对其求微分,整理可得:

$$\dot{Y} = \alpha_1 (I/Y) + \alpha_2 \dot{L} + (\theta' - \beta) \dot{G} (G/Y) + \beta \dot{G}。 \quad (7)$$

式中: $\alpha_1 = C_K$ , 代表非政府部门投入资金要素的边际产出; $I/Y$  代表投资要素占总产出比重; $\alpha_2$  代表非政府部门劳动投入要素的产出弹性; $\beta$  代表政府部门的产出对非政府部门的产出的弹性。 $\theta' = \frac{\theta}{1+\theta}$ ,  $\theta$  是政府部门与非政府部门投入的边际产出差异。公式(7)对于经济增长有影响的变量是投资要素占总产出比重  $I/Y$ 、劳动投入的增长率  $\dot{L}$ 、政府投入资金增长率  $\dot{G}$ 、政府财政投入占总经济产出的比重 ( $G/Y$ )。通过估算  $\theta'$  和  $\beta$  的数值大小可以推出政府公共投资对经济增长的作用。

当  $\beta = \theta'$  时,公式变形为:

$$\dot{Y} = \alpha_1 (I/Y) + \alpha_2 \dot{L} + \beta \dot{G}。 \quad (8)$$

式中: $\beta$  代表着政府部门对非政府部门的外部作用,由于假设  $\beta = \theta'$ , 所以通过估算  $\beta$  可以算出  $\theta$  的估计值,由此满足约束条件的情况可以得出两部门各要素的边际产出的估计值,并估算出政府投资对经济增长的影响。

假设政府部门对非政府部门的边际产出影响为一个固定值,但政府部门的产出弹性  $\beta$  是可变的。则公式变形为:

$$\dot{Y} = \alpha_1 (I/Y) + \alpha_2 \dot{L} + (\theta' + C_G) \dot{G} (G/Y)。 \quad (9)$$

当然,公式(9)相比公式(7)中的  $\dot{G} (G/Y)$  系数要大一些,通过公式(9)中  $\dot{G} (G/Y)$  系数可以了解政府部门对经济增长的影响,但很难估计出政府部门的外部性以及部门间生产要素的差异性。

结合理论框架,本研究想解决的问题是政府部门投资对于林业经济增长产生的正负影响,参数  $\beta$  的正负,以及部门之间要素  $\theta$  的边际产出差异。通过公式(9)可以计算得出政府对经济增长的作用大小,且根据公式(7)和公式(8)可以计算得出  $\beta$ 、 $\theta$ 。在实证过程中,对于模型的选取特意进行比对选择中间最优模型,并根据 1952—2013 年的数据选取不同阶段进行对比分析。

### 1.2 半参数混合模型及其估计

广义线性模型 (GLM) 分为随机、系统和连接函数三部分,而广义可加性模型 (GAM) 是 GLM 的扩展。由于 GLM 中的系统部分属于线性模式,而连接函数部分是非线性的,因此如果单独采用线性函数模型,对于模型本身是非线性的函数就不能很好的拟合其非线性部分。但是使用非参数的方法可以解决这个问题,由此得到 GAM 模型,其公式为:

$$\eta = g\{\mu(X)\} = s_0 + s_1(x_1) + \dots + s_p(x_p)。 \quad (10)$$

$s_j(j=1, \dots, p)$  是  $X_j$  的光滑函数。如果有自变量不需要光滑函数或者为二变量的分量,那么公式可以改为:

$$\eta = g\{\mu(X)\} = \beta_0 + \beta_1(x_1) + \dots + \beta_q(x_q) + s_{q+1}(x_{q+1}) + \dots + s_p(x_p)。 \quad (11)$$

该等式中前面  $\beta_0 + \beta_1(x_1) + \dots + \beta_q(x_q)$  为参数部分,后面属于非参数,此模型称为广义可加性混合模型 (GAMM)。该模型的参数估算方法与一般线性相同,都是采用最小二乘法计算。但是对于非参数部分可以选择光滑样条、贝叶斯、罚光滑样条方法拟合来实现。本研究的半参数混合模型中非参数部分采用光滑样条法进行估算。

