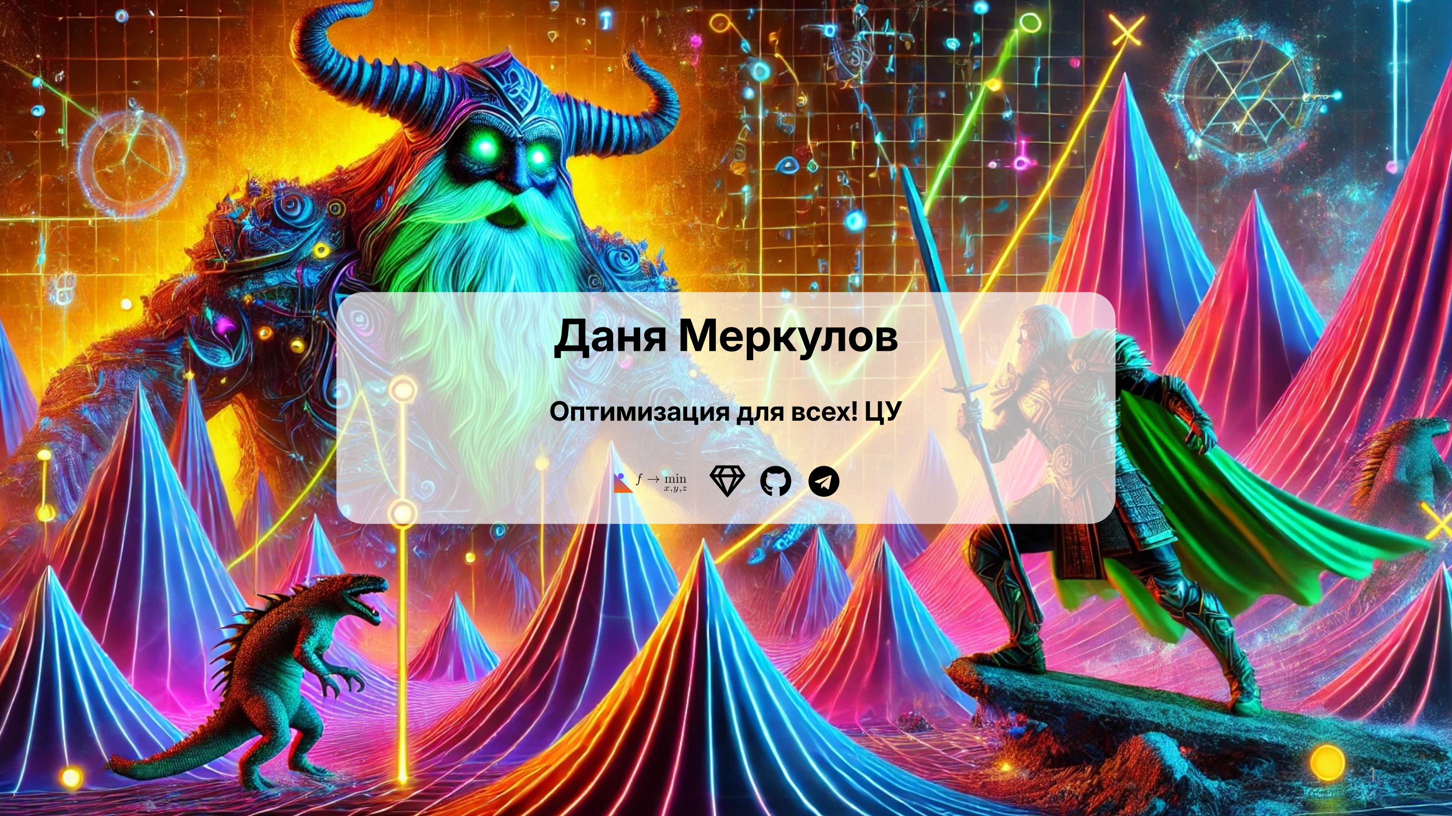
ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

Вспоминаем линейную алгебру. Скорости сходимости

МЕТОДЫ ВЫПУКЛОЙ ОПТИМИЗАЦИИ

НЕДЕЛЯ 1

Даня Меркулов



Даня Меркулов

Оптимизация для всех! ЦУ

$f \to \min\limits_{x,y,z}$

Вспоминаем линейную алгебру

Векторы и матрицы

Мы будем считать, что все векторы являются столбцами по умолчанию. Пространство векторов длины n обозначается $\mathbb{R}^n$, а пространство матриц размера $m \times n$ с вещественными элементами обозначается $\mathbb{R}^{m \times n}$. То есть [1]:

$$
x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad x^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \quad x \in \mathbb{R}^n, x_i \in \mathbb{R} \tag{1}
$$

[1] Подробный вводный курс по прикладной линейной алгебре можно найти в книге Introduction to Applied Linear Algebra – Vectors, Matrices, and Least Squares - книга от Stephen Boyd & Lieven Vandenberghe, которая указана в источнике. Также полезен материал по линейной алгебре приведенный в приложении A книги Numerical Optimization by Jorge Nocedal Stephen J. Wright.

Векторы и матрицы

Мы будем считать, что все векторы являются столбцами по умолчанию. Пространство векторов длины n обозначается $\mathbb{R}^n$, а пространство матриц размера $m \times n$ с вещественными элементами обозначается $\mathbb{R}^{m \times n}$. То есть [1]:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad x^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \quad x \in \mathbb{R}^n, x_i \in \mathbb{R} \tag{1}$$

Аналогично, если $A \in \mathbb{R}^{m \times n}$ мы обозначаем транспонирование как $A^T \in \mathbb{R}^{n \times m}$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix} \quad A \in \mathbb{R}^{m \times n}, a_{ij} \in \mathbb{R}$$

Мы будем писать $x \geq 0$ и $x \neq 0$ для обозначения покомпонентных неравенств

[1] Подробный вводный курс по прикладной линейной алгебре можно найти в книге Introduction to Applied Linear Algebra – Vectors, Matrices, and Least Squares - книга от Stephen Boyd & Lieven Vandenberghe, которая указана в источнике. Также полезен материал по линейной алгебре приведенный в приложении A книги Numerical Optimization by Jorge Nocedal Stephen J. Wright.

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

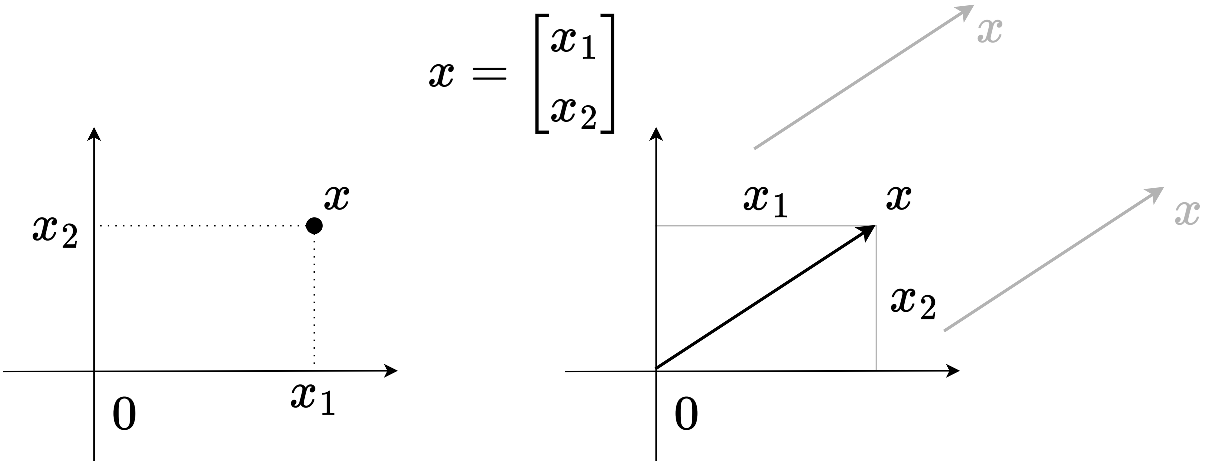


Рисунок 1. Эквивалентные представления вектора

Матрица A называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности n). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица A называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности n). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) определенной**, если для всех $x \neq 0 : x^T Ax > (<) 0$. Обозначается как $A \succ (\prec) 0$. Множество таких матриц обозначается как $\mathbb{S}^n_{++} (\mathbb{S}^n_{--})$

Матрица A называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности n). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) определенной**, если для всех $x \neq 0 : x^T Ax > (<)0$. Обозначается как $A \succ (\prec)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{++}(\mathbb{S}^n_{--})$

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) полуопределенной**, если для всех $x : x^T Ax \geq (\leq)0$. Обозначается как $A \succeq (\preceq)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{+}(\mathbb{S}^n_{-})$

Question

Верно ли, что положительно определенная матрица имеет все положительные элементы?

