

集体建设用地市场改革对我国经济发展的影响评估

摘要

2015 年以后, 广东南海等地 (33 个试点地区) 先后建立了集体产权交易中心, 进行集体建设用地交易, 引导集体建设用地合法流转入市, 标志着城乡建设用地市场进入全面试点阶段, 集体建设用地市场化改革对我国经济发展有着重要的影响。

对于问题一: 根据搜集的数据, 首先选取 31 城市的 13 个指标评估农村集体建设用地市场化开放的相关指标, GDP 表示我国经济的发展, 考虑到变量太多会增加计算和分析问题的复杂性, 因此我们将采用主成分分析的方法, 然后对数据标准化处理、共线性诊断, 求出条件指数为 $68.5025 > 10$, 说明自变量之间存在严重的多重共线性, 对数据进行差分, 根据累积贡献率选择 2 个主成分, 其中包含了原来 13 个指标 89.22% 的信息, 最后根据 MATLAB 编程求出回归方程, 找到影响经济发展的主要评估指标, 分别为建筑业, 交通运输、仓储、邮政, 住宿、餐饮, 就业人数 4 个指标。

对于问题二: 根据获得的数据, 选取北京市大兴区进行评估, 由于搜集到该地区 GDP 数据量较少, 且 GDP 具有指数增长趋势, 所以我们选用灰色预测模型, MATLAB 编程, 预测得出 2015-2017 年的 GDP 为 1656.60、1803.50、1963.42 万元, 后验差系数为 0.0043 小于 0.35, 可知模型精度较优。将 GDP 实际值与预测值差值的平方做为该地区的农村集体建设用地市开放度的程度大小, 开始时开放程度较小, 2017 年时, 开放程度急剧上升; 基于第一问中得出的四个主要影响指标以及经济指标 (GDP 与区域税收), 进行数据包络分析, 构建开放程度的评价指数 E, 得出 2015、2016、2017 年的相对效率值: $E_{11}=0.8761$, $E_{22}=0.9623$, $E_{33}=1$, 所以前两年是弱有效的, 2017 年是有效的, 在一定时期内, 农村集体建设用地市场化开放程度越高对于经济的影响是越来越好的。由松弛变量 s^- 和 s^+ 进行适应性检验, 可知 2017 年影响是相对有效的。

对该地区的特点, 采用 MATLAB 编程以图的形式表现, 根据数值图像可以看出建筑业的数值处于轻微波动, 根据比值图像可以看出建筑业的所占比在逐年减小, 故此地区特点为建筑业的所占比逐年减小, 加大了其他产业的发展力度。

对于问题三: 首先从各省份统计年鉴中收集 33 个试点城市的经济指标即 GDP, 并且收集了对应省份非试点地区的往年 GDP 作为对照。利用各地区 2009-2017 年数据做灰色预测, 预测到目前 (2018 年) 的 GDP 值, 然后使用泰尔指数来判断各地区土地市场改革对经济的影响差异程度, 计算可得试点与试点区域之间的泰尔指数为 0.0714, 组内差异较小, 即试点与试点区域改革对其经济影响的差异较小, 非试点与试点区域之间的泰尔指数为 0.3638、各地区之间的总的经济差距测度的泰尔指数为 0.4464, 说明组间差异较大, 即试点与非试点区域改革对其经济影响的差异较大。我们对前十年的 GDP 进行统计, 测算出从 2009 年到 2018 年试点地区之间和试点与非试点地区之间市场开放对其经济的影响差异程度, 并且我们观察到组间、试点组内差异随着年份增大而逐渐增加, 说明农村建设用地市场开放对经济产生了显著性的影响, 导致试点与非试点地区间的经济影响差距增加。

最后, 我们对建立的模型的优缺点进行了评价, 并探讨了模型的推广前景和改进方向。

关键词: 数据包络分析 泰尔指数 主成分回归分析 土地市场化改革

一、问题的背景与重述

1.1 问题的背景

1987 年中国建立了城镇建设用地市场,随着城乡一体化建设进程的推进,集体建设用地的需求与日俱增。全国的集体经营性建设用地的面积约为 4200 万亩,约占全国建设用地总量的 13.3%;宅基地的总面积约为 1.7 亿亩,约占集体建设用地的 54%。农村集体建设用地包括宅基地、公益性公共设施用地和经营性用地,一级市场即政府出让市场,是指政府有偿、有限期的出让土地使用权的市场;二级市场是指土地的使用权转让市场。

长期以来要素市场的扭曲特别是竞争性土地市场的缺失,已使产品、货币、劳动、资本市场和地方财政体系均发生扭曲进而引起经济结构的扭曲,市场化开放迫在眉睫。在北京市大兴区 33 个农村“三块地”改革试点,调整实施土地管理法、城市房地产管理法关于农村土地征收、集体经营性建设用地入市、宅基地管理制度的有关规定,引导集体建设用地合法流转入市,允许农村集体经营性建设用地出让、租赁、入股,实行与国有土地同等入市、同权同价,标志着城乡建设用地市场进入全面试点阶段。通过宏观调控农村经济的发展方向,力争为新形势下的农村经济发展注入新的活力。

土地是农民重要的生产生活资料之一,土地制度是国家的基本制度。农村经济发展作为我国经济发展的薄弱环节,很大程度上限制了我国经济发展的速度,党的十八届三中全会已从宏观角度明确了农村土地制度改革的方向和任务,此次审议,明确了要坚持土地公有制性质不改变、耕地红线不突破、农民利益不受损三条底线,在 33 个试点基础上有序推进,并适当调整政府征地行为,这不仅是国家土地管理制度和城乡间土地资源配置的改革,也将对农民土地财产权利和农村金融空间产生积极影响。

眼下,人们的关注点主要在土地征收、经营性建设用地和宅基地制度等。2015 年至今,已有三年有余,集体建设用地市场化改革对我国经济发展有怎样的影响,需选择对经济影响的某个侧面,定量评估研究农村集体建设用地市场化开放对我国经济发展的影响。

1.2 问题的重述

广东南海、重庆等地(33 个试点地区)先后建立了集体产权交易中心,进行集体建设用地交易。选择对经济影响的某个侧面(比如对地方财政、GDP、消费水平、生产力、市场配置效率等等的影响,以及在时间和空间上的扩散路径等等),建立数学模型,利用互联网数据、中国统计年鉴等,定量评估农村集体建设用地市场化开放对我国经济发展的影响。

本文建立了合适的数学模型,并解决了如下的问题:

问题一:搜集数据,选出一些评估农村集体建设用地市场化开放的相关指标,对数据进行预处理或标准化处理,挑选其主要评估农村集体建设用地市场化开放对我国经济发展的影响的指标。

问题二:就某一地区而言,根据所获得的数据,评估该地区的农村集体建设用地市场化开放程度和特点对该地区经济发展的影响,要有问题的分析过程以及模型的解释,并给出模型适应性验证。

问题三:请根据掌握的数据,分析目前各地(省、或县市)之间,包括试点与未试点地区间的,农村集体建设用地市场化开放对经济发展影响的差异性。

二、模型的假设

假设一:假设 2015 年后该地区除题中政策无其他对经济影响的措施。

假设二:假设 2015 年后该地区无重大自然灾害。

假设三：假设不考虑热钱的流出，对外投资对宏观经济产生的影响。
假设四：假设忽略跨年代的通胀，汇率，科技差异。

三、主要符号说明

符号	符号说明
CI_j	方差膨胀因子
n	决策单元 (DMU_j)
E_n	DMU_j 的效率评价指数
f	目标函数的系数向量
A	不等式约束的系数矩阵
A_{eq}	等式约束的系数
s^-	m 项输入的松弛变量
s^+	s 项输出的松弛变量
T_w	总的组内差距为各分组组内差距的加权和
T	泰尔指数
D_b	组间贡献率
D_w	各分组内的组内贡献率
T_b	组间差距
Z	各组的组内差距

*注：其余符号详见文中说明。

四、模型的建立与求解

4.1 问题一

4.1.1 数据的选取

利用互联网数据、各地方统计年鉴，通过中国统计年鉴找出 31 个城市 2016 年经济指标 GDP 和地方财政，找出与农村集体建设用地相关的 13 个指标，分别为农林牧渔业，工业，建筑业，批发和零售，交通运输、仓储、邮政，住宿，餐饮，金融，房地产，农村消费，固定资产投资，就业人数，房地产开发投资额，城镇村建设用地，对 2016 年的 31 个城市各个指标进行整理，得出相关数据。

4.1.2 问题的分析

考虑到影响农村集体建设用地有众多变量，故选取 31 城市的 13 个指标评估农村集体建设用地市场化开放，GDP 表示我国经济的发展，变量太多会增加计算和分析问题的复杂性，每个指标在不同程度上反映了与经济的某些信息，且彼此之间可能具有多重共线性，在统计方法里面，可以采用主成分分析降低维数，做回归拟合方程找到主要的评估变量。

4.1.3 模型的建立

(1) 数据标准化处理。

由于每种指标的量纲和数据的数量级不相同，因此要对原始数据矩阵进行标准化处理。假设对 p 个变量进行 n 次观测得到的数据用如下矩阵表示：

$$X_{n \times m} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}$$

将其标准化为:

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} \quad (i=1,2,\cdots,n; j=1,2,\cdots,m)$$

(2) 求协方差矩阵 W 。

(3) 特征分解得 $W = U\Lambda U^T$ (相当于将原来的坐标轴旋转后得到新的坐标轴 U), 得 Z 的 m 个非负特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$, 这 m 个特征值为主成分的方差。

Λ —— Z 的特征值组成的对角阵

U —— Z 的特征向量按列组成的正交阵, 它构成了新的矢量空间, 作为新变量(主成分)的坐标轴。

确定主成分个数(根据累积贡献率):

$$(4) \quad \eta_p = \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_m}{\lambda_1 + \lambda_2 + \cdots + \lambda_p}$$

当 η_p 大于 80% 时, 可认为主成分数目为 p 。

(5) 写出主成分表达式:

$$Z_{n \times m} = X_{n \times p} U_{p \times m}$$

Z 阵的每一行相当于原数据矩阵的所有行(即原始变量构成的向量)在主成分坐标轴(载荷轴)上的投影, 这些新的投影构成的向量就是主成分得分向量。

(6) 构造评价函数:

$$F = \alpha_1 * Z_1 + \alpha_2 * Z_2 + \cdots + \alpha_m * Z_m$$

$$\alpha_i = \frac{\lambda_i}{\sum_i \lambda_i}, i = 1, 2, \cdots, m$$

将每个样本的主成分带入评价函数, 得到每个样本的综合得分, 依据一定的准则可对样本进行排序。

4.1.4 模型结果分析

(1) 将 13 个因变量直接对自变量进行多元回归。

表 1: 方差分析表

方差来源	自由度	平方和	均方	F 值	p 值
回归	13.0000	11642636156.8350	895587396.6796	687.4828	0.0000
残差	16.0000	20843282.3997	1302705.1500		
总计	29.0000	11663479439.2347			

