

基于机器学习的未来工业投资重点模型研究

张志成

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2025年5月23日; 录用日期: 2025年6月16日; 发布日期: 2025年6月25日

摘要

经济是国家的基础, 经济安全是国家安全体系的重要组成部分, 是国家安全的依托和基石。我国经济已由高速增长阶段转向高质量发展阶段, 正处在转变发展方式、优化经济结构、转换增长动力的关键期。中国GDP由各行业GDP所组成, 研究各行业相关性, 分析中国未来在各行业的最优投资以及如何刺激就业已成为当前重要课题之一。根据历年来各行业投入、产出和就业基本情况等数据, 并基于皮尔逊相关系数矩阵、线性回归分析、投入-产出模型等研究方法, 对各行业的相关关系, 投资与各行业GDP的相关关系进行了详细分析, 建立基于历史数据的各行业GDP与刺激就业的模型, 并使用拉格朗日乘数法对该模型进行了求解。首先搜集数据, 对数据进行预处理, 检查中国各行业历年来的GDP是否有缺失值、异常值等, 并对数据分类汇总处理。其次利用可视化处理、皮尔逊相关系数矩阵和线性回归分析等方法, 详细分析了各行业之间的相关关系以及各行业对中国经济的促进和制约。本文我们建立了投入-产出模型、CES生产函数、Solow增长模型和VAR模型, 并运用中国2010年之前的数据作为训练集, 预测了2011~2023年的投资和各行业GDP之间的关系。之后, 根据预测结果, 得出最优模型为投入-产出模型。最后, 根据各行业的回归系数详细分析了投资与各行业GDP的相关关系。我们从一万亿投资资金出发, 结合第二问的最优投资分配投入-产出模型, 采用Cobb-Douglas生产函数, 构建总GDP最大化的投资分配模型, 运用拉格朗日乘数法对该模型进行求解, 给出了在限制投资资金时的最优投资方案, 并详细分析了限制三个行业时的最优投资方案。结合前面的分析和建模过程, 提出了促进就业和改善工作质量的就业最大化模型。并运用拉格朗日乘数法对该模型进行求解, 得出提高就业率和工作质量的最优投资分配, 并详细分析了限制三个行业时的最优投资方案。

关键词

线性回归分析, 投入-产出模型, 最优投资, 拉格朗日乘数法

Research on the Key Model of Future Industrial Investment Based on Machine Learning

Zhicheng Zhang

School of Mathematics, Liaoning Normal University, Dalian Liaoning

文章引用: 张志成. 基于机器学习的未来工业投资重点模型研究[J]. 应用数学进展, 2025, 14(6): 375-390.
DOI: 10.12677/aam.2025.146327

Abstract

Economy is the foundation of a country, and economic security is an important component of the national security system, serving as the support and cornerstone of national security. China's economy has shifted from a stage of high-speed growth to one of high-quality development. It is currently in a critical period of transforming the development mode, optimizing the economic structure and converting the growth drivers. China's GDP is composed of the GDP of various industries. Studying the correlation among industries, analyzing the optimal investment in various industries in China in the future and how to stimulate employment has become one of the important current issues. Based on the data such as the basic situations of input, output and employment in various industries over the years, and based on research methods such as Pearson correlation coefficient matrix, linear regression analysis, and input-output model, a detailed analysis was conducted on the correlations of various industries and the correlations between investment and the GDP of various industries. A model of the GDP of various industries and employment stimulation based on historical data was established. The Lagrange multiplier method was used to solve the model. First, collect data, preprocess the data, check whether there are missing values, outliers, etc. in the GDP of various industries in China over the years, and classify, summarize and process the data. Secondly, by using methods such as visualization processing, Pearson correlation coefficient matrix and linear regression analysis, the correlations among various industries and the promotion and constraints of each industry on the Chinese economy were analyzed in detail. In this paper, we established the input-output model, the CES production function, the Solow growth model and the VAR model, and used the data of China before 2010 as the training set to predict the relationship between investment and the GDP of various industries from 2011 to 2023. Afterwards, based on the prediction results, the optimal model is derived as the input-output model. Finally, the correlation between investment and the GDP of each industry was analyzed in detail based on the regression coefficients of each industry. Starting from one trillion investment funds, combined with the input-output model of the optimal investment allocation in the second question, adopting the Cobb-Douglas production function, we constructed an investment allocation model that maximizes total GDP. We used the Lagrange multiplier method to solve this model and presented the optimal investment plan when limiting investment funds. The optimal investment plan when restricting three industries was analyzed in detail. Combined with the previous analysis and modeling process, an employment maximization model for penetrating employment and improving job quality was proposed. The Lagrange multiplier method was used to solve the model, obtaining the optimal investment allocation for improving the employment rate and job quality, and the optimal investment schemes when limiting three industries were analyzed in detail.

Keywords

Linear Regression Analysis, Input-Output Model, Optimal Investment, Lagrange Multiplier Method

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 背景介绍

在全球经济格局加速演变的背景下，中国经济正从高速增长阶段逐步转向高质量发展阶段。深入分

析各产业间的相互关系并合理配置政府投资资金，以促进经济增长、提升就业率及实现可持续发展，已成为当前亟待解决的重要课题。本文聚焦于五个问题，进行了系统性分析，旨在为政府在中国未来投资决策方面提供理论依据和切实可行的方案。

1.2. 问题重述

在本文中，问题是层层递进的，并且服务于同一主题——如何实现最优投资。我们需要从不同的角度出发，并解决以下五个问题：

- 1) 研究产业间的相互关系与对中国经济发展的促进和制约作用。
- 2) 建立投资与各行业 GDP 之间关系的模型。
- 3) 在政府 1 万亿投资资金的基础上，建立实现 GDP 最大化的投资方案。
- 4) 并考虑如果限制投资行业数量，应如何分配以及选择这些行业的原因。从提高就业率和改善工作质量的角度来调整投资计划。
- 5) 在以增加 GDP 总值和改善就业为目标的前提下，解决中国应大力发展哪些产业并且面向可持续发展进一步优化投资。