Матрица A называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности n). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) определенной**, если для всех $x \neq 0 : x^T Ax > (<)0$. Обозначается как $A \succ (\prec)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{++}(\mathbb{S}^n_{--})$

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) полуопределенной**, если для всех $x : x^T Ax \geq (\leq)0$. Обозначается как $A \succeq (\preceq)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{+}(\mathbb{S}^n_{-})$

> **i Question**
>
> Верно ли, что положительно определенная матрица имеет все положительные элементы?

> **i Question**
>
> Верно ли, что если матрица симметрична, то она должна быть положительно определенной?

Матрица A называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности n). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) определенной**, если для всех $x \neq 0 : x^T Ax > (<) 0$. Обозначается как $A \succ (\prec) 0$. Множество таких матриц обозначается как $\mathbb{S}^n_{++} (\mathbb{S}^n_{--})$

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) полуопределенной**, если для всех $x : x^T Ax \geq (\leq) 0$. Обозначается как $A \succeq (\preceq) 0$. Множество таких матриц обозначается как $\mathbb{S}^n_{+} (\mathbb{S}^n_{-})$

Question

Верно ли, что положительно определенная матрица имеет все положительные элементы?

Question

Верно ли, что если матрица симметрична, то она должна быть положительно определенной?

Question

Верно ли, что если матрица положительно определена, то она должна быть симметричной?

Матричное умножение (matmul)

Пусть A - матрица размера $m \times n$, а B - матрица размера $n \times p$, тогда их произведение AB равно:

$$C = AB$$

Тогда C - матрица размера $m \times p$, элемент (i, j) которой равен:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Эта операция в наивной форме требует $\mathcal{O}(n^3)$ арифметических операций, где n обычно считается наибольшей размерностью матриц.

Матричное умножение (matmul)

Пусть A - матрица размера $m \times n$, а B - матрица размера $n \times p$, тогда их произведение AB равно:

$$C = AB$$

Тогда C - матрица размера $m \times p$, элемент (i, j) которой равен:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Эта операция в наивной форме требует $\mathcal{O}(n^3)$ арифметических операций, где n обычно считается наибольшей размерностью матриц.

> **Question**
>
> Возможно ли умножить две матрицы быстрее, чем за $\mathcal{O}(n^3)$? Как насчет $\mathcal{O}(n^2)$, $\mathcal{O}(n)$?

Умножение матрицы на вектор (matvec)

Пусть A - матрица размера $m \times n$, а x - вектор длины n, тогда i-й элемент произведения Ax равен:

$$z = Ax$$

равен:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной форме требует $\mathcal{O}(n^2)$ арифметических операций, где n обычно считается наибольшей размерностью входов.

Отметим, что:

- $C = AB \quad C^T = B^T A^T$

Умножение матрицы на вектор (matvec)

Пусть A - матрица размера $m \times n$, а x - вектор длины n, тогда i-й элемент произведения Ax равен:

$$z = Ax$$

равен:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной форме требует $\mathcal{O}(n^2)$ арифметических операций, где n обычно считается наибольшей размерностью входов.

Отметим, что:

- $C = AB \quad C^T = B^T A^T$
- $AB \neq BA$

Умножение матрицы на вектор (matvec)

Пусть A - матрица размера $m \times n$, а x - вектор длины n, тогда i-й элемент произведения Ax равен:

$$z = Ax$$

равен:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной форме требует $\mathcal{O}(n^2)$ арифметических операций, где n обычно считается наибольшей размерностью входов.

Отметим, что:

- $C = AB \quad C^T = B^T A^T$
- $AB \neq BA$
- $e^A = \sum\limits_{k=0}^{\infty} \frac{1}{k!} A^k$

Умножение матрицы на вектор (matvec)

Пусть A - матрица размера $m \times n$, а x - вектор длины n, тогда i-й элемент произведения Ax равен:

$$z = Ax$$

равен:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной форме требует $\mathcal{O}(n^2)$ арифметических операций, где n обычно считается наибольшей размерностью входов.

Отметим, что:

- $C = AB \quad C^T = B^T A^T$
- $AB \neq BA$
- $e^A = \sum\limits_{k=0}^{\infty} \frac{1}{k!} A^k$
- $e^{A+B} \neq e^A e^B$ (но если A и B коммутируют, то есть $AB = BA$, то $e^{A+B} = e^A e^B$)

Умножение матрицы на вектор (matvec)

Пусть A - матрица размера $m \times n$, а x - вектор длины n, тогда i-й элемент произведения Ax равен:

$$z = Ax$$

равен:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной форме требует $\mathcal{O}(n^2)$ арифметических операций, где n обычно считается наибольшей размерностью входов.

Отметим, что:

- $C = AB \quad C^T = B^T A^T$
- $AB \neq BA$
- $e^A = \sum\limits_{k=0}^{\infty} \frac{1}{k!} A^k$
- $e^{A+B} \neq e^A e^B$ (но если A и B коммутируют, то есть $AB = BA$, то $e^{A+B} = e^A e^B$)
- $\langle x, Ay \rangle = \langle A^T x, y \rangle$

Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$

Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)

Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha| \|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)
3. Если $\|x\| = 0$, то $x = 0$

Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)
3. Если $\|x\| = 0$, то $x = 0$

Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)
3. Если $\|x\| = 0$, то $x = 0$

Расстояние между двумя векторами определяется как

$$d(x, y) = \|x - y\|.$$

Наиболее широко используемой нормой является **Евклидова норма**:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2},$$

которая соответствует расстоянию в нашей реальной жизни. Если векторы имеют комплексные элементы, мы используем их модуль. Евклидова норма, или 2-норма, является подклассом важного класса p-норм:

$$\|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p}.$$

p-норма вектора

Существуют два очень важных частных случая. Бесконечность-норма, или норма Чебышева, определяется как максимальное абсолютное значение элемента вектора:

$$\|x\|_\infty = \max_i |x_i|$$

p-норма вектора

Существуют два очень важных частных случая. Бесконечность-норма, или норма Чебышева, определяется как максимальное абсолютное значение элемента вектора:

$$\|x\|_\infty = \max_i |x_i|$$

l_1 норма (или **манхэттенское расстояние**) определяется как сумма модулей элементов вектора x:

$$\|x\|_1 = \sum_i |x_i|$$

p-норма вектора

Существуют два очень важных частных случая. Бесконечность-норма, или норма Чебышева, определяется как максимальное абсолютное значение элемента вектора:

$$\|x\|_\infty = \max_i |x_i|$$

l_1 норма (или **манхэттенское расстояние**) определяется как сумма модулей элементов вектора x:

$$\|x\|_1 = \sum_i |x_i|$$

l_1 норма играет очень важную роль: она все связана с методами **compressed sensing**, которые появились в середине 00-х как одна из популярных тем исследований. Код для изображения ниже доступен *здесь:*. Также посмотрите *это* видео.

