

基于多元统计分析的中国区域经济高质量发展水平评价研究

高 宁

首都经济贸易大学统计与数据科学学院, 北京

收稿日期: 2026年9月27日; 录用日期: 2026年10月19日; 发布日期: 2026年10月29日

摘 要

本文基于创新、协调、绿色、开放、共享五个维度, 构建包含多个一级指标和二级指标的区域高质量发展评价指标体系, 并选取2021~2022年我国31个省级行政区的相关数据作为研究样本。首先对数据进行缺失值处理、异常值检测及标准化处理, 在此基础上采用因子分析、主成分分析、多维尺度分析和典型相关分析等多元统计方法, 对各地区高质量发展水平、结构差异及维度协同关系进行系统研究。研究结果表明: 我国区域高质量发展存在明显的空间差异, 东部沿海地区在创新能力、开放水平和综合竞争力方面总体具有优势, 而中西部部分地区在绿色发展、城乡协调和共享发展等方面仍存在短板; 因子分析提取出产业创新开放、民生共享、内需协调、创新协调及外资绿色约束等公共因子, 能够较好反映区域高质量发展的结构特征; 主成分分析结果显示广东、江苏、浙江、上海和北京等地区综合发展水平位居前列; 多维尺度分析进一步揭示了不同省级行政区在发展结构上的相似性与差异性; 典型相关分析表明创新、协调与绿色、开放、共享之间存在显著的整体协同关系。基于上述结果, 本文提出针对不同发展模式的差异化政策建议, 并从跨维度协同角度设计联动型政策工具, 为区域高质量发展提供数据支撑。

关键词

高质量发展, 因子分析, 主成分分析, 多维尺度分析, 典型相关分析

A Study on the Evaluation of China's Regional High-Quality Economic Development Based on Multivariate Statistical Analysis

Ning Gao

School of Statistics and Data Science, Capital University of Economics and Business, Beijing

文章引用: 高宁. 基于多元统计分析的中国区域经济高质量发展水平评价研究[J]. 统计学与应用, 2026, 15(10): 169-185. DOI: 10.12677/sa.2026.1510237

Abstract

This paper constructs an evaluation indicator system for regional high-quality development with multiple primary and secondary indicators from five dimensions: innovation, coordination, green development, openness, and sharing. Relevant data of 31 provincial-level administrative regions in China from 2021 to 2022 are selected as the research sample. Missing value imputation, outlier detection and standardization are conducted on the dataset. On this basis, multivariate statistical methods including factor analysis, principal component analysis, multidimensional scaling and canonical correlation analysis are adopted to systematically investigate the high-quality development level, structural disparities and dimensional synergies across regions. The results show that notable spatial disparities exist in regional high-quality development across China. The eastern coastal regions generally have advantages in innovation capacity, openness and comprehensive competitiveness, while some central and western regions still face weaknesses in green development, urban-rural coordination and shared development. Factor analysis extracts common factors such as industrial innovation and openness, people's livelihood and sharing, domestic demand coordination, innovation coordination, and foreign capital and green constraints, which adequately capture the structural features of regional high-quality development. Principal component analysis indicates that Guangdong, Jiangsu, Zhejiang, Shanghai and Beijing rank at the top in terms of comprehensive development level. Multidimensional scaling further reveals the similarities and differences in development structure among provinces. Canonical correlation analysis verifies significant overall synergies between innovation-coordination and green development-openness-sharing. Drawing on these findings, this paper puts forward differentiated policy recommendations tailored to diverse development patterns and designs linked policy instruments from the perspective of cross-dimensional coordination, so as to provide empirical support for regional high-quality development.

Keywords

High-Quality Development, Factor Analysis, Principal Component Analysis, Multidimensional Scaling, Canonical Correlation Analysis

Copyright © 2026 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. 研究背景及意义

近年来, 高质量发展评价逐渐成为经济学、管理学和统计学等领域的重要研究方向。现有研究主要围绕高质量发展的内涵界定、指标体系构建以及区域比较评价展开。魏婕、任保平[1]从经济增长质量测度角度开展区域比较研究, 为高质量发展的量化评价提供了经验基础; 朱启贵[2]系统讨论了高质量发展的内涵、统计指标与评价体系, 指出多维指标的选择与综合仍需要进一步论证。相关研究表明, 高质量发展具有显著的多维属性, 仅依赖单一指标难以全面刻画区域发展水平。

在区域高质量发展评价方面, 任保平[3]从区域经济发展路径与政策转型角度讨论了区域协调、创新与绿色发展的实现机制; 金碚[4]强调高质量发展的系统性与可持续性, 认为发展质量需要通过多方面因

素的协同改善加以体现；成功、刘筱涵等[5]基于多元统计分析研究区域经济高质量发展优化策略，发现创新能力、产业结构升级以及开放水平是影响区域发展质量的重要因素。总体而言，已有研究在评价指标体系和综合评价方法方面取得了较为丰富的成果，但对于多种多元统计方法之间的交叉验证以及不同评价维度之间协同关系的系统分析仍有进一步拓展空间。

总体来看，现有研究在高质量发展评价指标体系构建和综合评价方法应用方面取得了较大进展，但多数研究仍以单一评价模型为主，对于因子分析、主成分分析、多维尺度分析和典型相关分析等多种统计方法的综合运用相对不足，对不同评价维度之间关联关系的研究也有待进一步深化。因此，本文以我国 31 个省级行政区(不含港澳台地区)为研究对象，综合运用多元统计分析方法，对区域经济高质量发展水平进行系统评价，以期为促进区域协调发展和制定相关政策提供参考依据。

