

国家经济与互联网企业营收的代数模型

林霞^{1*}, 刘雨喆^{1#}, 马欣^{2#}

¹贵州大学数学与统计学院, 贵州 贵阳

²河南工程学院理学院, 河南 郑州

收稿日期: 2025年2月10日; 录用日期: 2025年2月24日; 发布日期: 2025年3月18日

摘要

向量空间模型由Salton-Wong-Yang在20世纪70年代提出, 并被用于文献的查重和检索, 由于其基本原理简单而受到广泛应用。本文以国内三个大型电子网络企业为例, 将此数学模型应用到实测数据, 并用来衡量这些实测数据之间的线性相关程度。研究发现, 上述三个企业的营收与我国消费者的消费能力呈现极强的关联, 且数据中没有出现表现异常的个例。

关键词

宏观经济, 互联网企业, 向量内积, 向量夹角

Algebraic Model of National Economy and Internet Enterprise Revenue

Xia Lin^{1*}, Yuzhe Liu^{1#}, Xin Ma^{2#}

¹School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

²School of Science, Henan University of Engineering, Zhengzhou Henan

Received: Feb. 10th, 2025; accepted: Feb. 24th, 2025; published: Mar. 18th, 2025

Abstract

Salton, Wong, and Yang proposed the vector space model in the 1970s, and have been widely used for literature plagiarism detection and retrieval due to its simple basic principles. This paper takes three large domestic electronic network enterprises as examples, applies this mathematical model to turnovers, and uses it to measure the degree of linear correlation between these turnovers. This study found that the revenue of the three companies mentioned above is strongly correlated with

*第一作者。

#通讯作者。

文章引用: 林霞, 刘雨喆, 马欣. 国家经济与互联网企业营收的代数模型[J]. 电子商务评论, 2025, 14(3): 609-615.
DOI: 10.12677/ecl.2025.143746

Chinese consumers' consumption ability, and there are no abnormal cases in the data.

Keywords

Macro Economy, Internet Corporate, Inner Products (Dot Products), Included Angles of Vectors

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

代数在现实应用中非常广泛，其涉及的领域包括密码学算法(例如 Shamir 秘密共享算法[1])，几何与设计(例如作图法[2] [3])，相似文本检测(例如文献重复率检测，搜索引擎等[4])。在数据统计中，可将多个经济指标看作向量的各个分量，分析其相关性。

在当今数字基础设施逐步加快建设的基础上，我国互联网经济蓬勃发展，腾讯、阿里巴巴和哔哩哔哩等大型企业作为我国互联网企业的典型代表，在市场中的扮演重要角色。在各大企业数字化能力飞跃提升的背景下，互联网企业经营状况不仅受内部管理和市场竞争的影响，宏观经济也发挥着关键作用。而国家的经济形势与企业经营之间的关系一直是应用数学、经济学、以及管理学等多学科之间的共同议题，受到国内外许多学者关注(例如[5]-[9])。

本文聚焦于 2019~2023 年这一具有代表性的时间段，通过代数视角，构建描述宏观经济的向量和反映互联网大型企业经营状况(如收入、毛利等)的向量，计算二者的夹角的余弦值，深入剖析宏观经济与企业营收之间的内在关系。代数视角是独特的，以往该视角仅用于文本检测，本文首次将其运用于经济学领域。

本文行文结构如下：本文的第 2 节包含两个小节，分别论述了本文的代数理论依据以及向量空间模型的建模方法；在第 3 节中，本文利用向量空间模型计算了我国整体经济形势与所选研究对象的营收之间的相关性；第 4 节是本文的初步研究成果，表明了宏观经济形势与互联网企业经营数据之间的强相关性。

2. 向量空间模型

本节的第一部分从高等代数的角度回忆向量空间、内积以及范数的定义，并简单介绍向量空间模型(vector space model, VSM)。本节的第二部分将提供 2019~2023 年期间我国国民的经济数据以及一些企业的营销数据，并将它们实现到上述向量空间模型中。

2.1. 向量的共线与内积

向量空间在实际社会研究和数据分析中主要使用的是 Euclid 空间 $\mathbb{R}^n$ 。从抽象代数的观点看，实数集 $\mathbb{R}$ 上的向量空间(vector space)指的是 $\mathbb{R}$ 到某个 Abel 群 M 的自同态环 $\text{End}M$ 上的环同态 $\mathbb{R} \rightarrow \text{End}M$ 。具体地说，它是一个以形如 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 的 n 元实数组为元素的集合，并且具有加法(addition)运算

$$\mathbf{x} + \mathbf{y} = (x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

