

ВИКИПЕДИЯ

Умножение матриц

Материал из Википедии — свободной энциклопедии

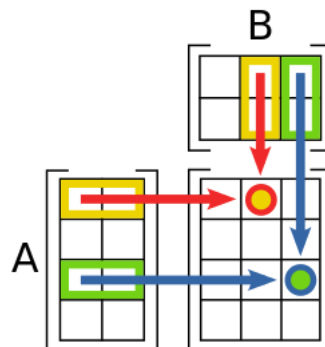
Умноже́ние ма́триц — одна из основных операций над матрицами. Матрица, получаемая в результате операции умножения, называется **произведе́нием ма́триц**. Элементы новой матрицы получаются из элементов старых матриц в соответствии с правилами, проиллюстрированными ниже.

Матрицы *A* и *B* могут быть перемножены, если они **совместимы** в том смысле, что число столбцов матрицы *A* равно числу строк *B*.

Матрицы обладают многими алгебраическими свойствами умножения, присущими обычным числам, за исключением коммутативности.

Для квадратных матриц, помимо умножения, может быть введена операция возведения матрицы в степень и обратная матрица.

Тогда как матрицы используются для описания, в частности, преобразований математических пространств (поворот, отражение, растяжение и другие), произведение матриц будет описывать композицию преобразований.



Содержание

Определение

Иллюстрация

Обсуждение

Свойства

Обратная матрица

Алгоритмы быстрого перемножения матриц

Степени матриц

См. также

Литература

Примечания

Определение

Пусть даны две прямоугольные матрицы *A* и *B* размерности $l \times m$ и $m \times n$ соответственно:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1} & a_{l2} & \cdots & a_{lm} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}.$$

Тогда матрица C размерностью $l \times n$:

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{l1} & c_{l2} & \cdots & c_{ln} \end{bmatrix},$$

в которой:

$$c_{ij} = \sum_{r=1}^m a_{ir} b_{rj} \quad (i = 1, 2, \dots, l; j = 1, 2, \dots, n).$$

называется их *произведением*.

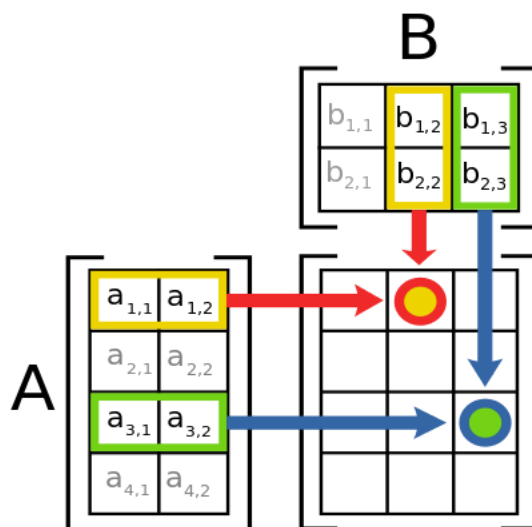
Операция умножения двух матриц выполнима только в том случае, если число столбцов в первом сомножителе равно числу строк во втором; в этом случае говорят, что матрицы *согласованы*. В частности, умножение всегда выполнимо, если оба сомножителя — квадратные матрицы одного и того же порядка.

Таким образом, из существования произведения AB вовсе не следует существование произведения BA .

Иллюстрация

Произведение матриц AB состоит из всех возможных комбинаций скалярных произведений вектор-строк матрицы A и вектор-столбцов матрицы B . Элемент матрицы AB с индексами i, j есть скалярное произведение i -ой вектор-строки матрицы A и j -го вектор-столбца матрицы B .

Иллюстрация справа демонстрирует вычисление произведения двух матриц A и B , она показывает как каждое пересечение в произведении матриц соответствует строкам матрицы A и столбцам матрицы B . Размер результирующей матрицы всегда максимально возможный, то есть для каждой строки матрицы A и столбца матрицы B есть всегда соответствующее пересечение в произведении матрицы.