2. 数据介绍

2.1. 数据来源及介绍

2.1.1. 数据来源

数据来自 EPS 数据平台中国宏观经济数据库、中国高技术产业数据库、中国卫生数据库、中国能源数据库、中国对外经济数据库以及中国统计年鉴。

2.1.2. 指标体系

在数据库中搜集并整理了 2021 年、2022 年 31 个省级行政区(不含港澳台地区)的有关中国高质量的各项指标数据，并对数据作了预处理。将有关高质量发展的指标总结为 5 个一级指标：创新、协调、绿色、开放、共享，其中每个一级指标下有二级指标，将有关变量总结，见表 1。

Table 1. High-quality development indicator system of a region
表 1. 某地区高质量发展指标体系

一级指标	二级指标
创新	GDP 增长率
	研发投入强度
	投资效率
	技术交易活跃度
	研究与试验发展(R&D)人员(人)
协调	专利申请数(件)
	需求结构
	城乡结构
	产业结构
	政府债务负担
绿色	每千人中卫生技术人员(人)
	第三产业增加值(亿元)
	能源消费弹性系数
	单位产出的废水
	单位产出的废气
	二氧化硫排放量(万吨)
	城市每万人拥有公共交通工具(标台)

续表

开放	对外贸易依存度
	外商投资比重
	市场化程度
	合同数(份)
非金融类对外直接投资流量(万美元)	
共享	劳动者报酬比重
	居民收入增长弹性
	城乡消费差距
	民生性财政支出比重

2.2. 数据预处理

为了确保高质量发展的指标数据分析结果的准确性和科学性，需要对数据进行预处理。主要包括缺失值处理、异常值检测、标准化以及多重共线性检验等步骤，具体如下：

在实际数据收集过程中，可能存在部分指标数据缺失的情况。为了保证分析的完整性与结果的可靠性，采用以下方法对缺失值进行处理：对于少量缺失值(缺失率 <10%)：采用均值填充法；对于部分指标缺失值较多(缺失率 10%~30%)：使用多重插补法结合数据分布情况进行插补，尽可能减小偏差；采用箱线图法检测并处理异常值；后续需要进行因子分析和主成分分析，采用 Z-Score 标准化方法对数据进行标准化处理。

3. 描述性统计

为了全面了解高质量发展指标数据的基本特征，对所有变量进行描述性统计分析，主要包括均值、中位数、标准差、偏度和峰度等，见表 2。描述性统计为后续的因子分析和主成分分析等提供了基础参考。

Table 2. Descriptive statistics of each indicator
表 2. 各指标描述性统计表

	Mean	SD	Median	Skewness	Kurtosis
GDP 增长率	0.96	0.06	0.97	-2.30	8.51
研发投入强度	0.01	0.01	0.01	0.33	-0.82
投资效率	0.51	1.03	0.72	-2.25	5.89
研究与试验发展人员	276801.90	305699.04	188413.00	1.72	2.42
专利申请数	12823.35	27535.11	4493.00	3.85	15.57
技术交易活跃度	0.03	0.04	0.02	2.58	7.88
需求结构	0.31	0.08	0.32	-0.33	-0.31
城乡结构	0.64	0.12	0.63	0.25	1.47
政府债务负担	0.20	0.16	0.16	1.97	4.51
每千人中卫生技术人员	8.19	1.15	8.03	2.55	8.97

续表

第三产业增加值	13842.88	11547.41	9564.03	1.41	1.50
产业结构	54.34	8.21	52.10	1.97	4.31
能源消耗弹性系数	2.00	8.89	0.37	0.36	0.45
单位产出的废水	4.43	4.91	4.46	0.85	1.88
二氧化硫排放量	8.86	5.76	8.49	0.30	-0.65
城市每万人拥有公共交通工具	10.97	2.20	10.75	0.21	0.20
单位产生的废气	0.00	0.00	0.00	1.65	2.77
对外贸易依存度	307.74	296.53	228.94	1.51	2.13
合同数	282.10	508.18	79.00	2.31	4.81
非金融类对外直接投资流量	282899.94	416153.07	120585.00	1.71	1.64
外商投资比重	905.73	704.44	650.07	0.90	-0.35
市场化程度	7.90	2.79	8.10	-0.21	-0.37
劳动者报酬比重	0.51	0.05	0.51	0.28	-0.46
居民收入弹性	0.04	0.08	0.02	4.18	18.20
城乡消费差距	1.64	0.30	1.58	0.84	0.06
民生性财政支出比重	0.48	0.05	0.48	-0.47	0.27

通过对各一级指标下的二级指标进行汇总统计，可以得出以下主要特征：

GDP 增长率、研发投入强度和专利申请数等，其均值和标准差表明区域间创新能力差异显著，部分变量如专利申请数的极值较大，表明存在区域创新成果的显著集中；需求结构、城乡结构和第三产业增加值的均值和中位数接近，但城乡消费差距和每千人卫生技术人员数的标准差较高，表明部分地区的资源配置和发展均衡性仍存在差异；单位产出的废气排放量和二氧化硫排放量的均值较低，但分布右偏，反映部分地区的环境治理成果显著，而部分高污染地区对总体平均值的拉升较大；对外贸易依存度和合同数等，其标准差较高，最大值与最小值的差距显著，显示了地区间对外经济开放程度的不均衡性；劳动者报酬比重和居民收入增长弹性的分布相对集中，说明大部分地区的共享发展特征较为一致，但城乡消费差距的分布存在长尾，反映部分区域共享发展效果不显著。

4. 因子分析

因子分析法是一种降维技术，通过提取公共因子来减少变量数量，同时保留原始数据的主要信息。在区域经济高质量发展评价中，因子分析法能够将多个经济指标综合为少数几个公共因子，从而简化评价过程。因子分析法的应用步骤通常包括数据标准化、因子提取、因子旋转和因子解释等环节。

在因子分析之前，先做 KMO 检验和 Bartlett 检验，检验数据是否适合展开因子分析。在对含有 28 个变量的数据做检验时，发现 KMO 的值小于 0.5，故不适合因子分析。为了深入研究区域经济发展的多维特征，在每个一级指标下删除 2 个二级指标，重新做 KMO 检验和 Bartlett 检验，得到下面的结果，见表 3。