以及纯量乘法(scalar multiplication)运算

$$r\mathbf{x} = r(x_1, x_2, \dots, x_n) = (rx_1, rx_2, \dots, rx_n) \quad (\forall r \in \mathbb{R}),$$

可参考[3]的第3章,[10]的第3章,或[11]的第3章。对于两个向量 $\mathbf{x}=(x_1, x_2, \dots, x_n)$ 和 $\mathbf{y}=(y_1, y_2, \dots, y_n)$ 。如果存在 $k \in \mathbb{R}$ 使得 $k\mathbf{x} = \mathbf{y}$, 则称这两个向量是线性相关的或者共线的(colinear), 并用分式

$$F_{\mathbf{x}, \mathbf{y}} = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$$

来衡量二者的线性相关性, 其中:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i, \quad \|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{i=1}^n x_i^2}, \quad \|\mathbf{y}\| = \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle} = \sqrt{\sum_{i=1}^n y_i^2}.$$

$\langle \mathbf{x}, \mathbf{y} \rangle$ 被成为内积(inner product), 记号 $\|\cdot\|$ 表示对向量求范数(norm)。分式 $F_{\mathbf{x}, \mathbf{y}}$ 的值始终介于-1 和 1 之间, 因此也被用来定义为两个向量的夹角, 即:

$$\theta = \arccos F_{\mathbf{x}, \mathbf{y}}.$$

显然, 当两个向量夹角越接近 0° 或者 180° , 分式 $F_{\mathbf{x}, \mathbf{y}}$ 的绝对值越接近 1。上述性质常被用于构造“向量空间模型”。该模型由 Salton-Wong-Yang 在 20 世纪 70 年代提出, 并被用于信息检索[4]。如今可以看到该项应用非常成功, 许多搜索引擎和检索模型(包括文字检索与查重, 声乐检索等)都是基于此开发。

2.2. 国民经济数据与大型电子网络企业营收的向量空间模型

表 1 是 2019~2023 年期间我国 $F_{\mathbf{x}, \mathbf{y}}$ 的国民的经济数据, 由国家统计局公布, 它能很好地反映我国在 2019~2023 年的经济形势, 见[12]。可以看到, 从 2019 年到 2023 年期间, 尽管经历了疫情带来的经济下滑, 但是我国国民总收入(单位: 亿元)、人均国内生产总值(单位: 元)、城镇/农村居民人均可支配收入(单位: 元)、以及第三产业增加值(单位: 亿元)仍然呈现稳步上升的趋势。

Table 1. National economic data

表 1. 国民经济数据

年份	国民总收入(亿元)	人均国内生产总值(元)	城镇居民人均可支配收入(元)	农村居民人均可支配收入(元)	第三产业增加值(亿元)
2023	1249990.6	89,358	51,821	21,691	688238.4
2022	1194401.4	85,310	49,283	20,133	642727.1
2021	1141230.8	81,370	47,412	18,931	614476.4
2020	1005451.3	71,828	43,834	17,131	551973.7
2019	983751.2	70,078	42,359	16,021	535371

表 2~4 是本文作者分别从腾讯、阿里巴巴、哔哩哔哩三个企业获取的 2019 年~2023 年的部分经营数据, 其中包括收入、毛利、以及除阿里巴巴外另两个企业的增值服务收入(单位: 百万元), 见[13]-[15]。

Table 2. Turnovers of Tencent

表 2. 腾讯经营数据

年份	腾讯收入(百万元)	腾讯毛利(百万元)	腾讯增值服务收入(百万元)	占收入总额(百分比)
2023	609015	293,109	298,375	49%
2022	554552	238,746	287,565	52%
2021	560118	245,944	291,572	52%
2020	482064	221,532	264,212	55%
2019	377289	167,533	199,991	53%

Table 3. Turnovers of Alibaba
表 3. 阿里巴巴经营数据

年份	阿里巴巴收入(百万元)	阿里巴巴净利润(百万元)
2023	941,168	71,332
2022	868,687	65,573
2021	853,062	47,079
2020	717,289	143,284
2019	509,711	140,350