Значения на пересечениях, отмеченных кружочками, будут:

$$\begin{aligned}
 x_{1,2} &= (a_{1,1}, a_{1,2}) \cdot (b_{1,2}, b_{2,2}) \\
 &= a_{1,1}b_{1,2} + a_{1,2}b_{2,2} \\
 x_{3,3} &= (a_{3,1}, a_{3,2}) \cdot (b_{1,3}, b_{2,3}) \\
 &= a_{3,1}b_{1,3} + a_{3,2}b_{2,3}
 \end{aligned}$$

В общем случае, произведение матриц не является коммутативной операцией. К примеру:

$$\begin{array}{c} 3 \times 4 \text{ matrix} \\ \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \end{bmatrix} \end{array} \begin{array}{c} 4 \times 5 \text{ matrix} \\ \begin{bmatrix} \cdot & \cdot & \cdot & \mathbf{a} & \cdot \\ \cdot & \cdot & \cdot & \mathbf{b} & \cdot \\ \cdot & \cdot & \cdot & \mathbf{c} & \cdot \\ \cdot & \cdot & \cdot & \mathbf{d} & \cdot \end{bmatrix} \end{array} = \begin{array}{c} 3 \times 5 \text{ matrix} \\ \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \mathbf{x}_{3,4} & \cdot \end{bmatrix} \end{array}$$

Элемент $x_{3,4}$ произведения матриц, приведённых выше, вычисляется следующим образом

$$x_{3,4} = (\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}) \cdot (\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) = \mathbf{1} \cdot \mathbf{a} + \mathbf{2} \cdot \mathbf{b} + \mathbf{3} \cdot \mathbf{c} + \mathbf{4} \cdot \mathbf{d}$$

Первая координата в обозначении матрицы обозначает строку, вторая координата — столбец; этот порядок используют как при индексации, так и при обозначении размера. Элемент x_{ij} на пересечении строки i и столбца j результирующей матрицы является скалярным произведением i -й строки первой матрицы и j -го столбца второй матрицы. Это объясняет почему ширина и высота умножаемых матриц должны совпадать: в противном случае скалярное произведение не определено.

Обсуждение

Увидеть причины описанного правила матричного умножения легче всего, рассмотрев умножение вектора на матрицу.

Последнее естественно вводится исходя из того, что при разложении векторов по базису действие (любого) линейного оператора A даёт выражение компонент вектора $\mathbf{v}' = A\mathbf{v}$:

$$v'_i = \sum_j A_{ij} v_j$$

То есть линейный оператор оказывается представлен матрицей, векторы — векторами-столбцами, а действие оператора на вектор — матричным умножением вектора-столбца слева на матрицу оператора (это частный случай матричного умножения, когда одна из матриц — вектор-столбец — имеет размер $n \times 1$).

(Равно переход к любому новому базису при смене координат представляется полностью аналогичным выражением, только v'_i в этом случае уже не компоненты нового вектора в старом базисе, а компоненты старого вектора в новом базисе; при этом A_{ij} — элементы матрицы перехода к новому базису).

Рассмотрев последовательное действие на вектор двух операторов: сначала A , а потом B (или преобразование базиса A , а затем преобразование базиса B), дважды применив нашу формулу, получим:

$$v_i'' = \sum_j B_{ij} \sum_k A_{jk} v_k = \sum_j \sum_k B_{ij} A_{jk} v_k = \sum_k \sum_j (B_{ij} A_{jk}) v_k,$$

откуда видно, что композиции BA действия линейных операторов A и B (или аналогичной композиции преобразований базиса) соответствует матрица, вычисляемая по правилу произведения соответствующих матриц:

$$(BA)_{ik} = \sum_j B_{ij} A_{jk}.$$

Определённое таким образом произведение матриц оказывается совершенно естественным и очевидно полезным (даёт простой и универсальный способ вычисления композиций произвольного количества линейных преобразований).