由表 3 可以看出，KMO 的值为 0.565，大于 0.5，故适合展开因子分析；Bartlett 检验的 P 值小于 0.05，能够展开主成分分析。

Table 3. Results of KMO and Bartlett test
表 3. KMO 和 Bartlett 检验结果

检验	MSA 值	卡方(χ^2)值	自由度(df)	p 值
KMO 检验	0.565			
Bartlett 球形度检验		320.975	153	5.618×10^{-14}

Table 4. Factor loading matrix
表 4. 因子载荷矩阵

指标名称	ML4	ML1	ML5	ML2	ML3
GDP 增长率	-0.185	-0.119	-0.512	-0.111	0.129
研发投入强度	-0.190	0.669	0.279	-0.145	0.265
投资效率	-0.163	—	—	0.538	—
技术交易活跃度	0.854	0.124	0.209	—	—
需求结构	-0.184	0.168	—	-0.104	0.960
城乡结构	0.581	0.512	0.245	-0.197	-0.108
政府债务负担	-0.251	-0.654	-0.430	0.274	-0.305
产业结构	0.908	0.289	—	—	0.173
能源消耗弹性系数	-0.557	—	—	0.165	—
单位产出的废水	-0.171	—	0.125	0.249	-0.131
单位产出的废气	-0.212	-0.253	—	—	-0.423
对外贸易依存度	0.399	0.879	-0.188	—	0.162
外商投资比重	0.647	0.441	-0.120	—	—
市场化程度	0.374	0.783	0.317	-0.131	0.145
劳动者报酬比重	—	-0.298	—	0.925	0.188
居民收入弹性	—	—	—	-0.212	-0.239
城乡消费差距	—	—	-0.490	0.203	—
民生性财政支出比重	-0.111	—	0.860	0.225	0.171

表 4 展示了各变量的因子载荷矩阵。可以发现，劳动者报酬比重对因子 ML2 的载荷较高，技术交易活跃度对因子 ML4 的载荷较高，研发投入强度对因子 ML1 的载荷较高，民生性财政支出比重对因子 ML5 的载荷较高，需求结构对因子 ML3 的载荷最高。上述结果表明，不同指标在公共因子上的贡献存在明显差异，为后续识别区域发展模式提供了依据。

Table 5. Variance explanation results
表 5. 方差解释结果

因子	特征根	方差解释率	累计方差解释率	提取方差解释率	累计方差提取解释率
ML4	3.2190	17.8833%	17.8833%	28.9504%	28.9504%
ML1	3.0283	16.8239%	34.7072%	27.2356%	56.1859%
ML5	1.7926	9.9589%	44.6666%	16.1221%	72.3080%
ML2	1.5589	8.6608%	53.3269%	14.0205%	86.3285%
ML3	1.5201	8.4451%	61.7772%	13.6771%	100.0000%

由表 5 可以看出每个因子的贡献率和累计贡献率,取因子 ML4、ML1、ML5 时的累计解释比例达到 72.31%,加上因子 ML2 时,累计解释比例达到 86.33%,解释比例较高,效果较好。

采用主成分分析法提取公共因子,共提取 5 个公共因子,五个因子累计方差解释率达到 86.33%。为简化变量归属识别,采用最大方差正交旋转处理因子载荷矩阵,并依据各变量在因子上的载荷绝对值大小界定各公共因子的经济内涵,具体概括为以下 5 个因子:

因子 1: 产业创新开放综合因子

技术交易活跃度(0.854)、产业结构高级化(0.908)和对外贸易依存度(0.879)在该因子上的载荷较高,分别对应创新成果转化、产业结构升级和对外经贸联系,共同刻画区域产业发展与开放竞争能力。该因子的方差贡献率较高,说明产业创新与开放是解释省级行政区之间高质量发展差异的重要维度。

因子 2: 民生共享发展因子

劳动者报酬比重(0.925)和民生性财政支出比重(0.860)在该因子上的载荷最高,均属于共享发展维度,主要反映居民收入分配和民生保障投入状况。该因子相对独立地反映民生共享水平,说明经济规模和产业竞争力较强的地区并不必然在居民共享方面同步占优。

因子 3: 内需结构协调因子

需求结构指标载荷高达 0.960,单独主导该公共因子,代表国内消费、投资、出口的需求均衡格局,侧重刻画宏观内需市场协调水平,是协调发展维度的细分分支。

因子 4: 创新与城乡协调因子

核心载荷指标为研发投入强度(0.669)、市场化程度(0.783)、城乡结构(0.512),综合反映区域创新资金投入、市场化改革深度、城乡发展均衡程度,聚焦创新驱动与城乡协调协同发展特征。

因子 5: 外资绿色约束因子

外商投资比重(0.647)和单位产出废气指标在该因子上的载荷较高,反映开放程度与环境约束之间的关联特征。绿色指标未形成完全独立的公共因子,说明在本研究样本中,绿色发展与开放、产业结构等维度存在一定程度的交叉联系。

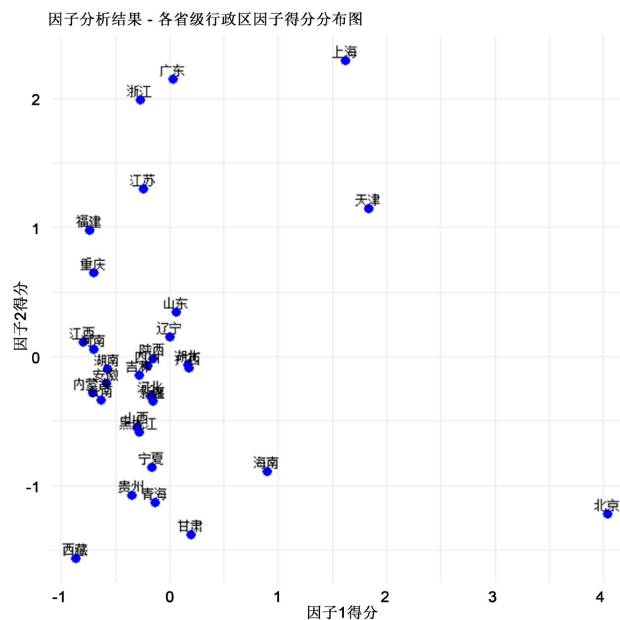


Figure 1. Distribution diagram of factor scores by provincial-level administrative region

