

Cointegration Analysis of Traditional Energy Consumption and Economic Development —Take Shandong Province as an Example

Haoda Wang, Fengjiao Yi

School of Mathematical Sciences, Ocean University of China, Qingdao Shandong
Email: wanghaodawhd@163.com, yifengjiao_yfj@163.com

Received: Jun. 27th, 2018; accepted: Jul. 6th, 2018; published: Jul. 13th, 2018

Abstract

According to the data provided by the National Bureau of Statistics, the GDP of Shandong Province has ranked third in the country for many years, and its corresponding energy consumption ranks high in all provinces and cities across the country. In order to quantitatively analyze the relationship between energy consumption and economic development in Shandong Province, this paper has collected relevant data for the past 20 years. Firstly, the long-term equilibrium relationship between the two is explored through cointegration analysis. In view of the existence of autocorrelation in the model, a hysteresis effect was introduced in the regression. At the same time, the sample size is too small, the parameter estimation accuracy is low, and bootstrap solves this problem. Finally, considering the problems that over-reliance on energy consumption and economic growth may affect the environment, economic and social development, we put forward corresponding suggestions.

Keywords

Energy Consumption, Economic Development, Principal Component Analysis, Cointegration Test, Bootstrap

传统能源消耗与经济发 展的协整性分析 ——以山东省为例

王浩达, 伊凤娇

中国海洋大学数学科学学院, 山东 青岛
Email: wanghaodawhd@163.com, yifengjiao_yfj@163.com

收稿日期: 2018年6月27日; 录用日期: 2018年7月6日; 发布日期: 2018年7月13日

摘要

根据国家统计局提供的数据显示山东省地区生产总值已经多年名列全国第三, 与之对应的能源消耗量在全国各省市中排名也很靠前。本文为了定量分析山东省能源消耗与经济发展的关系, 搜集了近20年的相关数据, 首先通过协整分析探寻二者的长期均衡关系。鉴于模型中存在自相关关系, 故在回归中引入了滞后效应。同时用自助法解决了样本容量偏小, 参数估计精度低的问题。最后考虑到过度依赖消耗能源拉动经济增长可能会带来很多环境、经济社会健康发展的问题, 提出了相应的建议。

关键词

能源消耗, 经济发展, PCA, 协整分析, Bootstrap

Copyright © 2018 by authors and Hans Publishers Inc.

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

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



Open Access

1. 引言

改革开放之后, 为了实现经济的快速发展, 我国的能源消耗量也迅速加大[1], 近年来由于能源过度使用以及低效率使用所带来的环境污染问题、能源紧张问题已经逐渐显现了出来。现在所推行的新能源汽车、绿色照明工程、建筑节能工程等都是贯彻国家节能减排政策、优化能源结构、既要金山银山又要绿水青山的思想的一系列有效的措施。GDP 每万元碳排放量等指标也成为国家关注发展的质量的一个重要指标, 但是仅仅通过这样一个比例关系并不能完全的反应二者之间存在关系, 代表性较弱。因此深入探索能源消耗与经济发展的关系, 可以为国家节能减排政策的制定提供科学的建议, 同时也可以对上述政策的实施效果做一个侧面的评价, 从而为我国经济可持续发展、经济结构的有效调整提供内在动力。

2. 数据来源于符号说明

本文从山东省统计局公布的经济发展与能源数据中选取了以下年度指标进行研究[2] (见表 1)。

对本文所出现的符号做统一的规定(见表 2)。

3. 研究方法

3.1. 主成分分析

主成分分析法的本质是降维, 将原来多个复杂的变量进行最佳综合, 简化为少数几个相互独立的综合指标, 这几个综合指标既可以包含原来多个复杂变量的绝大部分信息, 又去除了原变量信息间的重叠, 具有较强的客观性。