Table 4. Turnovers of Bilibili
表 4. 哔哩哔哩经营数据

年份	哔哩哔哩收入(百万元)	哔哩哔哩毛利(百万元)	哔哩哔哩增值服务收入(百万元)	占收入总额(百分比)
2023	22,528	5442	9910	44%
2022	21,899	3849	8715	40%
2021	19,384	4043	6935	36%
2020	11,999	2840	3846	32%
2019	6778	1190	1641	24%

接下来, 本文将按照年度来对各表所提供的数据的相关性展开分析。虽然各表所使用的单位不同, 但是这可以被忽略。这是因为, 从计算上, 有

$$F_{x,y} = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} = \left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle,$$

其中 $\frac{x}{\|x\|}$ 和 $\frac{y}{\|y\|}$ 是向量的单位化, 由于分子与分母单位一致, 所以在计算过程中单位被约分, 故 $\frac{x}{\|x\|}$ 和 $\frac{y}{\|y\|}$ 的单位为 1。本文选取了 2019 年~2023 年五个年度的数据, 因此选择向量空间 $\mathbb{R}^5$ 作为本文所要使用的向量空间模型, $\mathbb{R}^5$ 中的向量 $(x_1, x_2, x_3, x_4, x_5)$ 的分量 $x_1 \sim x_5$ 分别表示 2023, 2022, 2021, 2020, 2019 年度的数据。具体如下:

A: 国民总收入(见表 1 的第 2 列), 其中,

$$A = (124990.6, 1194401.4, 1141230.8, 1005451.3, 983751.2),$$

下文的向量类似, 为了写作简便起见, 本文省略;

B: 人均国内生产总值(见表 1 的第 3 列);

C: 城镇居民人均可支配收入(见表 1 的第 4 列);

D: 农村居民人均可支配收入(见表 1 的第 5 列);

E: 第三产业增加值(见表 1 的第 6 列);

F: 腾讯年度收入(见表 2 的第 2 列);

G: 腾讯年度毛利(见表 2 的第 3 列);

H: 腾讯增值服务收入(见表 2 的第 4 列);

I: 腾讯增值服务占收入总额的百分比(见表 2 的第 5 列);

J: 阿里巴巴年度收入(见表 3 的第 2 列);

K：阿里巴巴年度净利润(见表 3 的第 3 列)；

L：哔哩哔哩年度收入(见表 4 的第 2 列)；

M：哔哩哔哩年度毛利(见表 4 的第 3 列)；

N：哔哩哔哩增值服务收入(见表 4 的第 4 列)；

P：哔哩哔哩增值服务占收入额的百分比(见表 4 的第 5 列)。

以上数据在各企业的财务报告中有着非常详细的分析，本文不再过多赘述。表 5 给出了上述向量的单位化向量。例如向量 **A**，其范数为：

$$\|\mathbf{A}\| = \sqrt{124990.6^2 + 1194401.4^2 + 1141230.8^2 + 1005451.3^2 + 983751.2^2} = 2173320.92279,$$

于是，得：

$$\begin{aligned} \mathbf{A}' &= \frac{\mathbf{A}}{\|\mathbf{A}\|} = \frac{(124990.6, 1194401.4, 1141230.8, 1005451.3, 983751.2)}{2173320.92279} \\ &= (0.50, 0.48, 0.46, 0.40, 0.39). \end{aligned}$$

Table 5. Vector unitization

表 5. 向量单位化

原向量	单位化后对应向量
A	$\mathbf{A}' = (0.50, 0.48, 0.46, 0.40, 0.39)$
B	$\mathbf{B}' = (0.50, 0.48, 0.46, 0.40, 0.39)$
C	$\mathbf{C}' = (0.49, 0.47, 0.45, 0.42, 0.40)$
D	$\mathbf{D}' = (0.51, 0.48, 0.45, 0.41, 0.38)$
E	$\mathbf{E}' = (0.51, 0.47, 0.45, 0.41, 0.39)$
F	$\mathbf{F}' = (0.52, 0.47, 0.48, 0.41, 0.32)$
G	$\mathbf{G}' = (0.55, 0.45, 0.46, 0.42, 0.32)$
H	$\mathbf{H}' = (0.49, 0.47, 0.48, 0.44, 0.33)$
I	$\mathbf{I}' = (0.42, 0.45, 0.45, 0.47, 0.45)$
J	$\mathbf{J}' = (0.53, 0.49, 0.48, 0.40, 0.29)$
K	$\mathbf{K}' = (0.31, 0.29, 0.21, 0.63, 0.62)$
L	$\mathbf{L}' = (0.57, 0.56, 0.49, 0.30, 0.17)$
M	$\mathbf{M}' = (0.65, 0.46, 0.48, 0.34, 0.14)$
N	$\mathbf{N}' = (0.64, 0.56, 0.45, 0.25, 0.11)$
P	$\mathbf{P}' = (0.55, 0.50, 0.45, 0.40, 0.30)$