2. 假设与依据

为了简化问题，我们做出以下基本假设，每个假设都是合理的。

假设 1：本文所需的所有数据源都是真实可靠的。

论证：本文需要依靠各行业的国内生产总值和各行业就业情况等数据，分析各行业的相互关系、对中国经济发展的促进与制约作用以及研究中国未来对各行业最优的投资方案。因此，数据的可靠性非常重要。

假设 2：假设政府投资主要通过资本投入影响各行业的 GDP，其余因素的影响可以忽略不计。

论证：各行业的 GDP 并不只受到政府投资的影响，有时还会受到投资波动、政策调整或其他外部因素的影响，本文只考虑投资带来的主要影响，其余因素相对较小。

假设 3：忽略数据误差对结果的影响，即假设数据误差为 0。

3. 符号和数据

3.1. 符号

见表 1。

Table 1. Notation explanation
表 1. 符号说明

符号	描述	符号	描述
R	相关系数	α	资本的权重
a	回归系数	β	劳动的权重
b	线性回归方程的截距	ρ	替代弹性参数
Y_i	行业 i 的 GDP	Y_t	各行业的 GDP 和投资的向量
K_i	行业 i 的资本投入(即政府的投资)	C	常数项
L_i	行业 i 的劳动投入	$A_1, A_2, \dots, A_p$	滞后期的系数矩阵
A_i	行业 i 的全要素生产率(TFP)	ε_t	误差项

续表

α_i	资本的产出弹性	λ	拉格朗日乘数、投资总额约束的影响
β_i	劳动的产出弹性	E_i	第 i 个行业的就业人数
$\ln(Y_i)$	行业 i 的 GDP 的对数	I_i	第 i 个行业的投资
$\ln(K_i)$	行业 i 的资本投入的对数	γ_i	第 i 个行业的就业生产率常数
$\ln(L_i)$	行业 i 的劳动投入的对数	E_{total}	就业总人数
$\ln(A)$	行业 i 的全要素生产率的对数	Q_i	就业质量因子

注意：有些变量没有列出，它们的具体含义将在下面介绍。

3.2. 数据

3.2.1. 数据概览

通过对问题的分析，我们搜集了中国各行业历年来的国内生产总值，如农林渔牧业、工业、建筑业等。此外，还需要搜集各行业历年来的就业数据。

由于数据量庞大，难以全部列出，因此将数据可视化展示是一种很好的方法。

3.2.2. 数据搜集

除从参考文献中搜集的数据外，其他数据来源见下表 2。

Table 2. Databases and database networks
表 2. 数据库与数据库网

数据库名称	数据库网站
国家统计局	https://www.stats.gov.cn/
中国经济社会大数据研究平台	https://data.cnki.net/home
中国统计年鉴 2024	https://www.stats.gov.cn/sj/ndsj/2024/indexch.htm

注意：未列出的数据源将在引用时进行标记。

3.2.3. 数据的可视化

下图 1 展示了对搜集来的历年来中国各行业的国内生产总值(GDP)的可视化。

从图 1 中可以直观地看出各行业从 1978-2023 年的国内生产总值的变化。

4. 问题一的建模与求解

4.1. 问题分析

对于问题 1，首先对搜集的数据进行预处理，检查中国各行业历年来的国内生产总值是否有缺失值、异常值等。其次利用可视化处理、皮尔逊相关系数矩阵和线性回归分析等方法，研究主要产业的国内生产总值之间的相互关系以及影响。最后，以国内生产总值这个指标分析出不同产业之间的相互关系以及影响，并详细分析出，各行业对中国经济的促进和制约。

图 1. 历年来中国各行业的 GDP 变化

使用皮尔逊相关系数和 `fitlm` 线性回归分析来判断中国主要产业之间的相互性和各产业对中国经济的影响，并使用热力图将其可视化。

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

通过 Matlab 代码，得到以下热力图 2 以及结果。

结合热力图 2 可以洞察主要产业之间的相互关系, 以及相互影响, 促进或制约经济发展分析如下:

热力图直观地展示了不同行业之间的相关性，正相关系数(接近 1)表明两个行业可能存在协同发展的关系，即一个行业的增长可能带动另一个行业的增长。例如，工业与建筑业的相关系数非常高(0.9637)，这表明这呈极强正相关。工业为建筑业提供建筑材料、机械设备等，建筑业的发展又会拉动对工业产品的需求，如钢材、水泥等，两者相互依存、相互促进。

应用数学讲展

• 对经济发展的综合影响

这些产业之间的相互关系对经济发展有着重要影响。正相关的产业可以形成协同发展的效应，例如工业与建筑业的紧密结合，能够带动上下游产业的发展，促进就业和经济增长。而负相关的产业则需要产业结构调整中进行合理规划和引导，以实现经济的平稳转型和可持续发展[1]。

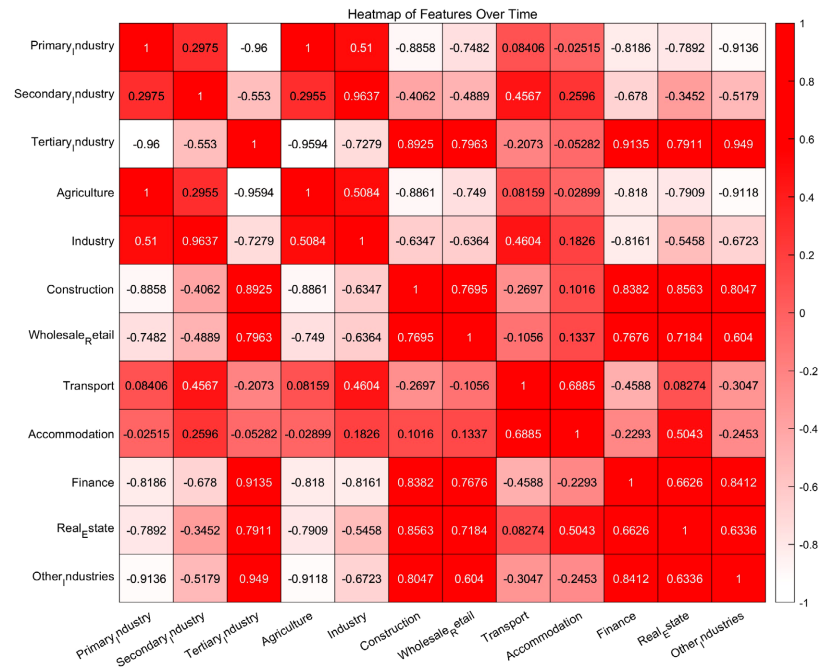


Figure 2. Heatmap of GDP correlation coefficients among various industries
图 2. 各行业之间 GDP 相关系数热力图

