

要素禀赋、产业距离与产业升级路径选择

陈 普

(华东交通大学 经济与管理学院, 南昌 330013)

摘 要:基于输入端的行业距离构造并利用自然条件、人力资源、物质资本、科技、经济结构和制度6大类共23种要素禀赋结构提出一种行业的“距离”度量。根据该度量绘制了行业的空间分布图,计算了10个低价值行业升级至3个高价值行业的最短路径。研究发现,资源型行业升级至高价值行业要经历更多的轮次,有时甚至缺乏升级至高价值行业的路径。有些中间行业如饮料制造业则始终出现在低价值行业升级至高价值行业的升级路径中,中西部省份产业升级路径选择若能趋近这些行业,产业升级将更为便利。

关键词:产品空间;要素禀赋;产业升级

中图分类号:F420 **文献标志码:**A **文章编号:**1002—980X(2020)6—0024—10

随着中国经济的不断发展,要素禀赋结构也在发生着巨大变化,供给侧改革给中国的经济发展提供了新的发展思路,而供给侧改革一个重要方面就是从要素禀赋层面破解供给约束^[1],如何从要素禀赋出发,为当前产业结构变化提供新的思路?或者,更为具体地,如何从一个低价值产业出发,选择何种产业发展路径,最终可以抵达高价值产业?针对上述问题展开研究,不仅进一步丰富了产品空间理论^[2-3],而且为供给侧改革提供了新的角度,并对地方政府的行业发展提供了重要的实践指导作用。

产业升级伴随着产业的结构演化^[4-5],而产业的结构演化又往往由要素禀赋结构的变化所推动。正如林毅夫^[6]指出的,“一个经济体在任何时点上的最优产业结构,是能够让该经济在国内市场和国际市场实现最强竞争力的产业结构。而这一最优产业结构是由该经济的比较优势决定的,而后者又由该时点上经济的禀赋结构决定”,即一个地区主要生产什么与当地的要素禀赋结构息息相关。因此,对于中西部地区而言,如果要素禀赋结构不匹配,期望从一个低价值行业直接跨越到高价值行业并不现实。这一点在产品空间理论中也有阐述,Hidalgo等^[3]在产品空间理论中曾对产业升级有一个形象的比喻:产品就像一片森林,这片森林有的地方茂密,有的地方稀疏,而国家就好像森林里面的一只猴子,产业升级就是这只森林中的猴子从这棵树往另一棵树的攀援过程。若树与树的距离太大,猴子攀不过去,产业升级也就中断了。张其仔和李颢^[7]根据Hidalgo等^[3]的距离算法,研究指出,要保持经济的稳定增长,具备跨越0.5的“距离”的能力是必要的。也即,距离如若过大,产业升级无法完成,经济增长也难以持续。

若产业升级只能从邻近的产业开始,那么其路径必然是循序渐进的。关于这一点,Gereffi^[8]在其全球价值链(global value chain, GVC)分析框架中有过初步的论述,他指出,如Porter^[9]所言,产业升级过程无非就是资本(包括物质和人力资本)禀赋相对劳动力禀赋日益富余的过程,但产业升级也不是从一个劳动密集行业随机地抵达某个资本或者技术密集的行业,其必然沿着全球价值链上的产品逐步爬升。若从要素禀赋结构的角度看,产业升级则是从具备相似要素禀赋结构的产业开始,逐步演化,最终抵达高价值行业。接下来的一个重要问题就是哪些行业间的距离更近,行业间的距离如何准确度量?

度量行业间的距离,据我们所知,Hidalgo等^[3]是最早进行研究的,他主要通过两步构造该距离,第一步计算显式比较优势(revealed comparative advantage, RCA);第二步基于显式比较优势,计算一个条件概率,即

收稿日期:2019—12—16

基金项目:教育部人文社科青年项目“长江经济带产业转型:迁移和升级”(15YJC790012);江西省社科规划青年项目“江西产业转型升级研究”(15YJ35);中国博士后科学基金面上项目“机器学习视角下要素禀赋与产业发展研究”(2019M662845);江西省高校人文社科研究项目“中部地区制造业与新兴产业集群的建设路径研究”(GL19103);江西省教改课题重点项目“遵循专业认证OBE理论的CIMA专业方向人才培养研究与实践”(JXJG-18-5-15)

作者简介:陈普(1982—),男,湖南邵阳人,博士,华东交通大学经管学院讲师,研究方向:时间序列分析、机器学习及其在产业经济的应用。

文献[3]中的 ϕ_{ij} 。因其度量方法的独特与有效,后续鲜有文献提出新的度量方式。相反的是,存在大量实证文献直接借用该方法计算行业距离。例如,邓向荣和曹红^[10]使用该距离实证检验了产业升级与比较优势的关系。张亭和刘林青^[11-12]也借用该距离讨论了中美以及中日产业升级中的作用机制。张其仔和李颢^[7]为度量持续的经济增长所应具备的行业跨越能力,也使用了该方法。不过,仔细观察Hidalgo等^[3]的计算步骤,这种度量产业距离的方法极大依赖于出口产品的数据,对于国家而言,出口的产品自然代表了该国的专业化所在。但正如Boschma等^[13]所言,产业更迭的考察更应从地区而非国家层面来考察。因此,当我们将研究视角投向省份时,发现一个省份所擅长生产的产品很可能仅仅销往省外而并不出口,此时若使用省份的出口数据来度量产品距离,则并不合适。

因此,本文特色在于:第一,不同于其他文献,本文从输入端要素禀赋结构入手,计算行业距离,提供了新的研究视角。第二,使用机器学习方法突破相关实证文献^[14-17]使用要素禀赋的数量较少(很少超过5个)的局限,本文使用了自然条件、人力资源、物质资本、科技、经济结构和制度6大类共23种要素禀赋,从而能够更为全面地捕捉相关禀赋特征。

一、研究设计

产业升级从来不是一蹴而就的,极端地看,一个主要种植苹果的地区无法跳跃式地直接生产计算机,需要有一个循序渐进的产业升级路径。与Hidalgo等^[3]从最终产品入手计算产品相似性不同,本文从输入端——要素禀赋入手,计算产业之间的相似性,从而为产业升级路径提供依据。具体地,每种要素禀赋对不同行业有着完全不同的重要性,这种不同的重要性在行业间的差异越大,也即“距离”越大,要实现行业间跨越的难度也越大。为得到低价值行业抵达高价值行业的最短演进路径,研究的主要步骤设计如下。

(1)量化两个基础指标。第一个是对于不同行业的不同要素禀赋的重要性。这里主要利用随机森林算法,以要素禀赋作为解释变量来预测各行业的规模,该算法同时会得到一个指标即各解释变量的重要性度量,以此作为该行业对各要素禀赋的依赖程度。第二个是不同行业间的距离。不同行业对要素禀赋有不同的依赖,譬如对于 i 行业, p 个要素禀赋的重要性可以记为 $w_i = \{w_{i1}, w_{i2}, \dots, w_{ip}\}$, $i \in \{1, 2, \dots, n\}$, w_i 间的欧氏距离则可作为行业距离的度量。