3. 相关性分析

将表 5 第 2 列中的向量分为两组： $\mathbf{A}', \mathbf{B}', \mathbf{C}', \mathbf{D}', \mathbf{E}'$ (宏观经济向量组)与 $\mathbf{F}', \mathbf{G}', \dots, \mathbf{P}'$ (企业经营向量组)。为了分析两组数据的相关性，本文需要计算所有的 $F_{X,Y}$ ，其中， $\mathbf{X} \in X := \{\mathbf{A}', \mathbf{B}', \mathbf{C}', \mathbf{D}', \mathbf{E}'\}$ ，

$\mathbf{Y} \in Y := \{\mathbf{F}', \mathbf{G}', \dots, \mathbf{P}'\}$ 。以 $\mathbf{X} = \mathbf{A}'$ 且 $\mathbf{Y} = \mathbf{F}'$ 的情形为例，有 $F_{X,Y} = \frac{\langle \mathbf{A}', \mathbf{F}' \rangle}{\|\mathbf{A}'\| \cdot \|\mathbf{F}'\|} = \langle \mathbf{A}', \mathbf{F}' \rangle \approx 0.997$ ，其中 $\mathbf{A}'$ 和

$\mathbf{F}'$ 已经由表 5 给出。重复上述计算步骤，就得到了表 6。

Table 6. The linear dependence of vectors
表 6. 向量的相关性

	A'	B'	C'	D'	E'
F'	0.997	0.997	0.996	0.998	0.997
G'	0.996	0.996	0.995	0.997	0.996
H'	0.997	0.997	0.997	0.998	0.997
I'	0.992	0.992	0.995	0.991	0.992
J'	0.994	0.994	0.992	0.995	0.994
K'	0.88	0.88	0.893	0.881	0.885
L'	0.965	0.965	0.958	0.966	0.963
M'	0.955	0.955	0.950	0.958	0.956
N'	0.937	0.937	0.928	0.939	0.936
P'	0.994	0.994	0.993	0.996	0.995

4. 结论

通过数量积和夹角余弦值的计算与分析,发现宏观经济因素与互联网企业营收有着密切的关联。其关联表现在以下几个方面:

1) 用于描述宏观经济的向量 $A' \sim E'$ 与用于描述本文所研究企业营业额的向量 $F' \sim P'$ 的线性相关性 $F_{X,Y}$ 表现非常显著。从表 6 第 2~6 行、第 8~11 行的数据来看, $F_{X,Y}$ 的数值介于 0.9 与 1 之间,总体接近 1。

2) 净利润与国家经济形势线性相关性显著。代表净利润的向量 K 与 $A' \sim E'$ 的线性相关性约为 0.88 (见表 6 第 7 行)。由于没有另外两家企业的净利润数据,因此该项数据暂时没有可以对比的案例。净利润的数值等于毛利润扣除上税的税额,其对公民的直接消费能力的反映弱于毛利润,这是因为毛利润直接反映了消费者的真实消费额。因此,它与 $A' \sim E'$ 线性相关性弱于其他数据具有一定合理性。

3) 表 6 所反映的数据整体趋于稳定,未出现反常的个例。与阿里巴巴和哔哩哔哩相比,腾讯的营业数据(见表 6 第 2~5 行)与国家经济形势的线性相关性的均值最高,为

$$\bar{F}_{X \times \{F', G', H', I'\}} = \frac{1}{20} \sum_{X \times Y \in \{A', \dots, E'\} \times \{F', G', H', I'\}} F_{X,Y} = 0.99565,$$

且就表 6 所给数值而言,线性相关性的数值表现非常稳定。初步推断这与该企业的增值服务营收占比(基本略高于 50%左右,见表 2)较高有关。与之相对的,哔哩哔哩的增值服务产生的营收占比较腾讯而言更低(见表 6 第 10 行),因此,其呈现线性相关性略微下降。根据表 6 所给的数据,可以算得此下降百分比为

$$\left(1 - \frac{1}{\bar{F}_{X \times \{F', G', H', I'\}}} \cdot \frac{1}{5} \sum_{X \in \{A', \dots, E'\}} F_{X, N'}\right) \times 100\% \approx 6.05\%。$$

