社会融资规模与区域经济发展实证分析

组号: 第 23 组

孔雨佳 学号: 7042621136

戴昀 学号: 7013320010

学校: 南昌大学科学技术学院

摘要

随着金融创新和金融体制改革的不断推进,发现社会融资规模已经逐步成为了解金融对实体经济的支持作用的重要指标。为了分析社会融资规模与区域经济增长,本文将建立多元线性回归模型来确定社会融资规模对区域经济增长实例的关系。

针对问题一,分析区域社会融资规模与经济增长之间的关系。由权威机构[□]得知,融资结构可分为四大类。我们将附件 1 中所给的融资规模指标分为:金融机构表内业务融资,金融机构表外业务融资,直接融资与其他融资,四大融资规模由其子融资规模之和所得。首先我们将 2017-2021 年各地区的社会融资规模与融资结构进行正态检验,检验结果都符合正态分布,正态概率图见模型建立与求解。然后将融资结构与社会融资规模进行相关性检验,得出的 Pearson 相关系数见模型建立与求解,相关系数表示社会融资规模与融资结构相关程度高。于是我们建立了多元线性回归模型来确定社会融资规模与融资结构的关系,模型检验得出拟合优度,残差个案次序图表示回归效果非常好。说明区域社会融资规模与融资结构之间有线性关系。

针对问题二,分析江西省社会融资规模与经济增长之间的关系。由附件 2 给出 2017-2021 年主要经济长指标,我们首先分析这些经济增长指标,将这些指标绝对值作 为数量型经济增长指标,增长率作为质量型经济增长指标。再把江西省 2017-2021 年的社会融资规模增量作为母序列,将经济增长指标作为子序列,分别进行灰色关联度分析,得到关联度高的经济增长指标。

将所得出的 11 项数量型和质量型经济增长指标进行 KMO 检验,然后对其进行主成分分析。得出主成分载荷系数都在 0.931-0.997 之间,累计方差解释率都在 99.971%-100%之间,所以这 11 项经济增长指标都是主成分。

结合上述的数量型指标和质量型指标,推出一个年度经济增长算数均值指标来代替数量型和质量型指标反应经济增长。将经济增长年度均值指标进行正态检验,再对社会融资规模增量与年度均值指标进行相关性检验,得出 Pearson 相关系数见模型建立与求解,根据相关系数得知社会融资规模与经济增长指标有很好的相关性,但是如果将经济增长年度均值指标与社会融资规模增量做回归分析发现拟合过度现象。

于是为了解决拟合过度现象。我们将年度均值指标分为符合金融规律的6类,分别为:地区生产总值,财政金融,消费市场,对外经济,金融机构人民币存贷款余额和居民人均可支配收入。再将这6类对社会融资规模增量进行回归分析,发现这6类年度均值指标均与社会融资规模增量可以有线性关系。所以建立多元线性回归方程组模型来确定社会融资规模增量与经济增长的关系。

进一步考虑到该模型中回归方程都是线性的,我们将以回归指标的关联度为权重,将多元线性回归方程组加权相加的到一个多元线性回归方程来确定社会融资规模额增量与经济增长的关系。说明社会融资规模增量与经济增长之间有线性关系。

针对问题三,根据前两问题的数学模型得出的数学结论,江西省经济经济发展应该做到如下几点:强化市场资源配置功能;扩大债券市场规模;进一步发挥票据服务实体经济优势;保持货币信贷的合理增长;对于进出口贸易,优化产业结构,提高企业自身效益,加大竞争力;出台更多适应居民的消费政策,进一步优化消费结构;要实施乡村振兴战略,推动城乡一体产业融合;发挥江西省的三大产业的推动作用增加江西省的GDP的增长;

关键字: 多元线性回归, Pearson 相关系数, 灰色关联度, 主成分分析

一、 问题重述

近年来,金融体系对我国实体经济的推动作用愈发凸显。2011 年我国宏观调控引入一个新的观测指标:社会融资规模。社会融资规模全面概括、衡量金融体系对实体经济发展贡献能力的总量指标。基于目前我国社会融资规模的发展现状,讨论社会融资规模会对我国经济增长产生怎样的影响,通过所给数据运用数学建模知识来解决以下问题:

- (1) 融资结构是从融资规模概念演变出来的,从结构角度来衡量金融体系内各部分对实体经济的贡献力度的指标。根据附件 1 中我国 2017-2021 年不同地区的不同的社会融资规模建立模型分析区域社会融资规模与结构的关系。
- (2) 社会融资规模对经济增长的影响效应表现也存在差异性。结合附件 2 中江西省 2017-2021 年主要经济指标,建立模型分析社会融资规模与经济增长之间的关系
 - (3) 基于问题一和问题二的分析结果,对江西省的经济发展提出意见和建议。

二、 问题分析

2.1 对于问题一的分析

由权威机构国家统计局可知,社会融资规模由间接融资、直接融资、表外融资和其他融资四大融资结构组成,而问题一中所给融资方式并无属于其他融资结构。

因此我们将先对历年各地区社会融资规模增量与间接融资、直接融资和表外融资这三大融资结构进行正态分布检验分析。再对社会融资规模数据及其间接融资、直接融资和表外融资这三项数据之间进行相关性检验,根据其 Pearson 相关系数得出社会融资规模与其三大组成部分的相关程度。根据所给的地区历年社会融资规模增量与融资结构表所做出的散点图,发现可以有线性关系,所以以间接融资、直接融资和表外融资为自变量,各地区社会融资规模增量为因变量进行多元线性回归分析。

得出如下线性回归分析结果: 拟合散点图, 残差个案次序图以及拟合优度, 作为判断依据。来判断社会融资规模与间接融资、直接融资和表外融资这三者之间是否存在显著的线性相关关系, 基于以上回归分析确定社会融资规模增量与融资结构之间的关系。

2.2 对于问题二的分析

针对社会融资规模与经济增长之间的关系。已知附录 2 中给出江西省历年生产总值(GDP)到农村居民可支配收入这 17 项经济指标的数量型指标绝对值和质量型指标增长率表示江西省历年的经济增长状况。

所以将所给的 17 项经济指标中的数量指标和质量指标与社会融资规模分别进行灰色关联度分析,得出关联度高的指标。首先将数量型 17 项指标中所给出的除规模以上工业增加值和固定资产投资这两项缺失数据的数量指标以外的 15 项数量指标进行灰色关联度分析,得出这 15 项经济指标的数量指标与社会融资规模灰色关联度值,将其进

行关联度排序。接下来又将质量型 17 项经济指标中所给出的除居民消费价格指数,工业生产者出厂价格指数和工业生产者购进价格指数这三项缺失数据的质量指标以外的 14 项质量指标进行灰色关联度分析,得出这 14 项经济指标的质量指标与社会融资规模 灰色关联度值,将其进行关联度排序。

综合数量型和质量型经济指标与社会融资规模增量的关联度排序结果,得到关联度高的数量型和质量型经济增长指标。

将所得出关联度高的数量型和质量型经济增长指标进行 KMO 检验,然后对其进行主成分分析。得出主成分载荷系数,判断关联度高的数量型和质量型经济指标中隐变量的重要性。得出对于社会融资规模增量主要的数量型和质量型经济指标来反应经济增长。

结合得出的数量型指标和质量型指标,推出一个年度经济增长算数均值指标来代替数量型和质量型指标反应经济增长。将经济增长年度均值指标进行正态检验,再对社会融资规模增量与年度均值指标进行相关性检验,根据 Pearson 相关系数得出年度均值指标对社会融资规模的相关程度。

如果将经济增长年度均值指标与社会融资规模增量做回归分析发现拟合过度显现即 $R^2=1$ 。

于是为了解决拟合过度现象。我们将年度均值指标分为符合金融规律的6类,再将这6类对社会融资规模增量进行回归分析,发现这6类年度均值指标均与社会融资规模增量可以有线性关系。所以建立多元线性回归方程组模型来确定社会融资规模增量与经济增长的关系。

进一步考虑到该模型中回归方程都是线性的,我们将以回归指标的关联度为权重,将多元线性回归方程组加权相加的到一个多元线性回归方程来确定社会融资规模额增量与经济增长的关系。