5. 问题二的建模与求解

5.1. 问题分析

对于问题二，目标是研究投资与各行业 GDP 增值之间的关系。我们通过收集各行业的 GDP、资本投入和劳动投入数据，量化了投资对 GDP 的影响。针对问题二，我们通过建立投入 - 产出模型、CES 生产函数、Solow 增长模型和 VAR 模型四个模型分别研究了投资与各行业 GDP 之间的关系。最后我们以中国 2010 年之前的数据作为训练集，预测 2011~2023 年的投资与各行业 GDP 之间的关系，将预测结果与真实情况进行比较，对模型进行评估和可视化。

5.2. 模型建立

本节使用四种不同的方法分析投资与各行业 GDP 之间的相互关系。分别是投入 - 产出模型、CES 生产函数、Solow 增长模型和 VAR 模型这四个角度分析两者之间的关系，并对这四种模型的预测准确度进行比较。

5.2.1. 投入 - 产出模型

本题要求我们建立一个投资与各行业国内生产总值(GDP)之间的关系模型，并对模型进行评估。为了实现这一目标，我们可以使用 Cobb-Douglas 生产函数[2]和多元回归分析来建立数学模型，具体步骤如下：

步骤 1: 本研究旨在探讨政府投资对各行业 GDP 的影响, 目标是通过构建数学模型表达投资与行业产出(GDP)之间的关系。这个问题可以转化为以下两个问题: 如何量化投资对 GDP 的影响和如何评估模型的准确性和效果。

步骤 2: 选择目标生产函数

针对本研究, 我们选择了能够较好地反映资本和劳动对产出的影响, 并且适用于各行业的生产过程的非线性 Cobb-Douglas 生产函数:

$$Y_i = A_i K_i^{\alpha_i} L_i^{\beta_i}$$

步骤 3: 将其转变为线性函数:

$$\ln(Y_i) = \ln(A_i) + \alpha_i \ln(K_i) + \beta_i \ln(L_i)$$

步骤 4: 回归分析: 根据收集到的数据, 我们可以对该模型进行线性回归分析。

步骤 5: 通过 Matlab 代码, 得到以下预测结果(如图 3):

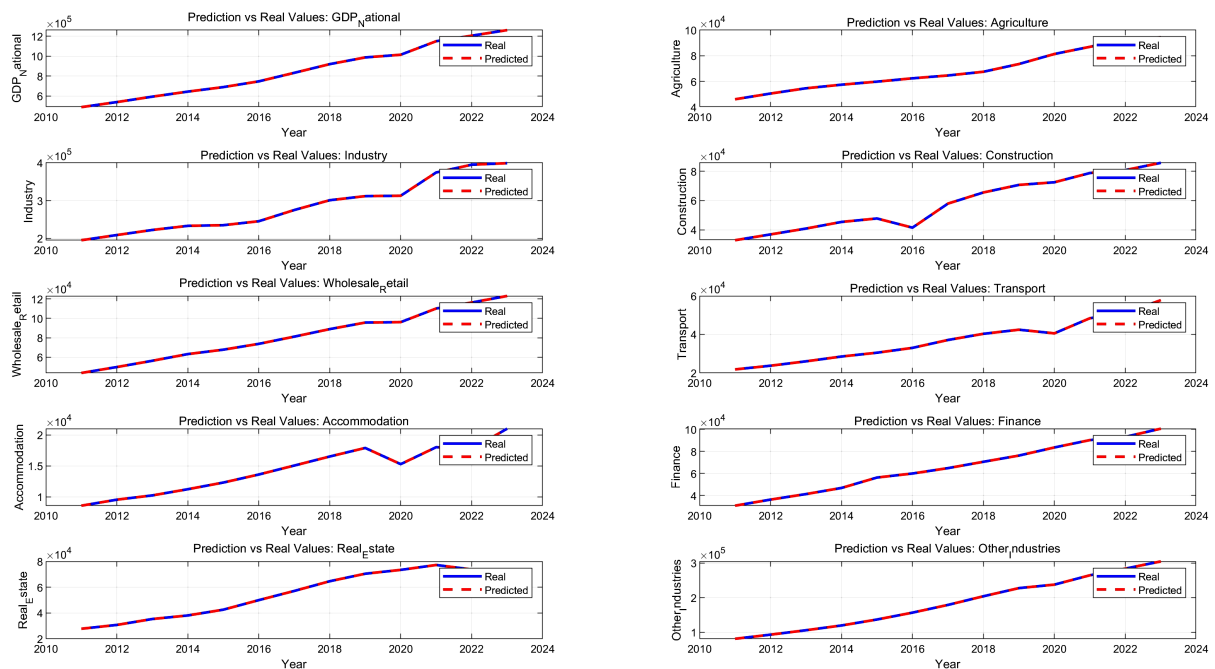


Figure 3. Input-output model: Comparative analysis of actual and predicted data for each industry from 2011 to 2023

图 3. 投入 - 产出模型: 2011~2023 年各行业实际与预测数据对比分析图

5.2.2. CES 生产函数(常弹性替代生产函数)

与投入 - 产出模型类似, 在这里选择的目标函数为能够更灵活地适应资本和劳动投入变化的 CES 生产函数[3], 它表示资本和劳动的替代关系具有弹性。具体步骤如下:

步骤 1: 选择目标生产函数

$$Y_i = A_i (\alpha K_i^\rho + \beta L_i^\rho)^{\frac{1}{\rho}}$$

步骤 2: 将其转变为线性函数

$$\ln(Y_i) = \ln(A_i) + \frac{1}{\rho} \ln(\alpha K_i^\rho + \beta L_i^\rho)$$

步骤 3: 使用回归分析估计参数, 收集各行业的 GDP、资本投入、劳动投入数据后, 采用非线性回归方法来估计 α, β, ρ 等参数。

步骤 4: 通过 Matlab 代码, 得到以下预测结果(如图 4):

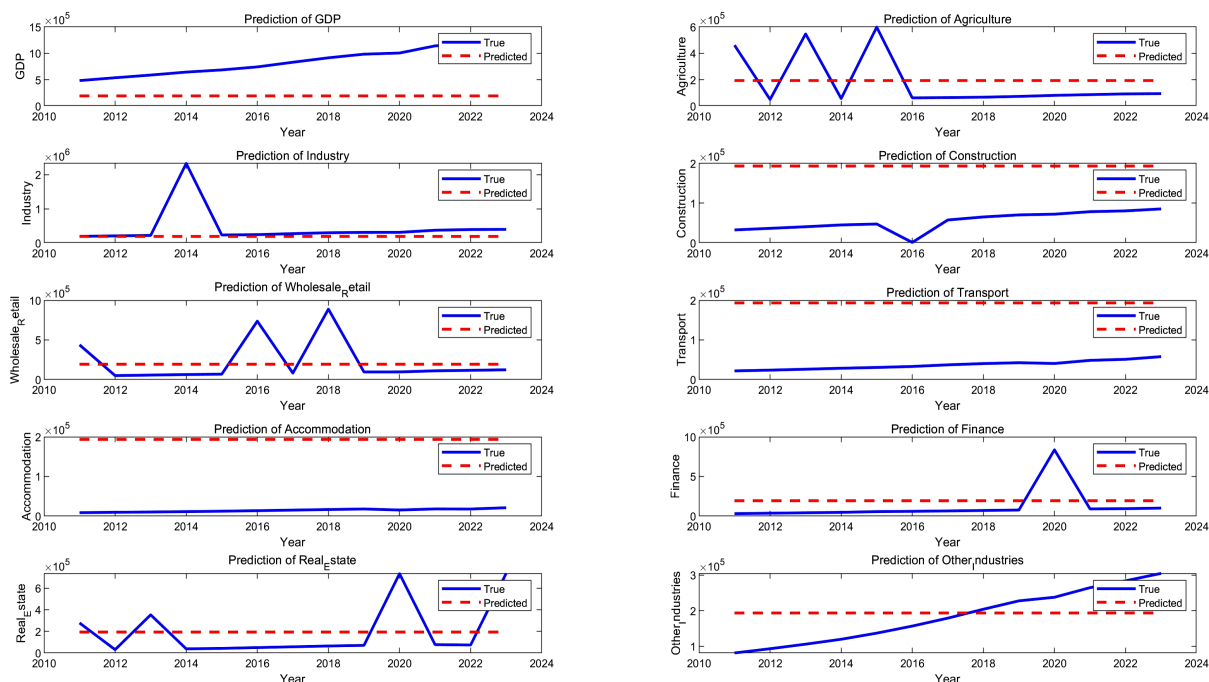


Figure 4. CES production function: Comparative analysis of actual and predicted data for each industry from 2011 to 2023
图 4. CES 生产函数: 2011~2023 年各行业实际与预测数据对比分析图

5.2.3. Solow 增长模型

Solow 增长模型是经典的宏观经济模型, 强调资本积累、劳动投入和技术进步对经济增长的影响。具体步骤如下:

步骤 1: 选择目标生产函数

$$Y_i = A_i K_i^\alpha L_i^{1-\alpha}$$

步骤 2: 将其转变为线性函数

$$\ln(Y_i) = \ln(A_i) + \alpha \ln(K_i) + (1-\alpha) \ln(L_i)$$

步骤 3: 使用回归分析估计参数 α

通过普通最小二乘法对转化后的线性模型进行回归分析, 得到参数 α 和 $1-\alpha$ 。

步骤 4: 通过 Matlab 代码, 得到以下预测结果(如图 5)。

5.2.4. VAR 模型(向量自回归模型)

VAR 模型是一个多变量时间序列模型, 用于分析多个经济变量之间的相互影响。如果有跨时间和跨行业的数据, VAR 模型能够帮助我们理解资本投入、劳动投入和 GDP 之间的动态关系。具体步骤如下:

步骤 1: 向量自回归模型

$$Y_t = C + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t$$

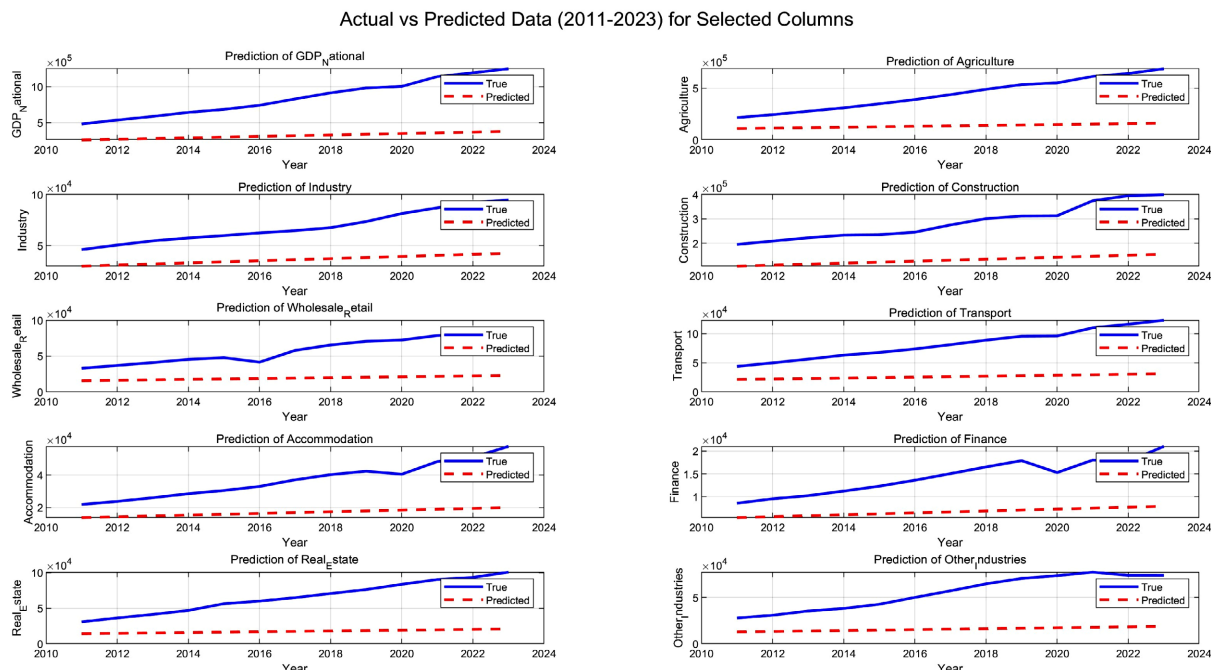


Figure 5. Solow growth model: Comparative analysis of actual and predicted data for each industry from 2011 to 2023

图 5. Solow 增长模型：2011~2023 年各行业实际与预测数据对比分析图

步骤 2：确定滞后期

使用信息准则(如 AIC、BIC)来选择滞后期，即模型中包含的历史期数。滞后期的选择决定了每个变量的历史数据对当前经济状态的影响。

步骤 3：回归估计

通过最小二乘法估计 VAR 模型中的参数，得到各个滞后期的系数矩阵。通过分析这些系数，可以确定不同变量之间的因果关系。

步骤 4：脉冲响应分析

VAR 模型的一个重要应用是脉冲响应分析(Impulse Response Function, IRF)，脉冲响应分析可以帮助我们了解投入对 GDP 的长期影响。

步骤 5：通过 Matlab 代码，得到以下预测结果(如图 6)。

5.2.5. 四个模型得出投资与各行业 GDP 的相关关系

考虑到投资波动、政策调整或其他外部因素对各行业 GDP 的影响，我们这里假设政府投资主要通过资本投资影响各行业的产出。在这里我们给出了四个经济模型在各个行业的回归系数的比较(如图 7)。

5.3. 模型评估

根据图 3~7 的结果表明，基于中国 2010 年之前的投资与产出数据，投入 - 产出模型对各行业的预测值与实际值高度吻合。尽管在某些年份中，实际结果与预测结果之间存在细微差异，这可能归因于投资波动、政策调整或其他外部因素的影响。此现象表明该模型在预测投资与各行业 GDP 关系方面具有较高的准确性和可行性。此外，VAR 模型(向量自回归模型)虽然在各行业的预测精度略逊于投入 - 产出模型，但其误差仍小于 CES 生产函数和 Solow 增长模型。相比之下，CES 生产函数在工业领域的预测表现较为精准，但在其他行业的预测结果与实际值存在显著偏差。值得注意的是，Solow 增长模型的预测误差最大，其应用可行性最低。

Actual vs Predicted Data (2011-2023) for Selected Columns

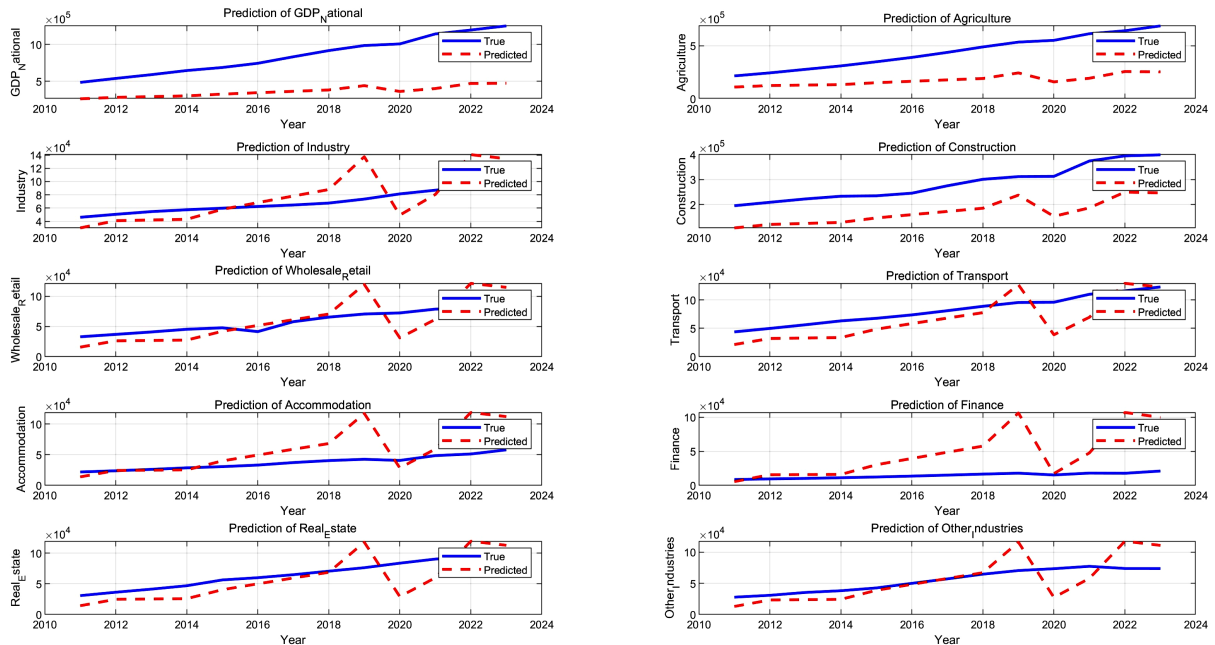


Figure 6. VAR model: Comparative analysis of actual and predicted data for each industry from 2011 to 2023