主成分是原始变量的一些线性组合, 从而将大量原始变量的信息提取出来。主成分的方差, 即 $\text{Var}(F)$ 反应了该主成分所包含的信息量, 显然应该选取方差最大的主成分做成分作为第一主成分[3]。通常情况下如果数据相关性较强选择第一主成分就可以对数据有一个较好的代表性, 当然也可以选择第二、三、四等主成分加强对原始数据的解释。但是为了有效的反映原始数据, 要保证主成分两两正交, 即即 F_1 已有的信息就不需要再出现在 F_2 中, 在确定第一主成分后找到满足 $\text{Cov}(F_1, F_2) = 0$ 的第二个线性组合 F_2 , 就是第二主成分。以此类推可以得到以下公式:

Table 1. Selection of indicators
表 1. 指标选取

经济发展(单位)	能源消耗(单位)
地区生产总值(亿元)	煤炭(吨)
居民消费水平(元)	焦炭(吨)
对外贸易出口总额(万美元)	原油(吨)
财政收入占地区生产总值的比重	燃料油(吨)
	汽油(吨)
	煤油(吨)
	柴油(吨)
	液化石油气(吨)
	电力消费量(万千瓦时)

Table 2. Explanation of symbols
表 2. 符号说明

符号	表示的意义
gdp	地区生产总值
con	居民消费水平
ex	对外贸易出口总额
pp	财政收入占地区生产总值的比重
F_1	第一主成分
F_2	第二主成分
λ_1	数据协方差阵的最大特征值, 表示第一主成分的贡献率。
λ_2	数据协方差阵的第二大特征值, 表示第二主成分的贡献率。

$$\begin{cases} F_1 = a_{11}z_1 + a_{21}z_2 + \dots + a_{n1}z_n \\ F_2 = a_{12}z_1 + a_{22}z_2 + \dots + a_{n2}z_n \\ \vdots \\ F_p = a_{1p}z_1 + a_{2p}z_2 + \dots + a_{np}z_n \end{cases}$$

F_1, \dots, F_p 代表前 p 个主成分, Z_1, \dots, Z_n 是原始变量矩阵标准化后的数据, $P < n$ 。
 $((a_{11}, a_{21}, \dots, a_{n1}))$ 是标准化后数据的协方差阵的最大特征值对应的特征向量。

3.2. 协整性检验

3.2.1. 单整

对于非平稳的时间序列数据, 通过运用差分的方式进行平稳化处理, 一阶单整指的是经过一次差分处理后非平稳时间序列便趋于平稳, 记为 $I(1)$ 。一般而言, 如果时间序列平稳, 记为 $I(0)$, 如果 d 阶差分后平稳, 称为 d 阶单整, 记为 $I(d)$ 。

3.2.2. 协整

如果时间序列 $y_{1t}, y_{2t}, \dots, y_{nt}$, 都是 d 阶单整, 即 $I(d)$, 存在一个向量 $\alpha = (\alpha_1 + \alpha_2 + \dots + \alpha_n)$ 使得 $\alpha y_t \sim I(d-b)$, 这里 $y_t = (y_{1t} + y_{2t} + \dots + y_{nt})$, $d \geq b \geq 0$

则称序列 $y_{1t} + y_{2t} + \dots + y_{nt}$ 是 (d, b) 阶协整, 记为 $\alpha y_t \sim CI(d, b)$, α 为协整向量。

协整在探索经济社会发展规律方面是具有显著意义的, 经济社会发展是一个具有长期规律性波动的

过程。但是其发展情况及其复杂, 通常受到外界各种因素的影响, 很有可能存在一个或多个与它保持长期稳定均衡关系的指标。这种协整关系的存在便可以通过其它变量的变化来影响另一变量的变化[4]。

3.2.3. 协整性检验(E-G 两步法)

两个或多个时间序列同阶单整时才可能协整, 否则直接判断为非协整。若同阶单整, 用以下两种方法检验是否协整。

E-G 两步法比较适合于两个变量之间协整关系的检验。 x_t 和 y_t 均为 d 阶单整的时间序列数据。

第一步, 做普通最小二乘回归。

$$\hat{y}_t = \hat{\alpha} + \beta x_t$$

得到模型的残差

$$\hat{\varepsilon}_t = y_t - \hat{y}_t$$

第二步, 检验残差是否平稳, 如果平稳, 则认为时间序列 x_t 和 y_t 是协整的。

3.3. Bootstrap 方法

Bootstrap 称为自助法, 是一种有放回的重复抽样。从容量为 n 的样本中进行有放回的抽样, 并不改变样本容量[5] [6]。

Bootstrap 优势: 一是可以减小估计量的偏差, 二是对很多估计能够获得精准的置信区间, 三是在某些情况下极限分布难以获得, bootstrap 可以用做统计推断, 四是存在严重的离群点或样本容量不够大时, 自助法非常实用。

