

Data
Envelopment
Analysis

数据包络分析

让数据自己说话

Let the **D**ata **S**peak for **T**hemselves

— [美] Joe Zhu 著 · 公彦德 李想 译 —



科学出版社

数据包络分析

让数据自己说话

[美] Joe Zhu 著
公彦德 李 想 译



本书由江苏省优势学科建设工程及江苏省高校青蓝工程项目资助

科学出版社

北京

图字：01-2016-1573 号

内 容 简 介

本书通过选购笔记本电脑的案例引入数据包络分析（data envelopment analysis, DEA）的基本概念和模型，并运用简单的 EXCEL 模型让读者更易理解和运用 DEA。本书着重阐述如何将 DEA 作为一种运营分析工具，进行绩效评估，找出基准标杆。探讨的主题包括：平衡的基准，最佳实践，相对效率的概念及绩效提升。特别地，本书探讨的数据分析方法能够帮助一个组织重新审视它对生产效率高低的已有观点是否合理，并为组织不断改进提供指导。本书是市面上唯一一本不要求读者具备线性规划和线性代数知识，就可以熟练应用 DEA 方法的 DEA 方法指导用书。

本书可供与管理决策、评价和优化有关的经济、管理、数学等领域的科研与应用工作者阅读，也可作为本科生、研究生和相关专业教师的参考用书。

Data Envelopment Analysis-let the Data Speak for Themselves,
Joe Zhu. Copyright © 2014 by Joe Zhu. All Rights Reserved.

图书在版编目（CIP）数据

数据包络分析：让数据自己说话/（美）朱乔（Zhu, J.）著；公彦德，李想译. —北京：科学出版社，2016.3

书名原文：Data Envelopment Analysis-let the Data Speak for Themselves

ISBN 978-7-03-047967-9

I. ①数… II. ①朱…②公…③李… III. ①统计数据—统计分析 IV. ①O212.1

中国版本图书馆 CIP 数据核字（2016）第 062231 号

责任编辑：胡 凯 王腾飞/责任校对：张凤琴

责任印制：张 倩/责任设计：许 瑞

科学出版社出版

北京东黄城根北街16号

邮政编码：100717

<http://www.sciencep.com>

文林印务有限公司印刷

科学出版社发行 各地新华书店经销

*

2016年3月第 一 版 开本：720×1000 1/16

2016年3月第一次印刷 印张：9 1/2

字数：120 000

定价：49.00 元

（如有印装质量问题，我社负责调换）

序

今日的大数据正伴随着云计算、移动互联网的发展，极大地影响着全球的经济和社会发展，改变着我们对世界的理解和观察方式。在这样的背景下，国家治理和审计科学等都需要以大数据为基础，而审计与大数据的高度互联让审计行业的升级转型迫在眉睫。

毫无疑问，为了实现审计与大数据的高度互联，我们必须探索和创新相应的理论方法和实用工具，审计人员和机构必须着力提高组织和个人在数据分析与挖掘方面的技能。因此，利用有效的数据分析方法提升审计能力已经成为审计创新的急迫任务。将已在数据分析领域得到成功应用的数据包络分析方法引入审计分析体系或许是我们义不容辞的责任。

基于此，为了满足大数据时代的审计要求，南京审计大学成立了“审计与评估国际中心”，着眼于审计和评估层面的运营分析，探索大数据背景下审计和评估研究的新方法和新应用，推动审计与评估的有效结合，并对运营管理在组织创新方面的定位建立全新的认识，促进审计行业的新发展。

《数据包络分析——让数据自己说话》一书是南京审计大学“审计与评估国际中心”主任 Joe Zhu 教授在研究与推广数据包络分析方法的过程中，积多年的研究心得凝结成的一本“普及型”专著。作为数据包络分析领域的国际顶级专家，Joe Zhu 教授善于用通俗的语言将高深的原理呈现于课堂之上，把自己高瞻远瞩的思考传授给年轻一代，值得我们学

习。该书虽然并未提及数据包络分析方法在现代审计中的具体应用，甚至很少提及“审计”这一名词，但它对推动国内审计新方法与新手段的探索与应用无疑提供了一条可行的途径。

还值得一提的是，为了让不同领域的学者和实务工作者都能尽其所需的掌握数据包络分析的核心方法，Joe Zhu 教授运用人们熟悉的 EXCEL 软件为依托，将数据包络分析作为一种运营分析工具来进行绩效评估，帮助组织重新审视自己的发展战略和改进方向。更重要的是，以往出版的有关数据包络分析的书籍几乎均要求读者具备较深的数学理论基础，特别是线性规划和线性代数的知识，该书恰恰相反，大概是目前唯一不要求读者具备专业数学知识就可以熟练掌握数据包络分析方法的专著。