### 1.3 数据说明

本研究以林业总产值  $\dot{Y}$  为被解释变量,劳动年增长率  $\dot{L}$ 、政府投资的增长率  $\dot{G}$ 、非政府部门固定投资占林业总产值比重  $I/Y$ 、政府部门投资占林业总产值比重  $G/Y$  为解释变量。林业总产值数据是通过林业总产值实际增长速度对名义林业总产值数据进行调整,得到以 1952 年不变价的林业总产值数据,计算林业总产值的实际增长率  $\dot{Y}$ ,劳动投入从各年的统计年鉴中的就业人数查出,由此计算出劳动年增长率  $\dot{L}$ ,资本投入由各年总投资林业的数据得出。同样也是以 1952 年为不变价格计算得出  $I/Y$  的数值。政府投资是利用 1952 年不变价格为基础,计算政府投资的实际增长率  $\dot{G}$  和政府部门投资占林业总产值比重  $G/Y$ 。所有变量数据都来源 1952—2014

年《中国林业年鉴 1949—1986》、以及各年的《中国林业统计年鉴》,都以 1952 年为基期进行折算。

2 政府公共投资与区域林业经济增长的实证分析

选取线性回归模型、广义线性模型以及半参数混合模型利用指标数据分别对公式(7)、公式(8)、公式(9)进行估算。首先假设随机误差项是满足经典条件,加上公式以及截距和随机误差就可以得到设定模型。先对 1953—2013 年整个时

间段建模,为了对比,在 1953—1980、1981—2013 年时间段分别进行建模。对于半参数模型,非参数项的选择一般取决于变量之间散点图以及实际拟合度来判断。本试验运用 R 软件构建 3 类 9 个模,最终结果显示:1953—2013 年,光滑样条为自变量  $I/Y$ 、 $\dot{G}(G/Y)$ ;而 1953—1980 年,光滑样条为自变量  $\dot{G}(G/Y)$ ;1980—2013 年,光滑样条为  $I/Y$ 、 $\dot{G}$ 、 $\dot{L}$ ,其他自变量都为线性形式。

表 1 1953—2013 年政府投资与林业经济增长关系的模型估计结果

参数类型	参数	截距	$I/Y$	$\dot{L}$	$\dot{G}(G/Y)$	$\dot{G}$	调整后 $R$
普通线性	7	0.073 350 4±0.041 977 7	0.000 100 1±0.000 121 1	0.068 754 3±0.099 805 5	0.000 193 1±0.000 163 8	0.312 390 2±0.102 510 8 **	0.178
	8	0.078 450 0±0.041 900 0	0.000 087 4±0.000 121 0	0.054 670 0±0.099 430 0		0.215 300 0±0.061 250 0 ***	0.173
	9	0.116 800 0±0.042 550 0 **	0.000 018 9±0.000 126 4	0.094 380 0±0.106 400 0	0.000 207 9±0.000 104 4		0.059
广义线性	7	0.073 350 4±0.041 977 7	0.000 100 1±0.000 121 1	0.068 754 3±0.099 805 5	0.000 193 1±0.000 163 0	0.312 390 2±0.102 510 8 **	0.215
	8	0.078 450 0±0.041 900 0	0.000 087 4±0.000 121 0	0.054 670 0±0.099 430 0		0.215 300 0±0.061 250 0 ***	0.253
	9	0.116 800 0±0.042 250 0 **	0.000 018 9±0.000 126 4	0.094 380 0±0.106 400 0	0.000 207 9±0.000 104 4		0.192
半参数	7	0.094 520 0±0.023 440 0 ***		0.098 780 0±0.087 860 0	***	0.250 440 0±0.097 740 0 *	0.433
	8	0.105 720 0±0.020 220 0 ***		0.118 220 0±0.106 610 0		0.183 940 0±0.060 710 0 ***	0.251
	9	0.137 870 0±0.015 660 0 ***		0.140 270 0±0.097 520 0	***		0.427