(2)计算行业价值,量化低价值行业、高价值行业,然后利用力引导Fruchterman-Reingold算法^[18]作为布局算法,绘制行业空间分布图。具体借鉴Hausmann等^[19]的方法,以行业规模的相对比例为权重对各地区的人均GDP进行加权,得到一个行业价值的估算。利用第一步得到的行业距离以及行业价值,再根据Fruchterman-Reingold算法就可以绘制一幅类似于Hidalgo等^[3]提出的产业空间分布图。

(3)利用行业空间分布图计算从低价值行业到高价值行业的最短路径。计算机算法领域有一种熟知的Dijkstra算法,利用网络连通图可以实现从一个顶点到另一个顶点的最短路径。利用该算法也能方便地计算从低价值行业往高价值行业升级的最短路径。

二、产业距离的度量

(一)变量选取与数据处理

直接通过要素禀赋来预测行业规模的文献十分鲜见,但使用要素禀赋进行经济类实证的文献存在不少^[6, 15-16]。由于这些文献使用的都是经典的计量经济学方法,所以实证中使用的要素禀赋一般较少。本文借鉴李敏纳等^[20]对要素禀赋的划分,基于数据的可得性,得到了23种要素禀赋,均总结在了表1中。

实证数据是分省面板数据,西藏地区和港澳台地区由于指标缺失严重,未纳入分析。时间跨度为2004—2011年,选择该区间的原因主要受3个条件的局限:①部分禀赋指标开始统计的时间较晚,比如林地面积和年产矿量2004年才开始统计,水资源总量和降水量是2003年才开始统计,铁路营业里程、高速公路里程和内河航道里程2002年才开始统计;②我国行业统计口径在2002年和2012年发生了两次变化,使得有关行业统计量如行业主营业务收入、行业就业人数等只在2001—2011年具有连续性和可比性;③若使用2012年以后的数据,则大部分可以公开获得的要素禀赋数据均止于2016年,总共只有5年的面板数据,总量不足

150 个,会给估计带来一定的局限。

因变量是各行业主营业务收入^①,数据源自 EPS 数据库的中国工业行业数据库。这些行业 2 位码的子行业共有 39 个。

相关自变量数据均源自 EPS 数据平台的不同数据库。本文将自然禀赋、人力资源禀赋、制度禀赋和结构禀赋的相关变量描述性统计见表 2。

表 1 要素禀赋分类

要素禀赋类别	一级测度指标	二级测度指标
自然禀赋	自然条件	水资源总量(<i>water</i> , 亿立方米)、降水量(<i>rain</i> , 亿立方米)
	土地资源	林地面积(<i>forest</i> , 万公顷)、农作物总播种面积(<i>CropArea</i> , 万公顷)
	矿产资源	年产量(<i>ore</i> , 万吨)
人力资源禀赋	教育结构	就业人员中未上过学和小学学历占比(<i>low</i> , %)、就业人员中初中和高中人员占比(<i>med</i> , %)、就业人员中专科和大学人员占比(<i>univ</i> , %)
	劳动力投入	城镇单位就业人员年末数(<i>CityEmp</i> , 千人)、城镇单位就业人员平均工资(<i>CityWage</i> , 元)、农村居民家庭人均纯收入(<i>CnWage</i> , 元)
物质资本禀赋	物质资本水平	资本存量(<i>K</i> , 亿元)
	道路里程	铁路营业里程(<i>railway</i> , 千米)、高速公路里程(<i>road</i> , 千米)、内河航道里程(<i>river</i> , 千米)
科技禀赋	科技投入	研究与试验发展人员全时当量(<i>RD</i>)、高技术产业企业数(<i>TecNum</i> , 个)、高技术产业从业人员年平均人数(<i>TecEmp</i> , 人)、高技术产业主营业务收入(<i>TecRetn</i> , 亿元)
结构禀赋	产业结构	非农产业占 GDP 比重(<i>StrGDP</i> , %)
制度禀赋	对外开放度	进出口总额占 GDP 比重(<i>io</i> , %)
	政府退市程度	地方财政支出占 GDP 比重(<i>fiscal</i> , %)
	经济民营化程度	非公有制单位职工数占职工总数(<i>NoPub</i> , %)

注:资料来源为作者整理。括号中第一项是后续公式中用到的变量名,第二项是变量的单位。

表 2 自然禀赋、人力资源禀赋、制度禀赋和结构禀赋指标的描述性统计

统计量	自然禀赋					制度禀赋		
	<i>water</i>	<i>rain</i>	<i>forest</i>	<i>CropArea</i>	<i>ore</i>	<i>io</i>	<i>fiscal</i>	<i>NoPub</i>
最小值	8.4	55.8	2.25	29	156.3	0.036	0.079	0.35
0.25 分位点	256.6	901.4	366.8	2516	9691	0.1	0.13	0.51
中值	508.6	1477	705.4	4792	19780	0.14	0.17	0.6
均值	722.4	1736	912.5	5231	23180	0.36	0.19	0.61
0.75 分位点	1015	2487	1067	7611	29410	0.38	0.22	0.69
最大值	2923	5111	4404	14260	129100	1.8	0.58	0.86
统计量	人力资源禀赋						结构禀赋	
	<i>low</i>	<i>med</i>	<i>univ</i>	<i>CityEmp</i>	<i>CityWage</i>	<i>CnWage</i>	<i>StrGDP</i>	
最小值	3.782	32.58	2.99	425	11690	1722	0.61	
0.25 分位点	22.92	52.43	5.66	2459	17610	3135	0.84	
中值	30.45	60.85	7.54	3612	25400	4379	0.87	
均值	31.39	58.7	9.47	4096	26650	4985	0.87	
0.75 分位点	38.96	66.45	10.33	5120	32940	6121	0.91	
最大值	63.54	76.43	43.48	12380	75590	16050	0.99	

注:资料来源为作者计算。

物质资本禀赋从物质资本水平和道路里程两个角度选择了 4 个指标,其中资本存量的计算主要依据贾润崧和张四灿^[21]的方法,即以永续盘存法为基础,使用里昂锡夫生产函数做回归分析,得到初始资本存量和因省而异的折旧率进行的估算。这些指标的描述性统计见表 3。

(二)产业距离的计算

为防止变量单位对结果的影响,本文首先对要素禀赋进行了如下的归一化处理:

$$x = \frac{x - \min x}{\max x - \min x} \tag{1}$$

其中: x 是要素禀赋; $\min x$ 和 $\max x$ 分别是该要素禀赋数据中的最小值和最大值。使用归一化数据,利用随机森林算法,得到各要素禀赋分别对各行业减少了多少的残差平方和,记为 RSS_p (即要素禀赋对各行业的重要性);其中, i 表示行业, p 表示要素禀赋。再计算每个要素禀赋减少的残差平方和占全部残差平方和的比例:

① 中小类是 3 位码的行业数据,本文重新进行了 2 位码的汇总。

$$w_{ip} = \frac{RSS_{ip}}{\sum_{p=1}^{23} RSS_{ip}} \quad (2)$$

该比例意味着各要素禀赋在行业规模预测中的重要程度。为节省篇幅,本文将每个行业中最重要5个要素禀赋总结见表4。

从表4可以看到每一个行业对应一个关于要素禀赋重要性的向量,这些向量间的欧式距离就可以作为行业“距离”的度量。距离越近,行业之间的转换相对容易,距离越远,转换相对困难。

三、产业升级的路径选择

(一)行业空间分布连通图

对于中西部省份而言,寻找一条从低价值行业到高价值行业的路径是很有必要的,而这又不能一蹴而就,必须依赖行业间的距离循序渐进地进行。

如何划分行业的价值高低?即行业质量如何定义?根据 Hausmann 等^[19]的方法,设*i*省人均GDP为 GDP_i ,*j*行业主营业务收入占*i*省全部行业主营业务收入的份额为 w_{ij} ,那么行业*j*的质量可定义为

$$q_j = \sum_i GDP_i \times \frac{w_{ij}}{\sum_i w_{ij}} \quad (3)$$

本质上这是以各省行业份额加权的人均GDP。计算中使用的都是2011年的数据,B9、C36、C43、D45、B7、B11因为2011年行业主营业务收入数据缺失,因此使用2008年的行业主营业务收入比例(即 w_{ij}),但使用2011年的人均GDP对其产品价值进行了计算。计算结果如图1所示,可以看到通信设备、计算机及其他电子设备制造业,仪器仪表及文化、办公用机械制造业,文教体育用品制造业等具有最高的行业价值;而烟草制品业,石油加工、炼焦及核燃料加工业,非金属矿物制品业等则有着较低的行业价值。

接下来,我们希望得到行业距离的一个直观体现。为此,本文分三步绘制了图2。第一步,计算每个行业对全部23个要素禀赋的依赖程度,可以将其记为, $w_i = (w_{i,1}, w_{i,2}, \dots, w_{i,23})'$,其中向量 w_i 即*i*行业对要素禀赋的依赖。

表3 物质资本禀赋和科技禀赋指标的描述性统计

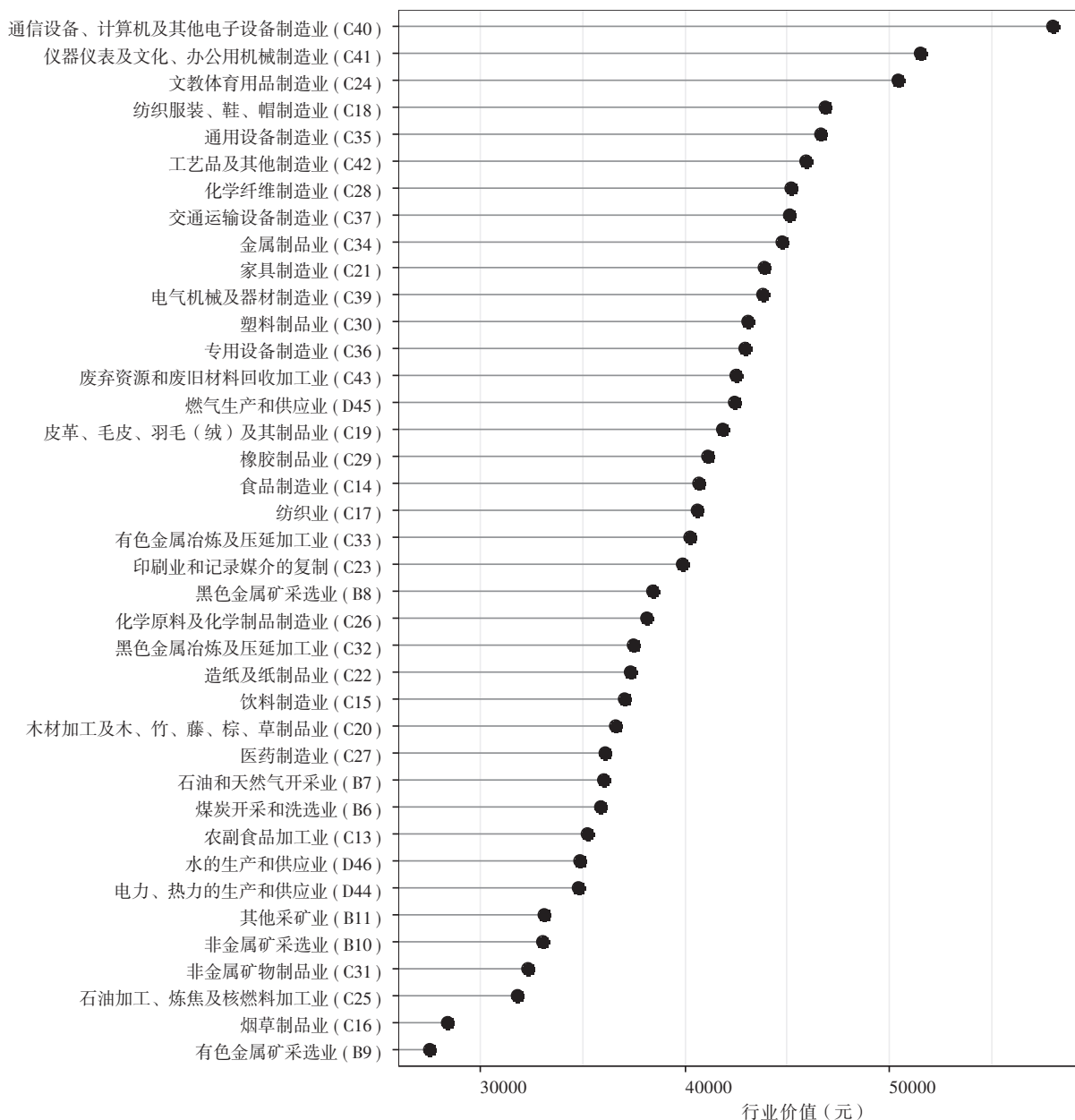
统计量	物质资本禀赋				科技禀赋			
	<i>K</i>	<i>railway</i>	<i>road</i>	<i>river</i>	<i>RD</i>	<i>TecNum</i>	<i>TecEmp</i>	<i>TecRetn</i>
最小值	211.6	211.7	171	0	1209	14	3678	5
0.25分位点	1978	1616	909.8	413	20140	146.5	48410	133.7
中值	3825	2546	1749	2314	44900	360	121170	427.9
均值	5623	2714	1912	4120	64340	745.5	290755	1793
0.75分位点	7681	3562	2691	5596	79170	743.5	238920	1721.4
最大值	26490	9162	5196	24350	410800	5774	3614903	23227.6

