Карчевский М.М.[1]

# Сингулярное разложение — это очень полезно!

В университетских курсах линейной алгебры и численных методов значительное внимание уделяется представлению матриц в виде произведения достаточно простых сомножителей. Укажем лишь некоторые примеры: $A = PLR$, $A$ — невырожденная матрица, $P$ — матрица перестановок, $L$ — нижняя треугольная матрица, $R$ — верхняя треугольная матрица[2]; $A = U\Lambda U^*$, $A$ — нормальная матрица, $U$ — унитарная матрица, $\Lambda$ — диагональная матрица; $A = UR$, $A$ — произвольная квадратная матрица, $U$ — унитарная матрица, $R$ — верхняя треугольная матрица.

В предлагаемой заметке речь идет о так называемом сингулярном разложении произвольной прямоугольной матрицы на множители. Оно играет очень важную роль как в теории, так и в разнообразных приложениях, однако редко упоминается в учебных курсах.

Более полно с затрагиваемыми здесь вопросами можно познакомиться по литературе, список которой приведен в конце статьи.

## 1. Сингулярные базисы и сингулярные числа матрицы

### 1. Основная теорема.

**Теорема 1.** *Пусть $A$ — произвольная матрица из $M_{m,n}$. Существуют такие ортонормированные базисы $\{e^k\}_{k=1}^n \subset \mathbb{C}^n$, $\{q^k\}_{k=1}^m \subset \mathbb{C}^m$ и положительные числа $\sigma_1 \geq \sigma_2 \geq \sigma_r$, $0 \leq r \leq \min\{m, n\}$, что*

$$Ae^k = \begin{cases} \sigma_k q^k, & k \leq r, \\ 0 \quad , & k > r, \end{cases} \tag{1}$$

Числа $\sigma_1, \sigma_2, \ldots, \sigma_r$ называют *сингулярными числами* матрицы $A$. Базисы $\{e^k\}_{k=1}^n$, $\{q^k\}_{k=1}^m$, обеспечивающие выполнение соотношений (1), называют *сингулярными базисами* матрицы $A$. Понятно, что $r$ есть размерность $\mathrm{Im}(A)$, т. е. $r$ — ранг матрицы $A$.

Доказательство теоремы 1. Матрица $A^*A$ самосопряжена и неотрицательно определена. Действительно $(A^*A)^* = A^*A$, $(A^*Ax, x) = (Ax, Ax) \geq 0$ для любого $x \in \mathbb{C}^n$. Поэтому существует ортонормированный базис собственных векторов $\{e^k\}_{k=1}^n$ матрицы $A^*A$. Все ее собственные числа неотрицательны. Таким образом,

$$A^*Ae^k = \sigma_k^2 e^k, \quad k = 1, 2, \ldots, n, \tag{2}$$

$\sigma_k^2 \geq 0$, $k = 1, 2, \ldots, n$. Будем считать при этом, что $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, $\sigma_{r+1} = \cdots = \sigma_n = 0$. Положим $z^k = Ae^k$ для $k = 1, 2, \ldots, r$ и заметим, что

$$(z^p, z^q) = (Ae^p, Ae^q) = (A^*Ae^p, e^q) = \sigma_p^2(e^p, e^q).$$

Значит,

$$(z^p, z^q) = \begin{cases} 0, & p \neq q, \\ \sigma_p^2, & p = q, \end{cases} \tag{3}$$

[1]Карчевский Михаил Миронович, mkarchev44@yandex.ru, кафедра вычислительной математики.

[2]Все не введенные в настоящей статье обозначения и определения можно найти, например, в [6].