4. 实证分析

4.1. 主成分分析

本文采用 R 语言进行山东省能源消耗与经济发展关系的实证分析, 首先利用主成分分析方法, 分别对反映经济发展的指标和反映能源消耗的指标进行降维, 确定两个综合指标得分。

首先提取能源消耗的主成分, 为了消除数据量纲的差异, 首先对数据进行了标准化处理。根据指标间的相关系数矩阵(见图 1)可以发现能源消耗变量间存在很强的相关性, 适用于进行主成分降维。

由主成分分析结果(见图 2)可以看出第一主成分的方差贡献率到达了 84.9%, 第二主成分的方差贡献率为 12.32%, 前两个主成分的累计方差贡献率高达 97.23%, 损失信息非常少, 所以选择前两个主成分便可以代表所有变量的信息, 通过碎石图可以更加直观的说明主成分的选择(见图 3)。

同理经济发展变量中我们提取了第一主成分的方差贡献率即可达到 97.72%, 即选取第一主成分就可以代表 97.72% 的信息

计算 1995 年到 2015 年各年份在第一主成分上的得分情况。下面展示了前 5 年的结果(见图 4、图 5):

	煤炭	焦炭	原油	燃料油	汽油	煤油	柴油	液化气	电力
煤炭	1.0000000	0.9907935	0.9494155	0.7115295	0.9803737	0.5118803	0.9769645	0.9244638	0.9550641
焦炭	0.9907935	1.0000000	0.9666458	0.7559819	0.9779830	0.5610731	0.9676120	0.9051902	0.9644930
原油	0.9494155	0.9666458	1.0000000	0.8790242	0.9369723	0.6997965	0.8932460	0.8521352	0.9944227
燃料油	0.7115295	0.7559819	0.8790242	1.0000000	0.7271995	0.8712243	0.6107367	0.6048653	0.8737084
汽油	0.9803737	0.9779830	0.9369723	0.7271995	1.0000000	0.5287483	0.9720642	0.9042611	0.9379428
煤油	0.5118803	0.5610731	0.6997965	0.8712243	0.5287483	1.0000000	0.3704111	0.3218299	0.6973262
柴油	0.9769645	0.9676120	0.8932460	0.6107367	0.9720642	0.3704111	1.0000000	0.9520375	0.8938562
液化气	0.9244638	0.9051902	0.8521352	0.6048653	0.9042611	0.3218299	0.9520375	1.0000000	0.8517068
电力	0.9550641	0.9644930	0.9944227	0.8737084	0.9379428	0.6973262	0.8938562	0.8517068	1.0000000

Figure 1. Correlation coefficient matrix for energy consumption

图 1. 能源消耗相关系数矩阵

Importance of components:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6
Standard deviation	2.6976441	1.0280239	0.34397969	0.244067659	0.192077917	0.107293828
Proportion of Variance	0.8490164	0.1232972	0.01380424	0.006949719	0.004304291	0.001343063
Cumulative Proportion	0.8490164	0.9723136	0.98611787	0.993067587	0.997371879	0.998714941
	Comp.7	Comp.8	Comp.9			
Standard deviation	0.0811025360	0.0611738987	2.636139e-02			
Proportion of Variance	0.0007673892	0.0004365954	8.107433e-05			
Cumulative Proportion	0.9994823303	0.9999189257	1.000000e+00			