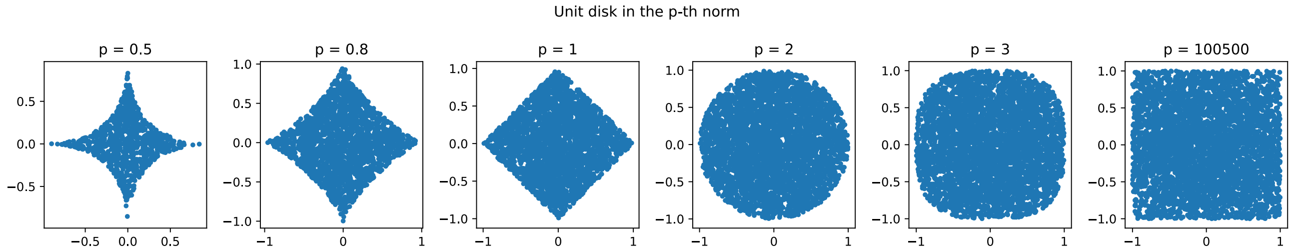


Рисунок 2. Шары в разных нормах на плоскости

Матричные нормы

В некотором смысле между матрицами и векторами нет большой разницы (вы можете векторизовать матрицу), и здесь появляется самая простая матричная норма **Фробениуса**:

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

Матричные нормы

В некотором смысле между матрицами и векторами нет большой разницы (вы можете векторизовать матрицу), и здесь появляется самая простая матричная норма **Фробениуса**:

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2\right)^{1/2}$$

Спектральная норма, $\|A\|_2$ является одной из наиболее широко используемых матричных норм (наряду с нормой Фробениуса).

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2},$$

Она не может быть вычислена непосредственно из элементов с помощью простой формулы, как в случае нормы Фробениуса, однако, существуют эффективные алгоритмы для ее вычисления. Она напрямую связана с **сингулярным разложением** (SVD) матрицы. Для неё справедливо:

$$\|A\|_2 = \sigma_1(A) = \sqrt{\lambda_{\text{max}}(A^T A)}$$

где $\sigma_1(A)$ - наибольшее сингулярное значение матрицы A.

Скалярное произведение

Стандартное **скалярное произведение** между векторами x и y из $\mathbb{R}^n$ равно:

$$\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i = y^T x = \langle y, x \rangle$$

Здесь x_i и y_i - i-ые компоненты соответствующих векторов.

> **Example**
>
> Докажите, что вы можете переставить матрицу внутри скалярного произведения с транспонированием: $\langle x, Ay \rangle = \langle A^T x, y \rangle$ и $\langle x, yB \rangle = \langle xB^T, y \rangle$

Скалярное произведение матриц

Стандартное **скалярное произведение** между матрицами X и Y из $\mathbb{R}^{m\times n}$ равно:

$$\langle X, Y\rangle = \text{tr}(X^TY) = \sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}Y_{ij} = \text{tr}(Y^TX) = \langle Y, X\rangle$$

Question

Существует ли связь между нормой Фробениуса $\|\cdot\|_F$ и скалярным произведением между матрицами $\langle\cdot,\cdot\rangle$?

Собственные вектора и собственные значения

Число λ является собственным значением квадратной матрицы A размера $n \times n$, если существует ненулевой вектор q такой, что

$$Aq = \lambda q.$$

Вектор q называется собственным вектором матрицы A. Матрица A невырожденная, если ни одно из её собственных значений не равно нулю. Собственные значения симметричных матриц являются вещественными числами, в то время как несимметричные матрицы могут иметь комплексные собственные значения. Если матрица положительно определена и симметрична, то все её собственные значения являются положительными вещественными числами.

Собственные вектора и собственные значения

Theorem

$$A \succeq (\succ) 0 \Leftrightarrow \text{все собственные значения } A \; \geq (>) 0$$

Proof

1. *$\rightarrow$ Предположим, что некоторое собственное значение λ отрицательно, и пусть x обозначает соответствующий собственный вектор. Тогда*

   $$Ax = \lambda x \rightarrow x^T A x = \lambda x^T x < 0$$

   что противоречит условию $A \succeq 0$.

Собственные вектора и собственные значения

Theorem

$$A \succeq (\succ) 0 \Leftrightarrow \text{все собственные значения } A \geq (>) 0$$

Proof

1. $\rightarrow$ *Предположим, что некоторое собственное значение λ отрицательно, и пусть x обозначает соответствующий собственный вектор. Тогда*
$$Ax = \lambda x \rightarrow x^T Ax = \lambda x^T x < 0$$
что противоречит условию $A \succeq 0$.
2. $\leftarrow$ *Для любой симметричной матрицы мы можем выбрать набор собственных векторов $v_1, \dots, v_n$, которые образуют ортонормированный базис в $\mathbb{R}^n$. Возьмем любой вектор $x \in \mathbb{R}^n$.*
$$\begin{aligned} x^T Ax &= (\alpha_1 v_1 + \ldots + \alpha_n v_n)^T A(\alpha_1 v_1 + \ldots + \alpha_n v_n) \\ &= \sum \alpha_i^2 v_i^T Av_i = \sum \alpha_i^2 \lambda_i v_i^T v_i \geq 0 \end{aligned}$$
Здесь мы использовали тот факт, что $v_i^T v_j = 0$, для $i \neq j$.

Спектральное разложение (eigendecomposition)

Пусть $A \in S_n$, т.е. A - вещественная симметричная матрица размера $n \times n$. Тогда A может быть разложена как

$$A = Q\Lambda Q^T,$$

[2]Хорошая шпаргалка с разложением матриц доступна на сайте курса по линейной алгебре website.

Спектральное разложение (eigendecomposition)

Пусть $A \in S_n$, т.е. A - вещественная симметричная матрица размера $n \times n$. Тогда A может быть разложена как

$$A = Q\Lambda Q^T,$$

где $Q \in \mathbb{R}^{n\times n}$ ортогональная, т.е. удовлетворяет $Q^T Q = I$, и $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Вещественные числа λ_i являются собственными значениями A и являются корнями характеристического полинома $\det(A - \lambda I)$. Столбцы Q образуют ортонормированный набор собственных векторов A. Такое разложение называется спектральным. [2]

[2] Хорошая шпаргалка с разложением матриц доступна на сайте курса по линейной алгебре website.

Спектральное разложение (eigendecomposition)

Пусть $A \in S_n$, т.е. A - вещественная симметричная матрица размера $n \times n$. Тогда A может быть разложена как

$$A = Q \Lambda Q^T,$$

где $Q \in \mathbb{R}^{n \times n}$ ортогональная, т.е. удовлетворяет $Q^T Q = I$, и $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Вещественные числа λ_i являются собственными значениями A и являются корнями характеристического полинома $\det(A - \lambda I)$. Столбцы Q образуют ортонормированный набор собственных векторов A. Такое разложение называется спектральным. [2]

Мы обычно упорядочиваем вещественные собственные значения как $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Мы используем обозначение $\lambda_i(A)$ для обозначения i-го наибольшего собственного значения $A \in S$. Мы обычно пишем наибольшее или максимальное собственное значение как $\lambda_1(A) = \lambda_{\text{max}}(A)$, и наименьшее или минимальное собственное значение как $\lambda_n(A) = \lambda_{\text{min}}(A)$.

[2] Хорошая шпаргалка с разложением матриц доступна на сайте курса по линейной алгебре website.