2.3 对于问题三的分析

根据问题一所做的回归分析得出社会融资规模与三大融资结构的回归方程可以得出三大融资结构对社会融资规模的贡献程度作为发展方向。再根据问题二所得出的江西省的社会融资规模与江西省经济增长指标的回归方程可以得到6类经济增长指标与社会融资规模的回归系数作为经济增长指标对社会融资规模的贡献程度作为发展方向。

三、 问题的假设

- 1、解释变量是非随机的或固定的,且相互之间互不相关, 随机扰动项具有 0 均值、同方差以及不存在序列相关。
- 2、解释变量与随机扰动项不相关,随机扰动项服从正态分布, 样本容量趋于无穷时, 各解释变量的方差趋于有界常数。

四、 符号说明

符号	说明
у	区域社会融资规模增量
x_1	金融机构表内业务融资金额
x_2	金融机构表外业务融资金额
x_3	直接融资金额
$x_{j \times i}$	融资结构数据矩阵
Y	社会融资规模向量
$oldsymbol{eta}_{\hbar}$	社会融资规模与融资结构的回归系数
ε	社会融资规模与融资结构回归方程的
	随机误差
$A_{k,k} \in [2017,2021]$	江西省 2017 年-2021 年社会融资规模
B_l	经济指标绝对值 <i>l</i> ∈ [1,17]
c_l	综合经济指标 $c_l, l \in [1,17]$
C_m	与江西省社会融资规模关联程度高的
	综合经济指标 C_m , $m \in [1,11]$
D_l	经济指标的增长率 $D_l, l \in [1,17]$
C_{mn}	与江西省社会融资规模关联程度高的
	经济指标数据矩阵 C_{mn} , $m \in [1,11]$ $n \in [$
	[1,6]
$\alpha_{mp,}$	江西省社会融资规模与各类经济指标
	的回归系数 $\alpha_{mp,m} \in [1,11] p \in [0,6]$
μ_q	江西省社会融资规模与各类经济指标
	的回归方程的随机误差 μ_q , q ∈ [1,6]

五、 模型的建立与求解

5.1 问题一模型的建立与求解

5.1.1 数据预处理

数据矩阵 X_{ij} 其中j=1,2,...,40分别对应表中从北京到新疆得各个区域,i=1,2,3分别对应三类融资结构。

为后续模型的建立,需对数据进行差异性检验,而常用的参数检验需要数据服从正态分布。因此我们先对 y 与 x_i (i=1, 2, 3, 4)用 MATLAB 进行正态检验。

得到其符合正态分布,再用 MATLAB 对 y 与 x_i 实现皮尔逊分析,得到两个变量之间的相关性。

5.1.2 模型的建立

1. 建立多元线性回归模型

以融资结构 x_{nm} 为自变量社会融资规模为因变量y,建立多元线性回归模型。

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$
 (1.1)

式中: $\beta_0\beta_1\beta_2\beta_3$ 为回归系数; ε 为随机误差,符合正态分布 N $(0, \sigma^2)$, σ 未知。由附件 1 中 2017-2021 年社会融资规模增量统计表得到

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$
 $\varepsilon = [\varepsilon_1 \ \varepsilon_2 & \cdots & \varepsilon_n]^T, \beta = [\beta_0 \ \beta_1 & \cdots & \beta_m]^T$ 式 (1.1) 可以表示为 $\begin{cases} Y = X\beta + \varepsilon \\ \varepsilon \sim N(0, \sigma^2 E_n) \end{cases}$

式中, E_a 表示为n阶单位矩阵

模型式(1.1)中参数 β_0 , β_1 ,…, β_m 用最小二乘法估计,即应选取估计值 b_j ,使当 $\beta_i = b_i, j = 0, 1, 2, \dots, m$ 时,误差平方和

$$Q = \sum_{i=1}^n arepsilon_i^2 = \sum_{i=1}^n (y_i - eta_0 - eta_1 x_{i1} - \dots - eta_m x_{im})^2$$
达到最小。

为此, 令
$$\frac{\alpha Q}{\alpha \beta_j} = 0, j = 0, 1, 2, \dots, m$$
. 得

$$egin{cases} rac{lpha Q}{lpha eta_0} = -2\sum_{i=1}^n (y_i - eta_0 - eta_1 x_{i1} - \cdots - eta_m x_{im}) = 0 \ rac{lpha Q}{lpha eta_j} = -2\sum_{i=1}^n (y_i - eta_0 - eta_1 x_{i1} - \cdots - eta_m x_{im}) x_{ij} = 0 \,, \; j = 1 \,, 2 \,, \cdots, m. \end{cases}$$

方程组的矩阵形式为

$$X^T X \beta = X^T Y \tag{1.2}$$

当矩阵 X 列满秩时, X^TX 为可逆方阵,式 (1.2) 的解为

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

将 $\hat{\beta} = [b_0, b_1, \dots, b_m]$ 代入式(1.1),得到y的估计值为

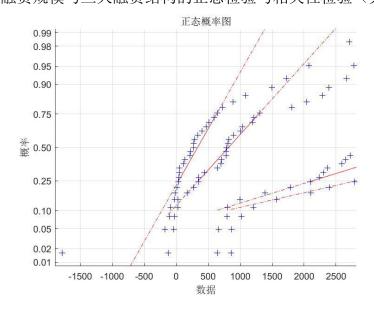
$$\hat{y} = b_0 + b_1 x_1 + \dots + b_m x_m$$
 .

而这组数据的拟合值为 $\hat{Y} = X\hat{\beta}$,拟合误差 $e = Y - \hat{Y}$ 称为残差,可作为随机误差 ε 的估

计。SSE=
$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$
为残差平方和(或剩余平方和)。

5.1.3 模型求解

1. 2017年社会融资规模与三大融资结构的正态检验与相关性检验(其他年份见附录)

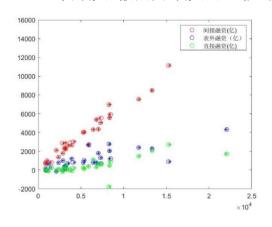


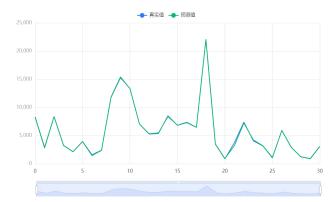
社会融资规模与各项融 资指标的 Pearson 相关性检验	假设性检验值	相关系数
结果		
社会融资规模	1	1
间接融资	0.000005	0.9879
表外融资	0.0007	0.7987
直接融资	0.1808	0.6892

3. 用 MATLAB 对多元回归方程的回归系数与拟合优度求解

	** / = / 1/ 1/ * 1	
年份	回归系数	拟合优度
2017	$\beta_0 = 35.3768 \ \beta_1 = 1.0618 \ \beta_2 = 0.9883 \ \beta_3 = 0.957$	0.99926
2018	$\beta_0 = 198.2513 \ \beta_1 = 0.9006 \ \beta_2 = 1.7805 \ \beta_3 = 1.1655$	0.99598
2019	$\beta_0 = -9.8033 \ \beta_1 = 0.973 \ \beta_2 = 1.3584 \ \beta_3 = 0.8685$	0.99764
2020	$\beta_0 = 71.1239 \ \beta_1 = 0.9949 \ \beta_2 = 1.1962 \ \beta_3 = 0.8855$	0.99938
2021	$\beta_0 = 86.5792 \ \beta_1 = 0.9207 \ \beta_2 = 1.3192 \ \beta_3 = 0.7597$	0.99937

4.2017年的拟合散点图与拟合曲线如下(其他年份拟合散点图见附录)





5.2 问题二的模型的建立与求解

5.2.1 数据预处理

1. 江西省 2017-2021 年社会融资规模 $A_{k,}$ $k \in [2017,2021]$,经济指标绝对值 $B_{l}, l \in [1,17]$,l为别为附件 2 中从生产总值依次到农村居民可支配收入的全省主要经济指标,如 l_1 代表江西省生产总值, l_{17} 为农村居民可支配收入。年度算数均值经济增长指标 $C_{l},\ l \in [1,17]$ 。

与江西省社会融资规模与经济增长指标的灰色关联度关联程度高且为主成分因子的综合经济指标 C_m , $m \in [1,11]$ m为指标数。与江西省社会融资规模关联程度高的经济指标数据矩阵 C_{mn} , $m \in [1,11]$ $n \in [1,6]$ n为分类数。江西省社会融资规模与各类经济指标的回归系数 α_{mp} , $m \in [1,11]$ $p \in [0,6]$ 江西省社会融资规模与各类经济指标的回归方

