

基于 Lasso 回归模型的 青岛海洋经济和海洋产业分析

张贺, 范梦萱

(中国海洋大学 青岛 266000)

摘要:为促进青岛海洋产业结构优化和海洋经济高质量发展,文章构建评价指标体系,采用 Lasso 回归模型分析对青岛海洋经济发展产生重要影响的海洋产业,并提出发展建议。研究表明:青岛海洋经济发展水平评价指标体系包括海洋渔业、海洋化学工业、海洋进出口总额、海洋环境、海洋交通运输业和滨海旅游业 6 个大类和 20 个特征变量;采用 10 折交叉验证确定调节系数的最优取值,并通过求解 Lasso 回归系数最终筛选 10 个特征变量进入回归模型;根据回归系数估计值,国内旅客数量的重要性最强,海洋渔业产值的重要性较强,而货物吞吐量的重要性较弱;Lasso 回归模型的预测准确率、模型泛化能力和可信度均较高,属于稳健的稀疏模型;未来青岛应进一步发挥滨海旅游业和海洋渔业等的产业优势,着力发展海洋交通运输业,保护海洋生态环境以及加大海洋科技研发投入。

关键词:海洋产业;海洋经济;交叉验证;模型泛化;调节系数

中图分类号:F062.9;P74

文献标志码:A

文章编号:1005-9857(2022)08-0022-07

The Marine Economy and Marine Industry in Qingdao Based on Lasso Regression Model

ZHANG He, FAN Mengxuan

(Ocean University of China, Qingdao 266000, China)

Abstract: In order to promote the optimization of marine industry structure and the high-quality development of marine economy in Qingdao, this paper established the evaluation index system, analyzed the marine industry which had an important impact on the development of marine economy in Qingdao by using Lasso regression model, and put forward development suggestions. The results showed that the evaluation index system of Qingdao's marine economic development level included 6 categories and 20 characteristic variables, which were marine fishery, marine chemical industry, total marine import and export, marine environment, marine transportation and coastal tourism. The optimal value of the adjustment coefficient was determined by 10 fold cross validation, and 10 characteristic variables were selected into the regression model by solving the Lasso

收稿日期:2021-12-14;修订日期:2022-07-22

基金项目:教育部人文社会科学规划基金项目“基于高维高频复杂数据的非等间隔日内市场风险评估与应用研究”(18YJA790113).

作者简介:张贺,硕士研究生,研究方向为金融学

regression coefficient. According to the estimated value of regression coefficient, the number of domestic passengers was the most important, the value of marine fishery output was important, and the value of cargo throughput was the least important. Lasso regression model had high prediction accuracy, model generalization ability and reliability, and belonged to robust sparse model. In the future, Qingdao should give full play to the industrial advantages of coastal tourism and marine fishery, focus on the development of marine transportation industry, protect the marine ecological environment and increase the investment in marine science and technology research and development.

Keywords: Marine industry, Marine economy, Cross validation, Model generalization, Adjustment coefficient

0 引言

我国是海洋大国,海洋资源极为丰富。随着海洋开发利用日益受到重视,推动海洋经济发展已成为全国共识。在建设海洋强国的背景下,实现海洋经济高质量发展和完成海洋产业转型升级对青岛整体经济的发展具有重要意义。

依海而生,向海图强。青岛地处山东东南沿海的丘陵地带,不适合发展大型机械化种植业;由于三面环海,海洋产业一直是青岛经济发展的重要组成部分,“蓝色硅谷”、西海岸新区和红岛经济区等建设对青岛海洋经济发展的引领作用日益显现^[1]。2021年上半年青岛海洋生产总值增长 21.9%,占 GDP 比重为 32%,海洋第一产业、第二产业和第三产业分别增长 7.7%、19.7%和 25.2%。目前青岛发展海洋经济的优势突出体现在产业体系完整、产业集群众多、基本实现产业全面覆盖、科研机构云集和港口辐射能力强,尤其在海水养殖业和滨海旅游业方面具有得天独厚的优势。