基于数据包络分析方法的特色优势，我可以乐见，该书所预示的方向将为大数据时代的国家治理和现代审计提供理论和方法性工具，从而提升治理和审计的效率效果，有助于解决当前国家治理和组织管理背景下审计供给的明显不足与审计需求的日益增长之间的矛盾。

是为序。



猴年初于润泽湖畔

中文版序

非常高兴得知我的这本小作将在大陆用中文出版。数据包络分析虽然是源自西方的方法，但一定会对改革、开放和发展中的中国的效率评价和数据分析有所帮助。

这本书当初是为希望知道并应用数据包络分析方法，但并非受过专业训练的读者所写，换句话说，写它的目的是为了推动数据包络分析方法的进一步普及。从 Charnes, Cooper 和 Rhodes 三位前辈于 1978 年共同创建这门当时主要用于评价具有多个输入和多个输出决策单元间的相对有效性的非参数方法以来，实际应用一直是它的主要目标。

遗憾的是，直到本世纪初，作为测量和分析同类组织之间相对绩效的一个工具，这一方法仍然还主要停留在学术研究的层面。其主要原因不外乎两个，其一是因为该方法在技术上主要以线性规划为基础，从而难住了一批人；其二是因为较长时期中主要用于决策单元间的效率评价，从而流失了一些人。

可喜的是，事情在新世纪中有了转机。首先是 2000 年之后不久，有关数据包络分析的算法已经能够在微软公司 Excel 这样的软件中直接运行，使得用户只需要很少的线性规划知识、甚至对此几乎一窍不通也能顺利使用 DEA。其次，人们发现其实这种方法可以扩展成更

广义的评估运营绩效的强有力的工具。在 2013 年的《麻省理工大学斯隆管理评论》（*MIT Sloan Management Review*）上，Sherman 教授和我将这种方法称之为“平衡的基准化（balanced benchmarking）”，因为通过对一组商店或办公室的多个绩效指标进行度量，该方法能够挖掘出隐藏其中的最佳实践单元；利用有关的分析结果，管理者能够显著提升各类经营单元的绩效或者指明它们绩效提升的方向。因而，它已经被成功地运用于世界各地不同类型的组织单元，包括医院、卫生维护组织、军事单位、大学、城市、法院、投资经理、零售商、金融机构、物流及制造企业等。

在中国经济的飞速发展中，有太多的需要提升经营绩效的决策单元了，因此我相信数据包络分析方法会有较大的应用市场。衷心希望中国的读者能喜欢我这本普及读本。

南京审计大学作为目前中国唯一以“审计”命名的高校（即使在全球以审计为核心专业组织的学校也不多见）肩负着发展现代审计理论与方法的历史使命。虽然数据包络分析目前还不是一种显性的审计评价工具，但我相信本书所介绍的方法至少能给审计科学的发展提供借鉴，甚至可行理论与方法基础。因此，我十分欣赏南京审计大学全力支持中译本的出版，特别感谢晏维龙校长的亲自作序。相信他的远见卓识定有丰硕成果。

要感谢的人还有很多，难以一一列举，包括东南大学的胡汉辉教授，他是一位一直努力推动将这种方法拓展至广义数据分析领域（例

如它在审计方面的应用)的学者;包括本书译者公彦德和李想,作为极有前途的年轻学者,他俩对本书的翻译尽心尽力,相信一定会准确达意;包括科学出版社的编辑们,一本书被读者所喜欢离不开他们这些幕后英雄。

谨以此书献给发展中的中国,我的同行和读者们。

A handwritten signature in black ink, appearing to read "Joe Zhu". The signature is fluid and cursive, with the first name "Joe" and the last name "Zhu" clearly distinguishable.

2016年2月于美国,麻省

序言(原版)

1978年, Charnes, Cooper 与 Rhodes 三位学者共同创建了“数据包络分析 (data envelopment analysis, DEA)”方法^①。直到 21 世纪初, 作为测量和分析同类组织之间相对绩效的一个工具, DEA 仍然还停留在学术研究的层面。这主要是因为, DEA 在技术上是以线性规划为基础的。用户必须非常熟悉线性规划, 才能有效求解 DEA 模型并合理解读计算结果。但 2000 年后, DEA 已经能够在微软公司 Excel 这样的软件中直接运行, 用户只需要很少的线性规划知识, 甚至对此一窍不通, 也能顺利使用 DEA。

与此同时, 众多领域的学者与实务工作者都认识到, DEA 是评估运营绩效的强有力工具。DEA 已经被成功地运用于世界各地不同类型的组织单元, 包括医院、卫生维护组织、军事单位、大学、城市、法院、投资经理、零售商、金融机构、物流及制造企业等。DEA 有助于识别出这些组织中表现最佳的单元, 并能够为其他单元指明绩效提升的方向。在 2006 年, 作为 DEA 的共同创建者, Charnes 和 Cooper 被 INFORMS 授予 Impact Prize^②, 以表彰他们在 DEA 上的开创性工作。

① Charnes A, Cooper W, Rhodes E. 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research*. 2(6),429-444.