由方差分析结果可知, $F=687.4828$, $p=0.0000$, 小于 0.05, 模型拟合较好。

表 2: 拟合优度检验

均方根误差	1141.3611	判定系数	0.9982
因变量均值	25867.2257	调整的判定系数	0.9968

由结果可知调整后 $R^2=0.9968$, 接近 1, 该模型总体拟合效果较好。

(2) 共线性诊断。

表 3: 参数估计

变量	估计值	标准误差	t 值	p 值	方差膨胀因子
常数项	-2258.3396	1268.6687	-1.7801	0.0941	0.0000
X1	1.3241	0.5059	2.6175	0.0187	12.1658
X2	0.7799	0.3575	2.1818	0.0444	199.9689
X3	1.3342	0.8739	1.5267	0.1464	14.7189
X4	1.1832	0.3975	2.9770	0.0089	18.7820
X5	-0.0095	0.9349	-0.0101	0.9920	12.6795
X6	1.0335	2.1419	0.4825	0.6360	16.1859
X7	1.4304	0.8235	1.7369	0.1016	38.4458
X8	3.1526	2.3604	1.3356	0.2004	217.6915
X9	0.1889	0.1066	1.7711	0.0956	8.6647
X10	0.1152	0.1033	1.1148	0.2814	41.1602
X11	3.5894	3.6258	0.9900	0.3369	53.1525
X12	-0.5075	0.3801	-1.3352	0.2005	21.1409
X13	0.0255	0.0697	0.3653	0.7197	2.7164

表 3 中方差膨胀因子大于 10，说明有较强的共线性问题。
共线性诊断的结果见附录二。

条件指数的定义： $CI_j = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}$ ，若 $CI_j > 10$ ，则存在严重多重共线性。

由附件二可知，条件指数为 68.5025，表明自变量之间存在严重的多重共线性。

表 4: 简单统计量

变量	均值	标准差
X1	2166.3517	1461.4092
X2	9489.8133	8384.2617
X3	1641.8223	930.4868
X4	2465.6633	2310.9840
X5	1163.8033	807.2640
X6	558.4400	398.1059
X7	1903.2567	1595.8351
X8	1257.3940	1324.8086
X9	13772.3000	5850.5296
X10	19906.5730	13159.3740
X11	592.8963	426.1658
X12	3359.7667	2563.8099
X13	6985.8270	5009.9683

13 个自变量间的相关系数矩阵见附录三。

表 5：相关矩阵特征值

序号	特征值	差分	贡献率	累积贡献率
1	9.8854	8.1721	0.7604	0.7604
2	1.7133	1.2614	0.1318	0.8922
3	0.4519	0.1216	0.0348	0.9270
4	0.3303	0.1563	0.0254	0.9524
5	0.1740	0.0454	0.0134	0.9658
6	0.1287	0.0226	0.0099	0.9757
7	0.1061	0.0267	0.0082	0.9838
8	0.0794	0.0277	0.0061	0.9899
9	0.0517	0.0210	0.0040	0.9939
10	0.0307	0.0036	0.0024	0.9963
11	0.0272	0.0082	0.0021	0.9984
12	0.0190	0.0169	0.0015	0.9998
13	0.0021	0.0000	0.0002	1.0000

选择主成分的个数为：2

从主成分的累计贡献率来看，前两个主成分包含了原来 13 个指标 89.22%的信息，因此，选择两个主成分代替原 13 个自变量。

特征向量结果见附录四；主成分的得分见附录五。

其中前两列数据为第一、二主成分得分，后 11 列为其余主成分得分情况。

(3) 标准化因变量 Y (GDP) 对主成分 Z1, Z2 进行多元回归，得如下结果。

表 6：方差分析表

方差来源	自由度	平方和	均方	F 值	p 值
回归	2.0000	28.7228	14.3614	1398.9680	0.0000
残差	27.0000	0.2772	0.0103		
总计	29.0000	29.0000			

表 7：拟合优度检验

均方根误差 (RootMSE)	0.1013	判定系数 (R-Square)	0.9904
因变量均值 (DependentMean)	-0.0000	调整的判定系 数 (AdjR-Sq)	0.9897

表 6，表 7 知 $F=1398.9680$ ， $p=0.0000<0.05$ ，调整 $R^2=0.9897$ ，说明提取的两个主成分模型总拟合效果较好。

表 8：参数估计

变量	估计值	标准误	t 值	p 值
常数项	-0.0000	0.0185	-0.0000	1.0000

Z1	0.3158	0.0060	52.7654	0.0000
Z2	0.0533	0.0144	3.7077	0.0010

常数项估计值为-0.0000, Z1 估计值为 0.3158, p 值为 0.0000<0.05, Z2 估计值为 0.0533, p 值为 0.0010<0.05, 变量 Z1, Z2 拟合较好。

(4) 回归方程的求解。

MATLAB 编程计算得出主成分回归方程:

①标准化数据的主成分回归方程为:

$$Y = -0.0000 + 0.0528 \times X_1 + 0.0974 \times X_2 + 0.0818 \times X_3 + 0.1018 \times X_4 + 0.0953 \times X_5 + 0.0916 \times X_6 + 0.1039 \times X_7 + 0.1028 \times X_8 + 0.0812 \times X_9 + 0.0749 \times X_{10} + 0.1010 \times X_{11} + 0.0995 \times X_{12} + 0.0457 \times X_{13}$$

②原始数据的主成分回归方程为:

$$y = -5574.0728 + 0.7247 \times x_1 + 0.2330 \times x_2 + 1.7629 \times x_3 + 0.8830 \times x_4 + 2.3683 \times x_5 + 4.6122 \times x_6 + 1.3058 \times x_7 + 1.5555 \times x_8 + 0.2782 \times x_9 + 0.1142 \times x_{10} + 4.7518 \times x_{11} + 0.7784 \times x_{12} + 0.1830 \times x_{13}$$

(5) 主要指标的选取。

根据原始数据的主成分回归方程选取系数大于 1.6 的指标纳入影响我国 GDP 的主要评估指标中, 可知 x_3, x_5, x_6, x_{11} 的系数大于 1.6, 分别为建筑业, 交通运输、仓储、邮政, 住宿、餐饮, 就业人数 4 个指标。因此该 4 个指标为主要评估农村集体建设用地市场化开放对我国经济发展的影响指标。

4.2 问题二

4.2.1.1 问题的分析

根据所搜集的数据, 选取北京市大兴区进行评估。我们搜集到该地区 GDP 数据量较少, 总体分布情况未知, 但可观察出 GDP 具有指数增长趋势, 所以首先用 2015 年前未改革的数据运用灰色预测模型对后三年的 GDP 进行预测, 将 GDP 实际值与预测值差值的平方的大小来当做该地区的农村集体建设用地市开放度的程度大小, 在此基础上进一步分析其用地市场化程度对地区 GDP 的影响。

其次市场开放程度与经济指标呈现正相关的关系, 根据第一问中得出的结论, 使用主要体现出农村集体建设用地市场开放对该地区经济影响的指标作为包络分析的输入变量, 经济指标 (GDP 与区域税收) 为输出变量, 再进行数据包络分析, 建立对应的输入输出权重向量, 构建开放程度的评价指数 E, 通过 E 值来评价对经济的相对有效性, 同时假设松弛变量, 对模型进行适应性检验分析, 最后得出农村集体建设用地市场化开放程度对经济的影响结果。

数据包络分析法, 是根据多项投入指标和多项产出指标, 应用线性规划模型, 对具有可比性的同类型单位做出相对有效评价的一种数量分析法, 又称为 DEA。

数据包络分析法的特点:

①多项投入与多项产出之间存在某种关系, 数据包络分析法不必确定存在的显性关系表达式。

②处理多目标决策问题, 不需要进行任何权重的假设, 有较强的客观性。

③ DEA 法 (数据包络分析法) 不是直接对数据进行评价, 因此不需要对数据进行量纲化处理。

4.2.1.2 模型的建立

设有 n 个决策单元 (DMU_j), 每一个单元有 m 项输入, s 项输出。

其中 $x_j = (x_{1j}, x_{2j}, \dots, x_{mj})^T$, $y_j = (y_{1j}, y_{2j}, \dots, y_{sj})^T$ 分别表示第 j ($j=1, 2, \dots, n$) 个决策单元 DMU_j 的输入量和输出量;

设 $v = (v_1, v_2, \dots, v_m)^T$, $u = (u_1, u_2, \dots, u_s)^T$ 分别表示 m 种输入和 s 种输出对应的权向量, 用总输出与总输入之比的大小来衡量 DMU_j 的有效性。

$$E_n = \frac{y_j^T \times u}{x_j^T \times v}, \quad E_n: DMU_j \text{ 的效率评价指数}$$

在上式中, 每一个分量都是非负的, 要求 E_n 达到最大值的权向量, 建立数据包络的 C^2R 模型:

$$E_n = \max \frac{y_j^T \times u}{x_j^T \times v}$$

$$s.t. \begin{cases} \frac{y_j^T \times u}{x_j^T \times v} \leq 1 (j=0, 1, 2, \dots, n) \\ v = (v_1, v_2, \dots, v_m)^T \geq 0 \\ u = (u_1, u_2, \dots, u_s)^T \geq 0 \end{cases}$$

将分式规划为一个等价线性问题, 令:

$$t = \frac{1}{x_j^T \times v}, \quad \omega = tv, \mu = tu$$

则目标函数为:

$$E_n = \max y_j^T * \omega \quad (P)$$

$$s.t. \begin{cases} y_j^T * \omega - x_j^T * \mu \geq 0, j=1, 2, \dots, n \\ \omega * x_0^T = 1 \\ \omega \geq 0, \mu \geq 0 \end{cases}$$

定义 假如线性规划 P 的解满足:

- (1) 若 $E_n = 1$, 则成为 DMU_j 为弱 DEA 有效 (C^2R) 的;
- (2) 若 $\omega > 0, \mu > 0$, 且 $E_n = 1$, 则成为 DMU_j 为 DEA 有效 (C^2R) 的;

线性规划的对偶规划 (D) 为:

$$\min \theta = \theta_0 - \varepsilon (e_1^T * s^- + e_2^T * s^+)$$

$$s.t. \begin{cases} \sum_{j=1}^n \lambda_j * x_j + s^- = \theta * x_j, \sum_{j=1}^n \lambda_j * y_j - s^+ = y_j \\ s^- \geq 0, s^+ \geq 0 \end{cases}$$

s^- : m 项输入的松弛变量; s^+ : s 项输出的松弛变量; $e_1^T = (1, 1, \dots, 1)_{1 \times m}$, $e_2^T = (1, 1, \dots, 1)_{1 \times s}$, ε 是一个很小的正数 (一般取 $\varepsilon = 10^{-6}$)。

定理 设线性规划最优解为 $*$, s^-, s^+ :

若线性规划 (D) 的解满足 $* = 1$, 则称为弱 DEA 有效 (C^2R) 的;

若 $* = 1$ 且 $s^- = 0, s^+ = 0$, 则称为 DEA 有效 (C^2R) 的;

MATLAB 编写程序, 求解的线性规划的极小值问题。

$$\begin{cases} \min f * w \\ s.t. A * w \leq b, Aeq * w = beq, LB \leq w \leq UB \end{cases}$$