Собственные значения

Наибольшее и наименьшее вещественныесобственные значения удовлетворяют

$$\lambda_{\text{min}}(A) = \inf_{x \neq 0} \frac{x^T A x}{x^T x}, \qquad \lambda_{\text{max}}(A) = \sup_{x \neq 0} \frac{x^T A x}{x^T x}$$

Собственные значения

Наибольшее и наименьшее вещественныесобственные значения удовлетворяют

$$\lambda_{\text{min}}(A) = \inf_{x \neq 0} \frac{x^T A x}{x^T x}, \qquad \lambda_{\text{max}}(A) = \sup_{x \neq 0} \frac{x^T A x}{x^T x}$$

и, следовательно, $\forall x \in \mathbb{R}^n$ (соотношение Рэлея):

$$\lambda_{\text{min}}(A) x^T x \leq x^T A x \leq \lambda_{\text{max}}(A) x^T x$$

Собственные значения

Наибольшее и наименьшее вещественныесобственные значения удовлетворяют

$$\lambda_{\text{min}}(A) = \inf_{x \neq 0} \frac{x^T A x}{x^T x}, \qquad \lambda_{\text{max}}(A) = \sup_{x \neq 0} \frac{x^T A x}{x^T x}$$

и, следовательно, $\forall x \in \mathbb{R}^n$ (соотношение Рэлея):

$$\lambda_{\text{min}}(A) x^T x \leq x^T A x \leq \lambda_{\text{max}}(A) x^T x$$

Число обусловленности невырожденной матрицы определяется как

$$\kappa(A) = \|A\| \|A^{-1}\|$$

Собственные значения

Наибольшее и наименьшее вещественныесобственные значения удовлетворяют

$$\lambda_{\text{min}}(A) = \inf_{x \neq 0} \frac{x^T A x}{x^T x}, \qquad \lambda_{\text{max}}(A) = \sup_{x \neq 0} \frac{x^T A x}{x^T x}$$

и, следовательно, $\forall x \in \mathbb{R}^n$ (соотношение Рэлея):

$$\lambda_{\text{min}}(A) x^T x \leq x^T A x \leq \lambda_{\text{max}}(A) x^T x$$

Число обусловленности невырожденной матрицы определяется как

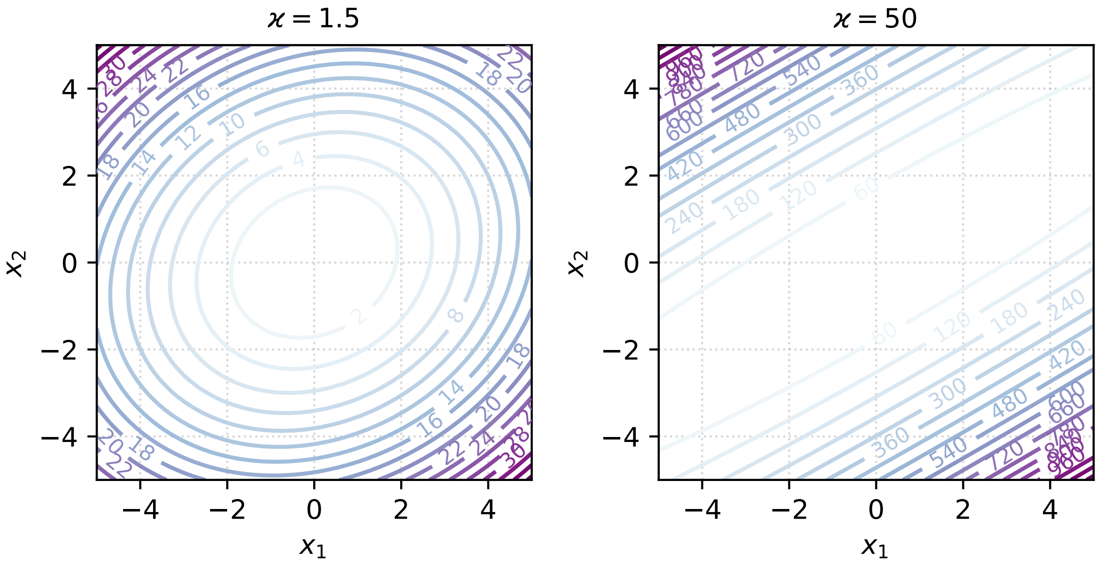
$$\kappa(A) = \|A\| \|A^{-1}\|$$

Если мы используем спектральную матричную норму, мы можем получить:

$$\kappa(A) = \frac{\sigma_{\text{max}}(A)}{\sigma_{\text{min}}(A)}$$

Если, кроме того, $A \in \mathbb{S}^n_{++}$: $\kappa(A) = \dfrac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)}$

Число обусловленности



Сингулярное разложение (SVD)

Пусть $A \in \mathbb{R}^{m \times n}$ с рангом $A = r$. Тогда A может быть разложена как

$$A = U\Sigma V^T$$

Сингулярное разложение (SVD)

Пусть $A \in \mathbb{R}^{m \times n}$ с рангом $A = r$. Тогда A может быть разложена как

$$A = U \Sigma V^T$$

где $U \in \mathbb{R}^{m \times r}$ удовлетворяет $U^T U = I$, $V \in \mathbb{R}^{n \times r}$ удовлетворяет $V^T V = I$, и Σ является диагональной матрицей с $\Sigma = \text{diag}(\sigma_1, ..., \sigma_r)$, такой что

Сингулярное разложение (SVD)

Пусть $A \in \mathbb{R}^{m \times n}$ с рангом $A = r$. Тогда A может быть разложена как

$$A = U \Sigma V^T$$

где $U \in \mathbb{R}^{m \times r}$ удовлетворяет $U^T U = I$, $V \in \mathbb{R}^{n \times r}$ удовлетворяет $V^T V = I$, и Σ является диагональной матрицей с $\Sigma =$ diag$(\sigma_1, ..., \sigma_r)$, такой что

$$\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0.$$

Сингулярное разложение (SVD)

Пусть $A \in \mathbb{R}^{m \times n}$ с рангом $A = r$. Тогда A может быть разложена как

$$A = U \Sigma V^T$$

где $U \in \mathbb{R}^{m \times r}$ удовлетворяет $U^T U = I$, $V \in \mathbb{R}^{n \times r}$ удовлетворяет $V^T V = I$, и Σ является диагональной матрицей с $\Sigma = \text{diag}(\sigma_1, ..., \sigma_r)$, такой что

$$\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0.$$

Это разложение называется **сингулярным разложением (SVD)** матрицы A. Столбцы U называются левыми сингулярными векторами A, столбцы V называются правыми сингулярными векторами, и числа σ_i являются сингулярными значениями. Сингулярное разложение может быть записано как

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T,$$

где $u_i \in \mathbb{R}^m$ являются левыми сингулярными векторами, и $v_i \in \mathbb{R}^n$ являются правыми сингулярными векторами.

Сингулярное разложение

Question

Пусть $A \in \mathbb{S}^n_{++}$. Что мы можем сказать о связи между его собственными значениями и сингулярными значениями?

Сингулярное разложение

Question

Пусть $A \in \mathbb{S}^n_{++}$. Что мы можем сказать о связи между его собственными значениями и сингулярными значениями?

Question

Как сингулярные значения матрицы связаны с её собственными значениями, особенно для симметричной матрицы?

Ранговое разложение (Skeleton decomposition)

Простое, но очень интересное разложение - это ранговое разложение, которое может быть записано в двух формах:

$$A = UV^T \quad A = \hat{C}\hat{A}^{-1}\hat{R}$$

Ранговое разложение (Skeleton decomposition)

Простое, но очень интересное разложение - это ранговое разложение, которое может быть записано в двух формах:

$$A = UV^T \quad A = \hat{C}\hat{A}^{-1}\hat{R}$$

Последнее выражение относится к забавному факту: вы можете случайным образом выбрать r линейно независимых столбцов матрицы и любые r линейно независимых строк матрицы и хранить только их с возможностью точно (!) восстановить всю матрицу.