已有学者采用多部门经济模型测度不同变量对海洋生产总值的贡献度,如王端岚^[2]研究福建海洋产业结构变化对海洋经济增长的促进作用,并提出海洋生态环境保护和资源可持续利用的重要价值。格兰杰因果方法在海洋产业结构和海洋经济发展的研究方面也有应用,结果表明二者具有相互促进的关系。部分学者从海洋资源、海洋生态环境、海洋产业、海洋文化和海洋制度等方面构建评价指标体系并加以分析,其中海洋资源和海洋生态环境属于基础性因素,海洋产业和海洋文化等属于

提升性因素^[3-4]。

目前我国海洋产业体系逐渐成熟,但也出现产业发展同质化、科技支撑能力弱和科研成果转化难等问题,阻碍海洋经济的高效发展^[5]。为实现青岛海洋经济的高质量发展,助力海洋强国目标的实现,本研究从海洋产业结构转型升级的角度出发,选取相关评价指标,采用 Lasso 回归模型分析对青岛海洋经济发展产生重要影响的海洋产业,并提出发展建议。

1 研究方法

当多个解释变量对被解释变量进行预测时,通常采用最小二乘估计法进行线性回归,但该方法对于基础回归存在一定的局限性。①出于最小化均方误差的目的而追求较低的偏差,导致方差变大,使得模型的泛化能力较差;②在线性回归中通常保留大量的解释变量,导致模型的可解释性降低。本研究应用具有较强影响力的解释变量子集,因此选择 Lasso 回归模型,从而克服普通最小二乘估计法的上述不足^[6]。一方面,Lasso 回归模型通过对线性回归系数施加约束和惩罚,将某些噪音变量的系数估计值压缩至 0,从而将其剔除并筛选出重点变量;另一方面,由于变量个数的减少,Lasso 回归模型往往表现出很强的泛化能力,预测效果通常优于线性回归。

通过求解目标函数的最小值得到 Lasso 回归模型的系数,目标函数的表达式为:

$$f(x) = \sum_{i=1}^n (y_i - \sum_{j=1}^k x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^k |\beta_j| = \text{RSS} + \lambda \sum_{j=1}^k |\beta_j|$$

式中： n 表示数据样本量； p 表示数据特征个数； y_i 表示第 i 个被解释变量的样本值； x_{ij} 表示第 i 个解释变量的第 j 个样本值 ($i=1,2,\dots,10; j=1,2,\dots,p$)； β_j 表示 p 维特征矩阵的待估系数； λ 表示调节系数；RSS 表示残差平方和。

除线性回归要求 RSS 达到最小外，Lasso 回归模型增加回归系数的惩罚项：

$$L_1 = \|\beta\|_1 = \sum |\beta_j|$$

式中： L_1 表示回归系数 $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ 的惩罚项^[7]。

λ 控制对回归系数的惩罚力度。随着 λ 的逐渐增大，系数估计值逐渐减小。当 λ 足够大时，为使目标函数达到最小值， L_1 能够将某些变量的系数估计值强制设定为 0，从而将这些不重要的变量从模型中剔除。因此，Lasso 回归模型能够通过 λ 的取值来确定留在模型中的变量个数，进而得到稀疏模型，使模型更易于解释。

通常采用坐标下降法求解 Lasso 回归系数，该方法可在每步迭代中沿某个特征（坐标）的方向进行搜索，通过循环使用不同特征以达到目标函数的局部最小值。具体算法包括 3 个步骤：

(1) 对 j 维向量 β 随机取初值，记为 $\beta^0 = (\beta_1^{(0)}, \beta_2^{(0)}, \dots, \beta_j^{(0)})$ ，上标括号里的数字代表当前迭代的轮数，此时为初始轮数 0。