此外,此数据呈现了一个有趣的现象:将哔哩哔哩的增值服务折算到总营业收入中后(见表 6 最后一行),与国家经济形势的线性相关性数值为 0.9944,呈现出与国家经济形势极高的相关性。

作者声明

本文三位作者对文章的贡献是等同的;本文作者承诺本文内容不涉嫌反动言论;本文的通讯作者为

上述承诺负第一责任人,且承诺本文无数据篡改、编造、和作假,无参考文献滥引现象,并对本文数据来源的真实性负第一责任。

基金项目

本文由河南工程学院教育教学改革研究项目(编号:2024JYZDX029)、贵州大学高等教育研究项目(编号:703217243301);贵州大学引进人才科研启动基金项目(贵大人基合字(2023)16号,(2022)53号,(2022)65号)、以及贵州省科技计划项目(黔科合基础-ZK[2024]一般 066)资助。

参考文献

- [1] Shamir, A. (1979) How to Share a Secret. *Communications of the ACM*, **22**, 612-613.
<https://doi.org/10.1145/359168.359176>
- [2] Leinster, T. (2024) Galois Theory. arXiv: 2408.07499.
<https://arxiv.org/abs/2408.07499>
- [3] 徐云阁, 章超, 廖军. 高等代数[M]. 北京: 科学出版社, 2021: 308-309.
- [4] Salton, G., Wong, A. and Yang, C.S. (1975) A Vector Space Model for Automatic Indexing. *Communications of the ACM*, **18**, 613-620. <https://doi.org/10.1145/361219.361220>
- [5] Heil, D. and Prieger, J.E. (2010) Encyclopedia of E-Business Development and Management in the Global Economy. In: Heil, D. and Prieger, J.E., Eds., *The Macroeconomic Impacts of E-Business on the Economy*, IGI Global, 1-11.
<https://www.igi-global.com/chapter/macro-economic-impacts-business-economy/41163>
- [6] 王帅, 张友祥. 实体企业金融投资收益对中国宏观经济波动的影响分析[J]. 贵州财经大学学报, 2020, 1(3): 40-48.
- [7] Welfens, P.J.J. (2022) The Internet in Macroeconomic Analysis and Growth Analysis. In: Welfens, P.J.J., Ed., *Inter-neteconomics.net: Macroeconomics, Deregulation, and Innovation*, Springer, 79-112.
https://doi.org/10.1007/978-3-540-24762-3_2
- [8] 裴琳. 营收水平视角下企业决策速度对消费者忠诚度的影响: 以网络零售企业为例[J]. 商业经济研究, 2022(2): 86-89.
- [9] Byrne, D.M. (2022) The Digital Economy and Productivity. *Finance and Economics Discussion Series (FEDS)*, **38**, 1-52. <https://doi.org/10.17016/feds.2022.038>
- [10] Olver, P.J. and Shakiban, C. (2022) Applied Linear Algebra. Springer. <https://doi.org/10.1007/978-3-319-91041-3>
- [11] Leon, S.J. (2015) Linear Algebra with Applications. Pearson.
<https://testbankdeal.com/sample/linear-algebra-with-applications-9th-edition-leon-solutions-manual.pdf>
- [12] 中华人民共和国国家统计局. 年度数据: 国民经济核算[EB/OL]. 2024-12-30.
<https://data.stats.gov.cn/easyquery.htm?cn=C01>, 2025-01-23.
- [13] 腾讯控股有限公司. 腾讯年度财务报告(2019年-2023年)[EB/OL].
<https://www.tencent.com/zh-cn/investors/financial-reports.html>, 2025-01-03.
- [14] 阿里巴巴集团控股有限公司. 阿里巴巴集团年度财务报告(2019年-2023年)[EB/OL].
<https://ali-home.alibaba.com/ir-financial-reports-financial-results>, 2025-01-03.
- [15] 上海宽娱数码科技有限公司. Bilibili 年度/中期报告(2019年-2023年)[EB/OL].
<https://ir.bilibili.com/cn/financial-information/#annual-and-interim-report>, 2025-01-03