程的随机误差 μ_a , $q \in [1,6]$ 经济指标的增长率 D_l , $l \in [1,17]$

2. 将表中数量型指标与质量型指标数据综合处理。

增长率 =
$$\frac{(现期 - 基期)}{\text{基期}} \times 100\%$$

我们取第
$$l$$
年的
$$\begin{cases} C_l = \frac{B_l + B_{l-1}}{2} \\ D_l = \frac{B_l - B_{l-1}}{B_{l-1}} \times 100\% \end{cases}$$
 推出 $\frac{B_l(D_l + 2)}{2(D_l + 1)} = C_l$

 C_l 为当年的经济指标指标与去年经济指标的算数平均值,并以其为描述经济指标的综合指标。

- (1) 对各项经济指标灰色关联度分析,得到与江西省社会融资规模关联程度高的经济指标,去除关联程度低的经济指标。分析后去除经济指标 C_2 , C_4 , C_{11} , C_{12} , C_{13} ,再按顺序对留下的指标进行排序从 C_1 到 C_{11} 。由分析结果可知,社会融资规模与经济指标里的工业生产类指标相关联程度低。再用主成分分析检验分析结果。
- (2)据我国学者常用经济指标分类,将江西省社会融资规模关联程度高的经济指标分为六类:

地区生产总值: GDP (C1);

财政金融:一般公共预算收入(C2);

消费市场: 社会消费品零售总额(C3), 限额以上消费品零售额(C4);

对外经济: 进出口总值(C5),出口(C6),实际利用外商直接投资(C7);

金融人民币存贷款余额:金融机构人民币存款余额(C8),金融机构人民币贷款余额(C9);

居民人均可支配收入:城镇居民可支配收入(C10),农村居民人均可支配收人(C11)。

5.2.2 建立多元线性回归方程组模型

社会融资规模分别与每类指标进行多元线性回归。模型建立原理与问题一中一致, 这里不再赘述。第一类和第二类为一元线性回归模型。

第一类: $A_k = \alpha_{10} + \alpha_{11}C_{11} + \mu_1$ (2.1)

第二类: $A_k = \alpha_{20} + \alpha_{21}C_{2i} + \mu_2$ (2.2)

第三类: $A_k = \alpha_{30} + \alpha_{33}C_{33} + \alpha_{34}C_{34} + \mu_3$ (2.3)

第四类: $A_k = \alpha_{40} + \alpha_{45}C_{45} + \alpha_{46}C_{46} + \alpha_{47}C_{47} + \mu_4$ (2.4)

第五类: $A_k = \alpha_{50} + \alpha_{58}C_{58} + \alpha_{59}C_{59} + \mu_5$ (2.5)

第六类: $A_k = \alpha_{60} + \alpha_{610}C_{610} + \alpha_{611}C_{611} + \mu_6$ (2.6)

其中, μ_n 为江西省社会融资规模与各类经济指标的回归方程的随机误差符合正态分布 N

 $(0, \sigma^2)$, σ 未知,n 为分类数; $\alpha_{mp,m} \in [1,11] p \in [0,6)]$ 为江西省社会融资规模与各

5.2.3 模型求解

类经济指标的回归系数;

1. 灰色关联度排行表 关联度结果

评价项	关联度	排名
金融机构人民币贷款余额	0.785	1
金融机构人民币存款余额	0.774	2
进出口总值	0.756	3
社会消费品零售总额	0.755	4
实际利用外商直接投资	0.738	5
农村居民人均可支配收入	0.732	6
限额以上消费品零售额	0.732	7
出口总值	0.729	8
城镇居民人均可支配收入	0.719	9
GDP	0.704	10
一般公共预算收入	0.667	11
工业生产者购进价格指数	0.599	12
工业生产者出厂价格指数	0.58	13
居民消费价格指数	0. 528	14
财政总收入	0.527	15

根据灰色关联度排名,我们选取关联度前 11 个经济增长指标 2. 将选取的 11 项经济增长指标,主成分分析因子载荷系数表

四」栽训尔奴农		
	因子载荷系 共同度	(公因子方
	数 主成分1	
GDP	0.993	0. 986
一般公共预算收入	0.981	0.962
社会消费品零售总额	0.991	0.982
限额以上消费品零售 额	0.996	0. 993
出口总值	0.974	0.949
金融机构人民币贷款 余额	0.993	0.985
进出口总值	0. 986	0.971
实际利用外商直接投资	0.993	0.985
城镇居民人均可支面 收入	C 0. 993	0.986
金融机构人民币存款 余额	0.992	0.984

因子载荷系数表

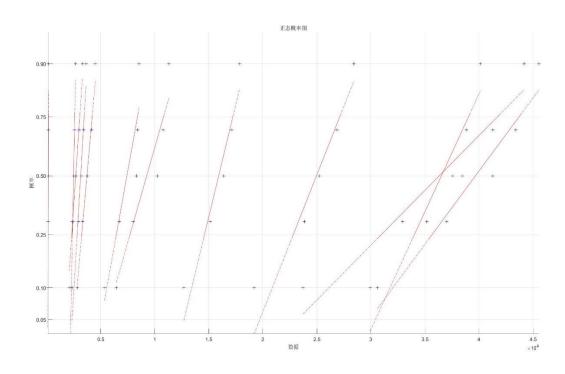
由此可见,11 项经济增长指标都为主要指标

3. 社会融资规模与六大类经济增长指标的正态性检验

农村居民人均可支配

0.997

0.993



由正态概率图可知,社会融资规模与六大类经济增长指标都符合正态分布 2.Pearson 相关系数、假设性检验值与拟合优度的求解如下表

	Pearson 相关	假设性检验	拟合优度
	系数	值	
C1 类	0.912	0.0307	0.83245
C2 类	0.9106	0.0316	0.82921
C3 类	0.9654	0.0077	0.93729
	0.9281	0.0229	
C4 类	0.8896	0.0433	0.9999
	0.8653	0.0581	
	0.9480	0.0141	
C5 类	0.9470	0.0145	0.89703
	0.9471	0.0145	
C6 类	0.9481	0.0141	0.9037
	0.9440	0.0158	

3.多元线性回归方程组的回归系数求解如下表

	回归系数
C1 类	$lpha_0 \! = \! -1.4644 lpha_1 \! = \! 0.3615$
C2 类	$lpha_0\!=\!-1.458 \; lpha_1\!=\!7.2481$
C3 类	$lpha_0 = 2.1639 lpha_1 = 0.8861 lpha_2 = -0.9549$
C4 类	$\alpha_0 = -2.0075 \ \alpha_1 = -37.7372 \ \alpha_2 = 37.7930 \ \alpha_3 = 485.8$
C5 类	$lpha_0 = 663.7611 lpha_1 = 0.0677 lpha_2 = 0.1142$
C6 类	$\alpha_0 = -7.1406 \alpha_1 = 0.8694 \alpha_2 = -1.0722$

5.3 问题三的解答

5.3.1 从前两问数学模型得到的数学结论

从问题一的求解中可知,社会融资规模和三大融资结构之间存在显著的线性相关关系,且社会融资规模与三大融资结构的相关性程度不同,由高到低分别为间接融资、表外融资和直接融资,又从问题二的求解中得到社会融资规模与反映经济增长的六大类指标同样存在显著的线性相关关系,六大类经济指标中的各项经济指标与社会融资规模关联程度由高到低分别为金融机构人民币贷款余额,金融机构人民币存款余额,进出口总值,社会消费品零售总额,农村居民人均可支配收入,限额以上消费品零售额,出口总值,城镇居民人均可支配收入,生产总值(GDP)和一般公共预算收入。基于以上的分析结果可以从融资结构和各项经济指标方面对江西省经济发展提出相应的意见和建议。

5.3.2 对江西省经济发展的意见与建议

1.由题目所提供的附件一中各地区的社会融资规模数据比较分析可知,江西省金融资源投入具有较大的空间,结合问题一分析结果,三大融资结构中的以下三个融资方式较其他融资方式存在更为明显的不足,首先与浙江、广东等东南沿海省份相比,江西省在非金融企业境内股票融资方面的投入明显不足,其次是企业债券融资,与浙江、广东等东南沿海省份差距明显,资本市场融资力度有待进一步提高,最后未贴现银行承兑汇票数据呈现负数,远远不及大部分东南沿海省份,票据业务有待进一步发展。因此,为促进江西省经济发展,提出的建议如下:

- (1) 从社会融资总量的角度,要保持合理的社会融资规模,强化市场资源配置功能,进一步提高江西省经济发展的内生动力。
- (2)进一步扩大债券市场规模,加快地方债、专项债发展,以进一步促进资本市场繁荣发展。
- (3)进一步发挥票据服务实体经济优势,综合运用票据承兑以及供应链票据等创新产品,缓解企业资金问题,助推区域经济发展。
- 2. 由题目所提供的附件二中 2017 年到 2021 年江西省全省的主要经济指标分析可以发现江西省的各项经济指标的绝对值相对较低且增速较慢,再结合对于问题二的求解结果,可以对江西省的经济发展提出以下建议:
- (1)金融经济是实体经济的助推器,主要服务于实体经济。因此要提高思想认识全面加强实体经济发展金融服务工作,把服务、发展实体经济作为金融部门自身的重要职责和使命,全面落实金融服务实体经济的各项措施。保持货币信贷的合理增长,满足实体经济合理资金需要。
- (2)属于中部地区的江西对于进出口贸易有着较差的竞争力但是江西省有着比较丰富的各类资源,所以可以立足资源优势,加大技术创新和改进,优化产业结构,提高企业自身效益,加大竞争力。
- (3)消费需求是经济增长的动力,针对社会消费品零售总额这方面,江西省要扩大江西省省内人人民的内生需求,出台更多适应居民的消费政策,进一步优化消费结构,最大限度地满足城乡居民日益增长的物质生活的需要,不断扩大消费规模,利用消费拉动江西省的经济发展。
- (4)从附件二给出的近几年城乡居民可支配收入数据看出,江西省的收入水平偏低,尤其是农村居民可支配收入,因此为推动江西省的经济,要实施乡村振兴战略,推动城乡一体产业融合,促进江西省人民充分就业,提高居民的收入,推动江西省经济发展,实现共同富裕。
 - (5) 根据江西省各年的 GDP 增长量与增长速度,为增加江西省的 GDP 的增长,

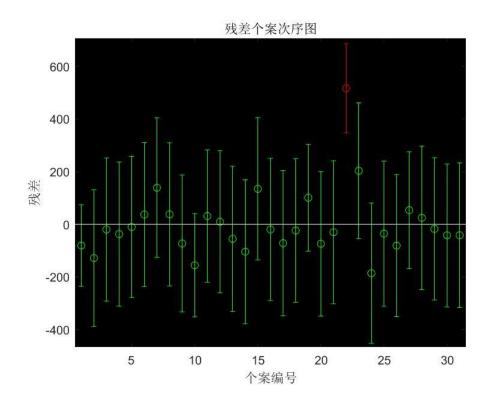
发挥江西省的三大产业的推动作用,不断调整产业结构,第一产业继续稳中求进,加强对第二产业的投入力度,扩大第三产业在 GDP 中所占的比重,利用三大产业的发展,促进江西省经济的高质量发展。

六、 模型的检验

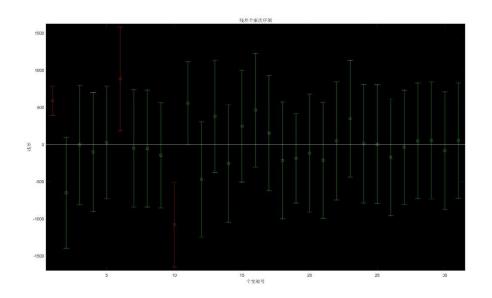
6.1 对于问题一的模型检验

对融资规模与融资结构多元线性回归的残差个案次序图

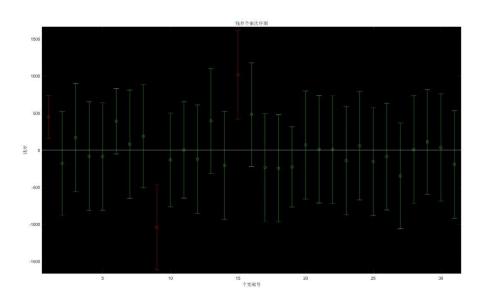
2017年:由残差个案次序图可知置信区间的绝大部分所在位置都很理想,拟合效果良好。



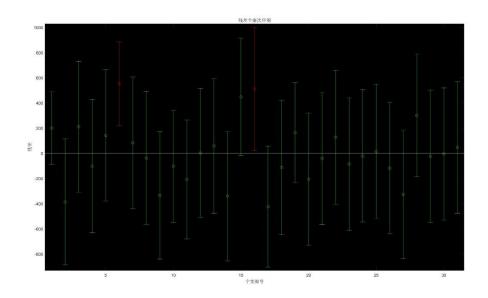
2018年:由残差个案次序图可知置信区间的绝大部分所在位置都很理想,拟合效果良好。



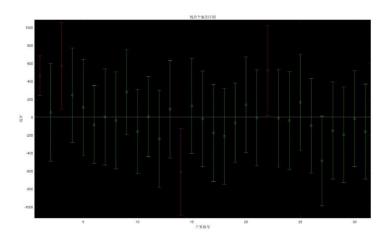
2019年:由残差个案次序图可知置信区间的绝大部分所在位置都很理想,拟合效果良好。



2020年:由残差个案次序图可知置信区间的绝大部分所在位置都很理想,拟合效果良好。

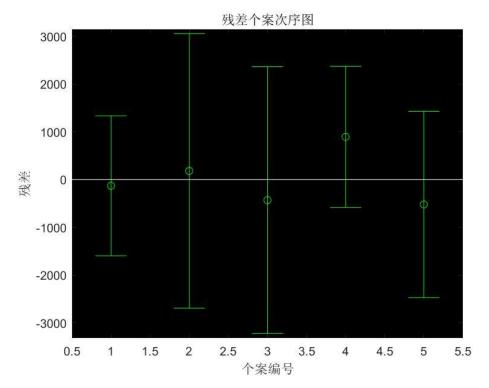


2021年:由残差个案次序图可知置信区间的绝大部分所在位置都很理想,拟合效果良好。



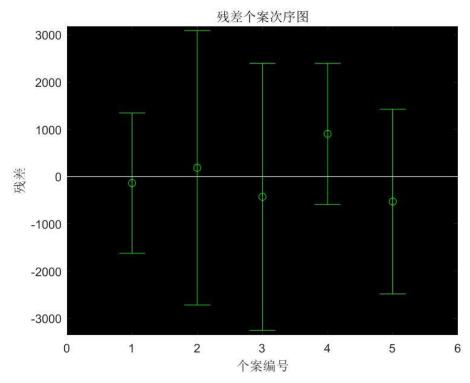
6.2 对于问题二的模型检验

对于社会融资规模与经济增长指标的多元线性回归的残差个案次序图 C1 类:



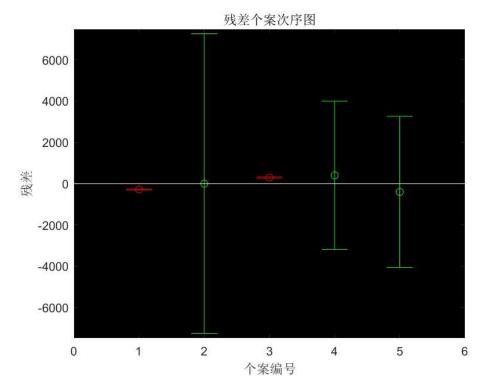
由残差个案次序图可知置信区间的所在位置都很理想,无红色线段,拟合效果优秀。

C2类:

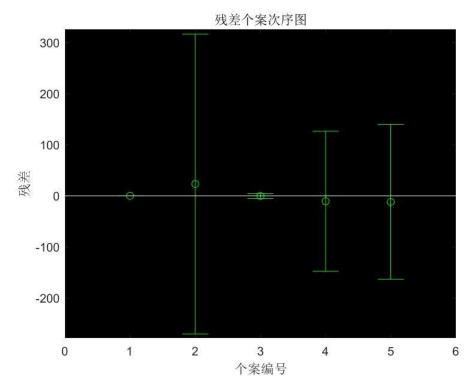


由残差个案次序图可知置信区间的所在位置都很理想,无红色线段,拟合效果优秀。

C3 类:

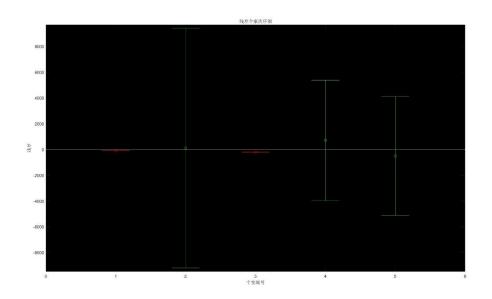


由残差个案次序图得知个别数据置信区间的位置接近理想状态,拟合效果良好。 C4类:

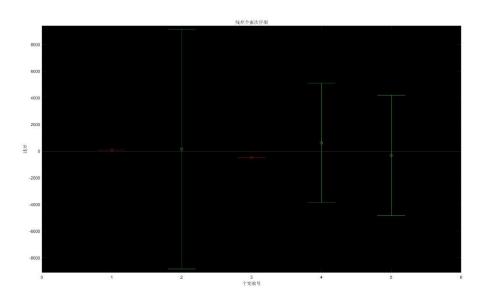


由残差个案次序图可知置信区间的所在位置都很理想,无红色线段,拟合效果优秀。

C5 类:



由残差个案次序图可知少部分数据接近理想状态,拟合效果较好。 C6类:



由残差个案次序图得知部分数据接近理想状态,拟合效果较为良好。

七、 模型的优缺点

- 一、优点:用多变量线性回归模型,通过多组数据,可直观、快速分析出数据之间的线性关系。
- 二、缺点:遇到拟合过度现象,无法直接进行多元线性回归,采用分类进行各类多元线性回归,这样描述数据间线性关系比较粗略,有待改进。其次该模型可能忽略了交互效应和非线性的因果关系,有待改进。

八、 参考文献

[1] 潘颖秋《社会融资由四大部分十个子项构成》

http://www.scio.gov.cn/xwfbh/xwbfbh/wqfbh/2014/20140115/zy30172/Document/13 60070/1360070.htm $\,$ 2022 年 11 月 28 日

九、附录

1.1 问题一关键代码求解模型

```
%% 变量说明 % M 为融资结构
```

%分为:间接融资,表外融资,直接融资,其他融资

% Sp 为各地区社会融资规模

%% 社会融资规模与融资结构变量正态检验

close all;

normplot(Sp); %社会融资规模正态检验图

hold on;

normplot(M(:,1));%间接融资正态正态检验图

normplot(M(:,2));%表外融资正态正态检验图

normplot(M(:,3));%直接融资正态正态检验图

%% Pearson 相关系数

close all;

[r,p]=corr(M,Sp);

disp(r);

disp(p);

%%多元线性回归

Y=Sp;

len=length(Y);

X=[ones(len,1),M];

 $[B,\!BINT,\!R,\!RINT,\!STATS]\!\!=\!\!regress(Y,\!X);$

%B 回归系数

%BINT 回归系数的区间估计

%R 残差

%RINT 置信空间

%STATS 检验回归模型统计量

%% 残差分析图

rcoplot(R,RINT);

%由图可知残差没有离群,回归效果良好

%%数据散点图与拟合效果

plot(Sp,M(:,1),'ro');

hold on;

plot(Sp,M(:,2),'bo');

plot(Sp,M(:,3),'go');

Y_fitting=B(1)+B(2).*M(:,1)+B(3).*M(:,2)+B(4).*M(:,3);

plot(Y_fitting,M(:,1),'r*');

plot(Y_fitting,M(:,2),'b*');

 $plot(Y_fitting,M(:,3),'g*');$

legend('间接融资(亿)','表外融资(亿)','直接融资(亿)');

 $R_2 = 1 - sum((Y_fitting - Y).^2)./sum((Y - mean(Y)).^2);$

str = num2str(R 2);

disp(['拟合优度为: ',str])

1.2 问题二关键代码求解模型

- %%变量说明
- % A 为江西省历年社会融资规模增量
- % C为11项经济增长指标
- % 把 C 分为 6 类
- % c_i 为这 6 类指标矩阵, i=1,2...6
- %地区生产总值 c 1 为
- %GDP
- % 财政金融 c 2 为
- %一般公共预算收入
- %消费市场 c_3 为
- %社会消费品零售总额 限额以上消费品零售额
- %对外经济 c 4 为
- %进出口总额 出口 实际利用外商直接投资
- % 金融机构人民币存贷款余额 c_5 为
- %金融机构人民币存款余额 金融机构人民币贷款余额
- %居民人均可支配收入 c_6
- %城镇居民人均可支配收入 镇居民人均可支配收入
- C=[c1,c2,c3,c4,c5,c6];
- %% 社会融资规模与 c 类正态检验
- %江西省社会融资规模正态检验图
- normplot(A);
- hold on;
- %c1 正态正态检验图
- normplot(c1);
- %c2 类正态检验
- normplot(c2);
- %c3 类正态检验
- normplot(c3(:,1));
- normplot(c3(:,2));
- %c4 类正态检验
- normplot(c4(:,1));
- normplot(c4(:,2));
- normplot(c4(:,3));
- %c5 类正态检验
- normplot(c5(:,1));
- normplot(c5(:,2));
- %c6 类正态检验

```
normplot(c6(:,1));
normplot(c6(:,2));
%% Pearson 相关系数
close all;
%c1
[r1,p1]=corr(c1,A);
disp('c1:');
disp(r1);
disp(p1);
%c2
[r2,p2]=corr(c2,A);
disp('c2:');
disp(r2);
disp(p2);
%c3
[r3,p3]=corr(c3,A);
disp('c3:');
disp(r3);
disp(p3);
%c4
[r4,p4]=corr(c4,A);
disp('c4:');
disp(r4);
disp(p4);
%c5
[r5,p5]=corr(c5,A);
disp('c5:');
disp(r5);
disp(p5);
%c6
[r6,p6]=corr(c6,A);
disp('c6:');
disp(r6);
disp(p6);
%%多元线性回归
len=length(A);
C1=[ones(len,1),c1];
C2=[ones(len,1),c2];
C3=[ones(len,1),c3];
C4=[ones(len,1),c4];
C5=[ones(len,1),c5];
C6=[ones(len,1),c6];
[B1,BINT1,R1,RINT1,STATS1]=regress(A,C1);
[B2,BINT2,R2,RINT2,STATS2]=regress(A,C2);
```

```
[B3,BINT3,R3,RINT3,STATS3]=regress(A,C3);
[B4,BINT4,R4,RINT4,STATS4]=regress(A,C4);
[B5,BINT5,R5,RINT5,STATS5]=regress(A,C5);
[B6,BINT6,R6,RINT6,STATS6]=regress(A,C6);
%B 回归系数
%BINT 回归系数的区间估计
%R 残差
%RINT 置信空间
%STATS 检验回归模型统计量
%% 残差分析图
rcoplot(R1,RINT1);
rcoplot(R2,RINT2);
rcoplot(R3,RINT3);
rcoplot(R4,RINT4);
rcoplot(R5,RINT5);
rcoplot(R6,RINT6);
%%回归方程组
%Y 为拟合值, x 为各类经济指标
close all;
Y_1=B1(1)+B1(2).*C1(:,2);
Y_2=B2(1)+B2(2).*C2(:,2);
Y_3=B3(1)+B3(2).*C3(:,2)+B3(3).*C3(:,3);
Y 4=B4(1)+B4(2).*C4(:,2)+B4(3).*C4(:,3)+B4(4).*C4(:,4);
Y_5=B5(1)+B5(2).*C5(:,2)+B5(3).*C5(:,3);
Y 6=B6(1)+B6(2).*C6(:,2)+B6(3).*C6(:,3);
R_1 = 1 - sum((Y_1 - A).^2)./ sum((A - mean(A)).^2);
str1 = num2str(R_1);
R_2 = 1 - sum((Y_2 - A).^2)./ sum((A - mean(A)).^2);
str2 = num2str(R 2);
R = 1 - sum((Y = 3 - A).^2)./ sum((A - mean(A)).^2);
str3 = num2str(R_3);
R = 1 - sum((Y = 4 - A).^2)./ sum((A - mean(A)).^2);
str4 = num2str(R_4);
R_5 = 1 - sum((Y_5 - A).^2)./sum((A - mean(A)).^2);
str5 = num2str(R 5);
R_6 = 1 - sum((Y_6 - A).^2)./ sum((A - mean(A)).^2);
str6 = num2str(R 6);
disp(['c1 类拟合优度为: ',str1])
disp(['c2 类拟合优度为: ',str2])
disp(['c3 类拟合优度为: ',str3])
disp(['c4 类拟合优度为: ',str4])
disp(['c5 类拟合优度为: ',str5])
disp(['c6 类拟合优度为: ',str6])
%权重 q
```