注:数值后面的\*、\*\*、\*\*\* 分别表示  $P<0.05$ 、 $P<0.01$ 、 $P<0.001$ 。下同。

表 2 1953—1980 年政府投资与林业经济增长关系的模型估计结果

参数类型	参数	截距	$I/Y$	$\dot{L}$	$\dot{G}(G/Y)$	$\dot{G}$	调整后 $R$
普通线性	7	0.016 140 0±0.043 580 0	0.000 174 9±0.000 153 6	0.175 900 0±0.105 900 0	0.000 011 9±0.000 175 3	0.211 000 0±0.141 800 0	0.472
	8	0.015 951 0±0.042 537 6	0.000 173 2±0.000 148 2	0.178 133 4±0.098 163 3		0.202 508 3±0.065 022 1 **	0.495
	9	0.006 576 0±0.044 230 0	0.000 133 8±0.000 155 1	0.259 300 0±0.092 090 0 **	0.000 218 5±0.000 084 3 *		0.444
广义线性	7	0.016 140 0±0.043 580 0	0.000 174 9±0.000 153 6	0.175 900 0±0.105 900 0	0.000 011 9±0.000 175 3	0.211 000 0±0.141 800 0	0.563
	8	0.015 951 0±0.042 537 6	0.000 173 2±0.000 148 2	0.178 133 4±0.098 163 3		0.202 508 3±0.065 022 1 **	0.601
	9	0.006 576 0±0.044 230 0	0.000 133 8±0.000 155 1	0.259 300 0±0.092 090 0 **	0.000 218 5±0.000 084 3 *		0.592
半参数	7	0.016 300 0±0.037 710 0	0.000 069 9±0.000 114 8	0.136 100 0±0.081 970 0	**	0.171 900 0±0.113 500 0	0.728
	8	0.032 790 0±0.020 160 0		0.135 620 0±0.087 970 0		0.166 350 0±0.060 500 0 *	0.630
	9	0.046 270 0±0.032 690 0	0.000 039 7±0.000 115 3	0.200 700 0±0.075 940 0 *	***		0.749

表 3 1980—2013 年政府投资与林业经济增长关系的模型估计结果

参数类型	参数	截距	$I/Y$	$\dot{L}$	$\dot{G}(G/Y)$	$\dot{G}$	调整后 $R$
普通线性模型	7	0.266 874 9±0.068 725 9 ***	0.000 234 9±0.000 175 1	0.090 071 0±0.388 114 7	0.000 758 8±0.000 563 9	0.277 671 8±0.206 917 5	0.015 6
	8	0.259 188 0±0.069 439 0 ***	0.000 219 4±0.000 177 2	0.009 440 2±0.386 294 8		0.049 806 5±0.120 5718 **	0.011 9
	9	0.287 183 6±0.067 958 2 ***	0.000 256 2±0.000 176 8	0.132 586 8±0.355 671 8	0.000 139 6±0.000 328 5		0.011 6
广义线性模型	7	0.073 350 4±0.041 977 7	0.000 100 1±0.000 121 1	0.068 754 3±0.099 805 5	0.000 193 1±0.000 163 8	0.312 390 2±0.102 510 8 **	0.032 4
	8	0.259 188 0±0.069 439 0 ***	0.000 219 4±0.000 177 2	0.009 440 2±0.386 294 8		0.049 806 5±0.120 571 8	0.054 3
	9	0.287 183 6±0.067 958 2 ***	0.000 256 2±0.000 176 8	0.132 586 8±0.355 671 8	0.000 139 6±0.000 328 5		0.013 3
半参数模型	7	0.244 696 7±0.042 051 7 ***			0.000 708 1±0.000 570 2 **	0.171 900 0±0.113 500 0 ***	0.081 4
	8	0.201 360 0±0.023 170 0 ***					0.105 0
	9	0.279 395 5±0.060 194 8 ***	0.000 256 0±0.000 174 1	0.074 138 1±0.355 150 8			0.022 1

由表 1、表 2 可知,半参数混合模型的拟合度要优于普通线性模型和广义线性模型。原因如下:首先从模型中的  $R^2(\text{adj})$  数值看,半参数混合模型的  $R^2(\text{adj})$  的数值在 3 个模型中以半参数混合模型中数值最高,相比普通线性和广义线性模型来说,半参数混合模型的整体拟合度最好。其次,在各个模型中的参数的显著性程度来看,普通线性和广义线性模型除了常数项外,其他自变量都不显著。但在半参数混合模型里面,常数项和各自变量在 95% 的置信区间是显著的。所以半参数混合模型对公式(7)、公式(8)、公式(9)拟合效果

最好。综合  $R^2(\text{adj})$  数值和自变量的显著性,1953—2013 年选择半参数混合模型(7)、(9)为该时段的最终模型,1953—1980 年选择半参数混合模型(7)、(9)为该时段的最终模型。结合  $R^2(\text{adj})$  数值在 1981—2013 年选择半参数混合模型(8)为该时段的最终模型。