Свойства

Сочетательное свойство, ассоциативность:

$$\begin{aligned} \mathbf{A}(\mathbf{BC}) &= (\mathbf{AB})\mathbf{C}; \\ \alpha(\mathbf{AB}) &= (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B}). \end{aligned}$$

Распределительное свойство, дистрибутивность относительно сложения:

$$\begin{aligned} \mathbf{A}(\mathbf{B} + \mathbf{C}) &= \mathbf{AB} + \mathbf{AC}; \\ (\mathbf{A} + \mathbf{B})\mathbf{C} &= \mathbf{AC} + \mathbf{BC}. \end{aligned}$$

Произведение матрицы на единичную матрицу \mathbf{E} подходящего порядка равно самой матрице:

$$\begin{aligned} \mathbf{EA} &= \mathbf{A}; \\ \mathbf{AE} &= \mathbf{A}. \end{aligned}$$

Произведение матрицы на нулевую матрицу $\mathbf{0}$ подходящей размерности равно нулевой матрице:

$$\begin{aligned} \mathbf{0A} &= \mathbf{0}; \\ \mathbf{A0} &= \mathbf{0}. \end{aligned}$$

Если \mathbf{A} и \mathbf{B} — квадратные матрицы одного и того же порядка, то произведение матриц обладает ещё рядом свойств.

Умножение матриц в общем случае некоммутативно:

$$\mathbf{AB} \neq \mathbf{BA}.$$

Если $\mathbf{AB} = \mathbf{BA}$, то матрицы \mathbf{A} и \mathbf{B} называются коммутирующими между собой.

Простейшие примеры коммутирующих матриц:

- любая квадратная матрица — с самой собой: $\mathbf{AA} = \mathbf{AA} = \mathbf{A}^2$ (возведение матрицы в квадрат);
- любая квадратная матрица — с единичной матрицей того же порядка: $\mathbf{AE} = \mathbf{EA} = \mathbf{A}$;

- любая квадратная матрица — с нулевой матрицей того же порядка: $\mathbf{A}\mathbf{0} = \mathbf{0}\mathbf{A} = \mathbf{0}$;
- любая невырожденная квадратная матрица — со своей обратной матрицей: $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{E}$.

Определитель и след произведения не зависят от порядка умножения матриц:

$$\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det \mathbf{A} \cdot \det \mathbf{B};$$

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA}).$$

Обратная матрица

Квадратная матрица \mathbf{A} называется неособенной (невырожденной), если она имеет единственную обратную матрицу \mathbf{A}^{-1} такую, что выполняется условие:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{E}.$$

В противном случае матрица \mathbf{A} называется особенной (вырожденной).

Матрица $\mathbf{A} = [a_{ik}]$ порядка n является невырожденной в том и только в том случае, если $\det \mathbf{A} = \det [a_{ik}] \neq 0$; в этом случае \mathbf{A}^{-1} есть квадратная матрица того же порядка n :

$$\mathbf{A}^{-1} = [a_{ik}]^{-1} = \left[\frac{A_{ki}}{\det \mathbf{A}} \right],$$

где A_{ik} — алгебраическое дополнение элемента a_{ik} в определителе $\det [a_{ik}]$.

Алгоритмы быстрого перемножения матриц

Сложность вычисления произведения матриц по определению составляет $O(n^3)$, однако существуют более эффективные алгоритмы^[1], применяющиеся для больших матриц. Вопрос о предельной скорости умножения больших матриц, также как и вопрос о построении наиболее быстрых и устойчивых практических алгоритмов умножения больших матриц остаётся одной из нерешённых проблем линейной алгебры.

Алгоритм Штрассена (1969)

Первый алгоритм быстрого умножения больших матриц был разработан Фолькером Штрассеном^[2] в 1969 году. В основе алгоритма лежит рекурсивное разбиение матриц на блоки **2X2**. Штрассен доказал, что матрицы **2X2** можно некоммутативно перемножить с помощью семи умножений, поэтому на каждом этапе рекурсии выполняется семь умножений вместо восьми. В результате асимптотическая сложность этого алгоритма составляет $O(n^{\log_2 7}) \approx O(n^{2.81})$. Недостатком данного метода является бóльшая сложность программирования по сравнению со стандартным алгоритмом, слабая численная устойчивость и большой объём используемой памяти. Разработан ряд алгоритмов на основе метода Штрассена, которые улучшают численную устойчивость, скорость по константе и другие его характеристики. Тем не менее, в силу простоты алгоритм Штрассена остаётся одним из практических алгоритмов умножения больших матриц. Штрассен также выдвинул следующую гипотезу Штрассена: для сколь угодно

малого $\epsilon > 0$ существует алгоритм, при достаточно больших натуральных n гарантирующий перемножение двух матриц размера $n \times n$ за $O(n^{2+\epsilon})$ операций.