图 6. VAR 增长模型：2011~2023 年各行业实际与预测数据对比分析图

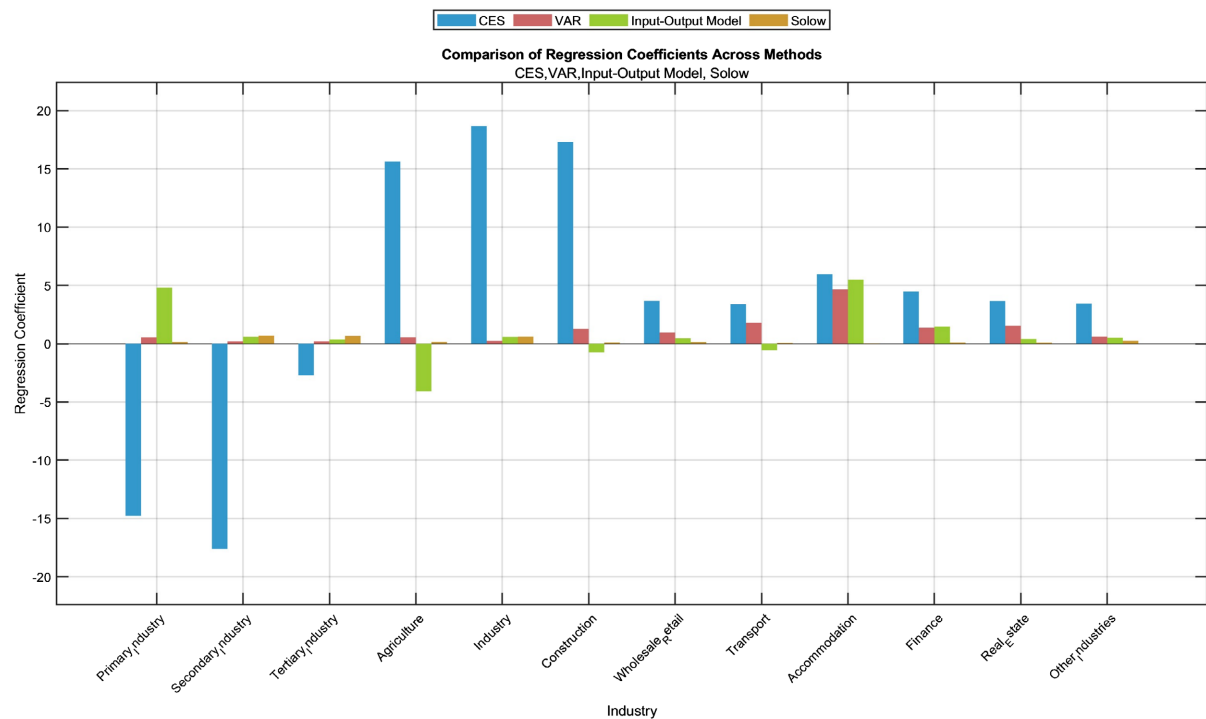


Figure 7. Regression coefficients of the four economic models in various industries

图 7. 四个经济模型在各个行业的回归系数图

综上所述，投入 - 产出模型的预测结果最为精确，并且相比于简单的线性回归，更真实地反映多要素共同决定的产出。

对投入 - 产出模型的分析框架, 结合 Cobb-Douglas 生产函数和多元回归方法的特性, 从适用性、局限性和数据特征三个维度展开:

一、模型适用性分析

理论契合性

Cobb-Douglas 生产函数的核心假设与宏观经济总量分析的简化需求高度契合, 适合研究政府投资对行业 GDP 的边际贡献。

通过对数线性化将非线性关系转化为线性回归模型, 降低了参数估计难度, 且能直接通过回归系数解释投资和劳动对产出的弹性。

行业适配性

该模型假设不同行业的资本 - 劳动弹性相似, 适用于同质性较强的行业(如制造业、农业), 但对异质性行业(如信息技术服务业、金融业)可能存在解释偏差, 需通过分行业回归或引入行业虚拟变量改进。

政策分析价值

通过估计投资弹性系数, 可量化政府投资对各行业 GDP 的拉动作用, 为财政政策制定(如基建投资优先领域选择)提供依据。

二、模型局限性分析

理论假设的局限性

固定替代弹性假设: Cobb-Douglas 函数要求投资与劳动的替代弹性恒为 1, 但现实中行业技术差异可能导致弹性动态变化。

规模报酬不变假设: 忽略行业规模效应(如互联网行业的边际成本递减), 高估或低估投资对产出的影响。

数据与估计问题

内生性问题: 投资与 GDP 可能存在双向因果关系(如高 GDP 行业吸引更多投资), 普通最小二乘法估计会产生偏误, 需采用工具变量法(如滞后投资、地理政策变量)修正。

遗漏变量偏差: 未控制技术进步、人力资本、政策环境等变量, 可能导致估计值虚高。

动态性缺失

模型仅反映当期投资与产出的静态关系, 忽略投资效果的滞后性(如基建投资对 GDP 的拉动需 1~2 年显现), 建议引入分布滞后模型。

三、数据特征的深入分析

数据结构特征

数据类型: 若使用面板数据(多行业多年份), 需检验行业间异质性和时间效应, 避免混合回归的误导性结论。

量纲差异: 不同行业 GDP 与投资额量级差异大(如重工业 vs 文创产业), 需对数据取对数或标准化以消除异方差。

统计问题诊断

多重共线性: 投资与劳动变量可能高度相关(如资本密集型行业同时投入高资本和高技术劳动), 需通过 VIF 检验或主成分分析降维。

异方差性: 行业波动性差异(如能源行业 GDP 受价格波动影响大)导致残差异方差, 需采用加权最小二乘法或稳健标准误修正。

数据代表性缺陷

行业覆盖不全: 若数据仅包含规模以上企业, 可能低估小微企业对 GDP 的贡献(如餐饮零售业)。

时间跨度不足：短期数据(如 5 年内)难以捕捉经济周期影响，建议补充长期数据或引入 HP 滤波法分离趋势项与周期项。

因此，投入 - 产出模型为最优模型。

6. 问题三的建模与求解

6.1. 问题分析

对于问题三，我们需要从一万亿单位投资资金出发，考虑中国政府投资在各行业的最优分配。假设政府有 1 万亿单位的投资资金，目标是通过投资最大化国内生产总值(GDP)。我们将使用基于投资回报率的优化方法，结合之前建立的生产函数和行业投资模型，来求解最优投资分配。