следовательно, векторы

$$q^k = \sigma_k^{-1} A e^k, \quad k = 1, 2, \dots, r, \tag{4}$$

образуют ортонормированную систему в пространстве $\mathbb{C}^m$. Если окажется, что $r$ меньше $m$, дополним ее произвольно векторами $q^k$, $k = r+1, r+2, \dots, m$, до ортонормированного базиса пространства $\mathbb{C}^m$. Из равенств (4) вытекает справедливость (1). $\square$

**2. Матричное представление сингулярного разложения.** Равенства (1) часто записывают в виде

$$A = V \Sigma W^*. \tag{5}$$

Здесь $V$ — матрица, столбцами которой служат векторы $\{q^k\}_{k=1}^m$; $W$ — матрица, столбцами которой являются векторы $\{e^k\}_{k=1}^n$; $\Sigma$ — матрица, у которой главный минор порядка $r$ диагональный, на его диагонали расположены сингулярные числа матрицы $A$, все остальные элементы матрицы $\Sigma$ — нули.

Именно в такой форме обычно реализуют сингулярное разложение в пакетах стандартных программ линейной алгебры. Например, в среде Matlab сингулярное разложение может быть выполнено при помощи функции `svd` (от английского SVD — Singular Value Decomposition).

**3. Тензорное представление сингулярного разложения.** Из равенств (1) следует, что

$$Ax = \sum_{j=1}^{r} \sigma_j (x, e^j) q^j \quad \forall x \in \mathbb{C}^n. \tag{6}$$

Введем в рассмотрение матрицы $q^j \otimes e^j$ порядка $m$ на $n$, определяемые тождествами

$$q^j \otimes e^j x = (x, e^j) q^j \quad \forall\, x \in \mathbb{C}^n. \tag{7}$$

Матрицу $q^j \otimes e^j$ называют тензорным произведением векторов $q^j$, $e^j$. Из тождеств (6), (7) получаем, что

$$A = \sum_{j=1}^{r} \sigma_j q^j \otimes e^j. \tag{8}$$

Равенство (8) есть *тензорное представление* матрицы $A$. Отметим, что для задания матрицы $A$, достаточно хранить в памяти компьютера Q=r(m+n+1) чисел. Элементарные оценки показывают, что, например, для квадратной матрицы экономия достигается, если ранг матрицы не превышает половины ее порядка.

**4. Оценки сингулярных чисел.** Вследствие (2) очевидным образом получаем, что $(A^*Ax, x) = \sum\limits_{j=1}^{r} \sigma_j^2 |(x, e^j)|^2$, поэтому

$$\sigma_1^2 = \max_{|x|=1} (A^*Ax, x). \tag{9}$$

Если $A$ — квадратная матрица, то из равенства (9) нетрудно вывести следующие оценки

$$\max_{1 \le i,j \le n} |a_{ij}| \le \sigma_1 \le n \max_{1 \le i,j \le n} |a_{ij}|. \tag{10}$$

Оценки (10) неулучшаемы. Левое неравенство превращается в равенство при $A = I$, где $I$ — единичная матрица, правое — при $A = E$, где $E$ — матрица, все элементы которой равны единице.

## 2. Некоторые применения сингулярного разложения

**1. Полярные разложения.** Пусть $A$ — произвольная квадратная матрица. Равенство (5) можно переписать в виде

$$A = US, \quad A = TU, \tag{11}$$

где $U = VW^*$ — унитарная матрица, $S = W\Sigma W^*$, $T = V\Sigma V^*$ — самосопряженные нетрицательные матрицы[3]. Формулы (11) определяют *полярные разложения* матрицы (левое и правое). Они находят разнообразные приложения, например, в механике деформируемого твердого тела.

**2. Решение систем линейных алгебраических уравнений.** Если сингулярное разложение матрицы $A$ построено, то решение системы линейных алгебраических уравнений

$$Ax = b, \tag{12}$$

где $b$ — заданный вектор из $\mathbb{C}^m$, $x \in \mathbb{C}^n$ — искомый вектор, строится без труда. В самом деле, представим вектор $b$ в виде $b = \sum\limits_{k=1}^{m} (b, q^k) q^k$, а вектор $x$ будем разыскивать в виде разложения $x = \sum\limits_{k=1}^{n} \xi_k e^k$. Для определения чисел $\xi_k$, используя (12), получим соотношения $\xi_k = (b, q^k)/\sigma_k$, $k = 1, 2, \dots, r$. Далее нужно различать два случая:

1) все числа $\eta_k = (b, q^k)$, $k = r+1, r+2, \dots, m$, — нули, тогда

$$x = \sum_{k=1}^{r} (b, q^k)/\sigma_k \, e^k + \sum_{k=r+1}^{n} \xi_k e^k, \tag{13}$$

где числа $\xi_k$, $k = r+1, r+2, \dots, n$, произвольны, есть общее решение системы уравнений (12);

2) если хотя бы одно из чисел $\eta_k$ отлично от нуля, то система (12) не имеет решений, вектор, определяемый формулой (13), называют в этом случае псевдорешением системы (12); легко проверить, что этот вектор дает не бóльшую длину вектору невязки $r = Ax - b$ по сравнению с любым другим вектором пространства $\mathbb{C}^n$.

Если сингулярное разложение матрицы $A$ построено, то вычисления по формулам (13) требуют дополнительно всего порядка $rm$ арифметических операций.

**3. Оценка возмущения решения системы линейных алгебраических уравнений за счет неточности задания правой части.** Пусть $A$ — квадратная невырожденная матрица, $x$ — решение системы уравнений (12), $\tilde{x}$ — решение системы уравнений $Ax = \tilde{b}$. Тогда $A(x - \tilde{x}) = b - \tilde{b}$. Величину $\delta_x = |x - \tilde{x}|/|x|$ называют величиной *относительного изменения решения* при изменении правой части. Выясним, как она зависит от $\delta_b = |b - \tilde{b}|/|b|$ — величины *относительного изменения правой части.* Представляя $x$, $\tilde{x}$ в виде разложений по базису $\{e^k\}_{k=1}^n$, а $b$, $\tilde{b}$ — в виде разложений по базису $\{q^k\}_{k=1}^n$, получим, что

$$\delta_x^2 = \frac{\sum\limits_{k=1}^{n} \dfrac{|\eta_k - \tilde{\eta}_k|^2}{\sigma_k^2}}{\sum\limits_{k=1}^{n} \dfrac{|\eta_k|^2}{\sigma_k^2}} \le \frac{\sigma_1^2}{\sigma_n^2} \frac{\sum\limits_{k=1}^{n} |\eta_k - \tilde{\eta}_k|^2}{\sum\limits_{k=1}^{n} |\eta_k|^2} = \frac{\sigma_1^2}{\sigma_n^2} \delta_b^2. \tag{14}$$

[3]Напомним, что $U$ не меняет длин векторов, а $S$ (как и $T$) растягивает $\mathbb{C}^n$ в $n$ попарно ортогональных направлениях.

Таким образом,

$$\delta_x \leq \frac{\sigma_1}{\sigma_n}\delta_b. \tag{15}$$

Величину $\sigma_1/\sigma_n$, характеризующую устойчивость решения уравнения (12) по отношению к изменению его правой части, называют *числом обусловленности* матрицы $A$ и обозначают через $\operatorname{cond}(A)$. Очевидно, $\operatorname{cond}(A) \geq 1$ для любой матрицы $A$.

Задача. Приведите примеры матриц $A$, для которых $\operatorname{cond}(A) = 1$. Докажите, что оценка (15) неулучшаема в том смысле, что для любой невырожденной матрицы $A$ можно указать такие $b$ и $\tilde{b}$, что неравенство (15) превращается в равенство.

**4. Малоранговые приближения матриц.** Как было отмечено выше, при использовании сингулярного разложения количество чисел, требуемых для задания матрицы, пропорционально ее рангу. В то же время, во многих случаях требуется лишь приближенное задание матрицы. В связи этим естественным образом возникает следующая задача. Пусть $A \in M_{m,n}$. Требуется построить матрицу $X \in M_{m,n}$, являющуюся наилучшим приближением к $A$ среди всех матриц, ранг которых не превосходит заданного числа $k$. Пусть $A$, матрица ранга $r \geq 2$, представлена в виде (8), и $k < r$. Положим