图 1. 各省级行政区因子得分分布图

基于旋转后因子得分函数测算全国各省级行政区因子 1、因子 2 得分，绘制二维散点分布图，见图 1。横轴为产业创新开放综合因子得分，纵轴为民生共享发展因子得分，点位坐标代表省级行政区在两大核心维度的发展水平，依据点位分布可将全国省级行政区划分为四类发展模式：

上海、广东、浙江、江苏、天津五个省级行政区集中分布于散点图右上方，因子 1、因子 2 得分均显著高于全国平均水平。以上海表现最为突出，两大因子得分均位居全国首位，产业创新开放实力强劲，同时民生分配、城乡均衡发展同步领跑，形成创新、开放、民生协同共进的高质量均衡发展模式；广东、浙江、江苏、天津紧随其后，双维度同步具备优势，是全国高质量发展标杆区域。

北京为图中孤立极值点，因子 1 得分断层领先全国所有省级行政区，创新资源、产业开放水平具备绝对优势；但因因子 2 得分显著为负，民生共享、城乡均衡维度存在明显短板。反映首都资源高度集聚，但收入分配、城乡发展差距突出，呈现“强创新、弱共享”的单一化发展结构，与上海均衡发展模式形成鲜明对比。

福建、重庆、山东、辽宁等省级行政区点位集中于坐标轴原点附近，两大因子得分均趋近于全国平均水平。福建、重庆作为沿海及中西部中心城市，产业开放具备微弱优势，但民生保障力度不及东部发达省级行政区；山东、辽宁北方传统工业省级行政区产业基础尚可，但民生均衡短板明显。

黑龙江、山西、宁夏、海南、西藏、贵州、青海、甘肃等省级行政区集中于散点图左下方，因子 1、因子 2 得分均显著低于全国均值。西部偏远省级行政区中，西藏、青海、甘肃、贵州两大发展维度基础均十分薄弱，产业创新动能不足、民生保障投入缺口较大；黑龙江、山西等北方传统资源型省级行政区产业转型升级缓慢，民生发展滞后；海南存在结构特殊性，依托旅游外贸使因子 1 小幅转正，但民生共享维度短板突出，发展结构单一。

从整体相关性规律来看，散点分布呈现弱正向相关特征：多数省级行政区产业创新开放(因子 1)得分越高，民生共享均衡(因子 2)得分同步提升，证明产业对外开放与民生改善存在天然协同效应；但北京的特殊分布也说明，二者不存在必然同步关系，城市发展定位、资源分配政策会打破二者协同规律。

5. 主成分分析

主成分分析是经典线性降维多元统计方法，核心目标是在尽可能保留原始指标信息的前提下，将存在多重共线性的多维度变量，线性组合为少数互不相关的综合主成分。本文构建的高质量发展评价体系包含多个存在量纲差异、内在关联的二级指标，直接综合评价易出现信息重叠、权重主观化问题。借助主成分分析能够依据数据自身方差贡献自动分配权重，通过提取累计方差贡献率达标主成分实现指标简化，依托各主成分加权测算各省级行政区高质量发展综合得分，客观划分区域发展层级，直观揭示各省级行政区高质量发展的差距与结构特征，为后续区域发展水平排序、特征识别提供量化支撑。

Table 6. Variance contribution rate and cumulative contribution rate of principal components
表 6. 主成分贡献率和累计贡献率

主成分	标准差	方差贡献率	累计方差贡献率
PC1	3.014238	33.6500%	33.6500%
PC2	1.937906	13.9090%	47.5600%
PC3	1.587987	9.3400%	56.8990%
PC4	1.401903	7.2790%	64.1780%
PC5	1.303987	6.2980%	70.4760%
PC6	1.132047	4.7460%	75.2220%

续表

PC7	1.012338	3.7960%	79.0180%
PC8	0.994920	3.6660%	82.6840%
PC9	0.910155	3.0680%	85.7520%
PC10	0.835108	2.6830%	88.4350%

由表 6 可知，标准差较高的主成分解释的数据变异更多，PC1 的标准差为 3.014238，远高于后续主成分，这表明它是最重要的主成分；PC1 的方差贡献率为 33.6%，PC2 为 13.9%。这说明 PC1 是最具代表性的主成分；前 6 个主成分累计贡献率就达到 75.22%。

通常情况下，当累计贡献率达到 75%时，说明主成分可以较好地代表原始数据的信息，在表 6 中可以看出，当选择 6 个主成分时，累计贡献率就达到 75%以上，说明 6 个主成分就可以较好地代表原始数据的信息，可以认为是主要的特征维度，其后主成分的重要性显著降低。

Table 7. Ranking results of comprehensive scores of provincial-level administrative regions
表 7. 省级行政区综合得分排名结果

排名	省级行政区	综合得分
1	广东	2.7144
2	江苏	2.0911
3	浙江	1.7777
4	上海	1.4368
5	北京	1.3177
6	天津	0.6519
7	山东	0.5982
8	湖北	0.3590
9	福建	0.2832
10	安徽	0.2055
11	河北	0.1086
12	辽宁	0.0478
13	湖南	0.0436
14	山西	0.0260
15	江西	0.0251
16	重庆	0.0122
17	四川	0.0029
18	河南	-0.0054
19	陕西	-0.1098
20	内蒙古	-0.2862
21	黑龙江	-0.7040
22	宁夏	-0.7073
23	甘肃	-0.7370
24	海南	-0.7617