Figure 2. Principal component analysis of energy consumption

图 2. 能源消耗主成分分析

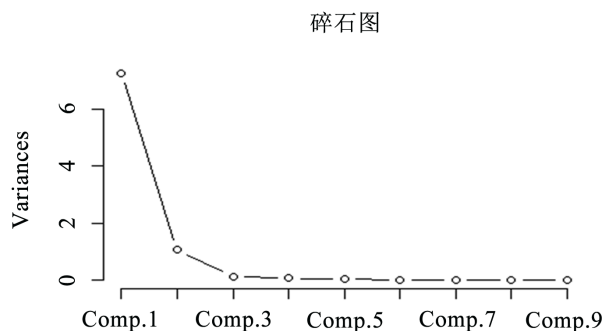


Figure 3. Gravel map

图 3. 碎石图

	Comp.1	Comp.2	Comp.3	Comp.4
[1,]	-2.553	-0.745	-0.061	0.025
[2,]	-2.300	-0.478	-0.072	0.012
[3,]	-2.049	-0.169	-0.046	0.020
[4,]	-1.906	0.049	-0.048	0.038
[5,]	-1.714	0.254	-0.044	0.019

Figure 4. Principal component scores of energy consumption indicators

图 4. 能源消耗主成分得分

	Comp.1	Comp.2
[1,]	3.050	-0.144
[2,]	2.951	-0.335
[3,]	2.720	-0.397
[4,]	2.726	-0.866
[5,]	2.789	-0.800
[6,]	2.562	-0.760
[7,]	2.485	-0.769

Figure 5. Principal component scores of economic indicators

图 5. 经济指标主成分得分

最后根据综合评价函数 $F = \frac{\lambda_1}{\lambda_1 + \lambda_2} F_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} F_2$ 计算主成分的综合得分如下(见表 3):

由综合得分时序图(见图 6)可以看出, 经济发展的主成分得分是逐年提高的, 这表明山东省的经济发展一直保持逐年增长的趋势, 而能源消耗的主成分得分却出现了一些小的波动, 特别是在 2011 年有了一个明显的下降。一方面, 受利比亚战乱、埃及局势动荡以及欧佩克组织限产等因素的影响, 世界石油价格高涨。另一方面, 全省节能降耗政策取得新成效, 如 351 户企业完成落后产能淘汰。因此, 燃料油、汽油、柴油和液化石油气的消耗量较 2010 年有较大下降, 能源消耗的主成分得分在 2011 年显著下降。

Table 3. Scores of comprehensive energy and economic indicators
表 3. 能源综合指标与经济综合指标得分

年份	经济得分	能源得分	年份	经济得分	能源得分
1995	-2.5534979	-2.0327646	2006	-0.2232484	0.9008678
1996	-2.3004994	-2.0351972	2007	0.2617071	1.1438427
1997	-2.0494325	-1.9069454	2008	0.6387333	1.401005
1998	-1.9056596	-2.0748051	2009	0.7333389	1.6382209
1999	-1.7135151	-2.0929226	2010	1.4037916	2.3117108
2000	-1.5471169	-1.931083	2011	2.1225991	1.386845
2001	-1.2425921	-1.8844139	2012	2.5618694	1.5956007
2002	-1.2554833	-1.8128894	2013	2.884805	1.8832169
2003	-1.1201941	-1.2961166	2014	3.3249955	2.284107
2004	-1.0332696	-0.8447114	2015	3.6679584	2.8495455
2005	-0.6552894	0.5168868			