注:资料来源为作者计算。

表4 行业对要素禀赋的依赖程度

行业	第一要素禀赋	第二要素禀赋	第三要素禀赋	第四要素禀赋	第五要素禀赋
B10	<i>road</i> (0.17)	<i>CityEmp</i> (0.10)	<i>fiscal</i> (0.10)	<i>CropArea</i> (0.08)	<i>river</i> (0.07)
B6	<i>med</i> (0.16)	<i>ore</i> (0.15)	<i>CropArea</i> (0.10)	<i>railway</i> (0.09)	<i>road</i> (0.07)
B8	<i>river</i> (0.14)	<i>K</i> (0.10)	<i>railway</i> (0.09)	<i>ore</i> (0.08)	<i>water</i> (0.06)
B9	<i>CropArea</i> (0.26)	<i>road</i> (0.17)	<i>CityEmp</i> (0.13)	<i>fiscal</i> (0.06)	<i>io</i> (0.04)
C13	<i>road</i> (0.18)	<i>CityEmp</i> (0.16)	<i>CropArea</i> (0.14)	<i>fiscal</i> (0.10)	<i>K</i> (0.04)
C14	<i>CityEmp</i> (0.23)	<i>road</i> (0.15)	<i>K</i> (0.10)	<i>fiscal</i> (0.07)	<i>CropArea</i> (0.06)
C15	<i>CityEmp</i> (0.25)	<i>TecEmp</i> (0.11)	<i>TecNum</i> (0.11)	<i>road</i> (0.09)	<i>TecRetn</i> (0.07)
C16	<i>K</i> (0.19)	<i>med</i> (0.15)	<i>river</i> (0.11)	<i>low</i> (0.08)	<i>TecNum</i> (0.06)
C17	<i>TecNum</i> (0.22)	<i>TecEmp</i> (0.14)	<i>CityEmp</i> (0.12)	<i>fiscal</i> (0.11)	<i>RD</i> (0.06)
C18	<i>TecNum</i> (0.28)	<i>TecEmp</i> (0.22)	<i>fiscal</i> (0.10)	<i>TecRetn</i> (0.07)	<i>river</i> (0.05)
C19	<i>fiscal</i> (0.23)	<i>TecEmp</i> (0.10)	<i>TecNum</i> (0.08)	<i>TecRetn</i> (0.06)	<i>water</i> (0.05)
C20	<i>TecNum</i> (0.18)	<i>CityEmp</i> (0.13)	<i>TecEmp</i> (0.12)	<i>fiscal</i> (0.10)	<i>K</i> (0.06)
C21	<i>CityEmp</i> (0.20)	<i>TecNum</i> (0.19)	<i>TecEmp</i> (0.19)	<i>TecRetn</i> (0.11)	<i>RD</i> (0.05)
C22	<i>CityEmp</i> (0.27)	<i>TecNum</i> (0.14)	<i>road</i> (0.13)	<i>TecEmp</i> (0.11)	<i>fiscal</i> (0.07)
C23	<i>TecNum</i> (0.22)	<i>TecEmp</i> (0.20)	<i>TecRetn</i> (0.18)	<i>CityEmp</i> (0.13)	<i>RD</i> (0.07)
C24	<i>TecNum</i> (0.27)	<i>TecEmp</i> (0.19)	<i>TecRetn</i> (0.11)	<i>CityEmp</i> (0.08)	<i>RD</i> (0.08)
C25	<i>railway</i> (0.09)	<i>K</i> (0.08)	<i>CityEmp</i> (0.07)	<i>RD</i> (0.06)	<i>fiscal</i> (0.06)
C26	<i>TecNum</i> (0.22)	<i>TecEmp</i> (0.20)	<i>CityEmp</i> (0.10)	<i>TecRetn</i> (0.08)	<i>K</i> (0.07)
C27	<i>TecEmp</i> (0.20)	<i>TecNum</i> (0.17)	<i>CityEmp</i> (0.10)	<i>TecRetn</i> (0.07)	<i>RD</i> (0.07)
C28	<i>TecNum</i> (0.23)	<i>TecEmp</i> (0.16)	<i>NoPub</i> (0.07)	<i>river</i> (0.07)	<i>fiscal</i> (0.06)
C29	<i>TecNum</i> (0.17)	<i>CityEmp</i> (0.15)	<i>TecEmp</i> (0.13)	<i>fiscal</i> (0.08)	<i>road</i> (0.06)
C30	<i>TecNum</i> (0.27)	<i>TecEmp</i> (0.20)	<i>RD</i> (0.07)	<i>CityEmp</i> (0.07)	<i>fiscal</i> (0.06)
C31	<i>road</i> (0.22)	<i>CityEmp</i> (0.21)	<i>TecNum</i> (0.10)	<i>fiscal</i> (0.07)	<i>CropArea</i> (0.06)
C32	<i>K</i> (0.16)	<i>road</i> (0.11)	<i>rain</i> (0.07)	<i>river</i> (0.07)	<i>fiscal</i> (0.07)
C33	<i>CityEmp</i> (0.26)	<i>TecNum</i> (0.12)	<i>road</i> (0.10)	<i>TecEmp</i> (0.08)	<i>fiscal</i> (0.07)
C34	<i>TecNum</i> (0.31)	<i>TecEmp</i> (0.20)	<i>TecRetn</i> (0.15)	<i>river</i> (0.11)	<i>RD</i> (0.04)
C35	<i>TecEmp</i> (0.28)	<i>TecNum</i> (0.20)	<i>K</i> (0.11)	<i>TecRetn</i> (0.07)	<i>RD</i> (0.07)
C36	<i>CityEmp</i> (0.19)	<i>road</i> (0.15)	<i>K</i> (0.11)	<i>TecEmp</i> (0.10)	<i>TecNum</i> (0.09)
C37	<i>TecEmp</i> (0.22)	<i>TecNum</i> (0.20)	<i>TecRetn</i> (0.13)	<i>RD</i> (0.09)	<i>CnWage</i> (0.04)
C39	<i>TecNum</i> (0.29)	<i>TecEmp</i> (0.23)	<i>TecRetn</i> (0.13)	<i>CityEmp</i> (0.07)	<i>RD</i> (0.06)
C40	<i>TecRetn</i> (0.24)	<i>TecNum</i> (0.18)	<i>TecEmp</i> (0.17)	<i>river</i> (0.11)	<i>io</i> (0.05)
C41	<i>TecRetn</i> (0.26)	<i>TecEmp</i> (0.23)	<i>TecNum</i> (0.17)	<i>river</i> (0.09)	<i>RD</i> (0.06)
C42	<i>TecEmp</i> (0.20)	<i>CityEmp</i> (0.15)	<i>TecNum</i> (0.14)	<i>fiscal</i> (0.12)	<i>RD</i> (0.05)
C43	<i>TecNum</i> (0.27)	<i>TecEmp</i> (0.17)	<i>NoPub</i> (0.10)	<i>river</i> (0.07)	<i>CityEmp</i> (0.04)
D44	<i>road</i> (0.23)	<i>CityEmp</i> (0.20)	<i>K</i> (0.12)	<i>RD</i> (0.09)	<i>TecNum</i> (0.06)
D45	<i>CityEmp</i> (0.18)	<i>TecNum</i> (0.17)	<i>TecRetn</i> (0.16)	<i>TecEmp</i> (0.13)	<i>RD</i> (0.07)
D46	<i>TecNum</i> (0.21)	<i>TecEmp</i> (0.20)	<i>TecRetn</i> (0.18)	<i>CityEmp</i> (0.18)	<i>RD</i> (0.04)
B7	<i>CropArea</i> (0.16)	<i>railway</i> (0.15)	<i>CityWage</i> (0.06)	<i>NoPub</i> (0.05)	<i>forest</i> (0.05)
B11	<i>K</i> (0.29)	<i>forest</i> (0.17)	<i>TecRetn</i> (0.08)	<i>river</i> (0.07)	<i>fiscal</i> (0.05)