Ранговое разложение (Skeleton decomposition)

Простое, но очень интересное разложение - это ранговое разложение, которое может быть записано в двух формах:

$$A = UV^T \quad A = \hat{C}\hat{A}^{-1}\hat{R}$$

Последнее выражение относится к забавному факту: вы можете случайным образом выбрать r линейно независимых столбцов матрицы и любые r линейно независимых строк матрицы и хранить только их с возможностью точно (!) восстановить всю матрицу.

Применения для рангового разложения:

- Сжатие модели, сжатие данных и ускорение вычислений в численном анализе: для матрицы ранга r с $r \ll n, m$ необходимо хранить $\mathcal{O}((n+m)r) \ll nm$ элементов.

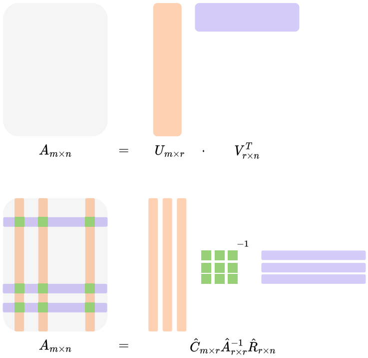


Рисунок 3. Иллюстрация рангового разложения

Ранговое разложение (Skeleton decomposition)

Простое, но очень интересное разложение - это ранговое разложение, которое может быть записано в двух формах:

$$A = UV^T \quad A = \hat{C}\hat{A}^{-1}\hat{R}$$

Последнее выражение относится к забавному факту: вы можете случайным образом выбрать r линейно независимых столбцов матрицы и любые r линейно независимых строк матрицы и хранить только их с возможностью точно (!) восстановить всю матрицу.

Применения для рангового разложения:

- Сжатие модели, сжатие данных и ускорение вычислений в численном анализе: для матрицы ранга r с $r \ll n, m$ необходимо хранить $\mathcal{O}((n+m)r) \ll nm$ элементов.
- Извлечение признаков в машинном обучении

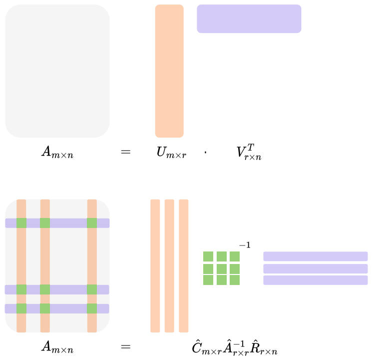


Рисунок 3. Иллюстрация рангового разложения

Ранговое разложение (Skeleton decomposition)

Простое, но очень интересное разложение - это ранговое разложение, которое может быть записано в двух формах:

$$A = UV^T \quad A = \hat{C}\hat{A}^{-1}\hat{R}$$

Последнее выражение относится к забавному факту: вы можете случайным образом выбрать r линейно независимых столбцов матрицы и любые r линейно независимых строк матрицы и хранить только их с возможностью точно (!) восстановить всю матрицу.

Применения для рангового разложения:

- Сжатие модели, сжатие данных и ускорение вычислений в численном анализе: для матрицы ранга r с $r \ll n, m$ необходимо хранить $\mathcal{O}((n+m)r) \ll nm$ элементов.
- Извлечение признаков в машинном обучении
- Все приложения, где применяется SVD, так как ранговое разложение может быть преобразовано в форму усеченного SVD.

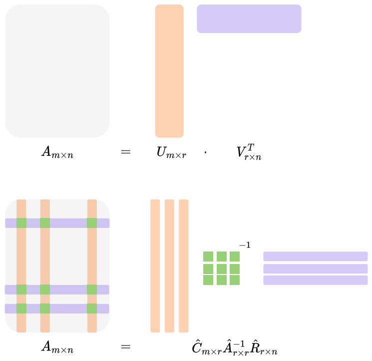


Рисунок 3. Иллюстрация рангового разложения

Каноническое тензорное разложение

Можно рассмотреть обобщение рангового разложения на структуры данных более высокого порядка, такие как тензоры, что означает представление тензора в виде суммы r простых тензоров.

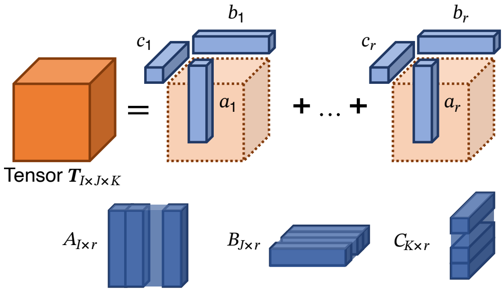


Рисунок 4. Иллюстрация канонического тензорного разложения

> **Example**
>
> Заметьте, что существует множество тензорных разложений: каноническое, Таккера, тензорный поезд (TT), тензорное кольцо (TR) и другие. В случае тензоров мы не имеем прямого определения *ранга* для всех типов разложений. Например, для разложения Тензорного поезда ранг является не скаляром, а вектором.

Определитель и след матрицы

Определитель и след матрицы могут быть выражены через собственные значения

$$\text{det}A = \prod_{i=1}^{n} \lambda_i, \qquad \text{tr}A = \sum_{i=1}^{n} \lambda_i$$

Определитель имеет несколько интересныхсвойств. Например,

- $\text{det}A = 0$ тогда и только тогда, когда A является вырожденной;

Определитель и след матрицы

Определитель и след матрицы могут быть выражены через собственные значения

$$\text{det}A = \prod_{i=1}^{n} \lambda_i, \qquad \text{tr}A = \sum_{i=1}^{n} \lambda_i$$

Определитель имеет несколько интересныхсвойств. Например,

- det$A = 0$ тогда и только тогда, когда A является вырожденной;
- det$AB = ($det$A)($det$B)$;

Определитель и след матрицы

Определитель и след матрицы могут быть выражены через собственные значения

$$\text{det}A = \prod_{i=1}^{n} \lambda_i, \qquad \text{tr}A = \sum_{i=1}^{n} \lambda_i$$

Определитель имеет несколько интересныхсвойств. Например,

- $\text{det}A = 0$ тогда и только тогда, когда A является вырожденной;
- $\text{det}AB = (\text{det}A)(\text{det}B)$;
- $\text{det}A^{-1} = \frac{1}{\text{det }A}$.

Определитель и след матрицы

Определитель и след матрицы могут быть выражены через собственные значения

$$\text{det}A = \prod_{i=1}^{n} \lambda_i, \qquad \text{tr}A = \sum_{i=1}^{n} \lambda_i$$

Определитель имеет несколько интересныхсвойств. Например,

- $\text{det}A = 0$ тогда и только тогда, когда A является вырожденной;
- $\text{det}AB = (\text{det}A)(\text{det}B)$;
- $\text{det}A^{-1} = \frac{1}{\text{det } A}$.

Определитель и след матрицы

Определитель и след матрицы могут быть выражены через собственные значения

$$\text{det}A = \prod_{i=1}^{n} \lambda_i, \qquad \text{tr}A = \sum_{i=1}^{n} \lambda_i$$

Определитель имеет несколько интересныхсвойств. Например,

- $\text{det}A = 0$ тогда и только тогда, когда A является вырожденной;
- $\text{det}AB = (\text{det}A)(\text{det}B)$;
- $\text{det}A^{-1} = \frac{1}{\text{det } A}$.

Не забывайте о циклическом свойстве следа для произвольных матриц A, B, C, D (предполагая, что все размерности согласованы):

$$\text{tr}(ABCD) = \text{tr}(DABC) = \text{tr}(CDAB) = \text{tr}(BCDA)$$

Определитель и след матрицы

Определитель и след матрицы могут быть выражены через собственные значения

$$\text{det} A = \prod_{i=1}^{n} \lambda_i, \qquad \text{tr} A = \sum_{i=1}^{n} \lambda_i$$

Определитель имеет несколько интересныхсвойств. Например,

- $\text{det} A = 0$ тогда и только тогда, когда A является вырожденной;
- $\text{det} AB = (\text{det} A)(\text{det} B)$;
- $\text{det} A^{-1} = \frac{1}{\text{det}\ A}$.