6.2. 模型建立

根据问题二，我们得到最优模型为投入 - 产出模型，本题我们将沿用以 Cobb-Douglas 生产函数为目标生产函数的投入 - 产出模型。

6.2.1. 确定约束条件

本目标为最大化 GDP，即使得总的国内生产总值 Y_{total} 最大化。

(1) 投资总额限制

政府总投资资金为 1 万亿单位，即总投资资金不能超过 1 万亿单位：

$$I_1 + I_2 + \cdots + I_n = 10000 \text{ hunderd million RMB}$$

其中 I_i 表示对第 i 个行业的投资额。

(2) 资本和劳动的边际产出弹性：

假设每个行业的投资和 GDP 之间的关系是通过生产函数反映的，且每个行业的投资回报率不同。为此，我们需要每个行业的资本弹性 α_i 和劳动弹性 $1 - \alpha_i$ 。

6.2.2. 目标函数的确立

我们需解决在一万亿投资基金下的最优投资分配，常见的生产函数是 Cobb-Douglas 生产函数，它的形式为：

$$Y_i = A_i K_i^{\alpha_i} L_i^{1-\alpha_i}$$

为了最大化总 GDP，我们需要通过投资来优化各行业的资本投入 K_i 。

6.2.3. 构建总 GDP 最大化模型

假设每个行业的 GDP Y_i 都符合上述的生产函数，且投资与资本投入 K_i 直接相关(假设投资资金直接用于资本投入)。因此，目标函数可以写为：

$$Y_{total} = \sum_{i=1}^n A_i K_i^{\alpha_i} L_i^{1-\alpha_i}$$

根据每个行业的资本弹性 α_i ，我们知道资本投入对 GDP 的影响。假设劳动投入是固定的，且各行业的劳动投入量 L_i 是政府无法直接调整的(即外部因素或固定资源)，那么我们可以将目标函数简化为：

$$Y_{total} = \sum_{i=1}^n A_i (I_i)^{\alpha_i}$$

其中： I_i 表示对第 i 个行业的投资， A_i 是行业的全要素生产率，反映了该行业的技术水平。

从而可知，最终建立的非线性规划模型[4]如下：

$$\begin{cases} \max Y_{total} = \sum_{i=1}^n A_i (I_i)^{\alpha_i} \\ \sum_{i=1}^n I_i = 10000 \text{ hunderd million RMB} \end{cases}$$

6.3. 模型求解

6.3.1. 使用拉格朗日乘数法进行求解

本题我们采用拉格朗日乘数法进行求解，为了求解这个优化问题，我们可以使用拉格朗日乘数法。定义拉格朗日函数：

$$L(I_1, I_2, \dots, I_n, \lambda) = \sum_{i=1}^n A_i (I_i)^{\alpha_i} + \lambda \left(10000 - \sum_{i=1}^n I_i \right)$$

其中， λ 是拉格朗日乘数，代表投资总额约束的影响。

对 I_i 和 λ 分别求偏导并令其为零，得到：

$$\begin{aligned} \frac{\partial L}{\partial I_i} &= A_i \alpha_i (I_i)^{\alpha_i-1} - \lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= 10000 - \sum_{i=1}^n I_i = 0 \end{aligned}$$

解得：

$$\lambda = A_i \alpha_i (I_i)^{\alpha_i-1}$$

将这个结果带回原约束条件中：

$$\sum_{i=1}^n I_i = 10000 \text{ hunderd million RMB}$$

通过求解上述方程组，我们可以得到每个行业的最优投资 I_i^* 。

6.3.2. 投资分配的具体计算

首先，需要收集每个行业的全要素生产率 A_i ，确定每个行业的资本弹性 α_i 。使用拉格朗日乘数法解上面的方程组，得到每个行业的投资分配比例。然后，计算每个行业的最优投资，根据拉格朗日乘数法的结果，计算出每个行业的最优投资额。

通过 Matlab 求解，我们得到各行业具体投资分配图 8 如下：

Industry	Investment Proportion (%)	Industry	Investment Proportion (%)
Agriculture, Forestry, Animal Husbandry, Fishery	4.4082	Finance	13.8415
Industry	26.7883	Real Estate	6.4376
Construction	6.0085	Other Industries	21.1468
Wholesale and Retail	8.7085	Accommodation and Catering	4.0681
Transportation, Storage, and Postal	8.5925		
Δ GDP	136%		

Figure 8. Investment allocation table without limiting the number of industries

图 8. 不限制行业个数的投资分配表

并且可以根据图 8 可以看出，我国应该优先大力发展 Industry、Finance、Other Industries 等行业，可以实现我国国内生产总值最大化。

7. 问题四的建模与求解

7.1. 问题分析

针对第四问，我们的目标是在保证 GDP 最大化的同时，还需要兼顾各行业就业率的提升和工作质量的改善。具体而言，我们将通过政府的投资来促进就业并改善工作质量，同时优化投资分配，以有效刺激就业市场。

7.2. 模型建立

7.2.1. 确立约束条件

(1) 总投资金额限制：

政府的总投资仍为 1 万亿单位，即：

$$\sum_{i=1}^n I_i = 10000 \text{ hundred million RMB}$$

其中， I_i 表示对第 i 个行业的投资。

(2) 每个行业的就业影响：

我们假设每个行业的投资对就业的影响是通过行业的就业弹性来表示的。即，某一行业的就业机会的变化可以通过该行业的投资水平来反映。

(3) 劳动市场的特点

每个行业的劳动需求与投资和生产量之间的关系，通过劳动弹性 β_i 来表达。高技术行业和服务业的投资往往能提高就业质量，而低技术和基础设施类行业则可能更多地创造基础就业机会。