W : 变量; f : 目标函数的系数向量; A : 不等式约束的系数矩阵; Aeq : 等式约束的系数矩阵; LB : 变量下界; UB : 变量上界。

MATLAB 中解决线性规划的语句:

$$w = LINPROG(f, A, b, Aeq, beq, LB, UB)$$

如果要解极大值问题 $\max f * w$, 只须转化为求极小值问题。

4.2.1.3 模型的求解

由 MATLAB 编程, 进行灰色预测得出 2015-2017 年的 GDP 为 1656.6013、1803.5019、1963.4290 万元, 预测值与实际值对比见图 1。

后验差系数为 0.0043 小于 0.35, 可知模型精度较优。

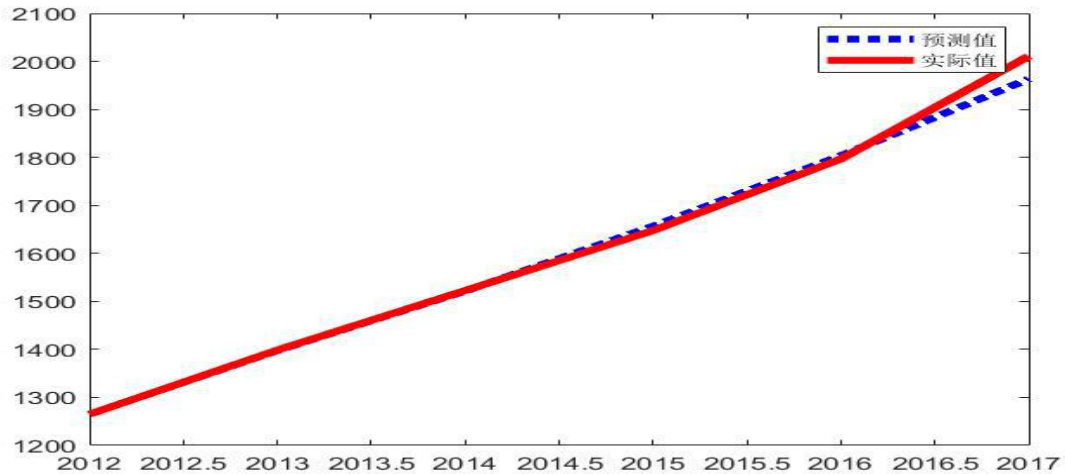


图 1: 预测 2015-2017 的 GDP 值与实际值对比情况

观察根据模型预测值和实际值绘制出的图像, 可以看出两条线之间有一定的差距, 假设 2015 年后该地区除题中政策无其他对经济影响的措施, 可将实际值与预测值差值的平方大小来当做该地区的农村集体建设用地市开放度的程度大小。

开放程度大小如下:

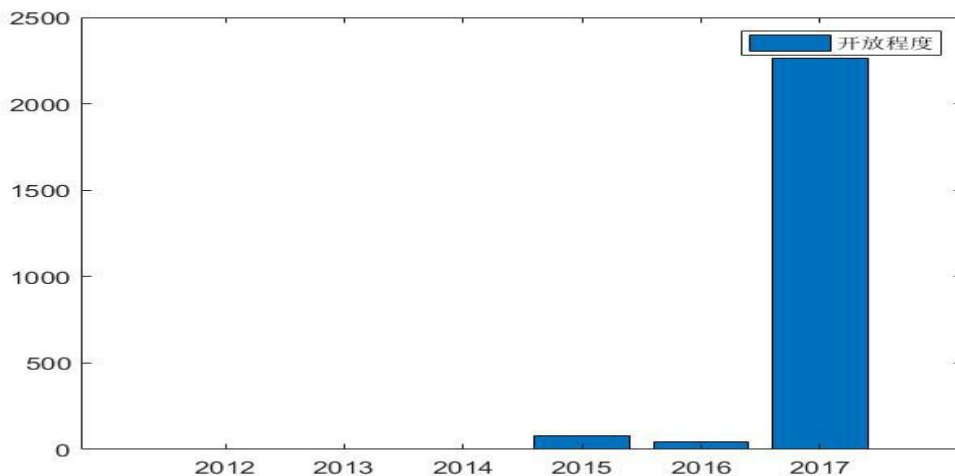


图 2: 用地市场化开放程度

在 2015 年以后, 广东南海、重庆、成都、上海等地 (33 个试点地区) 先后建立了集体产权交易中心, 进行集体建设用地交易, 开始时开放程度较小, 进行三年的改革之后, 2017 年的开放程度急剧上升。

根据第一问中得出的主要体现出农村集体建设用地市场开放对该地区经济影响的指标, 为包络分析的输入变量求解时, 为使得得到相对效率值较高, 经分析比较选取四个投入 (建筑业、交通运输, 仓储, 邮政、住宿, 餐饮、就业人数) 指标进行求解。

由附录四中程序 1，得到改革后即 2015、2016、2017 年的相对效率值：

$$E_{11}=0.8761, E_{22}=0.9623, E_{33}=1$$

表 9：各项投入，产出权向量值

		2015 年	2016 年	2017 年
投入权向量	建筑业	0	0	0
	交通运输，仓储，邮政	0.0061	0	0
	住宿，餐饮	0	0.0171	0
	就业人数	0	0	0
产出权向量	GDP（亿元）	0.0005	0.0005	0
	区域税收	0	0	0.0013

由定义得如下结论：

表 10：北京市大兴区某年份市场开放程度对经济影响结果

年份	E	结论
2015	0.8761	非 EDA 有效
2016	0.9623	非 EDA 有效，规模收益递增
2017	1	EDA 有效，规模收益递增

显而易见，市场开放农村集体建设用地市场化开放程度越高对于经济的影响是越来越好的。

由于定义可知前两年是非弱有效的，2017 年是有效的，为了确定有效和非有效的原因，需利用模型 (D) 编程求解。

表 11：2015-2017 松弛变量的结果

	2015 年	2016 年	2017 年
θ^*	0	0	0
(最佳权向量)	0	0	0
s^-	0.8193	0.8935	1.0000
(m 项输入的松弛变量)	72.5	50.0	0
s^+	0	9.5	0
(输出的松弛变量)	2.7	0	0
θ	7773.3	6738.0	0
	0	0	0
(输出的松弛变量)	53.8718	15.7509	0
theta	0.8718	0.9623	1.0000

由上面的表可以看出：

2017 年 $\theta^*=1$ 且松弛变量 $s^-=0$ ， $s^+=0$ ，结合上述定理可知，2017 年影响是相对有效的。

2015, 2016 是非有效性的, 对数据进行调整, 以 2015 为例, 根据有效性的经济意义, 在不减少各项输出的情下, 构造一个新的 2015 数据:

$$2015data=0.8193\times 2017data$$

同理构造一个新的 2016 数据:

$$2016data=0.8935\times 2017data$$

由附录程序 2 可得:

表 12: 调整后 2015, 2016 年薪、新数据

	建筑业	交通运输, 住宿, 餐饮 仓储, 邮政	就业人数	GDP(亿元)	区域税收
2015 年	439.8002	144.0329	186037.63	1647.612	622.1764
	4		17	3	
2016 年	479.6308	157.0773	202886.15	1796.828	678.5239
			15	5	

可知 2015 年的投入按比例减少到原投入的 75%, 并且由非零的松弛变量可知, 区域税收可以进一步增加 53.8764 亿元, 2016 年的投入按比例减少到原投入的 87%, 并且由非零的松弛变量可知, 区域税收可以进一步增加 15.7239 亿元。

4.2.2.1 问题的分析

还需对该地区的特点进行分析, 考虑画图的形式更为直观, 故采用 MATLAB 编程以图的形式表现。先将四个输入指标随年份变化的柱状图画出, 其中就业人数因量纲不同故单独一张图画出。再考虑其中三个经济指标在总 GDP 中的变化趋势, 根据变化趋势图可以直观的观测出该地区的特点。

4.2.2.2 问题的求解

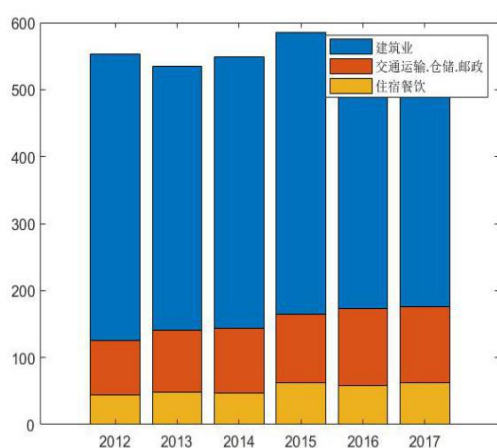


图 3: 2012-2017 年建筑业、交通运输等、住宿餐饮数量变化情况

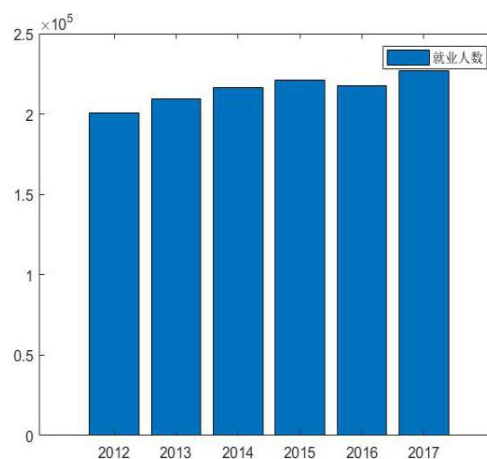


图 4: 2012-2017 年就业人数变化情况

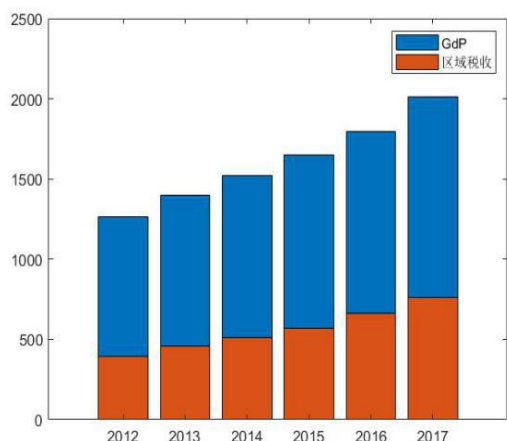


图 5: 2012-2017 年 GDP 与区域税收变化情况

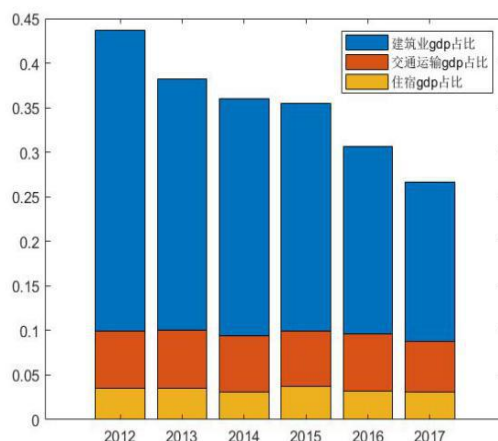


图 6: 2012-2017 年建筑业、交通运输、住宿分别占 GDP 的比例情况

4.2.2.3 结果的分析

函数中绘制了四个变量‘建筑业’，‘交通运输、仓储、邮政’，‘住宿餐饮’，‘就业人数’在 2012 年到 2017 年的数值变化、GDP 与区域税收在 2012 年到 2017 年的变化情况以及前三个变量在 GDP 中所占的比例变化情况。根据数值图像可以看出‘交通运输、仓储、邮政’，‘住宿餐饮’，‘就业人数’的数值在逐渐上升，建筑业的数值处于轻微波动，且有下降的状态。