注:资料来源为作者计算。括号内的数据为 w_{ip} 。



注:资料来源为作者通过 R 软件 ggplot2 包绘制。

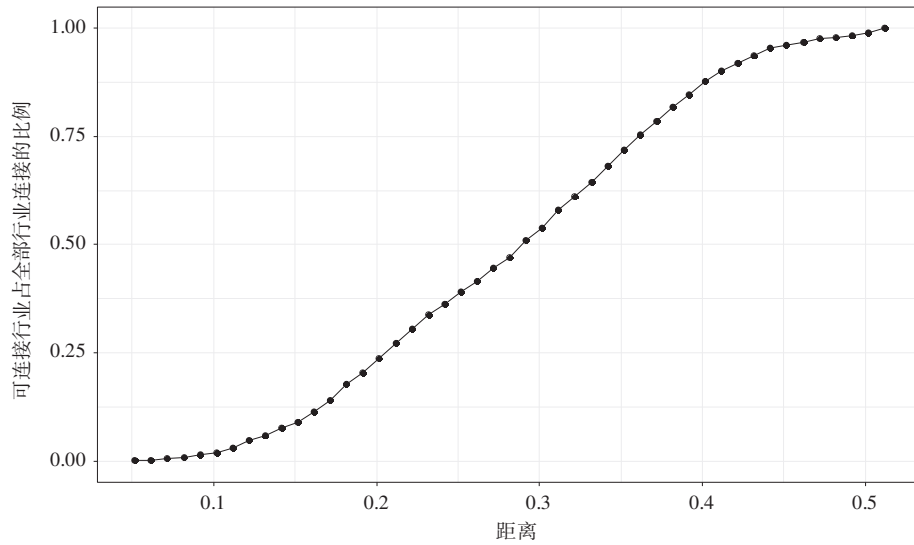
图 1 行业价值排序图

第二步,根据欧式距离的定义,就可以得到行业 i 与行业 j 之间的距离为 d_{ij} ,

$$d_{ij} = \sqrt{(w_{i,1} - w_{j,1})^2 + (w_{i,2} - w_{j,2})^2 + \cdots + (w_{i,23} - w_{j,23})^2} \quad (4)$$

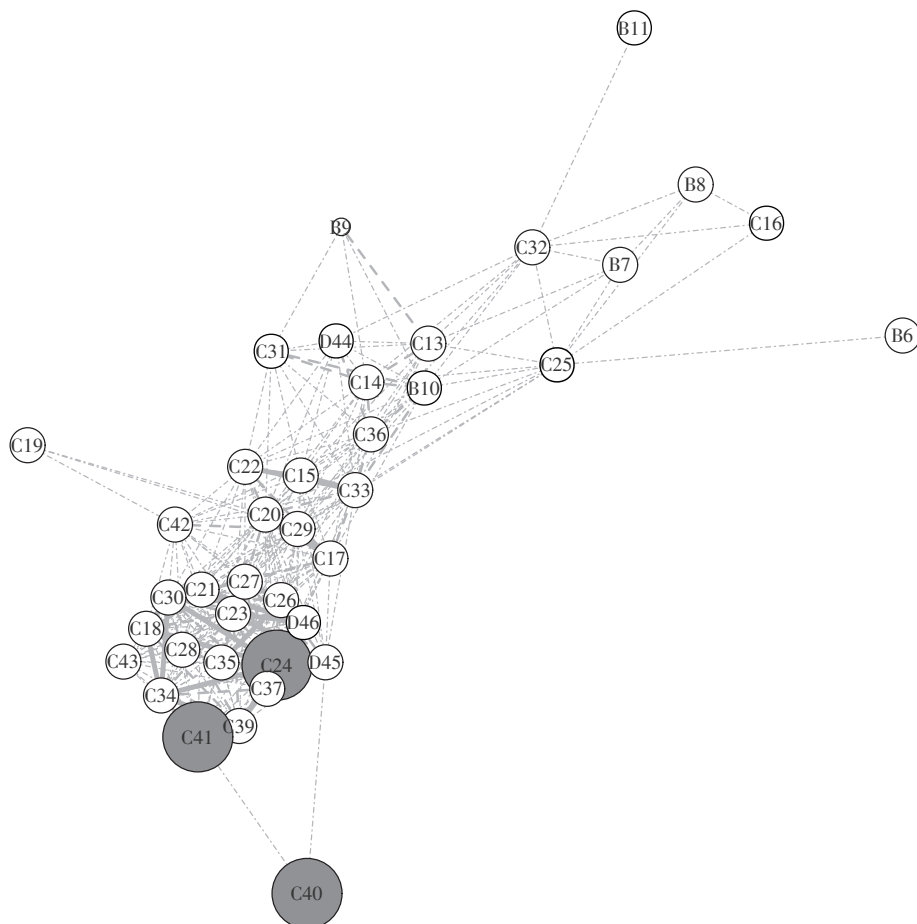
第三步,若定义当行业间距离小于某个值,即认为它们是联通的,或者可以跨越的。那么,令横轴表示行业间的距离,纵轴表示对应距离下联通的行业占最长距离下全部行业联通的比例,就可以得到图 2。通过计算,行业间最短的距离是 0,最长的距离是 0.52。距离越大意味着行业间跨越的难度越大。举例而言,从图 2 可以看到,在行业距离是 0.2 时,它对应的纵轴是 0.25,即距离小于 0.2 的总行业数目占距离小于 0.52 的总行业数目的比例为 0.25。