 $q1\!=\!0.093; q2\!=\!0.007; q3\!=\!0.164; q4\!=\!0.282; q5\!=\!0.185; q6\!=\!0.2;$

Y_1=B1(1)+B1(2).*C1(:,2);

Y_2=B2(1)+B2(2).*C2(:,2);

Y_3=B3(1)+B3(2).*C3(:,2)+B3(3).*C3(:,3);

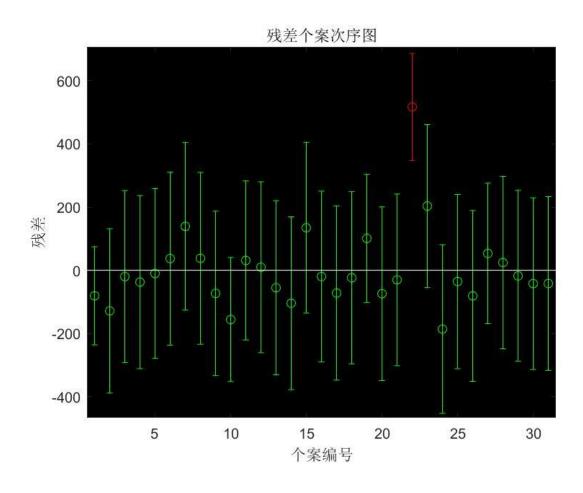
 $Y_4 = B4(1) + B4(2).*C4(:,2) + B4(3).*C4(:,3) + B4(4).*C4(:,4);$

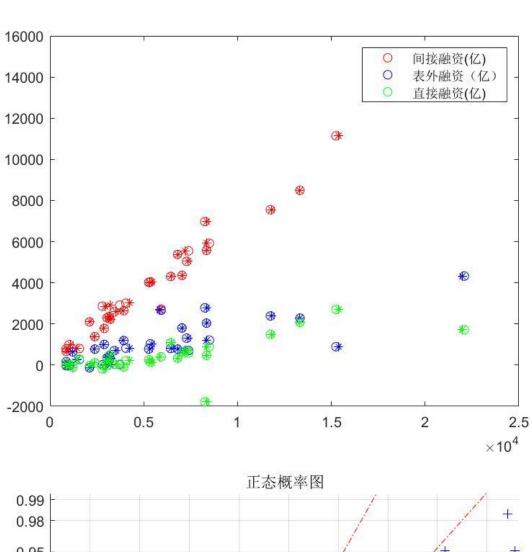
 $Y_5 = B5(1) + B5(2).*C5(:,2) + B5(3).*C5(:,3);$

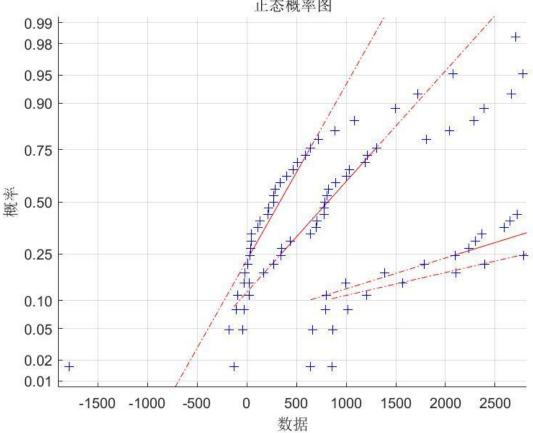
 $Y_6 = B6(1) + B6(2).*C6(:,2) + B6(3).*C6(:,3);$