通过对 1953—2013 年的半参数混合模型(7)、(9)的数据观察发现, $\dot{G}$  系数大于零,说明政府部门在 1953—2013 年的政府产出对于非政府部门产出的边际外部作用为正。而  $\dot{G}(G/Y)$  系数为光滑样条也就是说非线性的,对于政府部门要

素的生产率是否高于非政府部门要素生产率整个很难判断,所以对于政府投资对于林业经济增长的总效应不好断定。但总的来说政府部门投资对于林业经济增长有作用但效果有限,因为总的  $R^2(\text{adj})$  数值在半参数混合模型没有超过 50%,但是相对普通线性和广义线性还是高很多。通过对 1953—1980 年的半参数混合模型(7)、(9)的数据观察,发现  $\dot{G}$  系数大于零,说明政府部门在 1953—1980 年的政府产出对于非政府部门产出的边际外部作用为正。而  $\dot{G}(G/Y)$  系数为光滑样条也就是说非线性的,对于政府部门要素的生产率是否高于非政府部门要素生产率整个很难判断,所以对于政府投资对于林业经济增长的总效应不好断定。但是总的  $R^2(\text{adj})$  数值在半参数混合模型高达 70% 以上,说明拟合效果比较好,也说明这个时期政府投资对于林业经济增长有很大的影响作用,而且在此期间林业经济的发展多是依靠政府的扶持,多以林木采伐为主,以粗放的林业经济为主导产业。通过数据可以观察到,这个时期的林业经济产值比较低,且相对林业经济发展速度来说相对缓慢的发展。通过对 1980—2013 年半参数混合模型(8)的数据观察,发现  $\dot{G}$  系数为光滑样条也就是说非线性的,对于政府产出对非政府产出的边际外部作用不确定,且总的  $R^2(\text{adj})$  数值在半参数混合模型非常低,最高为 10.5%,虽然很低但对相对线性和广义线性来说效果还是好很多。其实也说明现在的林业经济增长对于政府投资的依赖度降低很多,政府投资在固定资产中的投资比重很大,但是林业经济增长现在更多的是靠发展新型林业,如休闲旅游、森林游憩等。另外观察  $\dot{L}$  系数,因为  $\dot{L}$  系数代表非政府部门劳动投入要素的产出弹性。1953—2013 年该数值为 0.140 27,1953—1980 年为 0.200 7,但 1981—2013 年该数值为 0.355 1。该数值的变化说明非政府部门劳动投入要素对产出的贡献比率,进一步说明随着技术的进步,非政府部门的劳动投入增加对于总产出的贡献越来越大,从另一个角度说明非政府投资对于林业经济增长的促进作用越来越大。

通过表 1 至表 3 的数据分析也可以看出,林业经济增长在最初的发展阶段是依靠政府投资,占比很大,而随着对于林业的生态效益、社会效益的关注以及对于生态环境的重视,使得政府投资对于林业经济发展的支持力度有限,非政府部门对于林业经济增长的促进作用加大。政府投资关注于固定资产的投资,提高林业从业人员的社会福利保障,以及提高林业对于生态环境保护的作用。而新型林业的发展不仅有利于林业经济的增长,也保障了林业的生态效益。此结果也很符合我国显著林业发展的国情有利于供给侧改革的进行。

### 3 结论

本研究以 1953—2013 年的林业数据为基础,结合两部门的生产函数理论框架,通过半参数混合模型分析政府投资对林业经济增长的作用。通过政府投资对林业经济增长的外部作用和两部门之间要素的生产率差异 2 个方面进行分析。结果表明:1953—2013 年我国政府投资对林业经济增长具有正向的外部作用,而由于两部门之间的要素生产率差异看起来,经过判断得出政府支出对林业经济增长产生正向总作用,但这个正向的总作用在统计上不显著。1953—1980 年我国的

政府投资对林业经济增长具有正向的外部作用,而由于两部门之间的要素生产率差异看不出来,使得政府投资对林业经济增长具有一定的正向总作用,且结合当时的国情来看政府投资对于林业经济增长正向作用很显著。1981—2013 年由于非线性的参数有 3 个,对于政府投资对于林业经济增长的作用就很难判断出来,只能说政府投资对于林业经济增长有一定的正向外部作用。由于本试验选择了 3 个模型分别分析拟合度,并且考虑了参数之间可能存在非线性的关系,相对来说得到的结论可靠性更高。总体来说,我国政府投资对林业经济增长的作用是正向的,有利于林业经济增长。由于本试验参考数据是从 1953—2013 年的林业数据,样本考察相对更详细,对比改革开放前后。但是由于采用仅为时间序列数据,所以相对样本量是小的,特别是 1980—2013 年由于都存在非线性的情况,所以参考时要慎重使用。而且结合我国的林业政策,现在政府投资重点是以生态保护为主,集体林权制度的改革更是改变了林业经营模型和产业结构,使得林业经济有了显著提高。