续表

25	广西	-0.7688
26	吉林	-0.8089
27	云南	-0.9111
28	贵州	-0.9493
29	新疆	-1.0929
30	青海	-1.2704
31	西藏	-2.5890

表 7 展示了基于主成分分析结果计算的综合得分以及各省级行政区排名。

综合得分是每个省级行政区在主成分维度上的综合表现，通常是按照主成分的贡献率加权计算的。广东省得分为 2.7144，排名第一，说明其在综合指标上的表现最优；省级行政区按照综合得分降序排列，得分越高，排名越靠前。江苏省得分 2.0911，排名第二，仅次于广东；西藏得分-2.5890，排名第 31，表现相对较差。

6. 多维尺度分析

多维尺度分析(MDS)通过将省级行政区之间的相似性或差异性关系映射到低维空间，以点之间的距离直观展示不同地区在综合指标上的接近程度。本文基于标准化指标构建省级行政区间距离矩阵，并采用二维空间进行可视化。需要强调的是，MDS 坐标轴本身不具有预先确定的经济含义，其主要用于识别省级行政区之间的相对接近关系和空间分化特征。

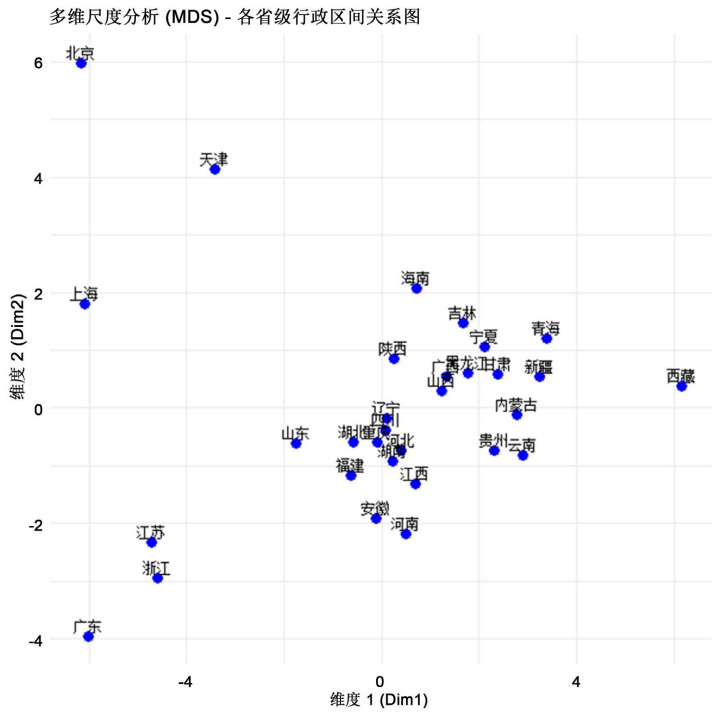


Figure 2. Multidimensional scaling analysis-relationship diagram among provincial-level administrative regions
图 2. 多维尺度分析 - 各省级行政区间关系图

如图 2，每个点代表一个省级行政区，点之间的距离反映了省级行政区在若干指标上的相似程度。图中如山东、河南、江苏等省级行政区分布较近，而青海、西藏、新疆则较为偏远。

江苏、浙江、广东：这些东部沿海省级行政区聚集在图的一个区域，说明它们在 GDP 增长率、研发投入强度等方面具有相似性；河南、安徽、湖南：这些中部省级行政区靠近另一个区域，表明它们在城乡结构、需求结构等指标上具有共同特征；青海、西藏、新疆：这些省级行政区在图上分布较远，表明其在能源消费弹性系数、绿色发展等指标上具有独特性。

7. 典型相关分析

典型相关分析用于考察两组变量之间的整体线性关联。本文将“创新 + 协调”划分为 X 组，将“绿色 + 开放 + 共享”划分为 Y 组，通过典型变量提取两组指标的综合信息，并利用 Wilks’ Lambda 检验典型相关关系的显著性。该方法能够从整体层面判断创新、协调与绿色、开放、共享之间是否存在协同变化，为后续设计跨维度的政策工具提供依据。

Table 8. Significance test of canonical correlation coefficients
表 8. 典型相关系数的显著性检验

典型变量	Wilks’ Lambda 统计量	近似卡方值	自由度 df1	自由度 df2	P 值
1	0.000	2.208	168	68.44	0.000
2	0.000	1.716	143	69.25	0.006
3	0.000	1.401	120	69.06	0.063
4	0.001	1.125	99	67.86	0.304
5	0.008	0.941	80	65.65	0.605
6	0.040	0.765	63	62.43	0.855
7	0.135	0.610	48	58.19	0.960
8	0.325	0.462	35	52.91	0.991
9	0.525	0.393	24	46.56	0.992
10	0.753	0.282	15	39.05	0.994
11	0.885	0.236	8	30.00	0.981
12	0.962	0.212	3	16.00	0.886

从表 8 可以看出，第 1 对典型变量的 Wilks’ Lambda=3.87273e-08，p 值显著小于 0.05，表明第 1 对典型变量是显著的，这对变量对 X 和 Y 组变量之间的关联具有统计学意义；第 2 对典型变量的 Wilks’ Lambda=2.70398e-06，p 值仍小于 0.05，表明第 2 对典型变量是显著的；第 3 对典型变量的 Wilks’ Lambda=6.94076e-05，p 值 =0.063，p 值大于 0.05，表明第 3 对及后续典型变量可能不显著；第 4 对及之后的典型变量的 Wilks’ Lambda 值逐步增大，p 值逐渐远大于 0.05；从第 4 对开始，典型变量不显著，说明后续变量在解释 X 和 Y 组变量之间的关联性时贡献较小。