根据图 6 的比值图像可以看出建筑业的所占比在逐年减小，其他比例有所增加，故北京市大兴区的地区特点为建筑业的所占比逐年减小，建筑业在该地区所占比重减小，加大了其他产业的发展力度。

4.3 问题三

4.3.1 问题的分析

包括试点与未试点地区间的,农村集体建设用地市场化开放对经济发展影响的差异性泰尔指数作为收入不平等程度的测度指标具备良好的可分解性质,泰尔指数可以分别衡量组内差距与组间差距的贡献。

由题干可知需要我们分析土地市场改革对经济的影响,首先从各省份统计年鉴中查找收集 33 个试点城市的经济指标即往年 GDP,由于部分地区的统计年鉴数据缺失或网站损坏,只能读取部分省份的试点地区的 GDP,并且在对应省份我们还收集了非试点地区的往年 GDP 作为对照。由于题干中要求分析目前土地市场开放对经济影响,所以我们利用可以收集到的各地区 2009-2017 年数据做灰色预测,预测到目前的 GDP 数目。

为了分析各个试点地区间土地市场开放对经济的影响差异,试点地区与非试点地区的经济影响的差异我们决定使用泰尔熵标准(Theil's entropy measure)作为衡量个人之间或者地区间收入差距(或者称不平等度)的指标,这一指数经常被使用。泰尔熵标准是由泰尔(Theil, 1967)利用信息理论中的熵概念来计算收入不平等而得名。

用泰尔熵指数来衡量不平等的一个最大优点是,它可以衡量组内差距和组间差距对总差距的贡献。泰尔熵标准只是普通熵标准(generalized entropy measures)的一种特殊情况。当普通熵标准的指数 $C=0$ 时,测量结果即为泰尔熵指数。取 $C=0$ 的优势在于分析组内、组间差距对总差距的解释力时更加清楚。

为了更好的了解目前农村建设用地市场化的影响,我们将每年的各个地区(包括试点地区与非试点地区)的泰尔指数,区域内的泰尔指数,区域间额泰尔指数的变化趋势

用折线图的形式画了出来，通过折线图看出过去，目前各区域间的差异，话可以通过折线图的变化趋势推测出将来各个地区农村土地市场开放的影响差异。并且我们测算出目前组间差异占总体差异的比值，试点区域内组内差异占总差异的比值大小，为了更直观的比较比值大小，我们将结果通过饼图的形式呈现出来。

4.3.2 模型的建立

假设包含 n 个个体的样本被分为 K 个群组，每组分别为 $g_k, (k=1, \dots, K)$ ，第 k 组 g_k 中的个体数目为 n_k ，则有 $\sum_{k=1}^K n_k = n$ ； y_i ：个体 i 的收入份额（占总收入的比例）； y_k ：第 k 组的收入份额（占总收入的比例）。

记 T_b 与 T_w 分别为组间差距和组内差距，则可将泰尔指数分解如下：

$$T = T_b + T_w = \sum_{k=1}^K y_k \ln(\frac{y_k}{n_k / n}) + \sum_{k=1}^K y_k (\sum_{i \in g_k} \frac{y_i}{y_k} \ln \frac{y_i / y_k}{1 / n_k})$$

中组间差距：

$$T_b = y_k \ln(\frac{y_k}{n_k / n})$$

组内差距：

$$T_w = \sum_{k=1}^K y_k (\sum_{i \in g_k} \frac{y_i}{y_k} \ln \frac{y_i / y_k}{1 / n_k})$$

其中，

$$T_k = \sum_{i \in g_k} \frac{y_i}{y_k} \ln \frac{y_i / y_k}{1 / n_k}, i \in g_k \text{ 为第 } k \text{ 组的组内差距 } (k=1, \dots, K)$$

再计算第 k 组组内差距的贡献率和组间差距的贡献率：

$$D_k = y_k * \frac{T_k}{T}, k=1, \dots, K$$

$$D_b = \frac{T_b}{T}$$

T_w ：总的组内差距为各分组组内差距的加权和； T：泰尔指数； D_b ：组间贡献率； D_w ：各分组内的组内贡献率；

4.3.3 模型的求解

表 13：泰尔指数及组间、组内贡献率

Tb:组间差距	Tw:组内差距	T:泰尔指数	Z: 各组的 组内差距	Db:组间贡献率	Dw:各分组 内的组内贡 献率
0.3639	0.0826	0.4353	0.0714	0.8150	0.1436
			0.1810		0.0413

试点区与试点区之间的泰尔指数为 0.0714，说明组内差异较小，即试点区域与试点区域改革对其经济影响的差异较小；非试点地区之间与试点地区之间的泰尔指数 0.3639 以及各地区之间的总的经济差距测度的泰尔指数为 0.4353，说明组间差异较大，即试点区域与非试点区域改革对其经济影响的差异较大。

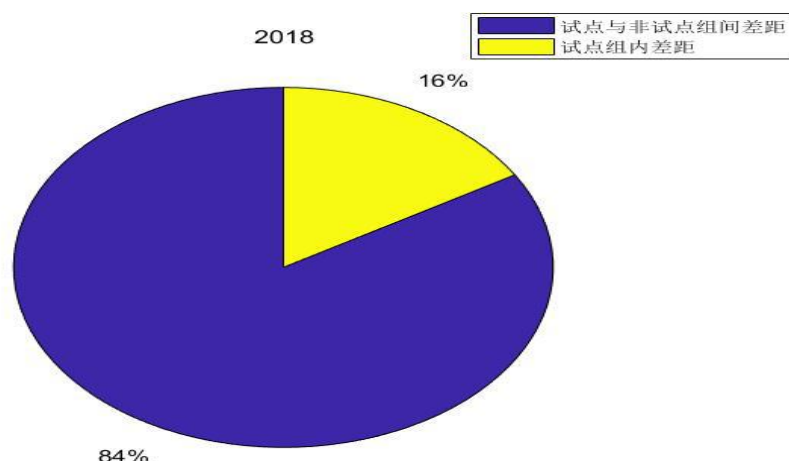


图 7：2018 年试点与非试点的组内、组间差距

由图 7 可知，较明显得出 2018 年组间差距占总体差距的 84%，占比较大，可得知试点与非试点之间差距较大。说明农村集体用地入市的政策在一定程度上带动了经济的发展，使得试点城市与非试点城市的经济差异性显著。

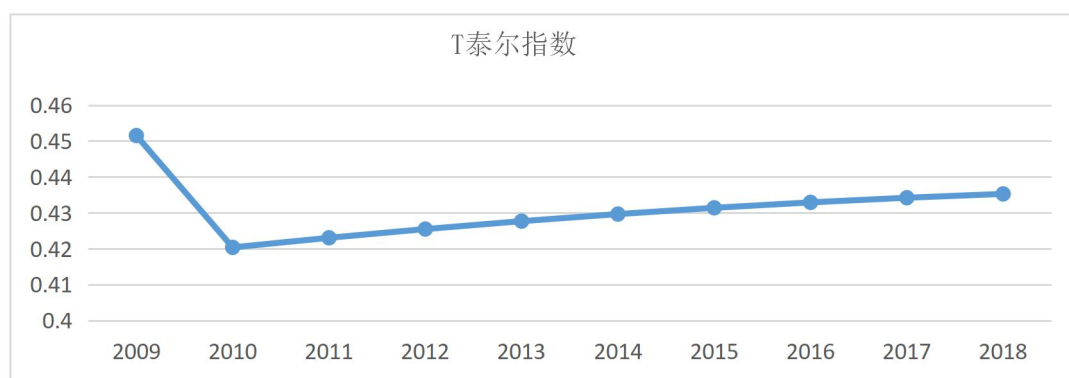


图 8：2009-2018 年泰尔指数变化情况

从总体上面来看，泰尔指数从 2010 年起呈逐年递增的趋势，可以得知该问题中选取的所有城市的农村集体建设用地入市对市场经济的影响的差异程度逐年增大，表明了该政策对地方经济产生了显著影响。由于总的泰尔指数不能体现出试点，非试点城市的具体差异程度，故对总的泰尔指数进行了分解，如下图所示：

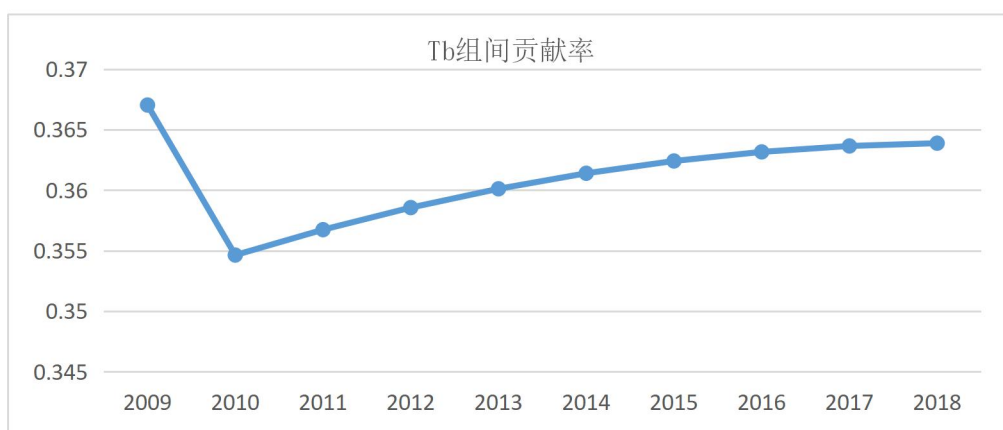


图 9：2009-2018 年组间贡献率变化情况

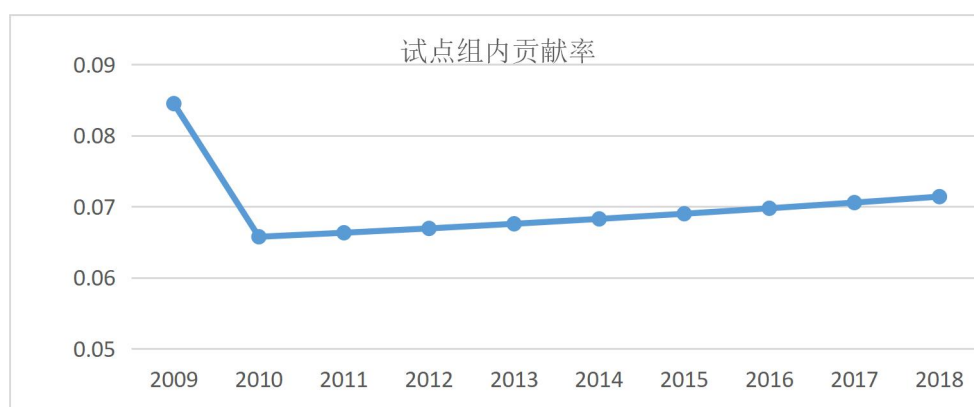


图 10：2009-2018 年试点组内贡献率变化情况

综合对于图 9 和图 10 的分析，组间和试点组内的图像都呈现缓慢递增趋势，并且组间差距明显大于组内差距，约为 5 倍，可以说明对于试点与非试点，进行农村集体建设用地市场化开放对经济发展有较显著性的差异。且组间的贡献率增长趋势逐渐趋于平缓，可得知试点与未试点地间的经济差异性仍然显著但改变趋势减小。

