

*Моргунов Евгений Павлович*¹,
доцент, канд. техн. наук, доцент;
*Моргунова Ольга Николаевна*²,
доцент, канд. техн. наук, доцент

МНОВОВАРИАНТНЫЕ ВЫЧИСЛЕНИЯ ПРИ ИСПОЛЬЗОВАНИИ МЕТОДА «АНАЛИЗ СРЕДЫ ФУНКЦИОНИРОВАНИЯ» В СРЕДЕ СУБД POSTGRESQL

^{1,2} Россия, Красноярск, Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева,
кафедра информатики и вычислительной техники,
¹ emorgunov@mail.ru, ² olgamorgunova@mail.ru

Аннотация. Рассмотрена ситуация, когда для проведения исследования эффективности группы объектов с помощью метода «Анализ Среды Функционирования» можно выбрать разные наборы показателей, описывающих эти объекты. В таком случае может оказаться целесообразным выполнить исследования с использованием различных наборов показателей, а затем оценить устойчивость получаемых результатов. В статье показана технология выполнения многовариантных вычислений в среде СУБД PostgreSQL. Для ее поддержки служит специально разработанное расширение этой СУБД.

Ключевые слова: эффективность, Анализ Среды Функционирования, АСФ, PostgreSQL, многовариантные вычисления.

*Evgeny P. Morgunov*¹,
Associate Professor, Candidate of Technical Sciences;
*Olga N. Morgunova*²,
Associate Professor, Candidate of Technical Sciences

MULTI-VARIANT CALCULATIONS WHEN USING THE DATA ENVELOPMENT ANALYSIS METHOD IN THE POSTGRESQL DBMS ENVIRONMENT

^{1,2} Computer Science Department, Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia,
¹ emorgunov@mail.ru, ² olgamorgunova@mail.ru

Abstract. The situation is considered when, to conduct a study of the efficiency of a group of objects using the Data Envelopment Analysis method, you can select different sets of indicators that describe these objects. In this case, it may be appropriate to conduct studies using different sets of indicators and then evaluate the robustness of the results obtained. The

article shows the technology for performing multi-variant calculations in the PostgreSQL DBMS environment. To support it, a specially developed extension of this DBMS is used.

Keywords: efficiency of systems, Data Envelopment Analysis, DEA, PostgreSQL, multi-variant calculations.

Введение

Для проведения системных исследований очень важен правильный выбор показателей, описывающих исследуемые объекты. Это справедливо и в отношении исследований эффективности объектов. Однако бывают ситуации, когда такой выбор сделать трудно. В этом случае можно попробовать выполнить исследование с использованием различных наборов показателей и сравнить получаемые результаты. Данный подход – многовариантные вычисления – можно применить и к методу Анализ Среды Функционирования (АСФ) [1]. Его оригинальное название – Data Envelopment Analysis (DEA) [6].

Проведение вычислений с разными наборами показателей требует дополнительной работы. Поэтому важно организовать ее выполнение с минимальными затратами времени исследователя.

Метод исследования

Метод АСФ (DEA) не предполагает использования какой-либо формулы, связывающей исходные показатели, и не требует использования весовых коэффициентов, что снижает степень субъективности полученных оценок эффективности систем. Однако он требует разделения показателей на входные и выходные. Входные показатели – это ресурсы, использованные для производства результатов, описываемых выходными показателями. Показатели выбираются с учетом специфики предметной области и заданной цели исследования.

Метод АСФ (DEA) основан на построении так называемой *границы эффективности* в многомерном пространстве входных и выходных переменных, описывающих объекты, эффективность которых требуется определить. Граница эффективности представляет собой гиперповерхность, огибающую (охватывающую) точки, соответствующие оцениваемым объектам. Степень эффективности конкретного объекта зависит от расстояния между ним и границей эффективности: чем дальше объект находится от границы, тем его эффективность ниже. Объекты, находящиеся на границе эффективности, считаются эффективными.

Представим формализованное описание метода на примере модели ВСС, ориентированной на выход (ее название образовано из первых букв фамилий ее авторов: Banker, Charnes и Cooper) [7, с. 93]. Пусть требуется определить показатель эффективности каждого из n объектов. Такими объектами могут быть, например, предприятия, организации, университеты, банки и т. д. Для описания каждого объекта $o_j, j = \overline{1, n}$, служит пара векто-