由表 9 可知，第一对典型变量得分较高的地区主要包括北京、上海、江苏、浙江和广东，第二梯队包括福建、天津、山东、湖北和安徽；河南、西藏等地区得分相对较低。与主成分综合排名相比，典型变量得分能够从 X 组(创新、协调)与 Y 组(绿色、开放、共享)的协同角度补充刻画区域发展水平。

其中，北京在 X 和 Y 两组典型变量上的得分均位居首位，表明其创新、协调与绿色、开放、共享之

间具有较强的综合协同表现；上海、江苏、浙江和广东同样处于较高水平。山西、海南、河南、西藏等地区的典型变量得分相对较低，说明其在不同维度之间仍存在较明显的提升空间。需要指出的是，典型相关分析反映的是变量组之间的统计关联，而非因果关系。

Table 9. Scores of canonical variables
表 9. 典型变量得分

排名	省级行政区	X 典型变量得分	Y 典型变量得分
1	北京	0.8678	0.8241
2	上海	0.7932	0.7715
3	江苏	0.6541	0.6329
4	浙江	0.6217	0.6014
5	广东	0.5893	0.5726
6	福建	0.4126	0.3985
7	天津	0.3874	0.3692
8	山东	0.2591	0.2470
9	湖北	0.1635	0.1518
10	安徽	0.0972	0.0863
11	湖南	0.0418	0.0351
12	重庆	0.0165	0.0107
13	江西	-0.0327	-0.0412
14	河北	-0.0491	-0.0584
15	辽宁	-0.0563	-0.0647
16	宁夏	-0.0790	-0.0665
17	青海	-0.0834	-0.0958
18	四川	-0.0649	-0.0942
19	陕西	-0.1053	-0.0489
20	内蒙古	-0.1098	-0.1176
21	甘肃	-0.1195	-0.1401
22	黑龙江	-0.1276	-0.1325
23	广西	-0.1342	-0.1537
24	贵州	-0.1396	-0.1725
25	新疆	-0.1442	-0.1615
26	云南	-0.1454	-0.1651
27	吉林	-0.1628	-0.1793
28	山西	-0.1715	-0.1862
29	海南	-0.1937	-0.2014
30	河南	-0.2106	-0.2279
31	西藏	-0.2894	-0.2732

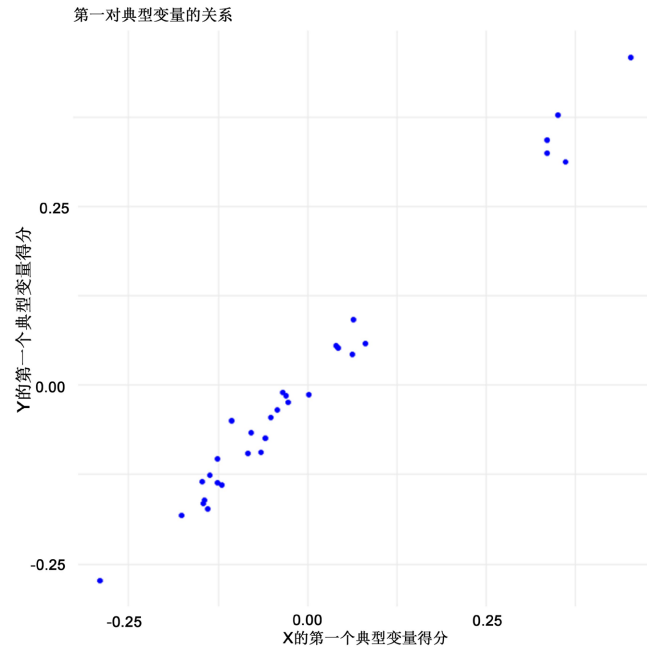


Figure 3. Relationship of the first pair of canonical variables
图 3. 第一对典型变量的关系

图 3 第一对典型相关变量的关系图中，横轴是 X 的第一个典型变量得分，表示 X 组变量通过第一个典型变量提取后的综合得分，X 组包括创新、协调两个指标；纵轴是 Y 的第一个典型变量得分，表示 Y 组变量通过第一个典型变量提取后的综合得分，Y 组变量包括绿色、开放、共享三个指标。

第一对典型变量的散点分布呈明显的正向关系，结合显著性检验结果可知，X 组与 Y 组之间存在较强的整体关联。由于 X 组由创新、协调构成，Y 组由绿色、开放、共享构成，因此一个地区在创新和协调方面的综合表现较高时，通常也伴随着绿色、开放和共享维度的较高表现。这一结果为构建跨维度联动政策提供了统计依据，但不意味着任一单项指标之间存在确定的因果关系。

进一步看，前两对典型变量通过显著性检验，说明本文所识别的协同关系并非仅由单一指标驱动，而是两组指标整体变化的结果。由此可将创新投入、产业结构调整、城乡协调、绿色转型、开放发展和民生改善视为相互联系的政策领域，在政策设计中应加强跨部门、跨区域和跨维度的联动。

8. 综合讨论

为避免单一统计方法可能造成的片面判断，本文进一步将因子分析、主成分分析、多维尺度分析和典型相关分析的结果进行交叉比对。因子分析主要用于识别区域发展的潜在结构，主成分分析用于给出综合水平排序，MDS 用于刻画省级行政区之间的发展结构相似性，典型相关分析则从变量组层面检验不同发展维度之间的协同关系。四种方法关注的对象不同，但在区域分化方向上具有较强的一致性。