从图 2 可以看到,行业距离为 0.2 时,大约只有四分之一(即纵轴的 0.25)的行业是连通的。即只有行业距离 ≤ 0.2 的行业,它们之间才能互相跨越,这样的行业占全部行业的 25%。而在行业距离为 0.3 时,则有一半的行业是连通的。为节省篇幅,使用力引导 Fruchterman-Reingold 布局算法,绘制了行业距离为 0.25 时的行业空间分布图如图 3 所示。



注：资料来源为作者通过 R 软件 ggplot2 包绘制。

图 2 行业距离与行业连通数目关系图



注：资料来源为作者通过 R 软件 igraph 包绘制。

图 3 行业距离 ≤ 0.25 时的行业空间连通分布图

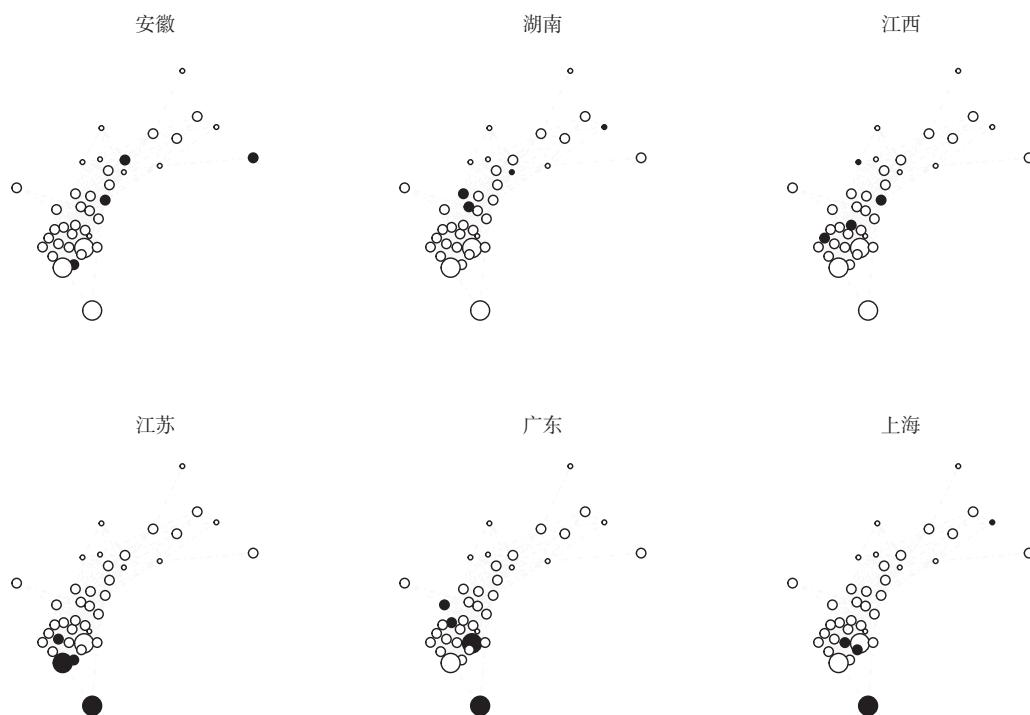
图 3 中连接的线段有粗实线、粗虚线和细虚线，分别表示连接行业间的距离最短（距离 <0.1 ）、中等（ $0.1 \leq$ 距离 <0.15 ）、最长（ $0.15 \leq$ 距离 ≤ 0.25 ）。距离超过 0.25 的，可以看作行业间距离较大，也即跨越愈为困难。图 3 中节点的大小表示行业价值的高低，我们划分了两个档次，行业价值超过 5 万元的为第一档，图上的节点也

较大,用深色进行了填充。行业价值处于 5 万元以下的为第二档,图上节点大小则次之,该节点则未填充颜色。从中可以看到高价值行业都聚集在图的下边缘,特别是具有最高价值的 C40 处于最边缘,而且除了与 C41 外,与其他行业的距离都是比较大的。同时,低价值行业(即小节点)处于与高价值行业相对的边缘地带,多数低价值行业与其他行业的连接也是较弱的,这意味着低价值行业的跨越从最开始就不是一件容易的事情。

最后,若 $x(c, i)$ 是 c 省 i 行业的行业主营业务收入,则根据 Hidalgo 等^[3]对显示性比较优势(revealed comparative advantage, RCA)的定义:

$$RCA_{c,i} = \frac{x(c, i) / \sum_i x(c, i)}{\sum_c x(c, i) / \sum_{c,i} x(c, i)} \quad (5)$$

可以为各省份最具比较优势的产业排序。为更具代表性,本文选择了 3 个中部省份安徽、湖南和江西,以及 3 个发达省份江苏、广东和上海中最具比较优势的 4 个行业在行业空间分布图中进行了黑色填充,绘制在图 4 中。为突出这些比较优势行业,我们淡化了连接线,并消去了行业标记(这些行业标记与图 3 是一致的)。图 4 的上面 3 个省份是中部省份,下面 3 个是发达省份或直辖市。从图可以发现,中部欠发达省份的黑色圆圈相对更加远离下边缘的高价值行业集聚区,而沿海发达地区的黑色圆圈则几乎集聚在下边缘的高价值区。也即,发达省份的比较优势行业均集聚在高价值行业周围,而中部省份的比较优势行业一般而言处在远离高价值行业的另一个边缘。



注:资料来源为作者通过 R 软件 igraph 包绘制。

图 4 中省份和发达省份的比较优势行业在行业空间连通分布图中的位置

(二)最优产业升级路径计算

为得到产业升级路径,利用计算机算法领域的 Dijkstra 算法,在产业距离分别 ≤ 0.3 (此时有约一半的行业是连通的), ≤ 0.25 以及 ≤ 0.2 (此时有约四分之一的行业是连通的)的情况下,计算了从价值最低的 10 个行业升级至 3 个 5 万元以上的高价值行业(即图 1 所表明的 C40、C41 和 C24)的最短路径。具体地,该算法的步骤如下。