五、模型的评价与推广

5.1 模型的优点

对于问题一：我们收集了各省份的多项数据，考虑到变量过多，我们采用主成分回归分析法使得变量降维，消除了变量之间的相关性，使得拟合方程的结果更加精确可靠，拟合优度更高。

对于问题二：我们根据问题一中的结果，收集了大兴区试点数据，对数据进行绘图分析并采用数据包络分析法对综合指标进行评价，不需要进行任何权重的假设，排除了一些主观因素的干扰，有较强的客观性；数据包络分析法而且不需要对数据进行量纲化处理。同时该模型可对松弛变量分析进而提出改进的方式，确定了模型的适应性。

对于问题三：我们尽可能的收集多的城市的信息，使问题的分析更为全面。将所收集城市分为试点与未试点两组，采用泰尔指数的方法来衡量数据间差异，且可以衡量组内差距和组间差距对总体差异的贡献，可以直观的看出差异性的大小程度。

5.2 模型的缺点

对于问题一：由于数据与时间有限，没有将农村集体建设用地市场化开放的指标全部列举出来做主成分。我们选取的指标总体方面较少，主要都是与经济相关，可以考虑更多方面的指标对地方财政的影响，并且主成分中选取评估指标的个数带有个人主观性，仍可进一步改进。

对于问题二：由于只对经济指标进行了简单的变换，并不能消除地区间其他因素带来的经济差异，只能根据市场开放程度越高 GDP 增长越大，呈现正相关的关系去考虑问题。由 GDP 增长速度可以看出市场开放对 GDP 的影响程度，所以模型有较大的误差干扰项。我们特点分析中所分析的指标还可增加，可更加精确地分析所选地区的特点，对于开放程度的表达式带有主观性制，仍可进一步改进。

对于问题三：我们建立的泰尔指数模型仍有部分不足，对于城市的选取和分组可以进一步的以区域进行细分，进而更加精确地分析政策对不同地区的影响的差异。

5.3 模型的推广和改进

① 找出更多的与集体建设用地开放化有关的指标，对其进行主成分回归分析。

重新定义市场开放程度：即根据土地市场化测算公式如下：

$$LM = LM1 * W1 + LM2 * W2 = \frac{H * W_{11} + X * W_{12} + Z * W_{13} + P * W_{14} + G * W_{15} + ZL * W_{16} + ZR * W_{21} + C * W_{22} + D * W_{23}}{H + X + Z + P + G + ZL + ZR + C + D}$$

其中 LM 为市场化程度，LM1 和 LM2 分别为土地一级，二级市场化程度，W1、W2 为对应的一级权重；H、X、Z、P、G、ZL 分别表示一级市场中土地划拨，协议，招标，拍卖，挂牌，出让机器租赁的宗数地，W11、W12、W13、W14、W15、W16 为对应的市场化二级权重；ZR、C、D 分别为土地二级市场中转让，出租，抵押交易的土地宗数，W21、W22、W23 为这些交易方式对应的市场化二级权重，求出该试点地区的历年土地市场化程度，与历年的经济指标做相关系数，通过分析找到市场化对经济影响的程度与趋势。

② 改进数据包络分析的输入值与输出值，将输入值改为 x1：代表协议出让面积土地出让总面积(%)；x2：代表挂牌出让面积土地招拍挂出让面积(%)；x3：代表本年土地违法立案数 / 上年土地违法立案数(%)。输出改为：GDP 增量（评估程度）或各行业的投资额该变量生产总值。

③ 对 33 个试点地区的 GDP 与该省份其他未地区的 GDP 作为对照。首先对每一省份内试点地区与非试点地区差异用泰尔指数进行测算，在对各试点与各非试点地区非为两组对其进行组间与组内泰尔指数进行测算。这样可以更加全面的了解各个省份内试点与非试点的差异，也可以了解全国范围内，试点与非试点的差异。分析面更加广阔。

本文对农村集体建设用地市场化开放对我国经济发展的影响进行评估，并建立了任务定模型，具有良好的稳定性和推广性，可以通过简单改变参数适用于其他相似的任务评估情况。

本文搜集到的数据有限，选取的指标较少具有主观性，还可以增加数据进行更多分析，新一轮土地改革的到来，将增加农民财产性收入，影响我国经济的发展。

六、参考文献

- [1] 崔凯. 城镇化进程中农地非农化与土地市场化的作用关系研究[D]. 中国农业科学院, 2016.
- [2] 张婉娜. 统一城乡建设用地市场的影响因素及其对策研究[D]. 哈尔滨商业大学, 2015.
- [3] 姜琳. 我国各地土地出让市场化程度评估[J]. 经济与管理研究, 2009(06):94-99.
- [4] 华秀萍, 熊爱宗, 张斌. 金融开放的测度[J]. 金融评论, 2012, 4(05):110-121+126.
- [5] 孙加森. 数据包络分析(DEA)的交叉效率理论方法与应用研究[D]. 中国科学技术大学, 2014.
- [6] 盛晨, 庞娟. 数据包络分析(DEA)方法综述[J]. 科技经济导刊, 2016(20):8-10+5.
- [7] 彭育威, 吴守宪, 徐小湛. MATLAB 在数据包络分析中的应用[J]. 西南民族学院学报(自然科学版), 2002(02):139-143.
- [8] 朱紫笛. 基于 PSR 模型的土地政策对城市地价影响研究[D]. 河北师范大学, 2019.
- [9] 杨美超. 区域金融发展与区域经济差距[D]. 安徽财经大学, 2015.
- [10] 石义寿. 交通运输发展的区域差异性及其对区域经济影响的研究[D]. 深圳大学, 2017.

七、附录

附录一：主成分分析回归 matlab 程序

```
clear all
clc
%%导入数据实验
data=xlsread('问题一.xlsx');
[m1,n1]=size(data);
X=data(:,1:n1-1);
Y=data(:,end);%提出因变量
xishu1=regression(Y,X);%进行多元线性回归
xishu2=gongxianxing(Y,X);%进行共线性诊断
xishu3=jiandan(Y,X);%计算简单统计量
xishu4=pca1(Y,X);%进行主成分分析
xishu5=zhuchengfen(Y,X);%主成分回归
```

附录二：共线性诊断结果

序 号	特征 值	条件 指数	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
1	9.88 54	1	0.00 05	0	0.00 06	0.00 05	0.00 07	0.00 06	0.00 02	0	0.00 03	0.00 02	0.00 02	0.00 04	0.00 17
2	1.71 33	2.402	0.00 92	0	0.00 09	0.00 07	0	0.00 02	0.00 21	0.00 01	0.02 3	0.00 1	0.00 01	0.00 01	0.03 37
3	0.45 19	4.677	0.01 25	0	0.03 46	0.00 01	0.00 08	0.00 04	0.00 09	0.00 13	0.02 44	0.00 54	0.00 13	0.00 09	0.25 94
4	0.33 03	5.470	0.00 56	0	0.00 08	0.00 01	0.00 03	0.00 63	0.00 07	0.00 1	0.11 78	0.00 2	0.00 36	0.00 12	0.47 18
5	0.17 4	7.536	0.00 23	0.00 12	0.01 54	0.12 1	0.00 04	0.05 09	0.01 42	0	0	0.00 14	0.00 13	0.06 99	0.00 1
6	0.12 87	8.764	0.16 99	0.00 08	0.03 3	0.00 28	0.24 61	0.01 83	0.00 78	0.00 01	0.01 41	0.00 79	0.00 01	0.03 71	0.01 09

7	0.10 61	9.652 8	0.07 51	0.00 58	0.00 99	0.01 1	0.13 68	0.19 34	0.01 56	0.00 12	0.04 14	0.00 66	0.00 67	0.00 31	0.08 56
8	0.07 94	11.15 78	0.01 0.01	0.00 04	0.28 87	0.05 85	0.05 43	0.07 74	0.03 37	0.00 05	0.04 08	0.05 19	0.00 05	0.04 01	0.12 27
9	0.05 17	13.82 52	0.07 02	0.00 27	0.19 12	0.01 93	0.24 24	0.10 48	0.08 07	0.00 15	0.00 57	0.12 98	0.00 69	0.03 69	0.00 64
10	0.03 07	17.93 44	0.62 33	0.00 38	0.06 76	0.05 8	0.19 18	0.08 97	0.12 25	0.00 4	0.20 46	0.08 97	0.01 2	0.30 91	0.00 3
11	0.02 72	19.07 79	0.00 8	0.03 01	0.03 96	0.44 66	0.06 77	0.00 9	0.05 63	0.02 05	0.23 47	0.00 19	0.08 33	0.35 76	0.00 06
12	0.01 9	22.80 85	0.00 01	0.02 73	0.01 9	0.26 87	0.01 1	0.37 8	0.07 43	0.00 95	0.00 64	0.03 09	0.55 13	0.00 16	0.00 14
13	0.00 21	68.50 25	0.01 32	0.92 79	0.29 87	0.01 28	0.04 77	0.07 12	0.59 1	0.96 04	0.28 69	0.67 14	0.33 27	0.14 2	0.00 17

附录三：相关系数矩阵

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
X1	1	0.74 33	0.82 82	0.58 69	0.738 5	0.778 3	0.370 7	0.51 77	-0.03 75	0.87 94	0.67 21	0.64 5	0.705 3
X2	0.743 3	1	0.87 77	0.93 32	0.938 5	0.931 2	0.784 4	0.91 47	0.449 9	0.88 68	0.93 86	0.92 3	0.660 4
X3	0.828 2	0.87 77	1	0.79 56	0.825 9	0.852 6	0.632 6	0.69 65	0.357 6	0.91 89	0.80 62	0.82 84	0.599 2

X4	0.586 9	0.93 32	0.79 56	1	0.882 8	0.89	0.844 8	0.90 47	0.612 5	0.75 29	0.89 95	0.85 57	0.514 2
X5	0.738 5	0.93 85	0.82 59	0.88 28	1	0.875 8	0.788 4	0.87 76	0.466 1	0.83 24	0.91 45	0.89 14	0.651 4
X6	0.778 3	0.93 12	0.85 26	0.89	0.875 8	1	0.732 8	0.84 36	0.386 9	0.85 96	0.92 26	0.87 43	0.630 4
X7	0.370 7	0.78 44	0.63 26	0.84 48	0.788 4	0.732 8	1	0.91 71	0.753 8	0.55	0.87 59	0.85 57	0.323 2
X8	0.517 7	0.91 47	0.69 65	0.90 47	0.877 6	0.843 6	0.917 1	1	0.551 9	0.66 36	0.95 77	0.91 95	0.510 5
X9	-0.03 75	0.44 99	0.35 76	0.61 25	0.466 1	0.386 9	0.753 8	0.55 19	1	0.24 7	0.48 83	0.51 46	-0.01 94
X10	0.879 4	0.88 68	0.91 89	0.75 29	0.832 4	0.859 6	0.55	0.66 36	0.247	1	0.78 39	0.80 98	0.715 7
X11	0.672 1	0.93 86	0.80 62	0.89 95	0.914 5	0.922 6	0.875 9	0.95 77	0.488 3	0.78 39	1	0.94 89	0.586 7
X12	0.645	0.92 3	0.82 84	0.85 57	0.891 4	0.874 3	0.855 7	0.91 95	0.514 6	0.80 98	0.94 89	1	0.587 5
X13	0.705 3	0.66 04	0.59 92	0.51 42	0.651 4	0.630 4	0.323 2	0.51 05	-0.01 94	0.71 57	0.58 67	0.58 75	1