首先，主成分分析与因子分析在高位地区的识别上具有较高一致性。主成分综合得分排名显示，广东、江苏、浙江、上海和北京位于前五位；因子散点图中，上海、广东、浙江、江苏和天津位于产业创新开放与民生共享两个维度的较高区域，北京则表现为产业创新开放维度显著突出而民生共享维度相对偏弱。由此可见，东部发达地区总体具有较强的综合发展基础，但其内部仍存在明显的结构差异：上海等地区表现为多维度协同优势，北京则呈现创新资源和产业开放优势突出、共享维度相对不足的特征。

其次，MDS 结果进一步验证了区域结构差异的客观存在。主成分分析给出的综合排名是单一尺度上

的线性排序，而 MDS 并不直接评价“高”或“低”，而是比较省级行政区之间发展结构的相似程度。图 2 显示，北京、上海、天津、广东、江苏、浙江和海南等省级行政区在二维空间中相对离散，说明这些地区具有较为独特的发展结构；河南、安徽、江西、湖北、湖南、重庆等省级行政区相互接近，说明其指标结构具有较强相似性；青海、新疆、宁夏、甘肃、贵州、云南等地区则表现出较为明显的西部异质性。因而，主成分分析的“水平差异”和 MDS 的“结构差异”共同构成了区域高质量发展的二维画像。

再次，典型相关分析为上述结果提供了维度协同层面的补充证据。第一、第二对典型变量的显著性检验均通过，说明由创新、协调构成的 X 组与由绿色、开放、共享构成的 Y 组之间存在显著的整体关联。典型变量得分靠前的地区与主成分综合排名靠前地区具有较高重合度，例如北京、上海、江苏、浙江、广东等均处于较高位置。这说明综合发展水平较高的地区通常也具有较强的跨维度协同能力，而排名靠后的地区则往往面临多维度同步提升的任务。

综合而言，本文不同方法的结论并非相互替代，而是形成互补：因子分析回答“区域高质量发展由哪些潜在维度构成”，主成分分析回答“各地区总体水平如何排序”，MDS 回答“哪些地区在发展结构上更相似或更特殊”，典型相关分析回答“不同维度之间是否存在整体协同”。因此，仅依据综合得分进行区域比较容易忽视内部结构，而将四类结果结合起来，可以形成“水平 - 结构 - 空间 - 协同”四个层面的立体评价框架(表 10)。

Table 10. Comparison of comprehensive portrait comparison of high-quality development among provincial-level administrative regions

表 10. 各省级行政区高质量发展综合画像比较

省级行政区	主成分排名	综合得分	因子发展模式	MDS 空间特征	典型变量排名
广东	1	2.7144	均衡协同型	左下沿海离散	5
江苏	2	2.0911	均衡协同型	左下沿海离散	3
浙江	3	1.7777	均衡协同型	左下沿海离散	4
上海	4	1.4368	均衡协同型	左上高位离散	2
北京	5	1.3177	创新优势型	左上高位离散	1
天津	6	0.6519	均衡协同型	左上高位离散	7
山东	7	0.5982	中等协调型	中部近邻集聚	8
湖北	8	0.3590	中等协调型	中部近邻集聚	9
福建	9	0.2832	中等协调型	中部近邻集聚	6
安徽	10	0.2055	中等协调型	中部近邻集聚	10
河北	11	0.1086	中等协调型	中部近邻集聚	14
辽宁	12	0.0478	中等协调型	中部近邻集聚	15
湖南	13	0.0436	中等协调型	中部近邻集聚	11
山西	14	0.0260	综合提升型	中部近邻集聚	28
江西	15	0.0251	中等协调型	中部近邻集聚	13
重庆	16	0.0122	中等协调型	中部近邻集聚	12
四川	17	0.0029	中等协调型	中部近邻集聚	18
河南	18	-0.0054	中等协调型	中部近邻集聚	30
陕西	19	-0.1098	中等协调型	中部近邻集聚	19

续表

内蒙古	20	-0.2862	中等协调型	右侧西部集聚	20
黑龙江	21	-0.7040	综合提升型	中部近邻集聚	22
宁夏	22	-0.7073	综合提升型	右侧西部集聚	16
甘肃	23	-0.7370	综合提升型	右侧西部集聚	21
海南	24	-0.7617	开放特色型	右上高位离散	29
广西	25	-0.7688	综合提升型	右侧西部集聚	23
吉林	26	-0.8089	综合提升型	中部近邻集聚	27
云南	27	-0.9111	综合提升型	右侧西部集聚	26
贵州	28	-0.9493	综合提升型	右侧西部集聚	24
新疆	29	-1.0929	综合提升型	右侧西部集聚	25
青海	30	-1.2704	综合提升型	右侧西部集聚	17
西藏	31	-2.5890	综合提升型	右侧高位离散	31

注：表中“因子发展模式”为依据图 1 中因子 1 (产业创新开放)与因子 2 (民生共享)相对位置形成的综合归纳标签，并非重新进行聚类分析；“MDS 空间特征”为依据图 2 点位相对位置进行的描述性归纳，主要用于辅助判断省级行政区之间的发展结构相似性。主成分排名与典型变量排名分别对应表 7 和表 8。

9. 结论和建议

9.1. 结论

本文运用因子分析、主成分分析、多维尺度分析和典型相关分析等多元统计方法，对我国 31 个省级行政区(不含港澳台地区)的区域经济高质量发展水平进行了综合评价，并进一步从发展水平、结构差异、空间相似性和维度协同四个方面进行交叉验证，得到以下主要结论。

(1) 我国区域高质量发展存在显著的省级行政区之间差异，东部地区总体处于较高水平。主成分分析显示，广东、江苏、浙江、上海和北京位居综合得分前五位，而西藏、青海、新疆、贵州、云南等地区排名相对靠后。由此可见，我国区域高质量发展在综合水平上仍存在明显梯度差异。

(2) 产业创新与开放是解释区域发展差异的重要维度，但不同地区的优势结构并不相同。因子分析中，产业创新开放综合因子的方差贡献率较高，技术交易活跃度、产业结构高级化和对外贸易依存度具有较高载荷；与此同时，民生共享因子相对独立，北京等地区表现出创新优势与共享短板并存的特征，说明提高综合发展水平还需要关注内部结构均衡。