第一步,假定一个初始节点,给每个节点按如下规则分配一个与初始节点的距离值:初始节点与自身的距离值为 0,与其他所有节点的距离值均为无穷大。

第二步,把初始节点看作当前节点,所有其他节点记为未访问节点,所有的未访问节点称为未访问节点集。

第三步,对于当前节点,计算所有相邻节点与它的距离,将该距离与之前分配的距离(第一次分配的距离就是无穷大)进行比较,然后取其小者作为当前节点与该点的距离值。

第四步,当计算完该点与相邻节点的距离时,应将这些相邻点从未访问节点集中剔除,并记为已访问节点,而且今后不会再次考虑该点。

第五步,当最终的目标节点被记为已访问节点或者在未访问集中所有节点最小距离值仍为无穷大,则停止算法。

第六步,或者将未访问集中具有最小距离值的点作为当前节点,然后回到第三步。

根据该算法,3种情况下的最短路径见表5。

表5 低价值产业的升级路径

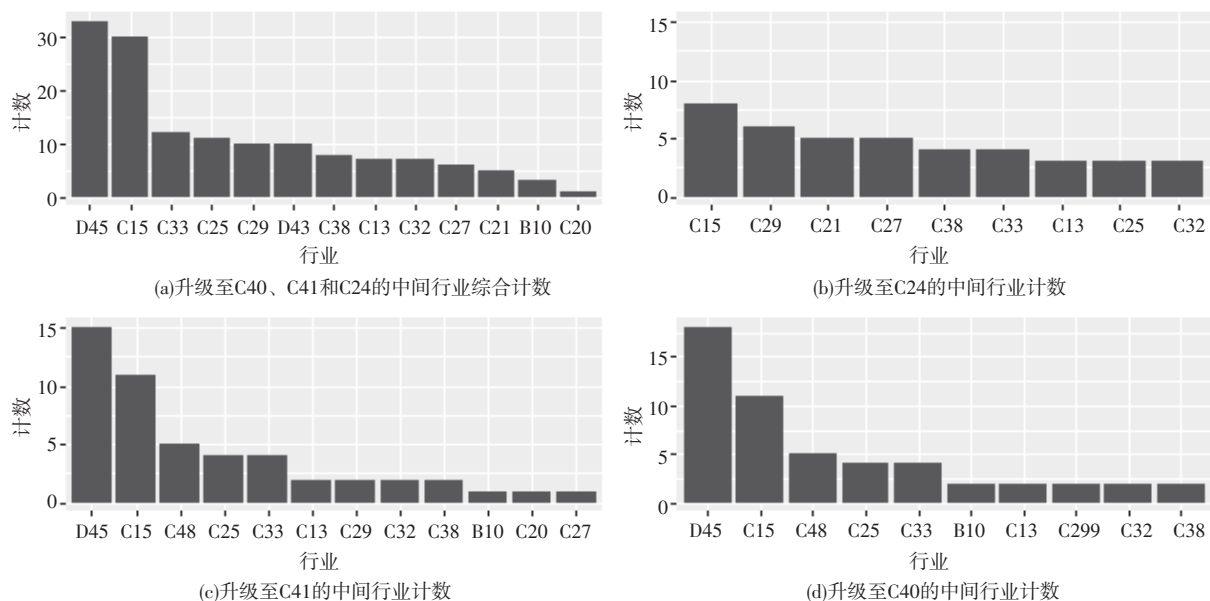
序号	A:距离 ≤ 0.3					B:距离 ≤ 0.25						C:距离 ≤ 0.2						
	D	V1	V2	V3	V4	D	V1	V2	V3	V4	V5	D	V1	V2	V3	V4	V5	V6
1	0.68	B6	C25	D45	C40	0.84	B6	C25	C29	D45	C40		B6					
2	0.60	C13	B10	D45	C40	0.60	C13	C15	D45	C40		0.68	C13	C33	C15	D46	C40	
3	0.20	D46	C40			0.20	D46	C40				0.20	D46	C40				
4	0.47	D44	D45	C40		0.55	D44	C15	D45	C40		0.57	D44	C15	D46	C40		
5	0.76	B11	C25	D45	C40	0.83	B11	C32	C36	D45	C40		B11					
6	0.49	B10	D45	C40		0.57	B10	C15	D45	C40		0.67	B10	C33	C15	D46	C40	
7	0.58	C31	C15	D45	C40	0.58	C31	C15	D45	C40		0.65	C31	C33	C15	D46	C40	
8	0.47	C25	D45	C40		0.63	C25	C29	D45	C40			C25					
9	0.72	C16	C25	D45	C40	0.81	C16	C32	C36	D45	C40		C16					
10	0.68	B9	B10	D45	C40	0.74	B9	C13	C15	D45	C40	0.82	B9	C13	C33	C15	D46	C40
11	0.70	B6	C25	D45	C41	0.84	B6	C25	C29	C24	C41		B6					
12	0.60	C13	C20	C27	C41	0.62	C13	C15	D45	C41		0.68	C13	C33	C15	D46	C41	
13	0.20	D46	C41			0.20	D46	C41				0.20	D46	C41				
14	0.49	D44	D45	C41		0.57	D44	C15	D45	C41		0.57	D44	C15	D46	C41		
15	0.77	B11	C25	D45	C41	0.85	B11	C32	C36	D45	C41		B11					
16	0.51	B10	D45	C41		0.59	B10	C15	D45	C41		0.67	B10	C33	C15	D46	C41	
17	0.60	C31	C15	D45	C41	0.60	C31	C15	D45	C41		0.65	C31	C33	C15	D46	C41	
18	0.49	C25	D45	C41		0.64	C25	C29	C24	C41			C25					
19	0.74	C16	C25	D45	C41	0.83	C16	C32	C36	D45	C41		C16					
20	0.70	B9	B10	D45	C41	0.76	B9	C13	C15	D45	C41	0.82	B9	C13	C33	C15	D46	C41
21	0.62	B6	C25	C27	C24	0.66	B6	C25	C29	C24			B6					
22	0.44	C13	C29	C24		0.44	C13	C29	C24			0.55	C13	C33	C15	C21	C24	
23	0.09	D46	C24			0.09	D46	C24				0.09	D46	C24				
24	0.39	D44	C21	C24		0.41	D44	C15	C24			0.44	D44	C36	C27	C24		
25	0.68	B11	C32	C36	C24	0.68	B11	C32	C36	C24			B11					
26	0.40	B10	C27	C24		0.43	B10	C15	C24			0.53	B10	C33	C15	C21	C24	
27	0.44	C31	C15	C24		0.44	C31	C15	C24			0.52	C31	C33	C15	C21	C24	
28	0.41	C25	C27	C24		0.45	C25	C29	C24				C25					
29	0.66	C16	C25	C27	C24	0.66	C16	C32	C36	C24			C16					
30	0.58	B9	C13	C29	C24	0.58	B9	C13	C29	C24		0.68	B9	C13	C33	C15	C21	C24

注:资料来源为作者计算。

表5中纵向分了3个区域,标志为A、B和C,分别是距离 ≤ 0.3 、距离 ≤ 0.25 和距离 ≤ 0.2 时从低价值行业升级至3个高价值行业C40(通信设备、计算机及其他电子设备制造业)、C41(仪器仪表及文化、办公用机械制造业)和C24(文教体育用品制造业)的路径。横向也分了3个区域,第1~10行、第11~20行和第21~30行分别是10个低价值行业升级至C40、C41和C24的具体路径。表中标识为“D”的列表示的是从低价值行业升级至高价值行业经过的总距离,该列中若无数据,说明从该行业出发是无法升级至高价值行业的。