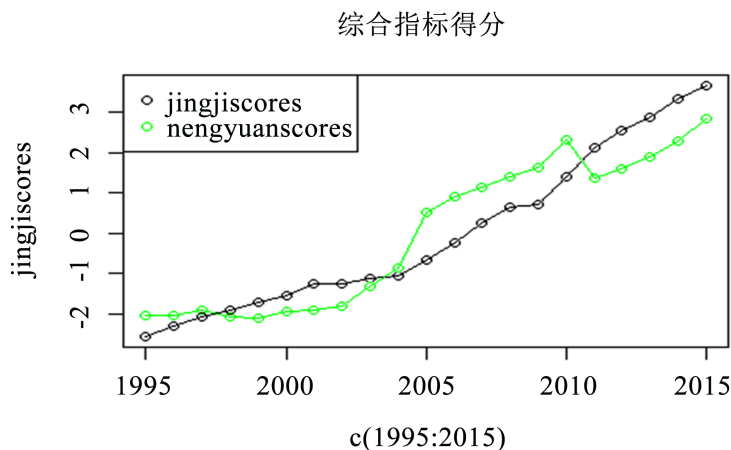


Figure 6. Composite Score Timing Chart
图 6. 综合得分时序图

以 1997 年、2004 年、2011 年为三个分割点，可以看出山东省经济发展受能源消耗的影响的发展趋势基本是相同的，但是其影响力却逐渐发生着变化。

4.2. 协整分析

由上图可以很清晰的得出结论，两个综合得分时间序列是不平稳的。分别作一阶差分后的时序图(见图 7)如下：

两项指标单位根检验的 P 值分别为 0.0252 和 0.0232 均小于 0.05 (见图 8)，说明经过一阶差分序列已经平稳。进而可以利用两变量的 Engle-Granger 检验进行协整性检验，首先建立回归方程并用最小二乘法估计参数，得到非均衡误差。

经过对误差项的单位根检验，可以确定误差项平稳，说明能源消耗与经济发展存在长期均衡关系，其回归方程为 $nengyuanscore = 0.86 * jingjiscore$ 。

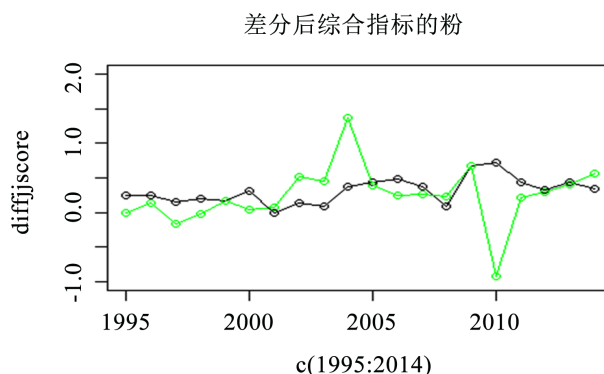


Figure 7. Composite Score After Difference

图 7. 差分后综合指标得分

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
jingjjscore  0.85962    0.07941   10.82 8.21e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7003 on 20 degrees of freedom
Multiple R-squared:  0.8542,    Adjusted R-squared:  0.8469
F-statistic: 117.2 on 1 and 20 DF,  p-value: 8.207e-10

```

Figure 8. Regression results

图 8. 回归结果

但是, 这里存在一个问题, 模型的 DW 检验未能通过($p\text{-value} = 3.288e^{-7}$), 说明模型中有严重的自相关问题

所以考虑把因变量的滞后一期和滞后二期的结果引入回归模型中重新建模(见图 9)。

所以经过修正后的模型变为 $Nengyuanscore = 0.3 * jingjjscore$, 即经济得分每提高一分, 能源得分提高 0.3 分。

4.3. Bootstrap 方法与最小二乘法的比较

模型诊断图(见图 10)可以看出, 误差项的正态性假定并不太满足, 样本量也不大, 根据误差正态性假定得出的回归系数的置信区间不具有说服力。

利用 bootstrap 方法, 重复抽样 2000 次, 对回归方程的系数进行估计。步骤如下:

- 1) 有样本容量为 n 的样本 $(x_1, y_1), \dots, (x_n, y_n)$, 从中有放回地重复抽样 n 次, 得到了第一个样本容量为 n 的新样本, 称为自助样本 1, 对于原始数据对 (x_i, y_i) , 在自助样本中可能出现多次, 可能出现零次。
- 2) 用自助样本 1 进行最小二乘法回归分析, 估计其回归参数, 不妨记为 t_1
- 3) 重复上述步骤 m 次, 得到 t_1, t_2, \dots, t_m 。取置信水平为 95%, 那么把 t_1, t_2, \dots, t_m 从小到大排序, 置信区间为这列数的(上 0.975 分位数, 上 0.025 分位数)。

通过比较可以发现, 采用 Bootstrap 所得到的回归方程, 其参数估计的置信区间要更具说服力(见表 4)。因此, 采用 bootstrap 法建立能源消耗与经济关系的回归模型可以更好的反映二者的关系, 从而有助于根据能源消耗量测算其所带来的经济增长量。经济得分每提高一分, 能源消耗得分提高 0.03 分到 0.5 分。

5. 结论建议

由能源消耗与经济关系的相互拉动关系可以看出, 能源消耗对于经济的贡献与在我国不同的发展时

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
beauty$jingjiscor  0.3007    0.1151   2.614 0.018817 *
beauty$ar.1.       0.9809    0.2203   4.453 0.000401 ***
beauty$ar.2.      -0.2906    0.1890  -1.537 0.143714
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4073 on 16 degrees of freedom
Multiple R-squared:  0.9508,    Adjusted R-squared:  0.9416
F-statistic: 103.1 on 3 and 16 DF,  p-value: 1.121e-10
    
```