附录四：特征向量结果

Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9	Z10	Z11	Z12	Z13
----	----	----	----	----	----	----	----	----	-----	-----	-----	-----

X1	0.24 13	-0.43 89	0.262 3	-0.15	0.070 1	0.515 8	0.311 4	0.098 4	0.210 1	0.482 8	0.051 5	0.004 6	0.018 4
X2	0.31 26	-0.02 37	-0.01 98	-0.03 66	-0.20 14	-0.14 59	-0.35 02	-0.07 62	0.168 5	0.152 3	0.404 4	0.322	-0.62 52
X3	0.28 49	-0.15 3	0.48	0.061 4	0.198 6	-0.24 99	-0.12 46	0.580 8	-0.38 15	-0.17 48	0.125 8	0.073	0.096 2
X4	0.29 7	0.149 7	0.033 7	0.018 4	-0.62 89	-0.08 25	-0.14 77	0.295 5	0.136 9	0.182 9	-0.47 73	-0.30 97	0.022 5
X5	0.30 35	-0.00 93	-0.06 62	0.037 4	-0.02 81	0.633 8	-0.42 89	-0.23 38	-0.39 87	-0.27 34	-0.15 26	0.051 6	0.035 7
X6	0.30 15	-0.06 82	0.051 6	-0.18 35	-0.37 88	-0.19 52	0.576 4	-0.31 54	-0.29 61	-0.21 12	-0.06 28	0.341	0.049 3
X7	0.26 62	0.372 6	-0.12 38	-0.09 24	0.308 7	0.196 2	0.252 2	0.320 7	0.400 7	-0.38 04	-0.24 24	0.233	-0.21 88
X8	0.29 24	0.195 8	-0.35 14	-0.26 28	-0.01 29	-0.05 19	-0.16 43	0.097 1	0.129 2	0.164 1	0.348 1	0.197 9	0.663 6
X9	0.15 83	0.584 9	0.308 9	0.580 5	0.003 7	0.125 4	0.195 1	-0.16 74	-0.05 05	0.233 4	0.235	-0.03 25	0.072 4
X10	0.28 18	-0.26 37	0.316 2	0.163 1	0.101 6	-0.20 4	-0.16 95	-0.41 17	0.525 7	-0.33 69	-0.04 58	-0.15 54	0.241 3
X11	0.30 73	0.073 8	-0.17 58	-0.25 3	0.107 9	-0.02 23	0.194 7	-0.04 62	-0.13 8	-0.14 02	0.346 8	-0.74 62	-0.19 3
X12	0.30 28	0.073 2	-0.09 31	-0.09 14	0.507 3	-0.31 76	-0.08 31	-0.25 95	-0.20 1	0.448 1	-0.45 31	0.025	-0.07 95

X13	0.21 16	-0.39 58	-0.56 43	0.650 7	0.021 7	-0.06 19	0.157 1	0.162 7	-0.03 01	-0.01 58	0.006 5	0.008 6	-0.00 31
-----	------------	-------------	-------------	------------	------------	-------------	------------	------------	-------------	-------------	------------	------------	-------------

附录五：主成分得分结果

序 号	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9	Z10	Z11	Z12	Z13
--------	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----

1	0.234 6	-1.52 69	-0.77 32	0.791 7	0.560 7	-0.55 15	0.050 3	0.143 6	0.307 8	0.234 3	-0.14 3	0.030 4	0.006 9
2	-0.34 42	3.320 3	-0.10 72	0.274 2	0.605 4	0.245 4	0.496 5	-0.06 36	-0.01 77	-0.19 33	-0.02 59	-0.26 54	0.003 4
3	0.760 6	0.254 2	0.725 7	0.317 4	0.588 4	0.228 6	-0.55 08	0.011 6	-0.43 35	0.141 9	0.203 4	-0.09 05	-0.02 54
4	-3.16 74	-0.20 84	-0.07 81	-0.61 67	-0.11 43	-0.22 12	-0.10 23	0.023 1	0.149 5	-0.11 89	-0.01 19	-0.11 01	0.028
5	8.050 4	0.852 9	-2.32 27	-1.22 58	-0.29 38	-0.06 18	-0.10 54	0.013 4	-0.06 82	0.111 2	0.129 2	0.053 1	0.005 3
6	-1.00 6	-1.09 09	-0.14 62	0.050 9	0.172 7	0.190 4	0.116 5	0.178 3	0.176 1	0.185 9	0.041 4	0.081 2	0.024 4
7	-1.64 13	-1.12 5	-0.92 17	0.832 6	0.044 5	0.209 6	0.247 5	-0.08 83	-0.25 13	0.000 2	-0.18 48	0.075 6	-0.03 52
8	1.414 2	-1.43 25	-0.30 85	0.273 6	0.431 3	0.870 2	-0.89 32	-0.29 52	0.022 1	-0.00 21	-0.24 26	-0.03 09	0.038 6
9	3.220 9	-1.87 42	-0.31 15	-0.13 87	0.32	-0.02 89	0.427 8	-0.63 36	0.005	-0.12 98	0.053 3	-0.20 24	-0.10 03

10	-1.75 06	-0.28 85	0.161 4	-0.49 37	-0.57 68	0.581 7	0.535 6	-0.05 19	0.072 3	0.257 8	0.074	-0.10 47	0.106 9
11	1.114 1	-0.63 38	0.866 3	-0.64 87	0.296 3	0.147 9	0.192 9	-0.13 15	0.282 5	0.062 3	0.010 5	0.072	-0.07 65
12	0.701 9	-1.61 49	0.089	0.568 5	-0.12 07	0.254 6	0.198 5	0.056 7	0.101 3	0.074 3	0.071 3	0.019 7	-0.01 32
13	-2.27 77	-0.20 63	0.214	-0.59 67	-0.38 28	-0.07 52	-0.14 24	-0.15 2	0.199	-0.06 52	0.136 5	0.021 9	-0.07 38
14	7.597 6	1.106 4	0.839 7	0.331 1	0.382 7	-0.10 38	-0.20 02	0.345 1	0.262 3	-0.12 31	0.308 8	0.183 2	0.016 5
15	-0.79 54	-1.06 47	-0.66 07	1.110 5	0.029 3	-0.20 97	0.232 6	0.305 6	0.011 4	-0.19	0.333 2	-0.04 5	0.053 1
16	-0.65 11	0.346 8	0.258 3	-0.38 52	-0.26 75	0.522 6	0.007 2	0.711 4	-0.51 53	0.121 8	0.017 8	-0.05 96	-0.06 96
17	-1.09 09	-0.29 02	-0.03 15	0.640 8	-0.69 98	0.166 5	0.278 8	-0.17 58	-0.34 66	-0.29 27	0.056	0.382 3	-0.02 01
18	-3.85 08	0.553 6	-0.15 96	0.038 7	-0.10 44	-0.15 52	-0.23 55	-0.11 48	0.018 6	0.160 7	0.061 6	0.065 7	-0.00 04
19	-4.05 12	0.641 2	-0.05 35	-0.09 57	-0.19 45	-0.16 92	-0.20 76	-0.13 04	0.119 6	0.162	0.102 2	0.035 3	-0.01 37
20	6.458 3	-1.05 53	1.061 6	0.430 6	-1.20 85	-0.06 09	-0.19 52	0.168	0.202 7	-0.09 69	-0.27 82	-0.18 27	0.006 6
21	-1.93 83	0.265 3	-0.23 03	-0.22 88	-0.10 8	0.013 5	-0.17 55	-0.30 9	-0.05 31	-0.43 62	0.012	-0.09 26	0.067 7

22	-0.95 52	-0.41 57	0.492	-0.46 36	0.148 2	-0.46 29	-0.23 16	0.139 7	-0.06 71	-0.24 69	0.057 1	-0.08 51	-0.00 15
23	0.074 1	3.457 7	-0.48 25	0.388 7	0.005 9	0.375 2	0.144 4	0.271	0.327	-0.07 75	-0.32 09	0.070 1	-0.03 83
24	1.557 1	-0.45	1.062 2	-1.03 64	0.581 6	0.194 6	0.493 1	-0.20 6	-0.09 73	0.000 4	-0.10 96	0.253 7	0.072 9
25	-1.67 06	2.226 8	0.548 6	0.716	-0.26 44	-0.06 21	-0.15 05	-0.50 63	0.193 6	0.236 3	0.116	0.056 5	-0.00 74
26	-4.44 26	0.175 2	-0.29 62	-0.63 43	-0.26 78	-0.27 98	-0.36 17	0.083 4	0.093 5	-0.03 86	-0.05 77	0.077 4	-0.02 48
27	-2.58 77	-0.45 28	-0.11 28	-0.05 47	0.042 5	0.173 8	-0.08 12	0.214 3	0.030 6	0.040 9	0.159 9	-0.18 67	0.020 5
28	-1.45 28	-0.76 42	0.345 9	-0.26 08	0.196 1	-0.49 83	0.415 4	0.405 3	-0.05	0.066 6	-0.18 83	-0.08 07	-0.01 46
29	3.543 4	1.328 5	0.363 8	0.355 3	-0.23 87	-0.79 95	0.060 4	-0.39 6	-0.49 37	0.271 1	-0.11 59	-0.08	0.034 8
30	-1.05 35	-0.03 44	-0.03 21	-0.24 1	0.436 5	-0.43 46	-0.26 43	0.183 7	-0.18 09	-0.11 64	-0.26 54	0.138 2	0.029 3