1.3 问题一的各个年份模型求解结果

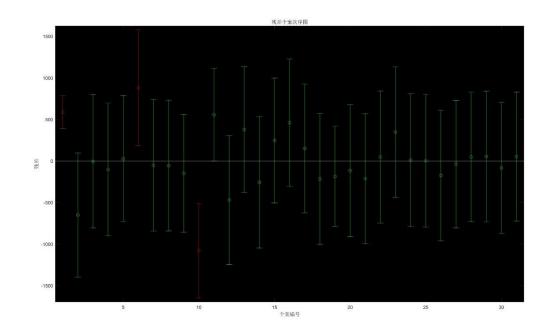
1.3.1 2017年

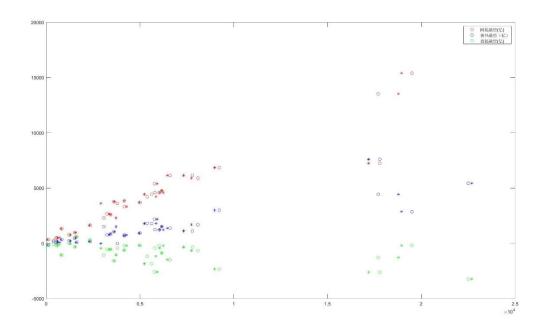


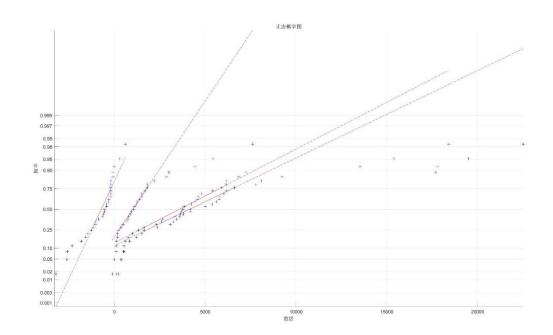




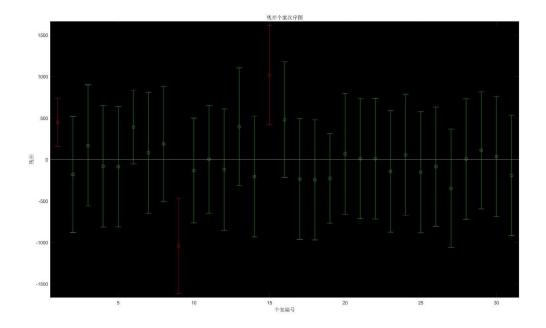
1.3.2 2018年

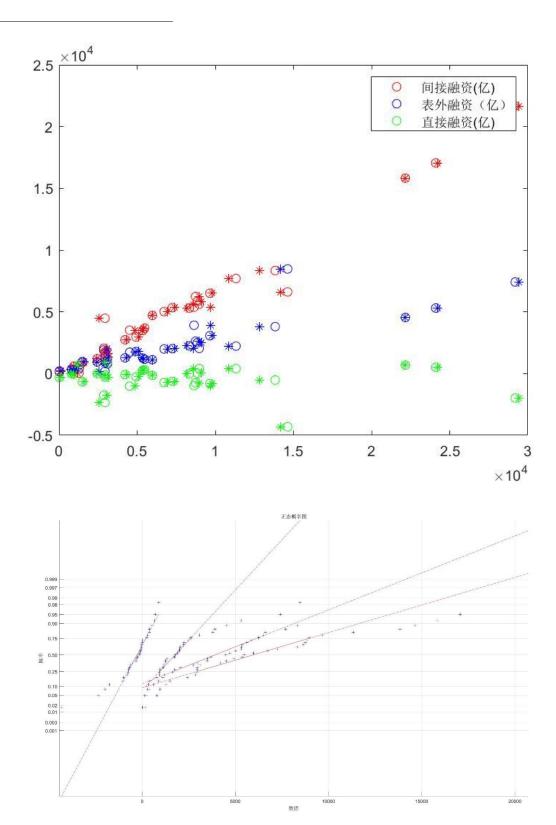




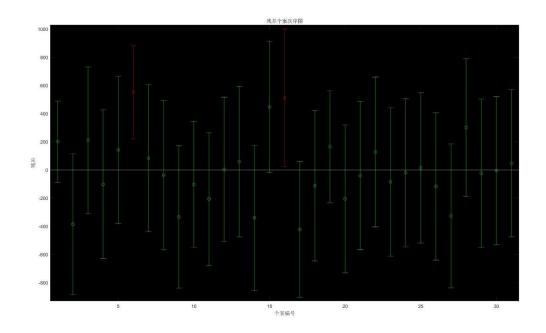


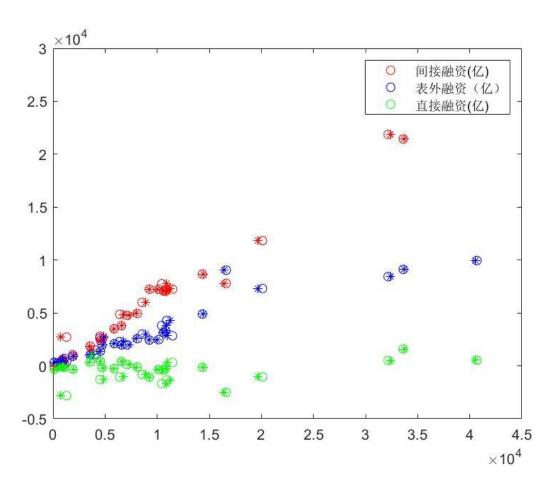
1.3.3 2019年

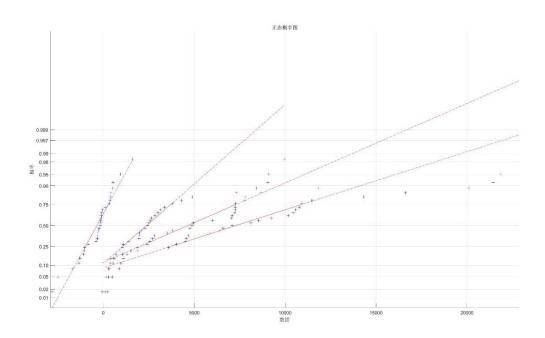




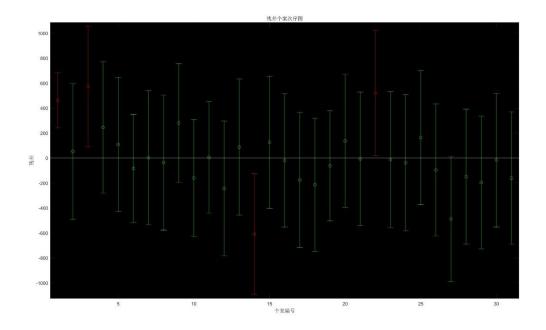
1.3.4 2020年

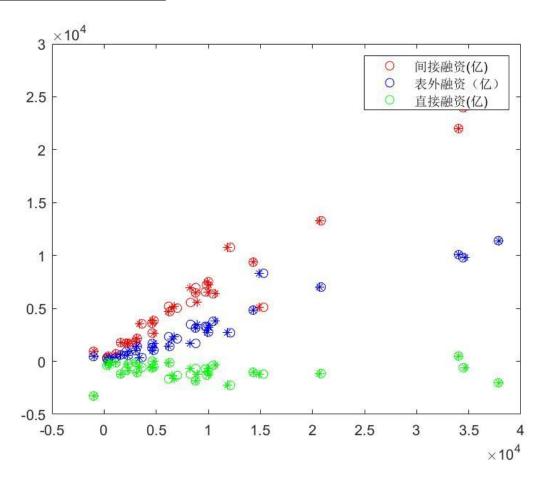


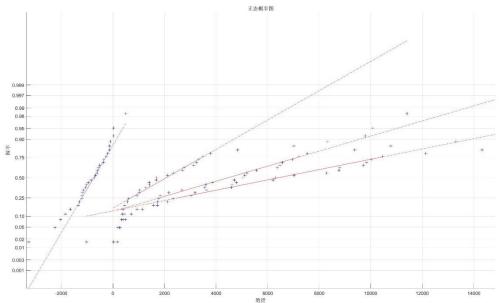




1.3.5 2021年







1.4 对问题二的数据处理细节

1.4.1 灰色关联度细节

输出结果 1: 灰色关联系数

关联系数结果

数数 资 余额 余额 收入 收入 0. 0. 0. 0.3 0.3 0.3 0. 0 $711967\ 895179\ 493891\ 887237\ 799846\ 826327\ 804311\ 632517\ 649264\ 363168\ 401558\ 730016\ 760886\ 592078\ 656990$ 953345 057614 058856 626340 444279 517470 553585 452186 289671 548469 287912 566671 142554 877435 669461 1 761776 568485 697448 771748 839052 937195 730424 779022 182190 111364 495055 104721 013355 133927 612169 355589 787661 041753 669165 888630 864893 985582 290021 192593 870443 262169 018275 179706 7976 744 88 789 3107 5842 5562 5899 0. 0. 0.6 0.7 0.7 0. 0. 0. 0. 0. 9128 968 2795 6791 3368 8294 0.4 0. 0. 0. 0. 707685 180765 681511 795064 814817 877443 787455 732036 038448 686397 894067 823382 920299 700718 747518 389085 966887 564554 460507 771884 735482 643627 701874 814536 676913 265388 131263 619510 279712 150671 02 540317 841910 507341 649319 532334 538392 530319 596482 401293 496581 359559 631221 684463 577300 585208 099778 120545 419698 373094 989344 008262 920219 870388 713362 714143 256493 911120 633632 281817 222638 1778 283 3992 7062 8027 147 8364 5204 396 251 2073 8908 1903 7857 8452 0 0.3 0. 0. 0. 0. 0. 0. 0.5 0.7 0.8 0. 0 0 772881 388898 982572 552953 645849 552369 525376 967765 746357 773440 537195 845419 620052 994102 889892 008884 685351 762642 191028 084152 335310 594828 023591 601130 130341 666187 928253 129631 407255 890360 1 8651 736 1719 6787 6693 8886 6836 1618 986 686 817 3676 7771 2025 8098

图表说明:

以上表格为预览结果,全部数据请点击下载按钮导出。

关联系数代表着该子序列与母序列对应维度上的关联程度值(数字越大,代表关联性越强)。

智能分析:

从上表可知,针对 15 个评价项(GDP(PCA)、财政总收入、(PCA)一般公共预算收入、(PCA)社会消费品零售总额、(PCA)限额以上消费品零售额、(PCA)进出口总值、(PCA)出口总值、(PCA)实际利用外商直接投资、居民消费价格指数、工业生产者出厂价格指数、工业生产者购进价格指数、(PCA)金融机构人民币存款余额、(PCA)金融机构人民币贷款余额、(PCA)城镇居民人均可支配收入、(PCA)农村居民人均可支配收入)以及 5 项数据进行灰色关联度分析,并且以社会融资规模作为"参考值"(母序列),研究 15 个评价项(GDP(PCA)、财政总收入、(PCA)一般公共预算收入、(PCA)社会消费品零售总额、(PCA)限额以上消费品零售额、(PCA)进出口总值、(PCA)出口总值、(PCA)实际利用外商直接投资、居民消费价格指数、工业生产者出厂价格指数、工业生产者购进价格指数、(PCA)金融机构人民币存款余额、(PCA)金融机构人民币贷款余额、(PCA)城镇居民人均可支配收入、(PCA)农村居民人均可支配收入与社会融资规模的关联关系(关联度),并基于关联度提供分析参考,使用灰色关联度分析时,分辨系数取 0.5,结合关联系数计算公式计算出关联系数值,并根据关联系数值,然后计算出关联度值用于评价判断。

PS: 分辨系数 $\rho \in (0, \infty)$, ρ 越小,分辨力越大,一般 ρ 的取值区间为 (0, 1), 具体取值可视情况而定。当 $\rho \leq 0.5463$ 时,分辨力最好,通常取 $\rho = 0.5$ 。

输出结果 2: 关联系数图



图表说明:

关联系数代表着该子序列 GDP(PCA)、财政总收入、(PCA)一般公共预算收入、(PCA)社会消费品零售总额、(PCA)限额以上消费品零售额、(PCA)进出口总值、(PCA)出口总值、(PCA)实际利用外商直接投资、居民消费价格指数、工业生产者出厂价格指数、工业生产者购进价格指数、(PCA)金融机构人民币存款余额、(PCA)金融机构人民币贷款余额、(PCA)城镇居民人均可支配收入、(PCA)农村居民人均可支配收入对与母序列对应维度上的关联程度值(数字越大,代表关联性越强)。