Figure 9. Regression results with the introduction of hysteresis
 图 9. 引入滞后效应的回归结果

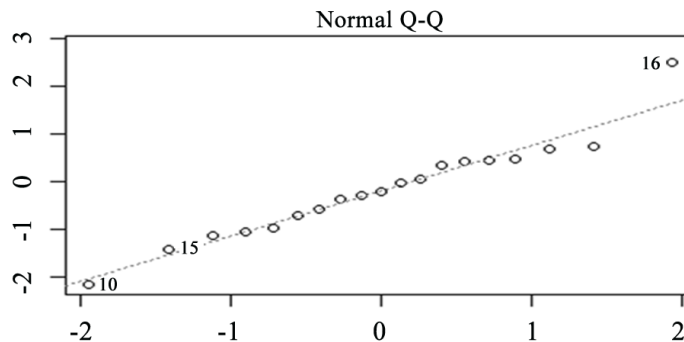


Figure 10. Residual Q-Q diagram
 图 10. 残差 Q-Q 图

Table 4. Comparison of Least Squares with Bootstrap
 表 4. 最小二乘与 Bootstrap 比较

	普通最小二乘法		Bootstrap 方法	
	置信下限	置信上限	置信下限	置信上限
Jingjiscor	0.0568	0.5446	0.0306	0.5437
Ar(1)	0.5139	1.4478	0.6187	1.5880
Ar(2)	-0.6912	0.1101	-0.7594	0.0579

期具有明显的差异, 在 2005 年到 2010 年, 能源消耗的贡献率明显低于 2010 年以后, 这说明我国近年来的能源结构调整和经济发展战略都取得了显著的成效, 必须将这种改革深入贯彻落实下去。同时必须采取积极有效的措施解决现有经济-能源-环境关系, 为此可以采取以下措施:

- 1) 在大力倡导使用天然气等清洁能源的同时着力解决其使用过程中存在的问题, 例如天然气汽车稳定性、加气便利性的问题都亟待解决, 只有让消费者感受到清洁能源要更加的便捷经济, 才能够从根本上刺激消费转型, 从而在不影响经济发展的基础上减少能源消耗带来的环境问题。
- 2) 继续推进产业结构升级, 寻找新的经济增长的拉动因素, 降低能源消耗特别是不可再生能源的消耗对于经济发展的贡献率, 进而达到提高单位能源消耗的经济转化率。
- 3) 建立能源消耗对环境影响的评价指标体系, 直接将能源消耗与环境问题挂钩, 不仅可以提高相关部门、企业的重视程度, 也可以提供一个更直观的评价标准。

致 谢

感谢在本文完成过程中家人和老师对我的帮助, 同时也感谢所有参考或引用的文献作者。

参考文献

- [1] 王彦彭. 我国能源环境与经济可持续发展[D]. 首都经济贸易大学, 2010.
- [2] 王杨, 韩凤歌. 湖北省能源消耗与经济发展关系研究[J]. 中国集体经济, 2016(18): 17-19.
- [3] 丁雪琪. 各省份发展水平的主成分及聚类分析[J]. 学理论, 2013(12): 62-64.
- [4] 宋琴. 人民币汇率变动对地产指数影响的实证分析[J]. 海南金融, 2009(2): 4-7.
- [5] Efron, B. (1992) Bootstrap Methods: Another Look at the Jackknife. Breakthroughs in Statistics. Springer New York, 565-568.
- [6] 孙明燕. 基于投资风格角度的基金业绩分解研究[D]. 哈尔滨工业大学, 2016.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2334-332X, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: cce@hanspub.org