② Impact Prize 奖项由运营研究与管理科学研究所 (Institute for Operations Research and the Management Sciences, 简称 INFORMS) 每两年颁发一次, 以表彰在该领域内产生广泛影响的学术成果。

2013 年《麻省理工大学斯隆管理评论》（*MIT Sloan Management Review*）期刊上，Sherman 和 Zhu 将 DEA 称之为“平衡的基准化（balanced benchmarking）”。他们指出，通过 DEA 对一组商店或办公室的多个绩效指标进行度量，能够挖掘出隐藏其中的最佳实践单元；利用 DEA 的分析结果，管理者能够显著提升各类商业单元的绩效。

DEA 是一种以数据导向的运营分析方法。它能够分析包含多重绩效指标在内的情形，将多个维度的数据整合为一个综合指标，并指出绩效提升的方向。目前已经有很多书籍介绍传统或新近发展的 DEA 模型。但是，其中很多仍然需要读者具备线性规划的数学基础。本书是目前唯一从应用角度阐述 DEA 方法的书，不要求读者具备线性规划和线性代数的知识。

关于本书

本书主要讲解 DEA 的基本概念和模型，以便读者理解和运用 DEA。本书着重阐述如何将 DEA 作为一种运营分析工具，进行绩效评估，并找出基准标杆。

本书探讨的主题包括：平衡的基准化，最佳实践，相对效率的概念，以及绩效提升。相关案例主要来源于生产与服务行业。此外，本书探讨的数据分析方法能够帮助一个组织重新审视对生产效率高低的已有观点是否合理，并为组织地不断改进提供指导。

通过学习这本书，读者有望达到以下目标：

(1) 面对具有多个甚至相互冲突的关键绩效指标时，学会利用数据分析的方法，实现平衡的基准化。

(2) 针对特定的应用问题，建立平衡的基准化模型。

(3) 在传统的数据分析工具难以奏效的情况下，找出最佳实践单元。

(4) 将不同的度量指标整合，而不需要将它们转化为货币价值。

(5) 有效处理具有不同度量单位的数据以及非经济化因素。

(6) 进行生产效率与市场分析，以及运营绩效评估。

本书包括以下 7 章内容。

第 1 章介绍 DEA 作为基准化工具如何用于运营分析。并通过一个选购笔记本电脑的例子，说明 DEA 如何将笔记本的多种属性整合起来，从而实现，让数据自己说话。

第 2 章阐述 DEA 为什么是平衡的基准方法。我们的分析表明 DEA 能够自行灵活调整，以适应特定评估目标的需要。因为需要很少甚至不需要任何的用户介入，所以 DEA 是一种客观的分析方法。与此同时，DEA 也允许用户加入价值判断，对分析结果进行优化和精炼。这些价值判断反映用户对多重绩效指标的主观偏好。

第 3 章给出 DEA 的相关术语和名词，介绍如何将绩效指标进行分类。

第 4 章介绍如何利用 DEA 提升低效率单元绩效。

第 5 章给出多个案例，介绍 DEA 的实践方法。

第 6 章介绍一种基本的 DEA 模型。利用该模型，用户能够找出参照基准及相应的基准份额。讨论 DEA 最佳实践前沿的形状。

第 7 章介绍能够加载到 Excel 进行使用的 DEA 软件——DEAFrontier。尽管有很多软件包可以执行 DEA 分析，本书使用的是微软公司的 Excel

及其内含的规划求解器 (Excel Solver)。DEAFrontier 软件调用了 Excel 的规划求解器, 作为求解 DEA 模型的引擎。本书 DEA 模型的建立和求解都在 Excel 及其规划求解器中完成的, 读者只需要编写一些非常简单的 Visual Basic for Application (VBA) 代码即可熟练的运用 DEA 方法。此外, 本书还在附录部分添加了 DEA 方法常用术语的英汉对照

致 读 者

以下内容可以从网站 www.deafrontier.net 获取。

(1) 本书使用的数据以及用 Excel 建立的 DEA 模型 (excel 文件格式)。

(2) DEAFrontier 的免费版本及其使用指南。

(3) 讲解关键概念以及如何在 Excel 中建立 DEA 模型的视频。

要学会 DEA, 不能只简单地阅读本书。为真正掌握书中讲解的绝大部分内容, 你需要自己操练其中的例子。我们强烈建议, 你在电脑上把本书的所有案例都演练一遍, 看看能否独立计算出本书和相应 Excel 文件显示的结果。网上的视频可作为重要的参考指南。只有这样, 你才能学会如何正确和有效地使用 DEA, 才能更好地理解其中的重要概念, 也才能更顺利地运用 DEA 解决你所关注的实际问题。

对于希望更深入学习 DEA 的读者, 可以参考以下两本书:

① Cook W D, Zhu J. 2013. Data Envelopment Analysis: Balanced Benchmarking.