■ Дальнейшие улучшения показателя степени ω для скорости матричного умножения

В дальнейшем оценки скорости умножения больших матриц многократно улучшались. Однако эти алгоритмы носили теоретический, в основном приближённый характер. В силу неустойчивости алгоритмов приближённого умножения в настоящее время они не используются на практике.

■ Алгоритм Пана (1978)

В 1978 году Пан^[3] предложил свой метод умножения матриц, сложность которого составила $\Theta(n^{2.78041})$.

■ Алгоритм Бини (1979)

В 1979 году группа итальянских учёных во главе с Бини^[4] разработала алгоритм умножения матриц с использованием тензоров. Его сложность составляет $\Theta(n^{2.7799})$.

■ Алгоритмы Шёнхаге (1981)

В 1981 году Шёнхаге^[5] представил метод, работающий со скоростью $\Theta(n^{2.695})$. Оценка получена с помощью подхода, названного *частичным матричным умножением*. Позже ему удалось получить оценку $\Theta(n^{2.6087})$.

Затем Шёнхаге на базе *метода прямых сумм* получил оценку сложности $\Theta(n^{2.548})$. Романи сумел понизить оценку до $\Theta(n^{2.5166})$, а Пан — до $\Theta(n^{2.5161})$.

■ Алгоритм Копперсмита — Винограда (1990)

В 1990 году Копперсмит и Виноград^[6] опубликовали алгоритм, асимптотическая сложность которого составляла $\Theta(n^{2.3755})$. Этот алгоритм использует идеи, схожие с алгоритмом Штрассена. На сегодняшний день модификации алгоритма Копперсмита-Винограда являются наиболее *асимптотически* быстрыми. В последней модификации Джоша Алмана и Вирджинии Вильямс (2020) алгоритм умножает матрицы со скоростью $\Theta(n^{2.3728596})$ ^[7]. Тот факт, что последние полученные улучшения ничтожны, позволяет говорить о существовании «барьера Копперсмита-Винограда» в асимптотических оценках скорости алгоритмов. Алгоритм Копперсмита-Винограда эффективен только на матрицах астрономического размера и на практике применяться не может.

■ Связь с теорией групп (2003)



Хронология улучшения оценок показателя степени ω для вычислительной сложности матричного умножения $O(n^\omega)$.

В 2003 году Кох и др. рассмотрели в своих работах^[8] алгоритмы Штрассена и Копперсмита-Винограда в контексте теории групп. Они показали, что гипотеза Штрассена справедлива (т.е. минимальная сложность ограничена $O(n^{2+\epsilon})$ для любого ϵ), если выполняется одна из гипотез теории групп^[9].

Степени матриц

Квадратные матрицы можно многократно умножать сами на себя так же, как обычные числа, так как у них одинаковое число строк и столбцов. Такое последовательное умножение можно назвать **возведением матрицы в степень** — это будет частный случай обычного умножения нескольких матриц. У прямоугольных матриц число строк и столбцов разное, поэтому их *никогда* нельзя возводить в степень. Матрица **A** размерности $n \times n$, возведённая в степень, определяется формулой

$$\mathbf{A}^k = \underbrace{\mathbf{A} \mathbf{A} \cdots \mathbf{A}}_k$$

и обладает следующими свойствами (λ — некоторый скаляр):

Нулевая степень:

$$\mathbf{A}^0 = \mathbf{E}$$

где **E** - единичная матрица. Это аналог того факта, что нулевая степень любого числа равна единице.

Умножение на скаляр:

$$(\lambda \mathbf{A})^k = \lambda^k \mathbf{A}^k$$

Определитель:

$$\det(\mathbf{A}^k) = \det(\mathbf{A})^k$$

Наиболее простой способ вычисления степени матрицы — это умножать k раз матрицу **A** на результат предыдущего умножения, начиная с единичной матрицы, как это часто делают для скаляров. Для диагонализируемых матриц существует лучший метод, основанный на использовании спектрального разложения матрицы **A**. Ещё один метод, основанный на теореме Гамильтона — Кэли, строит более эффективное выражение для \mathbf{A}^k , в котором в требуемую степень возводится *скаляр*, а не вся *матрица*.