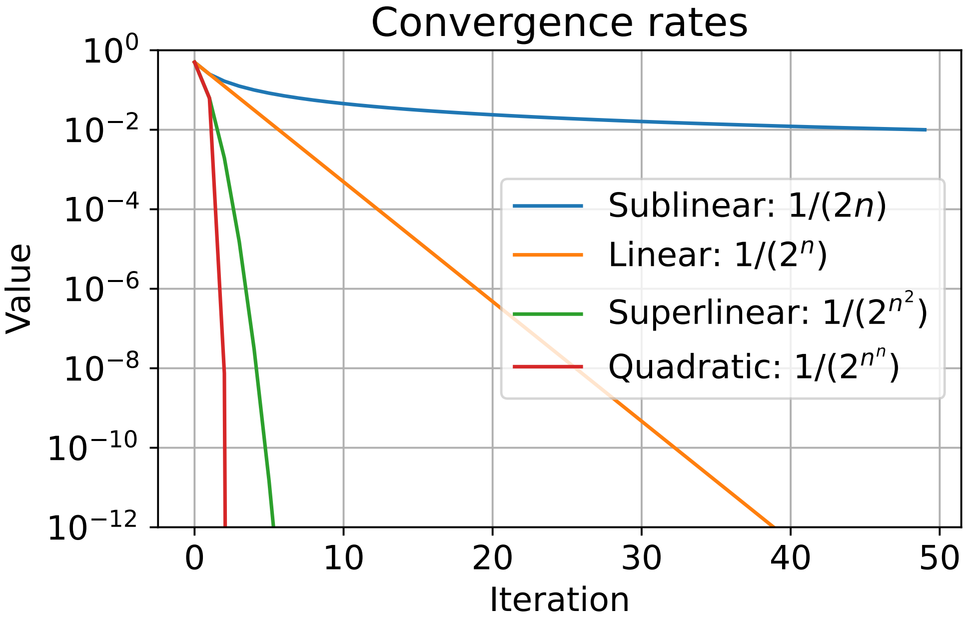
Не забывайте о циклическом свойстве следа для произвольных матриц A, B, C, D (предполагая, что все размерности согласованы):

$$\text{tr}(ABCD) = \text{tr}(DABC) = \text{tr}(CDAB) = \text{tr}(BCDA)$$

> **Question**
>
> Как определитель матрицы связан с её обратимостью?

Скорости сходимости

Скорость сходимости



Линейная сходимость

Чтобы сравнить производительность алгоритмов, мы должны определить термины для различных типов сходимости. Пусть r_k - последовательность неотрицательных вещественных чисел, которая сходится к нулю. Обычно мы имеем итерационный метод, который производит последовательность итераций x_k, приближающихся к оптимальному решению x^*, и $r_k = \|x_k - x^*\|_2$.

Линейная сходимость последовательности r_k определяется следующим образом:

Последовательность $\{r_k\}_{k=m}^\infty$ сходится линейно с параметром $0 < q < 1$, если существует константа $C > 0$ такая, что:

$$r_k \leq Cq^k, \quad \text{for all } k \geq m.$$

Если такое q существует, то последовательность называется линейно сходящейся. **Точная нижняя граница** всех q, удовлетворяющих неравенству, называется **скоростью линейной сходимости** последовательности.

Question

Предположим, у вас есть две последовательности с линейными скоростями сходимости $q_1 = 0.1$ и $q_2 = 0.7$, какая из них быстрее?

Линейная сходимость

Example

Предположим, у нас есть следующая последовательность:

$$r_k = \frac{1}{2^k}$$

Можно сразу заключить, что мы имеем линейную сходимость с параметрами $q = \dfrac{1}{2}$ и $C = 0$.

Question

Определите сходимость следующей последовательности

$$r_k = \frac{3}{2^k}$$

Сублинейная сходимость

Если последовательность r_k сходится к нулю, но не имеет линейной сходимости, то сходимость называется сублинейной. Иногда мы можем рассмотреть следующий частный случай сублинейной сходимости:

$$\|x_{k+1} - x^*\|_2 \leq Ck^q,$$

где $q < 0$ и $0 < C < \infty$. Интуитивно, сублинейная сходимость означает, что последовательность сходится медленнее любой геометрической прогрессии.

Сверхлинейная сходимость

Сходимость последовательности $\{r_k\}_{k=m}^{\infty}$ называется **сверхлинейной**, если она сходится к нулю быстрее любой линейно сходящейся последовательности. Проверьте, что последовательность $\{r_k\}_{k=m}^{\infty}$ является сверхлинейной, если она сходится линейно с параметром $q = 0$.

Для $p > 1$, последовательность имеет **сверхлинейную сходимость порядка** p, если существует $C > 0$ и $0 < q < 1$ такая, что:

$$r_k \leq Cq^{p^k}, \quad \text{for all } k \geq m.$$

Когда $p = 2$, это называется **квадратичной сходимостью**.

Сверхлинейная сходимость

Сходимость последовательности $\{r_k\}_{k=m}^{\infty}$ называется **сверхлинейной**, если она сходится к нулю быстрее любой линейно сходящейся последовательности. Проверьте, что последовательность $\{r_k\}_{k=m}^{\infty}$ является сверхлинейной, если она сходится линейно с параметром $q = 0$.

Для $p > 1$, последовательность имеет **сверхлинейную сходимость порядка** p, если существует $C > 0$ и $0 < q < 1$ такая, что:

$$r_k \leq C q^{p^k}, \quad \text{for all } k \geq m.$$

Когда $p = 2$, это называется **квадратичной сходимостью**.

Важный пример

Предположим, что $x^* = 1.23456789$ (истинное решение), и итерационная последовательность начинается с ошибки $r_k = 10^{-3}$, соответствующей 3 правильным значащим цифрам (1.234).

1. После первой итерации:

$$r_{k+1} \approx r_k^2 = (10^{-3})^2 = 10^{-6}.$$

   Теперь ошибка равна 10^{-6}, и мы имеем 6 правильных значащих цифр (1.23456).

Сверхлинейная сходимость

Сходимость последовательности $\{r_k\}_{k=m}^\infty$ называется **сверхлинейной**, если она сходится к нулю быстрее любой линейно сходящейся последовательности. Проверьте, что последовательность $\{r_k\}_{k=m}^\infty$ является сверхлинейной, если она сходится линейно с параметром $q = 0$.

Для $p > 1$, последовательность имеет **сверхлинейную сходимость порядка** p, если существует $C > 0$ и $0 < q < 1$ такая, что:

$$r_k \leq C q^{p^k}, \quad \text{for all } k \geq m.$$

Когда $p = 2$, это называется **квадратичной сходимостью**.

Важный пример

Предположим, что $x^* = 1.23456789$ (истинное решение), и итерационная последовательность начинается с ошибки $r_k = 10^{-3}$, соответствующей 3 правильным значащим цифрам (1.234).

1. После первой итерации:
$$r_{k+1} \approx r_k^2 = (10^{-3})^2 = 10^{-6}.$$
Теперь ошибка равна 10^{-6}, и мы имеем 6 правильных значащих цифр (1.23456).
2. После второй итерации:
$$r_{k+2} \approx r_{k+1}^2 = (10^{-6})^2 = 10^{-12}.$$
Теперь ошибка равна 10^{-12}, и мы имеем 12 правильных значащих цифр (1.234567890123).