### 参考文献:

- [1] 孔凡斌. 我国林业投资的机制转变和规模结构分析[J]. 农业经济问题, 2008(9): 91—96.
- [2] 刘珉. 林业投资研究[J]. 林业经济, 2011(4): 43—49.
- [3] 于江龙, 刘俊昌, 陈文汇. 我国林业投资与产业经济增长关系研究[J]. 价格理论与实践, 2011(9): 85—86.
- [4] 田淑英, 许文立. 基于 DEA 模型的中国林业投入产出效率评价[J]. 资源科学, 2012(10): 1944—1950.
- [5] 周家春, 陈岭. 林业固定资产投资与林业经济增长的相关性分析[J]. 中国林业经济, 2012(6): 41—42.
- [6] 孔凡斌, 苟龙巧, 廖文梅. 林业经济增长理论与实证研究: 综述与展望[J]. 林业经济问题, 2013, 33(2): 187—192.
- [7] 廖文梅. 南方集体林区林业经济增长的产业结构演变及其差异分析——基于 13 个省(区) 1995—2011 年的统计数据[J]. 林业科学, 2014(8): 131—140.
- [8] 才琪, 陈绍志, 赵荣. 中央林业投资与林业经济增长的互动关系[J]. 林业科学, 2015(9): 126—133.
- [9] Yuk W. Government size and economic growth: time-series evidence for the United Kingdom 1830—1993 [R]. Canada: Department of Economics, University of Victoria, 2005.
- [10] 刘洪, 金林. 基于半参数模型的财政支出与经济增长关系研究[J]. 财政研究, 2012(10): 65—68.
- [11] 陈高, 王朝才. 中国地方财政支出与经济增长关系研究——基于 1990—2012 年省级数据的线性混合模型分析[J]. 财政研究, 2014(8): 42—45.
- [12] 薛艳. 政府公共投资与区域经济增长的关系研究——基于半参数混合模型的分析[J]. 宏观经济研究, 2016(2): 81—88.
- [13] 张彩虹. 林业投资与林业经济增长的研究[M]. 北京: 中国林业出版社, 2001.
- [14] 冯达, 郑云玉, 温亚利. 改革开放以来我国林业经济增长的实证研究[J]. 安徽农业科学, 2010, 38(19): 10357—10359.
- [15] 国家林业局. 中国林业统计年鉴: 1949—2013 [M]. 北京: 中国林业出版社, 1949—2013.
- [16] 李晨婕, 温铁军. 宏观经济波动与我国集体林权制度改革——1980 年代以来我国集体林区三次林权改革“分合”之路的制度

彭 洁,徐剑晖,陈 超. 电子商务中基于潜在类回归模型的农产品个性化推荐方案[J]. 江苏农业科学,2017,45(12):274-278.  
doi:10.15889/j.issn.1002-1302.2017.12.068

# 电子商务中基于潜在类回归模型的农产品个性化推荐方案

彭 洁<sup>1</sup>, 徐剑晖<sup>1</sup>, 陈 超<sup>2</sup>

(1. 四川工程职业技术学院经济系, 四川德阳 618000; 2. 电子科技大学计算机科学与工程学院, 四川成都 611731)

**摘要:**针对现有电子商务中农产品个性化推荐方案精度较低的问题,提出一种基于潜在类回归模型(latent-class regression model,简称 LCRM)和组群偏好的个性化推荐方案。首先,收集农产品的评价信息,进行预处理,提取出每个评价者的特征-意见值对。然后,利用 LCRM 根据整体与特征评价,将具有相同爱好的评价者进行分组,构建组群偏好,并计算单个评价者的偏好。最后,通过计算用户与组群偏好的相似度来定位组群,通过计算用户与该组群中评价者偏好的相似度来定位农产品,最终列出推荐表。结果表明,该方案能够准确为用户推荐所需的农产品,推荐列表中农产品的命中率达到了 83%,同时具有较低的计算复杂度。