Особый случай составляют диагональные матрицы. Так как произведение диагональных матриц сводится к умножению соответствующих диагональных элементов, то k -ая степень диагональной матрицы **A** состоит из элементов, возведённых в требуемую степень:

$$\mathbf{A}^k = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}^k = \begin{pmatrix} a_{11}^k & 0 & \cdots & 0 \\ 0 & a_{22}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^k \end{pmatrix}.$$

Таким образом, возвести диагональную матрицу в степень несложно. При возведении произвольной матрицы (не обязательно диагональной) в степень часто полезным оказывается использовать сначала свойства диагонализируемых матриц.

Используя умножение матриц и возведение матриц в степень, можно определить другие операции над матрицами. Например, матричная экспонента может быть определена через степенной ряд, матричный логарифм — как обратная к матричной экспоненте функция и так далее.

См. также

- Произведение Кронекера
- Произведение Адамара

Литература

- Корн Г., Корн Т.* Алгебра матриц и матричное исчисление // Справочник по математике. — 4-е издание. — М.: Наука, 1978. — С. 392—394.

Примечания

- Кибернетический сборник. Новая серия. Вып. 25. Сб. статей 1983 — 1985 гг.: Пер. с англ. — М.: Мир, 1988 — В.Б. Алексеев. Сложность умножения матриц. Обзор.
- Strassen V.* Gaussian Elimination is not Optimal (<http://bioinfo.ict.ac.cn/~dbu/AlgorithmCourses/Lectures/Strassen1968.pdf>) (англ.) // *Numerische Mathematik* / F. Brezzi — Springer Science+Business Media, 1969. — Vol. 13, Iss. 4. — P. 354—356. — ISSN 0029-599X (<https://www.worldcat.org/issn/0029-599X>); 0945-3245 (<https://www.worldcat.org/issn/0945-3245>) — doi:10.1007/BF02165411 (<https://dx.doi.org/10.1007/BF02165411>)
- Pan V. Ya, Strassen’s algorithm is not optimal — trilinear technique of aggregating uniting and canceling for constructing fast algorithms for matrix operations. — Proc. 19th Annual Symposium on Foundations of Computer Science, Ann Arbor, Mich., 1978
- Bini D., Capovani M., Lotti G., Romani F. — $O(n^{2.7799})$ complexity for approximate matrix multiplication. — Inform. Process. Lett., 1979
- Schonhage A. Partial and total matrix multiplication. — SIAM J. Comput., 1981
- Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation*, 9:251-280, 1990.
- Quanta Magazine (<https://www.quantamagazine.org/mathematicians-inch-closer-to-matrix-multiplication-goal-20210323/>)
- Group-theoretic Algorithms for Matrix Multiplication (<http://www.cs.caltech.edu/~umans/papers/CKSU05.pdf>) . Дата обращения: 26 сентября 2009. Архивировано (<https://web.archive.org/web/20110806121526/http://www.cs.caltech.edu/~umans/papers/CKSU05.pdf>) 6 августа 2011 года.

9. Toward an Optimal Algorithm for Matrix Multiplication (<https://web.archive.org/web/20100331095603/http://www.siam.org/pdf/news/174.pdf>) . Дата обращения: 26 сентября 2009. Архивировано из оригинала (<http://www.siam.org/pdf/news/174.pdf>) 31 марта 2010 года.

Источник — https://ru.wikipedia.org/w/index.php?title=Умножение_матриц&oldid=130608370

Эта страница в последний раз была отредактирована 23 мая 2023 в 00:30.

Текст доступен по лицензии Creative Commons «С указанием авторства — С сохранением условий» (CC BY-SA); в отдельных случаях могут действовать дополнительные условия.

Wikipedia® — зарегистрированный товарный знак некоммерческой организации Фонд Викимедиа (Wikimedia Foundation, Inc.)