(2) 对于第 k 轮迭代，从第一个维度 $\beta_1^{(k)}$ 开始，到第 j 个维度 $\beta_j^{(k)}$ 为止。每次迭代仅更新 β 的 1 个维度，即将该维度视为变量，将剩下的 $(n-1)$ 个维度视为常量，通过最小化目标函数（求导或一维搜索）找到该维度对应的新值，得到 $\beta^k = (\beta_1^{(k)}, \beta_2^{(k)}, \dots, \beta_j^{(k)})$ 。

(3) 检查向量 β^k 和 β^{k-1} 在各维度的变化情况。如果其在所有维度上的变化均足够小并小于设定的阈值，那么 β^k 即最终结果；否则转入步骤 2，继续第 $(k+1)$ 轮迭代。

λ 的取值直接影响变量筛选的最终结果，对于 Lasso 回归至关重要。在确定 λ 的最优取值时，可采用赤池信息准则 (AIC)、贝叶斯信息准则 (BIC) 和交叉验证 (CV)。其中，AIC 和 BIC 同时考虑模型的拟合程度，并对模型中的变量系数进行惩罚；二

者的不同之处在于，BIC 将未知变量系数的惩罚权重由常数 2 变成样本容量的对数函数^[8]。

交叉验证可分为留一交叉验证和 K 折交叉验证，其中留一交叉验证为 K 折交叉验证中的折数 K 等于样本量时的特殊情况。最常用的是 10 折交叉验证 ($K=10$)，具体操作方法为将所有样本平均分为 10 份即 10 折，以其中的 9 折作为训练集进行模型训练，以剩余的 1 折作为测试集（或称验证集）计算训练模型预测的均方误差。以此类推，10 折中的每一折轮流作为训练集和测试集，计算 10 个均方误差的平均值，即最终的交叉验证误差。不同的 λ 会得到不同的交叉验证误差，当交叉验证误差最小时即得到 λ 的最优取值，并用其建立最终的模型。

与信息准则相比，交叉验证的计算量较少且操作方便，因此本研究采用经典的 10 折交叉验证进行 λ 的取值。

2 评价指标

2.1 指标选择

本研究参考和总结已有研究成果，综合青岛海洋经济发展现状，充分考虑评价指标及其数据的全面性、可获取性和代表性，选取海洋渔业、海洋化工业、海洋进出口总额、海洋环境、海洋交通运输业和滨海旅游业 6 个大类并细分 20 个特征变量，构建青岛海洋经济发展水平评价指标体系（表 1）^[9-10]。

表 1 青岛海洋经济发展水平评价指标体系

评价指标	特征变量
海洋渔业 (A)	海洋渔业产值 (A_1)
	海洋渔业增加值 (A_2)
	海水养殖面积 (A_3)
	海水产品产量 (A_4)
海洋化工业 (B)	烧碱产量 (B_1)
	纯碱产量 (B_2)
海洋进出口总额 (C)	进口额 (C_1)
	出口额 (C_2)
海洋环境 (D)	废水排放总量 (D_1)
	废水处理设施数量 (D_2)
	废水处理能力 (D_3)

续表 1

评价指标	特征变量
海洋交通运输业(E)	货运量(E ₁)
	客运量(E ₂)
	货物吞吐量(E ₃)
	旅客周转量(E ₄)
	货物周转量(E ₅)
滨海旅游业(F)	入境旅客数量(F ₁)
	入境旅游收入额(F ₂)
	国内旅客数量(F ₃)
	国内旅游收入额(F ₄)

其中,海洋渔业增加值是指海洋渔业及其服务业通过生产产品或提供服务而增加的价值;废水处理能力是指平均每日的废水处理总量;旅客周转量是指旅客数量与运送距离的乘积;货物周转量是指货物量与运送距离的乘积。

2.2 数据处理

为分析特征变量对青岛海洋经济发展水平的影响程度,本研究选取 2010—2019 年青岛海洋经济发展水平评价指标的变量数据作为建模样本,不仅满足数据的时效性要求,而且在一定程度上降低模型的信息冗余。选取的实证分析数据主要来自历年《青岛统计年鉴》中关于海洋产业的公开数据,程序实现采用 SPSS 和 R 等软件^[11]。