ров $(\mathbf{x}_j, \mathbf{y}_j)$. При этом вектор $\mathbf{x}_j = (x_{j1}, \dots, x_{ji}, \dots, x_{jm})^T$ содержит входные показатели (входы) для объекта o_j , а вектор $\mathbf{y}_j = (y_{j1}, \dots, y_{jr}, \dots, y_{js})^T$ содержит выходные показатели (выходы) для объекта o_j . Тогда матрица $\mathbf{X} = (\mathbf{x}_j)$, имеющая размерность $m \times n$, содержит вектор-столбцы с входными данными для всех n объектов, а матрица $\mathbf{Y} = (\mathbf{y}_j)$, имеющая размерность $s \times n$, содержит вектор-столбцы с выходными данными для всех n объектов. В основе метода АСФ (DEA) лежит метод линейного программирования, поэтому модель формулируется в таком виде:

$$\begin{aligned} \max_{\varphi, \lambda} (\varphi), \\ \mathbf{X}\lambda \leq \mathbf{x}_0, \\ \varphi \mathbf{y}_0 - \mathbf{Y}\lambda \leq \mathbf{0}, \\ \mathbf{e}\lambda = 1, \\ \lambda \geq \mathbf{0}, \end{aligned} \quad (1)$$

где $\mathbf{e}$ – это единичный вектор-строка. В данном случае показатель эффективности – скаляр $\varphi \in [1; \infty)$. Если $\varphi = 1$, то объект будет считаться эффективным *по сравнению* с остальными объектами исследуемой совокупности, если $\varphi > 1$, то неэффективным. На практике значение показателя эффективности зачастую переводится в диапазон $(0; 1]$ с помощью преобразования $1/\varphi$. Объекты, имеющие значение показателя $\varphi = 1$, считаются эффективными и находятся, как принято говорить, на *границе эффективности*. Аналогичная задача решается для каждого объекта, т. е. n раз. Таким образом, граница эффективности формируется путем многократного решения задачи линейного программирования.

Предлагаемый подход к решению поставленной задачи

Эта статья продолжает работу авторов, целью которой является облегчение использования метода АСФ (DEA) в прикладных программах [3, 4, 5]. Это достигается за счет встраивания метода в СУБД PostgreSQL в качестве *расширения* (extension) данной СУБД [9]. Существует целый ряд программных продуктов, реализующих данный метод, но, насколько известно авторам, не существует его реализации в виде, пригодном для интеграции в прикладные программы [2]. Встраивание метода в СУБД имеет важное достоинство: метод можно использовать в прикладных программах, написанных на любом языке программирования, имеющем интерфейс для работы с СУБД PostgreSQL.

Разрабатываемое расширение СУБД PostgreSQL включает не только программный код, но также ряд таблиц базы данных, объединенных в единый репозиторий [3]. Его наличие позволит исследователю упорядочить исходные данные, используемые в разных исследованиях, а также облегчит проведение многовариантных вычислений.

В качестве упрощенного примера многовариантных вычислений можно привести исследование эффективности научной работы кафедр. В качестве входных переменных, т. е. ресурсов, для первого варианта вычислений эффективности можно взять, скажем, численность сотрудников кафедры, а для второго варианта предусмотреть отдельные показатели численности кандидатов и докторов наук. Выходными переменными могли бы стать общее число опубликованных статей (первый вариант) либо число статей, опубликованных в журналах разного уровня (второй вариант).

Технология использования метода АСФ (DEA) в среде СУБД PostgreSQL представлена в работе [5]. При выполнении многовариантных вычислений следует для каждого варианта создать структуру данных, называемую этапом исследования, и выполнить расчеты по выбранной модели метода АСФ (DEA). Затем для сопоставления полученных результатов можно воспользоваться сервисными процедурами, входящими в состав расширения. Они позволят для каждого варианта вычислений рассчитать, например, средние значения показателей эффективности всех объектов в группе, про ранжировать исследуемые объекты на основе их показателей эффективности, а затем в удобной форме сопоставить места, занятые каждым объектом при выполнении каждого из вариантов вычислений, и т. д. Прикладной программе останется только визуализировать полученные результаты.

Конечно, число вариантов при выполнении таких вычислений может быть и больше двух.

Многовариантные вычисления могут рассматриваться еще с одной стороны. Речь идет о том, что центральной частью метода АСФ (DEA) является процедура решения задачи линейного программирования. Для ее программной реализации в рамках нашего расширения СУБД могут использоваться различные библиотеки. В настоящее время нами применяется `lp_solve` [8]. Одним из направлений модификации программного кода разрабатываемого расширения является поддержка нескольких способов решения задачи линейного программирования. Задать требуемый способ решения или библиотеку можно будет в специальной таблице базы данных, содержащей параметры конкретного исследования. Такой подход позволит сравнить быстродействие библиотек, а также результаты, получаемые с их помощью. Если они различаются на величину, превышающую некоторый заданный предел, это может послужить сигналом к более тщательному изучению особенностей использования библиотеки или даже поставить вопрос о корректности ее работы. Полученный в перспективе опыт использования разных библиотек, возможно, позволит увидеть преимущества той или другой библиотеки при решении задач, имеющих какую-либо специфику исходных данных.