输出结果 3: 灰色关联度

关联度结果		
评价项	关联度	排名
(PCA) 金融机构人民币贷款余额	0.785	1
(PCA) 金融机构人民币存款余额	0.774	2
(PCA) 进出口总值	0.756	3
(PCA) 社会消费品零售总额	0.755	4
(PCA) 实际利用外商直接投资	0.738	5
(PCA) 农村居民人均可支配收入	0.732	6
(PCA) 限额以上消费品零售额	0.732	7
(PCA) 出口总值	0.729	8
(PCA) 城镇居民人均可支配收入	0.719	9
GDP (PCA)	0.704	10
(PCA) 一般公共预算收入	0.667	11
工业生产者购进价格指数	0.599	12
工业生产者出厂价格指数	0.58	13
居民消费价格指数	0.528	14
财政总收入	0.527	15

关联度表示各评价项与"参考值"(母序列)之间的相似关联程度,其是由关联系数进行计算平均值得出,关联度值介于 0~1 之间,该值越大表示评价项与"参考值"(母序列)相关性越强,关联度越高,意味着评价项与"参考值"(母序列)之间关系越紧密,因而其评价越高。结合关联度值,针对所有评价项进行排序,得到各评价项排名。

智能分析:

结合上述关联系数结果进行加权处理,最终得出关联度值,使用关联度值针对 15 个评价对象进行评价排序;关联度值介于 0~1 之间,该值越大代表其与"参考值"(母序列)之间的相关性越强,也即意味着其评价越高。从上表可以看出:针对本次 15 个评价项,(PCA)金融机构人民币贷款余额评价最高(关联度为: 0.785),其次是(PCA)金融机构人民币存款余额(关联度为: 0.774)。

1.4.2 主成分分析细节

1.4.1 详细结论

输出结果 1: KMO 检验和 Bartlett 的检验

KMO 检验和 Bartlett 的检验		-
KMO 值		0.000
	近似卡方	-786.593
Bartlett 球形度检验	df	105
	P	2.000
注: ***、**、*分别代表 1%、5%、	10%的显著性水平	

图表说明:

上表展示了 KMO 检验和 Bartlett 球形检验的结果,用来分析是否可以进行主成分分析。

- 若通过 KMO 检验(KMO>0.6),说明了题项变量之间是存在相关性的,符合主成分分析要求。
- 若通过 Bartlett 检验: P<0.05, 呈显著性,则可以进行主成分分析。

智能分析

KMO 检验的结果显示,KMO 的值为 nan,同时,Bartlett 球形检验的结果显示,显著性 P 值为 2.000,水平上不呈现显著性,接受原假设,也就是变量间彼此独立,则无法从中提取公因子,主成分分析无效,建议调整数据质量,程度为极不适合。

输出结果 2: 方差解释表格

成分	总方差解释 特征根		
风分	特征根	方差解释率(%)	累积方差解释率(%)
1	11.522	76.813	76.813
2	3.385	22.564	99.377
3	0.094	0.623	100
4			100
5			100

6	100	
7	100	
8	100	
9	100	
10	100	
11	100	
12	100	
13	100	
14	100	
15	100	

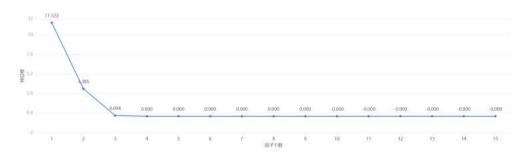
上表为总方差解释表格,主要是看主成分对于变量解释的贡献率(可以理解为究竟需要多少主成分才能把变量表达为100%),一般都要表达到90%以上才可以,否则就要调整因子数据。

•一般情况下,方差解释率越高,说明该主成分越重要,权重占比也应该越高。

智能分析

方差解释表中,在主成分3时,总方差解释的特征根低于1.0,变量解释的贡献率达到100.0,以上仅为参考,若特征根小于1.0临界值过大,也可以集合具体情况具体分析。

输出结果3:碎石图



图表说明:

- 碎石图是根据各主成分对数据变异的解释程度绘制的图。其作用是根据特征值下降的坡度 来确认需要选择的主成分个数,结合方差解释表可用于确认或调整主成分个数。
- ●每一个主成分为一个点,通过"坡度趋于平缓"的未知判断提取主成分的数量。

输出结果 4: 因子载荷系数表

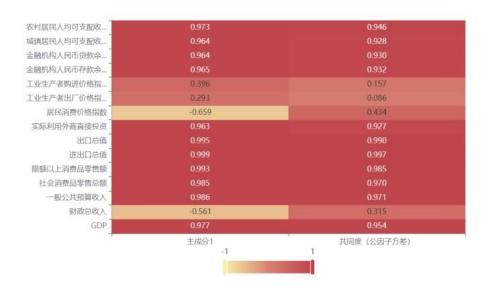
	因子载荷系数表 因子载荷系数 主成分1	共同度(公因子方差)
GDP	0.977	0.954
财政总收入	-0.561	0.315
一般公共预算收入	0.986	0.971
社会消费品零售总额	0.985	0.97

限额以上消费品零售额	0.993	0.985
进出口总值	0.999	0.997
出口总值	0.995	0.99
实际利用外商直接投资	0.963	0.927
居民消费价格指数	-0.659	0.434
工业生产者出厂价格指数	0.293	0.086
工业生产者购进价格指数	0.396	0.157
金融机构人民币存款余额	0.965	0.932
金融机构人民币贷款余额	0.964	0.93
城镇居民人均可支配收入	0.964	0.928
农村居民人均可支配收入	0.973	0.946

上表为因子载荷系数表,可以分析到每个主成分中隐变量的重要性。

 \bullet 假设前文确定得到 n 个因子,因子 i 中 a、b、c、d 的因子载荷系数较大,因此可将主成分 i 进行总结重命名。

输出结果 5: 因子载荷矩阵热力图



图表说明:

上图为载荷矩阵热力图,可以分析到每个主成分中隐变量的重要性。同时可结合具体业务进行各因子的隐变量分析。

输出结果 6: 成分矩阵表

成分矩阵表	
名称	成分 成分 1
GDP	0.085
财政总收入	-0.049
一般公共预算收入	0.086

0.085
0.086
0.087
0.086
0.084
-0.057
0.025
0.034
0.084
0.084
0.084
0.084

上表为成分矩阵表,意在说明各个成分的所包含的因子得分系数(主成分载荷),用于计算出成分得分,得出因子公式,其计算公式为:线性组合系数*(方差解释率/累积方差解释率),最后将其归一化即为因子权重得分。

• 线性组合系数,公式为:因子载荷系数除以对应特征根,即成分矩阵的系数。

智能分析:

模型的公式:

F1=0.08477638198935002×GDP-0.048686593439112955×财政总收入+0.08554421506267225×一般公共预算收入+0.08548008197697014×社会消费品零售总额+0.0861497577802133×限额以上消费品零售额+0.08667643254963134×进出口总值+0.0863679390061134×出口总值

+0.08355483721395086×实际利用外商直接投资-0.05716307852055251×居民消费价格指数

+0.025427624361768267×工业生产者出厂价格指数+0.03435016692107319×工业生产者购进价格指数+0.0837701545539884×金融机构人民币存款余额+0.08367613924562108×金融机构人民币贷款余额+0.08362829471865176×城镇居民人均可支配收入+0.08441780858714364×农村居民人均可支配收入

由上可以得到:

F=F1

输出结果 7: 因子权重分析

名称	方差解释率(%)	累积方差解释率 (%)	权重(%)
主成分 1	11.521958328770157	76.813	100%

图表说明:

上表为主成分分析的根据载荷系数等信息所做的主成分权重分析,其计算公式为:方差解释率/旋转后累积方差解释率。

智能分析:

主成分分析的权重计算结果显示,主成分1的权重为100%、

输出结果 8: 综合得分表

排名	行索引	综合得分	主成分 1
1	2021	1.2642472368835715	1.2642472368835715
2	2018	0.6200982139235535	0.6200982139235535
3	2020	-0.024050809036464494	-0.024050809036464494
4	2019	-0.5518414852143365	-0.5518414852143365
5	2017	-1.3084531565563224	-1.3084531565563224

上表格为预览结果,只显示综合排序的前 15 条数据,全部数据请点击下载按钮导出。 综合得分根据 F 值计算得到的综合得分进行降序排序,可得到各个样本的综合得分与排名情况, 同时输出因子浓缩后的结果。