附录六：多元线性回归所用程序

```
function xishu=regression(Y,X)
```

```
if nargin < 2
```

%当输入变量个数< 2 时，提示至少需要

输入两个变量

```
error('至少需要两个输入参数');
```

```
end
```

```
[n,p]=size(X);
```

%显示矩阵 X 的行数和列数

```
[n2,p2]=size(Y);
```

%显示向量 Y 的行数和列数

```
if n2~=n
```

%当自变量和因变量的观测值不一致时，


```

% 显示判定系数等统计量
fmt = '%22s%15.4f%25s%10.4f';
fprintf(fmt, '均方根误差(Root MSE)', sqrt(ST.mse), '判定系数'
(R-Square)', ST.rsquare);
fprintf('\n');
fprintf(fmt, '因变量均值(Dependent Mean)', mean(Y), '调整的判定系数(Adj
R-Sq)', ST.adjrsquare);
fprintf('\n');
fprintf('\n');

% 显示参数估计及 t 检验相关结果
fprintf('-----参数估计
-----');
fprintf('\n');
fprintf('%8s%17s%15s%13s%12s%12s', '变量', '估计值', '标准误差', 't 值', 'p 值', '
方差膨胀因子');
fprintf('\n');
for i = 1:size(t.beta,1)
    if i == 1
        fmt = '%8s%20.4f%17.4f%16.4f%12.4f%14.4f\n';
        fprintf(fmt, '常数项', t.beta(i), t.se(i), t.t(i), t.pval(i), vif(i));
    else
        fmt = '%10s%20.4f%17.4f%17.4f%12.4f%14.4f\n';

fprintf(fmt, varnames{i-1}, t.beta(i), t.se(i), t.t(i), t.pval(i), vif(i));
    end
end
fprintf('\n');
xishu=t.beta;
end
附录七：简单统计量所用 matlab 程序
function xishu=jiandan(Y,X)
if nargin < 2
    error('至少需要两个输入参数');
end
[n,p]=size(X);
[n2,p2]=size(Y);

if n2~=n
    error('自变量与因变量观测量不统一');
end

if p2~=1

```

```

        error('因变量个数必须为一')
    end

    xg=corr(X); %相关系数
    mx=mean(X); %计算平均值
    my=mean(Y);
    stdx=std(X,0); %计算标准差
    stdy=std(Y,0);
    varnames = strcat({'X'}, num2str([1:p]'));

    % 显示简单统计量
    fprintf('-----简单统计量-----');
    fprintf('\n');
    fprintf('%s%10s%10s', '变量', '均值', '标准差');
    fprintf('\n');
    for i=1:p
        fprintf('%3s%13.4f%12.4f\n', varnames{i}, mx(i), stdx(i));
    end
    fprintf('\n');

    %相关系数
    fprintf('-----相关矩阵-----');
    fprintf('\n');
    for i=1:p
        fprintf('%12s', strcat('X', num2str(i)));
    end
    fprintf('\n');
    for i=1:p
        fprintf('%s', strcat('X', num2str(i)));
        for j=1:p
            fprintf('%12.4f', xg(i, j));
        end
        fprintf('\n');
    end
    fprintf('\n');
    xishu=mx;
    end
附录八：主成分分析所用 matlab 程序
function xishu=pca(Y,X)
if nargin < 2
    error('至少需要两个输入参数');
end
[n,p]=size(X);

```

```

[n2, p2]=size(Y);

if n2~=n
    error('自变量与因变量观测量不统一');
end

if p2~=1
    error('因变量个数必须为一')
end
%主成分分析
PHO = corrcoef(X); %计算相关系数阵
[COEFF, LATENT, EXPLAINED] = pcacov(PHO); %主成分分析
LEXPLAINED=cumsum(EXPLAINED);
LATENT(p+1)=LATENT(p);

% 相关矩阵特征值
fprintf('-----相关矩阵特征值-----');
fprintf('\n');
fprintf('%5s%10s%9s%10s%10s','序号','特征值','差分','贡献率','累积贡献率');
fprintf('\n');
for i=1:p

    fprintf('%5s%14.4f%12.4f%12.4f%12.4f\n',num2str(i),LATENT(i),LATENT(i)-LATENT(i+1),EXPLAINED(i)/100,LEXPLAINED(i)/100);
end
fprintf('\n');
LATENT(p+1)=[];

COEFF(:,find(sum(COEFF)<0))=-COEFF(:,find(sum(COEFF)<0));

% 特征向量
fprintf('-----特征向量-----');
fprintf('\n');
for i=1:p
    fprintf('%12s',strcat('Z',num2str(i))); % 打印特征向量
end
fprintf('\n');
for i=1:p
    fprintf('%s',strcat('X',num2str(i)));
    for j=1:p
        fprintf('%12.4f',COEFF(i,j));
    end
    fprintf('\n');
end

```

```

end
fprintf('\n');

COEFF=COEFF';
Xlin=[];
for i=1:p
    Xlin(:,i)=sum(zscore(X).*COEFF(ones(n,1)*i,:),2);
end

% 主成分得分
fprintf('-----主成分得分
-----');
fprintf('\n');
fprintf('%3s','序号');
for i=1:p
    fprintf('%11s',strcat('Z',num2str(i)));
end
fprintf('\n');
for i=1:n
    fprintf('%3s',num2str(i));
    fprintf('%3s','');
    for j=1:p
        fprintf('%11.4f',Xlin(i,j));
    end
    fprintf('\n');
end
fprintf('\n');
xishu=Xlin;
end
附录九：主成分回归所用 matlab 程序
function [xishu,a]=zhuchengfen(Y,X)
if nargin < 2
    error('至少需要两个输入参数');
end
[n,p]=size(X);
[n2,p2]=size(Y);

if n2~=n
    error('自变量与因变量观测量不统一');
end

if p2~=1
    error('因变量个数必须为一')
end

```

```

mx=mean(X); %平均值
my=mean(Y);
stdx=std(X,0); %标准差
stdy=std(Y,0);

%主成分分析
PH0=corrcoef(X); %计算相关系数阵
[COEFF, LATENT, EXPLAINED] = pcacov(PH0); %主成分分析
COEFF(:,find(sum(COEFF)<0))=-COEFF(:,find(sum(COEFF)<0));
COEFF=COEFF';
Xlin=[];
for i=1:p
    Xlin(:,i)=sum(zscore(X).*COEFF(ones(n,1)*i,:),2);
end

%标准化因变量对主成分进行多元回归分析
i=input('选择主成分的个数为: '); %手动输入选取主成分的个数
jieyong=reglm(zscore(Y),Xlin(:,1:i)); %标准化因变量对主成分进行多元回归分析
a=jieyong.adjrsquare;
%计算标准化数据的回归方程
B=[];
B=[ones(n,1),Xlin(:,1:i)]'*[ones(n,1),Xlin(:,1:i)]^(-1)*[ones(n,1),Xlin(:,1:i)]'*zscore(Y);
B2=[B(1),B(2)*COEFF(1,:)];
for j=3:i+1
    B2(2:end)=B2(2:end)+B(j)*COEFF(j-1,:);%记录标准化数据的主成分回归方程的系数
end

%还原为原始数据的回归方程
B3=[((B2(1)-sum(B2(2:end)).*mx./stdx))*stdy+my)/stdy,B2(2:end)./stdx]*stdy;%记录原始数据的主成分回归方程的系数

%输出标准化数据回归方程
fprintf('\n');fprintf('\n');
fprintf('%s%1.4f','标准化数据的主成分回归方程为: Y=',B2(1)); %打印标准化数据的主成分回归方程
for i=2:p+1
    if B2(i)<0
        fprintf('%1.4f%s',B2(i),strcat('*X',num2str(i-1)));
    else
        fprintf('%s%1.4f%s','+',B2(i),strcat('*X',num2str(i-1)));
    end
end

```

```

        end
    end

%输出原始数据回归方程
fprintf('\n');
fprintf('%s%1.4f', '原始数据的主成分回归方程为: y=', B3(1)); %打印原始数据的主成分回归方程
for i=2:p+1
    if B3(i)<0
        fprintf('%1.4f%s', B3(i), strcat('*x', num2str(i-1))));
    else
        fprintf('%s%1.4f%s', '+', B3(i), strcat('*x', num2str(i-1))));
    end
end
end
fprintf('\n');
xishu=B3;
end
附录十：线性回归所用 matlab 程序
function stats = reglm(y,X,model,varnames)
% 多重线性回归分析或广义线性回归分析
%
% reglm(y,X)，产生线性回归分析的方差分析表和参数估计结果，并以表格形式显示在屏幕上。参
% 数 X 是自变量观测值矩阵，它是 n 行 p 列的矩阵。y 是因变量观测值向量，它是 n 行 1 列的列向量。
%
% stats = reglm(y,X)，还返回一个包括了回归分析的所有诊断统计量的结构体变量 stats。
%
% stats = reglm(y,X,model)，用可选的 model 参数来控制回归模型的类型。model 是一个字符串，
% 其可用的字符串如下
%      'linear'          带有常数项的线性模型（默认情况）
%      'interaction'     带有常数项、线性项和交叉项的模型
%      'quadratic'       带有常数项、线性项、交叉项和平方项的模型
%      'purequadratic'   带有常数项、线性项和平方项的模型
%
% stats = reglm(y,X,model,varnames)，用可选的 varnames 参数指定变量标签。
varnames
% 可以是字符矩阵或字符串元胞数组，它的每行的字符或每个元胞的字符串是一个变量的标签，它的行
% 数或元胞数应与 X 的列数相同。默认情况下，用 X1,X2,...作为变量标签。

```

```

if nargin < 2
    error('至少需要两个输入参数');
end
p = size(X,2);    % X 的列数, 即变量个数
if nargin < 3 || isempty(model)
    model = 'linear';    % model 参数的默认值
end

% 生成变量标签 varnames
if nargin < 4 || isempty(varnames)
    varname1 = strcat({'Z'}, num2str([1:p]'));
    varnames = makevarnames(varname1,model);    % 默认的变量标签
else
    if ischar(varnames)
        varname1 = cellstr(varnames);
    elseif iscell(varnames)
        varname1 = varnames(:);
    else
        error('varnames 必须是字符矩阵或字符串元胞数组');
    end
    if size(varname1,1) ~= p
        error('变量标签数与 X 的列数不一致');
    else
        varnames = makevarnames(varname1,model);    % 指定的变量标签
    end
end

ST = regstats(y,X,model);    % 调用 regstats 函数进行线性回归分析, 返回结构体
变量 ST
f = ST.fstat;    % F 检验相关结果
t = ST.tstat;    % t 检验相关结果

% 显示方差分析表
fprintf('\n');
fprintf('-----方差分析表\n');
fprintf('\n');
fprintf('%s%7s%15s%15s%15s%12s', '方差来源', '自由度', '平方和', '均方', 'F 值', 'p 值');
fprintf('\n');
fmt = '%s%13.4f%17.4f%17.4f%16.4f%12.4f';
fprintf(fmt, '回归', f.dfr, f.ssr, f.ssr/f.dfr, f.f, f.pval);
fprintf('\n');
fmt = '%s%13.4f%17.4f%17.4f';