Практические наблюдения о скоростях сходимости

- $\|x_{k+1} - x^*\|_2 \leq \frac{1}{k^{\frac{1}{p}}} \|x_0 - x^*\|_2$ означает сублинейную скорость сходимости

Практические наблюдения о скоростях сходимости

- $\|x_{k+1} - x^*\|_2 \leq \frac{1}{k^{\frac{1}{p}}}\|x_0 - x^*\|_2$ означает сублинейную скорость сходимости
- $\|x_{k+1} - x^*\|_2 \leq q\|x_k - x^*\|_2$ означает линейную скорость сходимости, где $q < 1$

Практические наблюдения о скоростях сходимости

- $\|x_{k+1} - x^*\|_2 \leq \frac{1}{k^{\frac{1}{p}}}\|x_0 - x^*\|_2$ означает сублинейную скорость сходимости
- $\|x_{k+1} - x^*\|_2 \leq q\|x_k - x^*\|_2$ означает линейную скорость сходимости, где $q < 1$
- $\|x_{k+1} - x^*\|_2 \leq q\|x_k - x^*\|_2^2$ означает квадратичную скорость сходимости, где $q\|x_0 - x^*\| < 1$

Тест корней

Theorem

Пусть $(r_k)_{k=m}^{\infty}$ - последовательность неотрицательных чисел, сходящаяся к нулю, и пусть $\alpha := \limsup_{k\to\infty} r_k^{1/k}$. (Заметим, что $\alpha \geq 0$.)
(a) Если $0 \leq \alpha < 1$, то $(r_k)_{k=m}^{\infty}$ сходится линейно с константой α.

Доказательство.

Тест корней

Theorem

Пусть $(r_k)_{k=m}^{\infty}$ - последовательность неотрицательных чисел, сходящаяся к нулю, и пусть $\alpha := \limsup_{k\to\infty} r_k^{1/k}$. (Заметим, что $\alpha \geq 0$.)
(a) Если $0 \leq \alpha < 1$, то $(r_k)_{k=m}^{\infty}$ сходится линейно с константой α.
(b) В частности, если $\alpha = 0$, то $(r_k)_{k=m}^{\infty}$ сходится сверхлинейно.

Доказательство.

Тест корней

Theorem

Пусть $(r_k)_{k=m}^{\infty}$ - последовательность неотрицательных чисел, сходящаяся к нулю, и пусть $\alpha := \limsup_{k\to\infty} r_k^{1/k}$. (Заметим, что $\alpha \geq 0$.)
(a) Если $0 \leq \alpha < 1$, то $(r_k)_{k=m}^{\infty}$ сходится линейно с константой α.
(b) В частности, если $\alpha = 0$, то $(r_k)_{k=m}^{\infty}$ сходится сверхлинейно.
(c) Если $\alpha = 1$, то $(r_k)_{k=m}^{\infty}$ сходится сублинейно.

Доказательство.

Тест корней

Theorem

Пусть $(r_k)_{k=m}^{\infty}$ - последовательность неотрицательных чисел, сходящаяся к нулю, и пусть $\alpha := \limsup_{k\to\infty} r_k^{1/k}$. (Заметим, что $\alpha \geq 0$.)

(a) Если $0 \leq \alpha < 1$, то $(r_k)_{k=m}^{\infty}$ сходится линейно с константой α.

(b) В частности, если $\alpha = 0$, то $(r_k)_{k=m}^{\infty}$ сходится сверхлинейно.

(c) Если $\alpha = 1$, то $(r_k)_{k=m}^{\infty}$ сходится сублинейно.

(d) Случай $\alpha > 1$ невозможен.

Доказательство.

Тест корней

Theorem

Пусть $(r_k)_{k=m}^{\infty}$ - последовательность неотрицательных чисел, сходящаяся к нулю, и пусть $\alpha := \limsup_{k\to\infty} r_k^{1/k}$. (Заметим, что $\alpha \geq 0$.)

(a) Если $0 \leq \alpha < 1$, то $(r_k)_{k=m}^{\infty}$ сходится линейно с константой α.

(b) В частности, если $\alpha = 0$, то $(r_k)_{k=m}^{\infty}$ сходится сверхлинейно.

(c) Если $\alpha = 1$, то $(r_k)_{k=m}^{\infty}$ сходится сублинейно.

(d) Случай $\alpha > 1$ невозможен.

Доказательство.

1. Покажем, что если $(r_k)_{k=m}^{\infty}$ сходится линейно с константой $0 \leq \beta < 1$, то $\alpha \leq \beta$. Действительно, по определению константы линейной сходимости, для любого $\varepsilon > 0$ такого, что $\beta + \varepsilon < 1$, существует $C > 0$ такое, что $r_k \leq C(\beta+\varepsilon)^k$ для всех $k \geq m$. Отсюда, $r_k^{1/k} \leq C^{1/k}(\beta+\varepsilon)$ для всех $k \geq m$. Переходя к пределу при $k \to \infty$ и используя $C^{1/k} \to 1$, мы получаем $\alpha \leq \beta + \varepsilon$. Учитывая произвольность ε, получаем $\alpha \leq \beta$.

Тест корней

Theorem

Пусть $(r_k)_{k=m}^{\infty}$ - последовательность неотрицательных чисел, сходящаяся к нулю, и пусть $\alpha := \limsup_{k\to\infty} r_k^{1/k}$. (Заметим, что $\alpha \geq 0$.)

(a) Если $0 \leq \alpha < 1$, то $(r_k)_{k=m}^{\infty}$ сходится линейно с константой α.

(b) В частности, если $\alpha = 0$, то $(r_k)_{k=m}^{\infty}$ сходится сверхлинейно.

(c) Если $\alpha = 1$, то $(r_k)_{k=m}^{\infty}$ сходится сублинейно.

(d) Случай $\alpha > 1$ невозможен.

Доказательство.

1. Покажем, что если $(r_k)_{k=m}^{\infty}$ сходится линейно с константой $0 \leq \beta < 1$, то $\alpha \leq \beta$. Действительно, по определению константы линейной сходимости, для любого $\varepsilon > 0$ такого, что $\beta + \varepsilon < 1$, существует $C > 0$ такое, что $r_k \leq C(\beta + \varepsilon)^k$ для всех $k \geq m$. Отсюда, $r_k^{1/k} \leq C^{1/k}(\beta + \varepsilon)$ для всех $k \geq m$. Переходя к пределу при $k \to \infty$ и используя $C^{1/k} \to 1$, мы получаем $\alpha \leq \beta + \varepsilon$. Учитывая произвольность ε, получаем $\alpha \leq \beta$.
2. Таким образом, в случае $\alpha = 1$ последовательность $(r_k)_{k=m}^{\infty}$ не может иметь линейной сходимости в соответствии с приведенным выше результатом (доказано от противного). Тем не менее, $(r_k)_{k=m}^{\infty}$ сходится к нулю, поэтому она должна сходиться сублинейно.

Тест корней

Theorem

1. Теперь рассмотрим случай $0 \le \alpha < 1$. Пусть $\varepsilon > 0$ - произвольное число такое, что $\alpha + \varepsilon < 1$. Согласно свойствам limsup, существует $N \ge m$ такое, что $r_k^{1/k} \le \alpha + \varepsilon$ для всех $k \ge N$. Отсюда, $r_k \le (\alpha + \varepsilon)^k$ для всех $k \ge N$. Следовательно, $(r_k)_{k=m}^{\infty}$ сходится линейно с параметром $\alpha + \varepsilon$ (не имеет значения, что неравенство выполняется только для числа N). Учитывая произвольность ε, это означает, что константа линейной сходимости $(r_k)_{k=m}^{\infty}$ не превышает α. Поскольку, как показано выше, константа линейной сходимости не может быть меньше α, это означает, что константа линейной сходимости $(r_k)_{k=m}^{\infty}$ точно равна α.