本研究对原始数据的预处理包括 2 个方面。

①采用时间序列趋势预测的方法插补缺失数据。由于部分年份的数据记录问题,海洋化工业大类下的烧碱和纯碱产量存在不同程度的数据缺失。为充分利用现有数据并最大限度地提高模型回归结果的准确性,本研究利用 1949—2013 年的烧碱产量和 1965—2015 年的纯碱产量对其进行趋势预测,从而完成对缺失数据的插补。②采用数据标准化的方法消除指标量纲的影响。由于评价指标反映产业种类、发展潜力和优势特点等各方面的海洋经济发展水平,各评价指标下的特征变量量纲存在巨大差异。例如:海洋渔业产值以人民币计量,而海洋进出口总额则以美元计量,存在汇率的差异,此外数量和长度等指标也不具有可比性。为消除指标量纲的影响,本研究对原始数据

进行标准化处理。

3 实证分析

3.1 特征变量的描述性统计和相关性

各评价指标特征变量的描述性统计如表 2 所示。

表 2 特征变量的描述性统计

特征变量	平均值	标准差	最小值	最大值
A ₁	169.73	23.33	118.19	193.69
A ₂	4 803.97	5 198.92	0.83	11 608.91
A ₃	3.46	0.28	3.15	4.01
A ₄	108.68	3.97	100.79	113.41
B ₁	18.79	11.12	9.78	35.36
B ₂	67.22	2.93	61.70	72.07
C ₁	2 414.04	2 250.31	567.21	5 779.56
C ₂	3 022.53	2 940.20	598.83	6 944.38
D ₁	5.03	0.75	4.04	6.34
D ₂	430.20	48.74	346.00	486.00
D ₃	100.84	56.84	37.48	172.10
E ₁	2 310.10	1 131.70	1 313.00	4 407.00
E ₂	423.90	357.70	161.00	1 302.00
E ₃	4.72	0.72	3.50	5.77
E ₄	70.99	91.22	0.30	196.73
E ₅	1 541.81	987.47	427.00	3 237.40
F ₁	135.01	18.17	108.05	170.26
F ₂	61.76	20.35	39.97	108.40
F ₃	1 185.82	432.12	540.00	1 897.20
F ₄	7 273.55	2 159.16	4 396.00	11 132.58

由于特征变量的个数较多,变量之间易出现信息重复、相关或冗余的现象,从而导致多重共线性等问题。基于此,本研究计算特征变量的斯皮尔逊线性相关系数,并对其进行相关性分析。经计算和分析,海洋经济发展水平的特征变量之间存在强相关性,其中海洋渔业产值与海水养殖面积、客运量和货物吞吐量之间的相关系数均超过 0.90,进口额与出口额和旅客周转量之间的相关系数甚至超过 0.99,表明海洋经济发展水平评价指标体系中的某些变量存在冗余信息,应予以剔除。因此,本研究剔除特征变量中的重复变量,筛选重要变量并构建 Lasso 回归模型。

3.2 模型构建

3.2.1 λ 的最优取值

在 Lasso 回归模型中, λ 直接决定进入模型的变量个数, 进而影响模型回归的准确性。随着 λ 的不断增大, 本研究中 20 个特征变量的系数依次缩减到 0, 即相应变量可从模型中剔除。其中, 系数最晚缩减到 0 的特征变量对被解释变量预测的重要性最强, 以此类推。

为使 Lasso 回归模型的结果更加准确, 采用 10 折交叉验证确定 λ 的最优取值。本研究中每个特征变量的可用样本量仅有 10 个 (2010—2019 年), 即每折仅有 1 个样本, 因此 10 折交叉验证等价于留一交叉验证, 即每次留 1 个样本作为测试集, 其他 9 个样本作为训练集, 进而得到训练模型在测试集中的均方误差。最后, 取 10 个测试集均方误差的平均值作为交叉验证误差, 以最小交叉验证误差对应的 λ 为最优取值。