② Zhu J. 2014. Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets. 3rd edition.

New York: Springer Science.

致 谢

在这里，我要对 Yao Chen 教授、Jessica Colon、Steven Jiahua Tang 和 Alec Zhu 在文稿撰写期间给予的建议和评论表示诚挚的谢意。当然，由于本书涉及术语和数据较多，编写过程中难免出错，如果读者发现书中存在错漏，真心希望读者提出批评意见，以便在今后的修订中改正过来。

Joe Zhu, 2014 年 9 月
美国，麻省

目 录

序

中文版序

序言(原版)

第 1 章 运营分析和基准	1
1.1 运营分析和基准	1
1.2 回归分析	3
1.3 基准	7
1.3.1 构建模型	8
1.3.2 规划求解器 (Solver) 的调用	14
1.3.3 规划求解的构建	15
1.3.4 模型求解	18
1.3.5 最优电脑的识别	19
第 2 章 平衡的基准	26
2.1 综合指标	26
2.2 平衡的基准	28
2.3 最佳权重	38
2.4 交叉评价	44
第 3 章 数据包络分析	47
3.1 数据包络分析概述	47
3.2 DEA 是一种平衡基准	49

3.3	绩效指标的分类	51
3.4	DEA 的输入和输出指标	55
3.4.1	美国银行和储蓄机构首席执行官的绩效	55
3.4.2	菲律宾农业	56
3.4.3	配电设计的改进措施	56
3.4.4	制造业绩效	57
3.4.5	休斯顿小学	58
3.4.6	旅行社	59
3.4.7	安达卢西亚橄榄种植农场	59
3.4.8	渔船效率评估	60
3.4.9	职业网球运动员	60
3.4.10	施工安全	61
3.4.11	对冲基金	62
3.4.12	家庭共同基金	62
3.4.13	快餐连锁店	63
3.4.14	疗养院	64
3.4.15	警察局	65
3.5	DEA 的图形表述	65
第 4 章	绩效提升	70
4.1	输入导向	70
4.2	输出导向	74
4.3	只有输入（或输出）的 DEA	79
第 5 章	案例	85

5.1	案例 1: 生活幸福质量	85
5.2	案例 2: 航空公司绩效评价	88
5.2.1	航空公司指标	88
5.2.2	DEA 分析	92
5.3	案例 3: 数据包络分析在招聘中的应用	99
5.3.1	DEA 辅助候选人甄选过程	100
5.3.2	输入和输出指标的选择	101
5.3.3	练习	105
第 6 章	基准份额	107
6.1	合作与虚拟单元构造	107
6.2	基准份额的确定	111
6.3	DEA 最佳实践前沿	119
第 7 章	利用 DEA Frontier 软件求解 DEA	123
7.1	DEA Frontier 软件	123
7.2	数据表单的格式	125
7.3	DEA 模型: 乘子模型	127
7.4	DEA 模型: 包络模型	129
7.5	CRS 与 VRS	130
附录: DEA	术语中英对照	133

第 1 章 运营分析和基准

1.1 运营分析和基准

根据维基百科,数据分析是检查、清理、转换和模型化数据的过程,用以发现有用信息、提供建议并为决策者提供决策支持。即用适当的数据分析方法对收集的大量数据进行分析,提取有用信息,并对有用信息概括总结的过程。在现实生活中,数据分析可帮助人们成功做出判断,以便采取适当行动。数据分析在不同的商业、自然科学和社会科学领域具有不同的技术方法和名称。

数据分析的类型包括 4 种,描述性分析、预测性分析、规范性分析和决策性分析。描述性分析是运用图表等形式对数据进行归纳;预测性分析是运用统计和机器学习技术,通过分析历史数据来预测未来;规范性分析是运用最优化技术规范决策;决策性分析是运用可视化工具为决策者提供支持。

利用以上方法,分析运营数据,对运营过程进行优化,以改善运营现状。运用各种数据分析工具,将数据转换为决策行动,从而为组织决策服务^①。

本书提出一种称为“平衡基准”的特殊数据分析方法,运用此方法可以改进运营过程。这种基准平衡方法也被称为数据包络分析 (data

^① Liberatore M J, Luo W. 2010. The Analytics Movement: Implications for Operations Research. *Interfaces*, 40: 313-324.