**关键词:**电子商务;农产品个性化推荐;潜在类回归模型;组群偏好

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 1002-1302(2017)12-0274-05

现今,利用网络进行日常商业交易的互联网用户越来越多,许多公司利用网络来销售他们的商品和服务。由于冷藏运输条件的改善,水果、蔬菜等农产品也开始融入到电子商务中<sup>[1]</sup>。在电子商务中,对于一个特定的商品,顾客面临多个选择,常处于困惑和迷失状态。对于网站管理员而言,评估提供的商品和服务是否迎合用户,为用户提供感兴趣的个性化商品推荐单至关重要<sup>[2]</sup>。

目前学者提出了多种电子商务推荐方案,例如 Huang 提出了一种基于知识决策支持的推荐方案,将推荐问题转化成约束满意问题,通过知识库检测商品和用户偏好的匹配度,查找与用户首选最接近的商品来生成推荐列表<sup>[3]</sup>。然而,这种方案仅依靠评价特征词出现的频率来定位商品,准确率较低。Krohn-Grimberghe 等提出了一种基于评价特征分析的推荐方案,从众多评价中提取特征,采用多关系矩阵分解(multi-relational matrix factorization,简称 MRMF)来搭建用户对商品和特定特征观点之间相关性的模型,从而预测客户所需商品的可能性<sup>[4]</sup>。然而,这种方案的局限性在于并没有强调新用户“不完全偏好”现象。Jain 等提出了一种基于线性回归模

型(linear regression model,简称 LRM)的推荐方案,利用评价者评论形成评价者偏好,根据用户和该偏好的相似度来定位商品<sup>[5]</sup>。然而,该方案仅匹配用户与单个评价者的偏好,没有考虑其他用户评价中的商品潜在信息,一定程度上影响了推荐精度。另外,现有的推荐方案主要是应用在电影、图书、电子产品等商品,对农产品的个性化推荐研究较少。郑云飞等设计了一种农产品协同过滤推荐系统<sup>[6]</sup>,但主要侧重于软件系统的构建,对推荐方案的描述较少,且效果不佳。将高效的个性化推荐技术应用到农产品推荐中,将会有助于农产品电子商务和农业地区经济的发展,具有重要的意义<sup>[7-8]</sup>。

偏好模型基于多属性效用理论(multi-attribute utility theory,简称 MAUT)<sup>[9]</sup>,根据用户偏好,利用匹配工具将所有商品进行排序从而给出推荐。然而,传统偏好模型尽管可以基于交互式偏好技术来了解买家的需求,但所得出的偏好不完整且不准确。另外,现有基于偏好模型的推荐方案中,大多仅考虑根据单个评价者对商品的评价信息建立偏好,没有充分挖掘商品评价中其他客户有价值的评价信息,不能很好地为新用户进行推荐。

为此,本研究针对农产品的个性化推荐应用,提出一种基于潜在类回归模型(latent-class regression model,简称 LCRM)的推荐方案<sup>[10]</sup>。利用 LCRM 根据整体与特征评价,将具有相同爱好的评价者进行分组,构建组群偏好,并计算单个评价者的偏好。通过计算用户与组群偏好的相似度来定位组群,通过计算用户与该组群中评价者偏好的相似度来定位农产品,最终给出推荐列表。结果表明,本研究方案能够准确地

收稿日期:2016-10-18

基金项目:四川省科技支撑计划(编号:2014RZ0048)。

作者简介:彭 洁(1981—),男,四川绵阳人,讲师,主要从事电子商务、物流信息化等研究。E-mail:pjason618@163.com。

通信作者:徐剑晖,博士,讲师,主要从事电子商务、供应链等研究, E-mail:1837273167@qq.com;陈 超,博士,讲师,主要从事电子商务等研究, E-mail:chenchaouestc@163.com。

变迁分析[J]. 中国软科学,2009(6):33-42,127.

[17]李 平,张俊飏. 中国财政科技投入与林业经济增长互动关系的实证研究[J]. 中国科技论坛,2011(9):134-138.

[18]孔凡斌,廖文梅. 中国林业市场化进程的林业经济增长效应[J]. 中国农村经济,2013(9):87-96.

[19]黄烈亚,翟印礼,梁 霁. 产业结构变动与区域林业经济增长及其空间差异[J]. 中国农业资源与区划,2008(6):27-31.

[20]肖 政,陈奕钢,周 悦. 公共服务投入和林业经济增长动态研究——基于初始模型 OLS 回归[J]. 林业经济问题,2012,32(2):177-184.