10 折交叉验证误差随 λ 取值变化的曲线如图 1 所示。其中, 上下闭口的竖直线表示交叉验证误差的 95% 置信区间, 第一条虚线表示当交叉验证误差最小时对应的 λ 和模型自由度 (变量个数), 第二条虚线表示交叉验证误差增速由低到高的转折点。

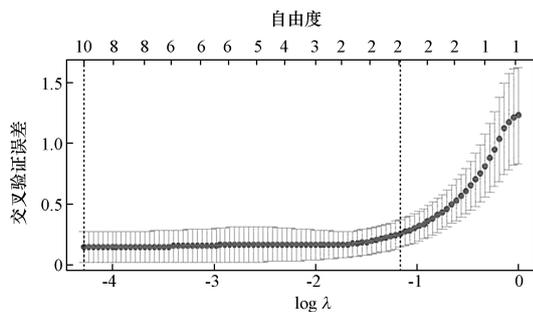


图 1 交叉验证误差和 λ 的取值

由图 1 可以看出: 随着 $\log \lambda$ 的增大, Lasso 回归模型的交叉验证误差缓慢增大; 当 $\log \lambda$ 增至 -1 附近时, 交叉验证误差增速骤升; 当 $\lambda = 0.013\ 76$ 即 $\log \lambda = -4.286\ 0$ 时, 交叉验证误差达到最小值即 0.001 0。因此, λ 的最优取值为 0.013 76, 此时模型中含有 10 个特征变量。

数值形式的模型自由度、交叉验证误差和 λ 如表 3 所示。

表 3 模型的自由度、交叉验证误差和 λ

自由度	交叉验证 误差	λ	自由度	交叉验证 误差	λ
10	0.001 0	0.013 76	4	0.014 0	0.097 05
9	0.001 1	0.014 41	4	0.022 5	0.128 30
10	0.001 2	0.015 10	3	0.024 4	0.134 40
10	0.001 3	0.015 82	3	0.030 7	0.154 50
9	0.001 4	0.016 57	3	0.033 3	0.161 90
8	0.001 5	0.017 36	2	0.036 0	0.169 60
8	0.002 6	0.025 18	2	0.149 3	0.374 00
7	0.002 7	0.026 38	2	0.213 7	0.450 50
7	0.003 1	0.030 34	2	0.336 3	0.568 40
6	0.003 3	0.033 29	1	0.367 9	0.595 50
6	0.003 5	0.034 88	1	0.633 0	0.787 20
6	0.006 8	0.060 95	1	0.832 5	0.905 10
5	0.007 3	0.063 85	1	0.912 3	0.948 20
5	0.007 8	0.066 89	0	1.000 0	0.993 30
4	0.010 4	0.080 57			

根据 λ 由小变大和由大变小 2 个维度进行分析。①当 λ 由小变大时, 自由度不断减小, 当 $\lambda = 0.993\ 30$ 时的自由度为 0, 此时模型中不再含有变量, 且交叉验证误差达到最大值即 1.000 0; ②当 λ 由大变小时, 交叉验证误差整体不断减小, 且当 $\lambda = 0.013\ 76$ 时, 交叉验证误差达到最小值即 0.001 0, 由此得到 λ 的最优取值即 0.013 76, 此时模型中含有 10 个变量, 这与图 1 的验证结果相对应。值得注意的是, 当 λ 由 0.015 10 减至 0.014 41 时, 交叉验证误差由 0.001 2 减至 0.001 1, 自由度由 10 减为 9; 随着 λ 的继续减小, 交叉验证误差随之减至 0.001 0, 但自由度又增至 10。这种现象表明 λ 的取值依赖于交叉验证误差, 既考虑模型的拟合程度, 又考虑模型中的变量个数。

基于上述分析, 本研究采用 $\lambda = 0.013\ 76$ 建立 Lasso 回归模型, 并采用坐标下降法求解 Lasso 回归系数。

3.2.2 Lasso 回归系数

对 20 个原始特征变量的待估系数 $\beta = (\beta_1, \beta_2, \dots, \beta_{20})$ 随机取初值, 采用 R 软件进行正态模拟, 生成正态随机数并记为 β^0 (表 4)。