```

```

fprintf(fmt, '残差', f.dfe, f.sse, f.sse/f.dfe);
fprintf('\n');
fmt = '%s%13.4f%17.4f';
fprintf(fmt, '总计', f.dfe+f.dfr, f.sse+f.ssr);
fprintf('\n');
fprintf('\n');

% 显示判定系数等统计量
fmt = '%22s%15.4f%25s%10.4f';
fprintf(fmt, '均方根误差(Root MSE)', sqrt(ST.mse), '判定系数'
(R-Square)', ST.rsquare);
fprintf('\n');
fprintf(fmt, '因变量均值(Dependent Mean)', mean(y), '调整的判定系数(Adj
R-Sq)', ST.adjrsquare);
fprintf('\n');
fprintf('\n');

% 显示参数估计及 t 检验相关结果
fprintf('-----参数估计
-----');
fprintf('\n');
fprintf('%8s%18s%15s%15s%12s', '变量', '估计值', '标准误', 't 值', 'p 值');
fprintf('\n');
for i = 1:size(t.beta,1)
    if i == 1
        fmt = '%8s%20.4f%17.4f%17.4f%12.4f\n';
        fprintf(fmt, '常数项', t.beta(i), t.se(i), t.t(i), t.pval(i));
    else
        fmt = '%10s%20.4f%17.4f%17.4f%12.4f\n';
        fprintf(fmt, varnames{i-1}, t.beta(i), t.se(i), t.t(i), t.pval(i));
    end
end

if nargout == 1
    stats = ST; % 返回一个包括了回归分析的所有诊断统计量的结构体变量
end

% -----子函数-----
function varnames = makevarnames(varname1,model)
% 生成指定模型的变量标签
p = size(varname1,1);
varname2 = [];
for i = 1:p-1
    varname2 = [varname2;strcat(varname1(i),'*',varname1(i+1:end))];
end

```



```

end
varname3 = strcat(varname1,'*',varname1);
switch model
    case 'linear'
        varnames = varname1;
    case 'interaction'
        varnames = [varname1;varname2];
    case 'quadratic'
        varnames = [varname1;varname2;varname3];
    case 'purequadratic'
        varnames = [varname1;varname3];
end
附录十一：简单统计量所用 matlab 程序
function xishu=jiandan(Y,X)
if nargin < 2
    error('至少需要两个输入参数');
end
[n,p]=size(X);
[n2,p2]=size(Y);

if n2~=n
    error('自变量与因变量观测量不统一');
end

if p2~=1
    error('因变量个数必须为一')
end

xg=corr(X); %相关系数
mx=mean(X); %计算平均值
my=mean(Y);
stdx=std(X,0); %计算标准差
stdy=std(Y,0);
varnames = strcat({'X'},num2str([1:p]'));

% 显示简单统计量
fprintf('-----简单统计量-----');
fprintf('\n');
fprintf('%s%10s%10s','变量','均值','标准差');
fprintf('\n');
for i=1:p
    fprintf('%3s%13.4f%12.4f\n',varnames{i},mx(i),stdx(i));
end

```

```

fprintf(' \n');

%相关系数
fprintf('-----相关矩阵-----');
fprintf(' \n');
for i=1:p
    fprintf(' %12s', strcat(' X', num2str(i)));
end
fprintf(' \n');
for i=1:p
    fprintf(' %s', strcat(' X', num2str(i)));
    for j=1:p
        fprintf(' %12.4f', xg(i, j));
    end
    fprintf(' \n');
end
fprintf(' \n');
xishu=mx;
end
附录十二：第二问数据包络所用 matlab 程序
%初始化
clc,clear
%此程序进行适应度分析
X=[584.7  550.4  536.8
    164.4  173.1  175.8
    62     58.5  63
    221217 217852 227069];%用户输入多指标输入矩阵 X
Y=[1647.6 1796.9 2011
    568.3  662.8  759.4];%用户输入多指标输出矩阵 Y
[m,n]=size(X);%读入输入矩阵大小
s=size(Y,1);%输出变量类数
A=[-X' Y'];
b=zeros(n,1);
LB=zeros(m+s,1);UB=[];
for i=1:n
    f=[zeros(1,m) -Y(:,i)'];
    Aeq=[X(:,i)' zeros(1,s)];beq=1;
    w(:,i)=linprog(f,A,b,Aeq,beq, LB, UB);%解线性规划，得最佳权向量
    E(i,i)=Y(:,i)'*w(m+1:m+s,i);
end
w
E
omega=w(1:m,:)%输出投入权向量
mu=w(m+1:m+s,:)%输出产出权向量

```

附录十三：第二问数据包络适应度分析所用 matlab 程序

```
%初始化
clc,clear
%此程序探讨对经济发展的影响
%导入数据 2015 至 2017 年四个影响指标的数据为输入矩阵 X
X=[584.7 550.4 536.8
    164.4 173.1 175.8
    62 58.5 63
    221217 217852 227069];
%GDP 和区域税收作为输出矩阵 Y
Y=[1647.6 1796.9 2011
    568.3 662.8 759.4];
%%多指标输入与多指标输出 考虑数据包络分析
[m,n]=size(X);%读入输入矩阵大小
s=size(Y,1);%输出变量类数
epsilon=10^-10;%定义非阿基米德无穷小
f=[zeros(1,n) -epsilon*ones(1,m+s) 1];
A=zeros(1,n+m+s+1,1);b=0;
LB=zeros(n+m+s+1,1);UB=[];
LB(n+m+s+1)=-inf;
for i=1:n
    Aeq=[X eye(m) zeros(m,s) -X(:,i)
        Y zeros(s,m) -eye(s) zeros(s,1)];
    beq=[zeros(m,1)
        Y(:,i)];
    w(:,i)=linprog(f,A,b,Aeq,beq,LB,UB);%解线性规划，得最佳权向量
end
w %输出最佳权变量
lambda=w(1:n,:) %输出.*
s_minus=w(n+1:n+m,:) %输出 s*+
s_plus=w(n+m+1:n+m+s,:) %输出 s*-
theta=w(n+m+s+1,:) %输出 .*
```

附录十四：灰色预测函数所用 matlab 程序

```
function [X,c,error1,error2]=GM11(X0,k)
% 其中 X0 为输入序列，k 为预测长度，
% X 为预测输出序列，c 为后验差检验数，error1 为残差，error2 为相对误差
format long;
n=length(X0);
X1=[];
X1(1)=X0(1);
for i=2:n
    X1(i)=X1(i-1)+X0(i); %计算累加生成序列
end
for i=1:n-1
```

```

B(i,1)=-0.5*(X1(i)+X1(i+1));    %计算 B, Yn
B(i,2)=1;
Y(i)=X0(i+1);
end
alpha=(B'*B)^(-1)*B'*Y';        %做最小二乘估计
a=alpha(1,1);
b=alpha(2,1);
d=b/a;                            %计算时间响应函数参数
c=X1(1)-d;
X2(1)=X0(1);
X(1)=X0(1);
for i=1:n-1
X2(i+1)=c*exp(-a*i)+d;
X(i+1)=X2(i+1)-X2(i);            %计算预测序列
end
for i=(n+1):(n+k)
X2(i)=c*exp(-a*(i-1))+d;         %计算预测序列
X(i)=X2(i)-X2(i-1);
end
for i=1:n
error(i)=X(i)-X0(i);
error1(i)=abs(error(i));         %计算残差
error2(i)=error1(i)/X0(i);      %计算相对误差
end
c=std(error1)/std(X0);           %计算后验差检验数
附录十五：灰色预测开放度所用 matlab 程序
%初始化
%此程序计算开放度
clear,clc
%导入数据
[data,text]=xlsread('beijing.xlsx');
X=data(8,[1,2,3]);
k=3;%预测数据个数
%%灰色预测
[X1,c,error1,error2]=GM11(X,k);
X=data(8,:);
%绘制预测值与实际值图像
figure(1)
plot(2012:2017,X1,'b:', 'linewidth',4)
hold on
plot(2012:2017,X,'r', 'linewidth',4)
legend('预测值','实际值');
%用 A 来表示开放度程度大小
A=abs(X-X1).^2;

```

```

%绘制开放度程度的柱状图
figure(2)
bar(2012:2017,A)
legend('开放程度')
附录十六：第二问特点分析所用 matlab 程序
%初始化
clear,clc
%此程序绘制各变量图像，分析特点
[data,text]=xlsread('beijing.xlsx');%导入数据
%开始绘制图像
figure(1)
bar(data(1,:),data(2,:));hold on
bar(data(1,:),data(3,:));hold on
bar(data(1,:),data(4,:));hold on;
legend('建筑业','交通运输.仓储.邮政','住宿餐饮');
figure(2);
bar(data(1,:),data(7,:));
legend('就业人数');
figure(3);
bar(data(1,:),data(8,:));hold on;
bar(data(1,:),data(9,:));
legend('GdP','区域税收');
%计算产业在 GDP 中所占比例
for i=1:6
    zb([1,2,3],i)=data([2,3,4],i)./data(8,i);
end
%绘制比例图像
figure(4);
bar(data(1,:),zb(1,:));hold on;
bar(data(1,:),zb(2,:));hold on
bar(data(1,:),zb(3,:));
legend('建筑业 gdp 占比','交通运输 gdp 占比','住宿 gdp 占比');%建筑业占比减小，
交通运输与住宿占比波动。
附录十七：泰尔指数所用 matlab 程序
function [Tb,Tw,T,z,Db,Dw]=TbTw(x,n)
%函数 TbTw() 计算泰尔指数分解
%返回 Tb 为组间差距，Tw 为组内差距，z 为各个组内差距，
%返回 T 为泰尔指数，Db 为组间贡献率，Dw 为各个组内贡献率
%泰尔指数 T=Tb+Tw
%x 为 N 个个体的收入向量，依次分为 K 个分组，n=[n1,...,nK]为各分组的个体数向量，
sum(n)=N
K=length(n);
s=[0,cumsum(n)];
for k=1:K

```

```

X{k}=x(s(k)+1:s(k+1))./sum(x); %X{k} 为第 k 个分组的 nk 个个体的收入份额(占总收入的比例)
y(k)=sum(X{k}); %y(k) 为第 k 组的收入份额(占总收入的比例)
end
Tb=sum(y.*log(y./(n./length(x)))); %组间差距
for k=1:K
    z(k)=sum((X{k}./y(k)).*log(n(k)*X{k}./y(k))); %第 k 组的组内差距
end
Tw=sum(y.*z); %总的组内差距为各分组组内差距的加权和
T=Tb+Tw; %泰尔指数
Db=Tb/T; %组间贡献率
Dw=y.*z/T; %各分组内的组内贡献率
附录十四：第三问差异性分析所用 matlab 程序
%初始化
clc,clear
%读入数据
[data,text]=xlsread('问题 3.xlsx');
x1=data(2:7,:); %试点地数据
x0=data(9:end,:); %非试点地数据
X=[x1;x0]; %将所有市的数据放在一起
[m,~]=size(X); %记录地区数量
%%灰色预测
%对每个地区都进行预测
k=1;%预测数据个数 2018 年
yuce=[]; c1=[]; %初始化
for i=1:m
    X1=X(i,:); %提取第 i 个省份历年 GDP
    [X2,c,error1,error2]=GM11(X1,k);
    c1(i)=c; %后验差检验数
    yuce(i,:)=X2; %记录预测值
end
%%对每个年份进行泰尔指数
[~,n1]=size(yuce); %计算年份数
n=[6,6]; %试点地与非试点地个数
for i=1:n1
    X0=yuce(:,i); %提取每年的数据
    [Tb(i),Tw(i),T(i),z(i,:),Db(i),Dw(i,:)]=TbTw(X0,n);
end
%%使用 excel 画出图线
%绘制比例图
T=z(:,1)'+Tb;
Y=[Tb./T;z(:,1)'./T];
pie(Y(:,end)); %画出 2018 年饼状图
title('2018')

```

```
legend('试点与非试点组间差距','试点组内差距');
```