表5表明有些行业要升级至高价值行业,必须经过较多的轮次,如有色金属矿采选业(B9)、农副食品加工业(C13)、非金属矿物制品业(C31),平均要三轮以上。特别地,当行业距离时(区域C),煤炭开采和洗选业(B6)、其他采矿业(B11)、石油加工、炼焦及核燃料加工业(C25)和烟草制品业(C16)至高价值行业根本就不存在升级的通路,若某省的支柱产业集聚在这些行业中,很可能就陷入了行业发展的低水平循环而无法升级。我们也注意到,这些升级至高价值行业比较困难的行业以资源类行业居多。

进一步,图 5 中统计了从 10 个低价值行业往 3 个高价值行业升级时,在距离 ≤ 0.3 、距离 ≤ 0.25 、距离 ≤ 0.2 三种情况下,哪些行业更容易出现在升级的路径上,这些行业作为承上启下的中间行业有着重要的指导意义。



注:资料来源为作者通过 R 软件 ggplot2 包绘制。

图 5 中间行业统计图

图 5(a)是对升级至 C40、C41 和 C24 三个高价值行业的全部中间行业计数(即对 90 条升级路径中,途径的中间行业的频数统计),图中可以看到燃气生产和供应业(D45)和饮料制造业(C15)在升级路径中出现多达 30 次以上。特别地,C15 在升级至 3 个高价值行业中都有着重要的中介作用,而 D45 则仅在升级至 C40 和 C41 的路径中十分重要,其在升级至 C24 的路径中一次也未出现。从图 5 也能观察到,C40 与 C41 相对 C24 处于更加边缘的位置,因此,升级至 C24 也许有更多的道路可以选择,但要升级至 C40 和 C41,则有些中间行业是无法避开的。这一点从图 5 也能看到,升级至 C24 的中间行业分布更为均匀,而升级 C40 和 C41 的中间行业分布则较为极端。

四、结论与启示

从要素禀赋结构入手,依托随机森林算法,本文提出了一种新的产业距离构造方式,基于该方法,绘制了产业空间分布图。与 Hidalgo 等^[3]高价值、低价值行业的中心-边缘布局不同,本文图中的高价值和低价值行业均聚集在边缘地带,只是处于不同的方向上而已。然后,基于该行业空间分布图,计算了从低价值行业升级逐步演化至高价值行业的最短路径。计算中发现了一些频繁出现的中间行业,它们是行业升级的必经之路,也发现资源型行业升级至高价值行业存在更多的困难。

上述结论对中西部地区结合自身的条件承接东部发达地区产业具有重要的借鉴意义。对中西部省份而言,在产业承接中,若能不断趋近这些高频出现的中间行业,其未来行业发展将愈为顺利,反之,则可能陷入低水平发展的陷阱而无法自拔。未来,如果数据可得,更加细分的行业研究更有意义。

参考文献

- [1] 贾康, 苏京春. 论供给侧改革[J]. 管理世界, 2016(3): 1-24.
- [2] HAUSMANN R, KLINGER B. Structural transformation and patterns of comparative advantage in the product space[R]. CID Working Papers, Cambridge: Harvard University, 2006: 128.
- [3] HIDALGO C A, KLINGER B, BARABASI A L, et al. The product space conditions the development of nations [J]. Science, 2007, 317(5837): 482-487.
- [4] 刘守英, 杨继东. 中国产业升级的演进与政策选择——基于产品空间的视角[J]. 管理世界, 2019(6): 81-94.
- [5] 韩自然, 芮明杰. 要素结构、信息化与地区产业体系优化[J]. 技术经济, 2019, 38(6): 46-57.
- [6] 林毅夫. 新结构经济学[M]. 北京: 北京大学出版社, 2014.

- [7] 张其仔, 李颖. 中国产业升级机会的甄别[J]. 中国工业经济, 2013, 302(5): 44-56.
- [8] GEREFFI G. International trade and industrial upgrading in the apparel commodity chain [J]. Journal of International Economics, 1999, 48(1): 37-70.
- [9] PORTER M E. The competitive advantage of nations[M]. New York: Free Press, 1990.
- [10] 邓向荣, 曹红. 产业升级路径选择: 遵循抑或偏离比较优势[J]. 中国工业经济, 2016(2): 52-67.
- [11] 张亨, 刘林青. 中美产业升级的路径选择比较——基于产品空间理论的分析[J]. 经济管理, 2016(8): 18-28.
- [12] 张亨, 刘林青. 产品复杂性水平对中日产业升级影响的比较研究[J]. 经济管理, 2017(5): 115-129.
- [13] BOSCHMA R, MINONDO A, NAVARRO M. The emergence of new industries at the regional level in Spain: A proximity approach based on product relatedness[J]. Economic Geography, 2013, 89(1): 29-51.
- [14] 余泳泽, 张先轸. 要素禀赋、适宜性创新模式选择与全要素生产率提升[J]. 管理世界, 2015(9): 13-31.
- [15] 刘胜, 顾乃华, 陈秀英. 全球价值链嵌入、要素禀赋结构与劳动收入占比[J]. 经济学家, 2016(3): 96-104.
- [16] 何强. 要素禀赋、内在约束与中国经济增长质量[J]. 统计研究, 2014, 31(1): 70-77.
- [17] 钱学锋, 王备. 中间投入品进口、产品转换与企业要素禀赋结构升级[J]. 经济研究, 2017(1): 58-71.
- [18] FRUCHTERMAN T M J, REINGOLD E M. Graph drawing by force-directed placement [J]. Software-Practice and Experience, 1991, 21: 1129-1164.
- [19] HAUSMANN R, HWANG J, RODRIK D. What you export matters[J]. Journal of Economic Growth, 2007, 12: 1-25.
- [20] 李敏纳, 蔡舒, 张慧蓉, 等. 要素禀赋与黄河流域经济空间分异研究[J]. 经济地理, 2011, 31(1): 14-20.
- [21] 贾润崧, 张四灿. 中国省际资本存量与资本回报率[J]. 统计研究, 2014, 31(11): 35-42.

Factor Endowments, Industry Distances and Choosing Industry Upgrading Path

Chen Pu

(School of Economics and Management, East China Jiaotong University, Nanchang 330013, China)

Abstract: Based on the industrial distance constructed by inputs, using the natural conditions, human resources, material capital, science and technology, economic structure and regime which include 23 kinds of factor endowments, we construct the "distance" measure of all manufacturing industry. According to the distance, the spatial distribution map of the industry is drawn, and then the Dijkstra algorithm is used to calculate the shortest path from 10 low-value industries upgrading to three high-value industries. The study finds that while resource-based industries upgrade to high-value industries, they need experience more rounds, and sometimes there even are no paths upgrading to high-value industries. Some of the intermediate industry has always appeared in the upgrading path from low-value industries to high-value industries. If the central and western provinces could approach these industries when they undertake eastern industries, their ways of industry development and upgrading is bound to become broader.

Keywords: product spaces; factor endowments; industries upgrading