(3) 不同地区具有明显的异质化发展模式。因子得分显示，上海、广东、浙江、江苏、天津呈现产业创新开放与民生共享同步较强的均衡协同特征；北京表现为创新优势型；福建、山东、湖北、安徽等地区整体处于中等协调状态；西部及部分东北地区的多个因子得分相对较低，表现为综合提升型。海南则呈现开放特色较强而共享维度相对偏弱的结构特征。

(4) 主成分分析与 MDS 从不同角度验证了区域差异。主成分分析反映地区总体水平，MDS 则揭示地区之间发展结构的相似性。北京、上海、天津、广东、江苏、浙江和海南等地区在 MDS 空间中相对离散，而河南、安徽、江西、湖北、湖南、重庆等地区较为接近，说明省级行政区之间不仅存在发展水平差距，也存在发展结构差异。

(5) 创新、协调与绿色、开放、共享之间存在显著的整体协同关系。典型相关分析表明，前两对典型变量通过显著性检验，且第一对典型变量呈明显正向关系；典型变量排名靠前的地区与综合排名靠前地

区具有较高重合度。这表明不同发展维度之间具有较强的协同变化特征，为实施跨维度联动政策提供了统计依据，但这种统计关联不应被解释为严格的因果关系。

9.2. 建议

结合本文的综合画像结果，区域政策应由“一套方案”转向分类施策，并将不同统计方法识别出的短板与优势落实到具体政策工具中。

(1) 针对均衡协同型地区，重点推动由“综合领先”向“多维引领”转变。上海、广东、浙江、江苏和天津在因子1、因子2上均处于较高水平，且主成分综合排名靠前，应减少单纯扩大投入的政策导向，重点提升科技成果转化效率、产业链高端化和绿色技术应用，同时利用其开放平台和产业基础向周边地区开展技术、人才和产业合作，发挥区域辐射带动作用。

(2) 针对创新优势型地区，重点补齐民生共享与协调发展短板。北京主成分排名第5、典型变量排名第1，但因因子1显著为正、因子2为负，表现出创新开放优势与民生共享维度不完全匹配的特征。因此，应在保持研发和创新优势的同时，更加关注劳动者报酬、公共服务均等化和城乡协调等领域，通过创新成果产业化和就业扩容将科技优势进一步转化为居民收入与公共服务改善。

(3) 针对中等协调型地区，重点实施“产业升级-创新投入-民生改善”联动政策。福建、山东、湖北、安徽、湖南、重庆、江西、河北、辽宁等地区综合得分处于中上游或中游，因子得分多接近全国平均水平。此类地区应围绕产业结构高级化提高研发投入效率，促进传统产业数字化、绿色化改造，同时将产业升级带来的增值收益更多转化为就业和居民收入增长，避免产业发展与民生改善脱节。

(4) 针对综合提升型地区，重点解决多维度共同短板。西藏、青海、新疆、贵州、云南、甘肃、宁夏、广西、吉林、黑龙江等地区主成分排名相对靠后，因子1和因子2整体偏低，MDS又呈现较强的区域异质性。因此，应优先加强交通、数字基础设施和公共服务等发展基础，在此基础上因地制宜发展特色产业，推动资源型产业绿色转型，增加研发和技术成果转化投入，并通过教育、医疗和社会保障投入提高居民共享水平。

(5) 针对开放特色型地区，重点提高开放优势向共享效应的转化效率。海南在MDS空间中表现为相对独立的高位点，在因子1上相对较高但因子2偏低，说明其开放和特色产业优势尚未完全转化为民生共享优势。应进一步推动旅游、现代服务业与先进制造业、数字经济等产业融合，完善就业和收入分配机制，提高产业链本地化程度，使开放发展更有效地带动居民收入和公共服务改善。

(6) 构建以典型相关关系为依据的跨维度联动政策工具。典型相关分析将创新、协调组成的X组与绿色、开放、共享组成的Y组联系起来，前两对典型变量均达到显著水平。因此，政策不宜将科技创新、产业升级、绿色转型、开放发展和民生保障割裂实施，可建立“创新投入-产业升级-绿色技术-开放合作-民生共享”的联动机制。例如，将科技创新项目与绿色技术改造、产业链招商和就业目标同步设置考核指标；将重大产业项目的准入与能耗、污染排放和本地就业贡献联动考察；将区域开放平台建设与技术转移、人才培养和公共服务改善结合起来，以一个关键环节的改善带动多个维度同步提升。

总体来看，区域高质量发展具有明显的多维性和异质性。不同地区应根据自身发展模式和结构性短板实施差异化政策，同时通过创新、协调、绿色、开放和共享之间的联动机制提高政策协同效应。

参考文献

- [1] 魏婕, 任保平. 中国各地区经济增长质量指数的测度及其排序[J]. 经济学动态, 2012(4): 27-33.
- [2] 朱启贵. 建立高质量发展评价体系[EB/OL]. 文汇报. <https://saif.sjtu.edu.cn/show-58-1429.html>, 2018-06-18.
- [3] 任保平. 新发展阶段我国区域经济高质量发展的理论逻辑、实践路径与政策转型[J]. 四川大学学报(哲学社会科学版), 2021(4): 1-12.

学版), 2023(3): 81-90, 193.

- [4] 金磊. 关于“高质量发展”的经济学研究[J]. 中国工业经济, 2018(4): 5-18.
- [5] 成功, 刘筱涵. 基于多元统计分析的区域经济高质量发展优化策略研究[C]//重庆市大数据和人工智能产业协会, 西南大学, 重庆工商大学, 重庆建筑编辑部. 人工智能与经济工程发展学术研讨会论文集. 2025.