表 4 正态随机数

序号	正态随机数	序号	正态随机数
1	0.918 977	11	1.358 680
2	0.782 136	12	-0.102 790
3	0.074 565	13	0.387 672
4	-1.989 350	14	-0.053 810
5	0.619 826	15	-1.377 060
6	-0.056 130	16	-0.414 990
7	-0.155 800	17	-0.394 290
8	-1.470 750	18	-0.059 310
9	-0.478 150	19	1.100 025
10	0.417 942	20	0.763 176

对于第一轮迭代,从第一个特征变量即海洋渔业产值对应的系数 $\beta_1^{(1)}$ 开始,到第二十个特征变量即国内旅游收入额对应的系数 $\beta_{20}^{(1)}$ 为止。每次迭代仅更新 1 个特征变量对应的待估系数,即将该系数视为变量,将剩余的 19 个系数视为常量,通过目标函数对该系数求偏导的方式得到其对应的极小值,如此可得 Lasso 回归系数迭代 1 次的解,记作 $\beta^1 = (\beta_1^{(1)}, \beta_2^{(1)}, \dots, \beta_{20}^{(1)})$; 然后继续迭代,此时出现当部分系数取值为 0 时目标函数才能取得极小值的情况,可将这些系数对应的变量从模型中剔除;持续迭代直至目标函数取值的变化足够小,可将此时的系数作为最终结果。

由于系数估计值较复杂且数量较多,本研究不再展示求解过程。结果表明有 10 个特征变量的系数估计值为 0,在求解过程中将其从模型中剔除。最终的 Lasso 回归模型中共含有 10 个特征变量,其系数估计值和进入模型的顺序如表 5 所示。

表 5 Lasso 回归模型中的特征变量

评价指标	特征变量	系数估计值	进入模型的顺序
海洋渔业	海洋渔业产值	0.129 368 0	3
	海洋渔业增加值	0.008 927 0	8
	海水养殖面积	-0.164 600 0	2
海洋化工业	烧碱产量	0.034 346 0	7
	纯碱产量	0.004 349 0	9
海洋进出口总额	出口额	0.049 385 0	6
海洋环境	废水排放总量	0.046 163 0	5
	废水处理能力	-0.056 210 0	4
海洋交通运输业	货物吞吐量	0.000 060 2	10
滨海旅游业	国内旅客数量	0.556 582 0	1

3.3 回归结果

采用 10 折交叉验证确定 λ 的最优取值为 0.013 76,图 1 和表 3 分别以曲线和数值的形式将确定 λ 最优取值的过程具体化。 λ 的取值取决于交叉验证误差,既考虑模型的拟合程度,又对模型中的变量个数加罚,能够显著提高模型的泛化能力。由于原始数据已进行标准化处理,整体变化幅度较小,模型中的变量系数也较小,但仍可根据系数估计值判别各变量的重要性。

由表 5 可以看出,Lasso 回归模型筛选变量的重要性存在差别。其中,国内旅客数量的系数估计值最大即重要性最强,而货物吞吐量的重要性较弱;海洋渔业产值的系数估计值明显大于绝大多数变量,表明其对青岛海洋经济发展具有显著影响,海洋渔业在所有海洋产业中表现较为突出也是青岛海洋经济发展的明显特征之一;在滨海旅游业中,国内旅客数量因系数估计值最大而被最先选入最终模型,成为关键性变量,且其在整个评价指标体系中的地位也十分显著;海洋化工业中的烧碱产量和纯碱产量均被保留,但烧碱产量的重要性明显高于纯碱产量;海洋交通运输业中仅有货物吞吐量被保留,但其系数估计值最小,表明其对青岛海洋经济发展的贡献较小,尚存在广阔的发展前景。