7.2.2. 目标函数的确立——行业就业函数

我们需要解决 GDP 最大化时，提高各行业就业率和工作质量的最优投资分配。我们可以采用如下的就业函数作为目标函数来表示行业的就业机会：

$$E_i = f(I_i) = \gamma_i (I_i)^{\beta_i}$$

其中 β_i 是第 i 个行业的劳动弹性。

总就业人数 E_{total} 可以表示为所有行业就业人数的加总：

$$E_{total} = \sum_{i=1}^n E_i = \sum_{i=1}^n \gamma_i (I_i)^{\beta_i}$$

劳动市场的约束条件为

$$\begin{aligned} \sum_{i=1}^n I_i &= 10000 \text{ 亿元} \\ 0 &\leq I_i \leq \text{上限} \quad \forall i \end{aligned}$$

7.2.3. 构建就业最大化模型

为了提高就业率并且有效改善工作质量，我们需要在总投资约束下，通过最优的投资分配来最大化就业人数 E_{total} 。因此，优化问题可以写作：

$$\max_{I_1, I_2, \dots, I_n} E_{total} = \sum_{i=1}^n \gamma_i (I_i)^{\beta_i}$$

约束条件：

$$\sum_{i=1}^n I_i = 10000 \text{ hundred million RMB}$$

$$0 \leq I_i \leq \text{上限} \quad \forall i$$

其中, γ_i 和 β_i 是由行业的劳动需求和就业弹性决定的参数。

7.2.4. 考虑就业质量的因素

为了进一步改善就业质量, 假设不同行业的就业质量不一样。我们可以引入一个就业质量因子 Q_i , 表示行业的平均工资水平、福利待遇、职业安全性等因素。在优化模型中, 可以将就业质量因子引入目标函数, 使得不仅仅考虑行业就业人数, 还要考虑就业质量:

$$E_{total} = \sum_{i=1}^n \gamma_i (I_i)^{\beta_i} Q_i$$

在这种情况下, 优化问题变为:

$$\max_{I_1, I_2, \dots, I_n} = \sum_{i=1}^n \gamma_i (I_i)^{\beta_i} Q_i$$

并且同样受到总投资约束:

$$\sum_{i=1}^n I_i = 10000 \text{ hundred million RMB}$$

通过引入就业质量因子, 可以使投资更多地倾向于那些能够提供高质量就业的行业(如高技术产业、金融业、教育行业等)。

7.3. 模型求解

与第三问类似, 我们可以使用拉格朗日乘数法[5]来求解这个优化问题。定义拉格朗日函数:

$$L(I_1, I_2, \dots, I_n, \lambda) = \sum_{i=1}^n \gamma_i (I_i)^{\beta_i} + \lambda \left(10000 - \sum_{i=1}^n I_i \right)$$

对 I_i 和 λ 分别求偏导并令其为零, 得到:

$$\frac{\partial L}{\partial I_i} = \gamma_i \beta_i (I_i)^{\beta_i - 1} - \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = 10000 - \sum_{i=1}^n I_i = 0$$

解得:

$$\lambda = \gamma_i \beta_i (I_i)^{\beta_i - 1}$$

将这个结果带回总投资约束条件中:

$$\sum_{i=1}^n I_i = 10000 \text{ hundred million RMB}$$

通过求解上述方程组, 我们可以得到每个行业的最优投资 I_i^* 以最大化就业[6]。

通过 Matlab 求解, 我们得到提高就业率的最优投资分配图 9 如下。

Industry	Investment Proportion (%)	Industry	Investment Proportion (%)
Agriculture, Forestry, Animal Husbandry, Fishery	3.2430	Finance	10.8190
Industry	20.7445	Real Estate	5.3000
Construction	5.1750	Other Industries	16.1295
Wholesale and Retail	7.3450	Accommodation and Catering	3.0110
Transportation, Storage, and Postal	7.3205		
Δ GDP	118%		

Figure 9. Optimal investment allocation table without limiting investment industry

图 9. 不限制投资行业的最优投资分配表

我们可以根据图 9 看出，在保证国内 GDP 最大化时，如果要促进就业，我国应该优先大力发展 Industry、Finance、Other Industries 等行业。

7.4. 针对行业数量限制的处理

根据最优解，得出每个行业的投资分配比例，并分析哪些行业应优先获得资金支持以最大化就业。如果需要，可以调整不同的投资组合，考察不同的就业质量因子对最终结果的影响(如图 10)。

Industry	Investment Proportion (%)
Agriculture, Forestry, Animal Husbandry, Fishery	38.8768
Finance	47.5912
Accommodation and Catering	13.5320
Δ GDP	105%

Figure 10. Optimal investment allocation table for limiting the activities of 3 investment industries

图 10. 限制 3 个投资行业的最优投资分配表

综合考虑前三问，并结合图 9、图 10 我们可以得出农林牧渔业、金融业、住宿和餐饮业在未来的中国投资里最具有优势，因此应该优先投资这三个行业。并且我们给出了在限制投资个数的情况下，各行业特定的投资比例：农林牧渔业 38.8762%，金融业 47.5912%，住宿和餐饮业 13.5320%。

参考文献

- [1] 刘雪敏. 有效市场、有为政府与产业结构升级[D]: [硕士学位论文]. 大连: 东北财经大学, 2019.
- [2] 刘云东, 邹涛, 郑沅维, 等. 基于 Cobb-Douglas 生产函数的广东省农业信息化实现路径研究[J]. 智慧农业导刊, 2025, 5(3): 11-15.
- [3] 李鹏, 于梦. 中国出口能源的供给价格弹性大小的一个估计——基于 CES 生产函数的估计[J]. 江苏商论, 2024(6): 55-58+87.
- [4] 徐石, 李璐. 铁路空车调整的非线性规划模型研究[J]. 铁道运营技术, 2025, 31(1): 9-12.
- [5] 赵莉莉, 王蕾. 拉格朗日乘数法及其推广[J]. 河南财政金融学院学报(自然科学版), 2024, 33(1): 50-54.
- [6] 司守奎, 孙玺菁. 数学建模算法与应用[M]. 北京: 国防工业出版社, 2011.