



Title	Computing the Convex Frontier for Large-Scale Data Envelopment Analysis
Author(s)	庄, 乾偉
Citation	大阪大学, 2025, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/101750
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

Abstract of Thesis

Name (Qianwei ZHUANG)	
Title	Computing the Convex Frontier for Large-Scale Data Envelopment Analysis (大規模データ包絡分析における凸フロンティア計算法)
<p>Abstract of Thesis</p> <p>Data Envelopment Analysis (DEA) is a widely used analytical tool that leverages linear programming (LP) to classify or benchmark decision-making units (DMUs) by establishing a convex frontier. Each DMU is represented as a point in \mathbb{R}^d, where d denotes the dimensionality. In DEA, lower values are preferred for input dimensions, while higher values are favored for output dimensions. The DEA frontier is defined as the oriented boundary of a polytope formed by extreme points.</p> <p>Effective identification of this frontier is critical, particularly in large-scale contexts where the number of DMUs n is substantial. However, this process is computationally expensive, requiring the solution of numerous LP problems, each involving a large number of variables, which leads to significant time and memory challenges.</p> <p>This study proposes methods to improve the computational performance of DEA frontier identification. First, we introduce a sequential categorization scheme that organizes extreme DEA points into initial blocks, enabling the early establishment of the frontier without processing the entire dataset. Next, we propose a novel reference set selection technique that minimizes the number of decision variables in the LP problems to the size of the dimension, thereby substantially reducing computational costs. Finally, we present an adaptive search and accumulation strategy that minimizes dependency on LP solutions, avoiding redundant computations. Regarding the number of dimension-sized LPs to be solved, these approaches collectively achieve a sublinear time complexity, ensuring computational efficiency superior to $O(n)$.</p> <p>Rigorous proofs are provided for each proposed method. Experimental results demonstrate the scalability and robustness of the techniques, yielding substantial reductions in computational time and memory requirements while maintaining accurate frontier identification.</p>	

論文審査の結果の要旨及び担当者

氏 名 (Qianwei Zhuang)			
	(職)	氏 名	
論文審査担当者	主 査	教授	森田 浩
	副 査	教授	鈴木 秀幸
	副 査	教授	藤崎 泰正
	副 査	准教授	山口 勇太郎

論文審査の結果の要旨

データ包絡分析（DEA）は、意思決定単位（DMU）の効率性を評価するための手法として広く用いられているが、大規模データセットへの適用に際しては、計算負荷が極めて高いという課題を抱えている。標準的なDEAの計算では、 n 個のDMUに対して n 回の線形計画問題を解く必要があり、このため大規模な問題においては計算コストが著しく増大する。本論文では、この計算負荷を軽減し、DEAの計算性能を向上させることを目的とし、フロンティア構築のための代替的手法を提案している。

まず、逐次的分類スキームを導入し、フロンティアに属するDMUを初期ブロックとして分類する手法を提案している。これにより、フロンティア構築に伴う計算の複雑性を低減することが可能となった。次に、新たな参照集合の選択技術を開発し、これによって線形計画問題の変数の数を削減することで計算コストを大幅に削減することを実現した。そして、適応的探索および蓄積スキームを提案することで、線形計画問題の解法への依存を最小限に抑えつつ、冗長な計算を排除する仕組みを提供している。

以上の手法について、その有効性とスケーラビリティを理論的分析および実験的結果を通じて明らかにした。提案手法は、従来のDEA計算手法に比して計算負荷を大幅に低減するのみならず、大規模データセットにおいても効率かつ実用的なアプローチが示されている。これにより、金融、医療、運輸など、幅広い分野における大規模データの効率性分析への応用可能性が拡大することが期待される。

以上より、本論文はDEAのさらなる応用を志向する研究者および実務者に対し、計算効率の向上に関する重要な知見を提供するものであり、DEAの理論および実践の両面における発展に寄与するものである。

この分野における課題解決に対する数理最適化の有効性を明らかにしており、情報科学や計算幾何学における大きな貢献が認められる。よって、博士（情報科学）の学位論文として価値のあるものと認める。