在模型回归效果方面,应用 Lasso 回归模型对 10 个关键性变量的原始数据进行拟合,并将拟合值与真实值进行对比,得到整体平均预测误差为 0.000 98,表明预测准确率较高,且模型泛化能力和可信度远高于普通最小二乘估计法的线性回归。在相同条件下,若采用多元线性回归,由于变量个数远大于样本量,得到的设计矩阵为非列满秩矩阵;尽管此时的平均预测误差接近于 0,但这样的回归结果很可能是不可靠的,甚至无法得到准确的系数估计值。因此,本研究接受 Lasso 回归模型的预测结果,并得到稳健的稀疏模型。

4 发展建议

4.1 进一步发挥滨海旅游业和海洋渔业等产业的优势

滨海旅游业是青岛的支柱产业,应充分挖掘海洋文化和丰富旅游项目,给各地旅客带来新鲜感和

体验感,尤其应完善相关基础设施建设,为旅客提供“一站式”服务,提高旅客黏性;严格实行禁渔制度,定期放苗,保护海洋生态环境,促进海洋渔业的可持续发展^[12]。

4.2 着力发展海洋交通运输业,提高青岛港口的影响力

货轮拥有装载量大和运输成本低等特点,适合发展国际贸易。青岛应抓住“后疫情时代”的发展机遇,提高青岛港口的装卸效率,通过智能化的码头和空轨集疏运系统,由平面交通向立体交通升级;积极“走出去”,与世界名企联合,建设世界一流的国际化港口。

4.3 保护海洋生态环境,提高废水处理能力

近年来我国大力倡导绿色发展,在发展经济的同时保护环境已是社会各界共识。青岛是依托海洋发展的城市,海水养殖业、海洋交通运输业和滨海旅游业等主要产业的可持续发展都对海水质量有很高的要求。因此,亟须完善“废水入海”的管理制度和标准,加大执法力度和科研投入,着力提高对工业和生活废水的处理能力,使海洋生态文明建设更上新台阶。

4.4 加大海洋科技研发投入,提高成果转化能力

海洋经济的跨越式发展须依靠海洋科技的促进和支撑。应制定科学的海洋经济发展规划,优化海洋产业结构,大力培养海洋人才,以科技促发展,

实现海洋经济发展方式的转型升级^[13]。

参考文献

- [1] 尤芳湖,王凤起.再论海上山东[M].青岛:青岛海洋大学出版社,2000.
- [2] 王端岚.福建省海洋产业结构变动与海洋经济增长的关系研究[J].海洋开发与管理,2013,30(9):85-90.
- [3] 于会娟.现代海洋产业体系发展路径研究:基于产业结构演化的视角[J].山东大学学报(哲学社会科学版),2015(3):28-35.
- [4] 任培民,赵树然,姜文远.基于结构方程-组套索的复杂专利组合测度研究[J/OL].科研管理:1-17 [2021-10-24].
- [5] 徐质斌,牛福增.海洋经济学教程[M].北京:经济科学出版社,2003.
- [6] 曹金磊.基于先验稀疏框架的 Lasso 回归模型的研究及其运用[D].杭州:浙江工商大学,2020.
- [7] TIBSHIRANI R. Regression shrinkage and selection via the Lasso[J]. Journal of the Royal Statistical Society. Series B (Methodological), 1996, 58(1): 267-288.
- [8] JAMES G, WITTEN D, HASTIE T, et al. An introduction to statistical learning[M]. New York: Springer, 2013.
- [9] 郭小军.山东省海洋主导产业的选择及升级研究[D].青岛:中国石油大学(华东),2018.
- [10] 孟月娇.海洋主导产业选择研究[D].青岛:中国海洋大学,2013.
- [11] 朱雪宁.R语言[M].北京:中国人民大学出版社,2018.
- [12] 徐质斌,孙吉亭.山东海洋产业结构与布局优化研究[M].北京:海洋出版社,1998.
- [13] 王玲玲,殷克东.我国海洋产业结构与海洋经济增长关系研究[J].中国渔业经济,2013,31(6):89-93.