Выполнять описанные процедуры можно как интерактивно в среде терминала `psql`, входящего в комплект поставки СУБД PostgreSQL, или

какой-либо графической утилиты, так и вызывая процедуры, входящие в состав расширения, из прикладных программ, написанных на различных языках (Python, PHP, Perl и др.). Конечно, первый способ требует некоторых знаний языка SQL, но это вполне достижимо.

Сервисные процедуры, используемые для сопоставления результатов многовариантных исследований, можно применять и для сопоставления результатов исследования одной и той же группы объектов, выполненных на основе одной и той же модели метода АСФ (DEA), но в разные временные периоды.

Заключение

Таким образом, многовариантные вычисления могут помочь в выборе набора показателей, наиболее адекватно описывающих объекты, эффективность функционирования которых нужно оценить. При этом предлагаемая технология их выполнения в среде СУБД PostgreSQL позволит избавить исследователя от выполнения рутинных операций, а также увеличить количество вариантов, из которых выбирается набор показателей, используемых в исследовании эффективности. Это сделает полученные оценки эффективности более обоснованными.

Список литературы

1. Кривоножко В. Е., Лычев А. В. Анализ деятельности сложных социально-экономических систем. – М.: Изд. отдел ф-та ВМиК МГУ; МАКС Пресс, 2010. – 208 с.
2. Моргунов Е. П., Моргунова О. Н. Обзор программного обеспечения, реализующего метод Data Envelopment Analysis [Электронный ресурс] // XXII Междунар. науч.-практ. конф. «Решетневские чтения», СибГУ им. М. Ф. Решетнева, г. Красноярск, 12–16 ноября 2018 г.: материалы: в 2 ч. – Красноярск, 2018. – Ч. 2. – С. 359–360. – URL: <https://reshetnev.sibsau.ru/page/materialy-konferentsii> (дата обращения: 25.08.2024).
3. Моргунов Е. П., Моргунова О. Н., Постойко А. Ю. Интеграция метода «Анализ Среды Функционирования» в СУБД PostgreSQL [Электронный ресурс] // Системный анализ в проектировании и управлении: сб. науч. трудов XXIII Междунар. науч.-практ. конф., Санкт-Петербургский политехн. ун-т Петра Великого, г. Санкт-Петербург, 10–11 июня 2019 г.: в 3-х частях. – СПб.: Изд-во Политех-Пресс, 2019. – Ч. 3. – С. 178–184. – URL: <https://elib.spbstu.ru/dl/2/k19-141.pdf/view> (дата обращения: 25.08.2024).
4. Моргунов Е. П., Моргунова О. Н., Постойко А. Ю. Реализация метода «Анализ Среды Функционирования» в виде расширения СУБД PostgreSQL [Электронный ресурс] // Системный анализ в проектировании и управлении: сб. науч. трудов XXIV Междунар. науч.-практ. конф., Санкт-Петербургский политехн. ун-т Петра Великого, 13–14 октября 2020 г.: в 3-х частях. – СПб.: Изд-во Политех-Пресс, 2020. – Ч. 3. – С. 146–152. – URL: <https://elib.spbstu.ru/dl/2/id20-207.pdf/view> (дата обращения: 25.08.2024).
5. Моргунов Е. П., Моргунова О. Н. Технология использования метода «Анализ Среды Функционирования» в виде расширения СУБД PostgreSQL [Электронный ресурс] // Системный анализ в проектировании и управлении: сб. науч. трудов XXV Междунар. науч. и учеб.-практ. конф., Санкт-Петербургский политехн. ун-т Петра Великого, 13–14 октября 2021 г.: в 3-х частях. – СПб.: Изд-во Политех-Пресс, 2021. – Ч. 3. – С. 134–139. – URL: <https://elib.spbstu.ru/dl/2/id21-362.pdf/view> (дата обращения: 25.08.2024).

6. Charnes A., Cooper W. W., Rhodes E. Measuring the efficiency of decision making units // European Journal of Operational Research. – 1978. – Vol. 2. – Pp. 429–444.
7. Cooper W. W., Seiford L. M., Tone K. Data envelopment analysis: A comprehensive text with models, applications, references, and DEA-Solver software. – 2nd ed. – New York: Springer, 2007. – xxxviii, 490 p.
8. lp_solve: a Mixed Integer Linear Programming (MILP) solver [Electronic resource]. – URL: <http://lpsolve.sourceforge.net/5.5/> (access date: 25.08.2024).
9. PostgreSQL [Electronic resource] // The PostgreSQL Global Development Group: website. – URL: <http://www.postgresql.org> (access date: 25.08.2024).