$$X = \sum_{j=1}^{k} \sigma_j q^j \otimes e^j. \tag{16}$$

Имеем $\operatorname{rank}(X) = k$, причем $A - X = \sum\limits_{j=k+1}^{r} \sigma_j q^j \otimes e^j$, следовательно, максимальное сингулярное число матрицы $A - X$ равно $\sigma_{k+1}$. Пусть теперь матрица $Y \in M_{m,n}$ имеет ранг, не превосходящий $k$. Тогда $\operatorname{def}(Y)$ не меньше, чем $n - k$. Поэтому существует вектор $z \neq 0$, $z \in \operatorname{span}\{e^1, e^2, \dots, e^{k+1}\} \cap \operatorname{Ker}(Y)$. Можно считать, что $|z| = 1$. Положим $z = \sum\limits_{j=1}^{k+1} \alpha_j e^j$. Элементарные вычисления дают, что

$$((A - Y)^*(A - Y)z, z) = (A^*Az, z) = \sum_{j=1}^{k+1} \sigma_j^2 \alpha_j^2 \geq \sigma_{k+1}^2.$$

Вследствие (9) отсюда вытекает, что максимальное сингулярное число матрицы $A - Y$ не меньше, чем $\sigma_{k+1}$. Таким образом, можно сказать, что в определенном здесь смысле матрица $X$ — наилучшее приближение к матрице $A$ на множестве всех матриц ранга, не превосходящего $k < r$. Полученный результат составляет содержание *теоремы Эккарта — Янга* [1].

Говорят, что матрица $X$ осуществляет сжатие информации, содержащейся в матрице $A$. Если $A$ — квадратная матрица, то, используя оценку (10), близость матриц $X$ и $A$ можно характеризовать неравенством $\max\limits_{1 \leq i,j \leq n} |a_{ij} - x_{ij}| \leq \sigma_{k+1}$.

**5. Сингулярное разложение и фотографии.** Используя функции среды Matlab, можно существенно сократить объем памяти компьютера для хранения фотографий без заметного ухудшения качества изображений. Продемонстрируем, как это делается, на примере конкретного файла `doll.jpg`. Чтение информации из этого файла и последующее преобразование ее в стандартный числовой формат выполняется функциями `A1=imread('doll.jpg')`, `DA1=im2double(A1)`. Получили трехмерный массив `DA1` с размерами `m=972, n=1296, p=3.` «Вертикальный» размер `p` для любой цветной фотографии равен 3. Размеры «горизонтальных» слоев, `m, n`, зависят от качества конкретной фотографии. Можно определить ранг каждого горизонтального слоя, используя функцию `rank`. В данном примере все ранги оказались равными `m=972`, то есть все горизонтальные слои — полноранговые матрицы. Экранный образ файла `DA1` можно получить при помощи функции `image(DA1)`. Он будет дан в виде фигуры Matlab. Задавая затем некоторое целое $1 \leq k < 972$, каждую матрицу `DA1(:,:,j)`, `j=1,2,3`, используя функцию `svd`, приблизим матрицей вида (16). На модифицированный таким образом массив `DA1` можно подействовать затем функцией `image`. Результаты работы описанного метода иллюстрирует рис. 1, с. 6. Поясним, что `q` обозначает объем необходимой памяти для хранения матриц вида (16) по отношению к количеству чисел в исходном массиве `DA1`.

a) исходный массив

b) ранг `k=60, q=10.8%`

c) ранг `k=30, q=5.4%`

d) ранг `k=15, q=2.7%`

Рис. 1:

# Список литературы


[1] **Eckart Carl, Young Gale.** The approximation of one matrix by another of lower rank//Psychometrika.— 1936.— Vol. 1, no. 3. — pp. 211–218.

[2] **Гантмахер Ф.Р.** Теория матриц. М.: Физматлит, 2010.

[3] **Хорн Р., Джонсон Ч.** Матричный анализ. М.: Мир, 1989.

[4] **Деммель Дж.** Вычислительная линейная алгебра. Теория и приложения. М.: Мир, 2001.

[5] **Beilina L., Karchevskii E., Karchevskii M.** Numerical Linear Algebra: Theory and Application. Springer, 2017.

[6] **Карчевский Е.М., Карчевский М.М.** Лекции по линейной алгебре и аналитической геметрии. СПб.: Издательство «Лань», 2018.