Тест корней

ℹ Theorem

1. Теперь рассмотрим случай $0 \leq \alpha < 1$. Пусть $\varepsilon > 0$ - произвольное число такое, что $\alpha + \varepsilon < 1$. Согласно свойствам limsup, существует $N \geq m$ такое, что $r_k^{1/k} \leq \alpha + \varepsilon$ для всех $k \geq N$. Отсюда, $r_k \leq (\alpha + \varepsilon)^k$ для всех $k \geq N$. Следовательно, $(r_k)_{k=m}^{\infty}$ сходится линейно с параметром $\alpha + \varepsilon$ (не имеет значения, что неравенство выполняется только для числа N). Учитывая произвольность ε, это означает, что константа линейной сходимости $(r_k)_{k=m}^{\infty}$ не превышает α. Поскольку, как показано выше, константа линейной сходимости не может быть меньше α, это означает, что константа линейной сходимости $(r_k)_{k=m}^{\infty}$ точно равна α.
2. Наконец, покажем, что случай $\alpha > 1$ невозможен. Действительно, предположим, что $\alpha > 1$. Тогда из определения limsup следует, что для любого $N \geq m$ существует $k \geq N$ такое, что $r_k^{1/k} \geq 1$, и, в частности, $r_k \geq 1$. Но это означает, что r_k имеет подпоследовательность, которая не ограничена от нуля. Следовательно, $(r_k)_{k=m}^{\infty}$ не может сходиться к нулю, что противоречит условию.

Тест отношений

Пусть $\{r_k\}_{k=m}^{\infty}$ - последовательность строго положительных чисел, сходящаяся к нулю. Пусть

$$q = \lim_{k \to \infty} \frac{r_{k+1}}{r_k}$$

- Если существует q и $0 \leq q < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой q.

Тест отношений

Пусть $\{r_k\}_{k=m}^{\infty}$ - последовательность строго положительных чисел, сходящаяся к нулю. Пусть

$$q = \lim_{k \to \infty} \frac{r_{k+1}}{r_k}$$

- Если существует q и $0 \le q < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой q.
- В частности, если $q = 0$, то $\{r_k\}_{k=m}^{\infty}$ имеет сверхлинейную сходимость.

Тест отношений

Пусть $\{r_k\}_{k=m}^{\infty}$ - последовательность строго положительных чисел, сходящаяся к нулю. Пусть

$$q = \lim_{k \to \infty} \frac{r_{k+1}}{r_k}$$

- Если существует q и $0 \leq q < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой q.
- В частности, если $q = 0$, то $\{r_k\}_{k=m}^{\infty}$ имеет сверхлинейную сходимость.
- Если q не существует, но $q = \lim\limits_{k \to \infty} \sup_k \dfrac{r_{k+1}}{r_k} < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой, не превышающей q.

Тест отношений

Пусть $\{r_k\}_{k=m}^{\infty}$ - последовательность строго положительных чисел, сходящаяся к нулю. Пусть

$$q = \lim_{k \to \infty} \frac{r_{k+1}}{r_k}$$

- Если существует q и $0 \leq q < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой q.
- В частности, если $q = 0$, то $\{r_k\}_{k=m}^{\infty}$ имеет сверхлинейную сходимость.
- Если q не существует, но $q = \lim_{k \to \infty} \sup_k \frac{r_{k+1}}{r_k} < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой, не превышающей q.
- Если $\lim_{k \to \infty} \inf_k \frac{r_{k+1}}{r_k} = 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет сублинейную сходимость.

Тест отношений

Пусть $\{r_k\}_{k=m}^{\infty}$ - последовательность строго положительных чисел, сходящаяся к нулю. Пусть

$$q = \lim_{k \to \infty} \frac{r_{k+1}}{r_k}$$

- Если существует q и $0 \le q < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой q.
- В частности, если $q = 0$, то $\{r_k\}_{k=m}^{\infty}$ имеет сверхлинейную сходимость.
- Если q не существует, но $q = \lim_{k \to \infty} \sup_k \frac{r_{k+1}}{r_k} < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой, не превышающей q.
- Если $\lim_{k \to \infty} \inf_k \frac{r_{k+1}}{r_k} = 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет сублинейную сходимость.
- Случай $\lim_{k \to \infty} \inf_k \frac{r_{k+1}}{r_k} > 1$ невозможен.

Тест отношений

Пусть $\{r_k\}_{k=m}^{\infty}$ - последовательность строго положительных чисел, сходящаяся к нулю. Пусть

$$q = \lim_{k \to \infty} \frac{r_{k+1}}{r_k}$$

- Если существует q и $0 \leq q < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой q.
- В частности, если $q = 0$, то $\{r_k\}_{k=m}^{\infty}$ имеет сверхлинейную сходимость.
- Если q не существует, но $q = \lim_{k \to \infty} \sup_k \frac{r_{k+1}}{r_k} < 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет линейную сходимость с константой, не превышающей q.
- Если $\lim_{k \to \infty} \inf_k \frac{r_{k+1}}{r_k} = 1$, то $\{r_k\}_{k=m}^{\infty}$ имеет сублинейную сходимость.
- Случай $\lim_{k \to \infty} \inf_k \frac{r_{k+1}}{r_k} > 1$ невозможен.
- В остальных случаях (т.е., когда $\lim_{k \to \infty} \inf_k \frac{r_{k+1}}{r_k} < 1 \leq \lim_{k \to \infty} \sup_k \frac{r_{k+1}}{r_k}$) мы не можем сделать никаких конкретных утверждений о скорости сходимости $\{r_k\}_{k=m}^{\infty}$.

Лемма о тесте отношений

Theorem

Пусть $(r_k)_{k=m}^{\infty}$ - последовательность строго положительных чисел. (Строгая положительность необходима для того, чтобы отношения $\frac{r_{k+1}}{r_k}$, которые появляются ниже, были определены.) Тогда

$$\liminf_{k\to\infty} \frac{r_{k+1}}{r_k} \leq \liminf_{k\to\infty} r_k^{1/k} \leq \limsup_{k\to\infty} r_k^{1/k} \leq \limsup_{k\to\infty} \frac{r_{k+1}}{r_k}.$$

Доказательство.

1. Среднее неравенство следует из того, что liminf любой последовательности всегда меньше или равен её limsup. Докажем последнее неравенство; первое доказывается аналогично.

Лемма о тесте отношений

Theorem

Пусть $(r_k)_{k=m}^{\infty}$ - последовательность строго положительных чисел. (Строгая положительность необходима для того, чтобы отношения $\frac{r_{k+1}}{r_k}$, которые появляются ниже, были определены.) Тогда

$$\liminf_{k\to\infty} \frac{r_{k+1}}{r_k} \leq \liminf_{k\to\infty} r_k^{1/k} \leq \limsup_{k\to\infty} r_k^{1/k} \leq \limsup_{k\to\infty} \frac{r_{k+1}}{r_k}.$$

Доказательство.

1. Среднее неравенство следует из того, что liminf любой последовательности всегда меньше или равен её limsup. Докажем последнее неравенство; первое доказывается аналогично.
2. Обозначим $L := \limsup_{k\to\infty} \frac{r_{k+1}}{r_k}$. Если $L = +\infty$, то неравенство очевидно, поэтому предположим, что L конечно. Заметим, что $L \geq 0$, поскольку отношение $\frac{r_{k+1}}{r_k}$ положительно для всех $k \geq m$. Пусть $\varepsilon > 0$ - произвольное число. Согласно свойствам limsup, существует $N \geq m$ такое, что $\frac{r_{k+1}}{r_k} \leq L + \varepsilon$ для всех $k \geq N$. Отсюда, $r_{k+1} \leq (L+\varepsilon)r_k$ для всех $k \geq N$. Применяя индукцию, получаем $r_k \leq (L+\varepsilon)^{k-N} r_N$ для всех $k \geq N$. Пусть $C := (L+\varepsilon)^{-N} r_N$. Тогда $r_k \leq C(L+\varepsilon)^k$ для всех $k \geq N$, откуда $r_k^{1/k} \leq C^{1/k}(L+\varepsilon)$. Переходя к limsup при $k \to \infty$ и используя $C^{1/k} \to 1$, получаем $\limsup_{k\to\infty} r_k^{1/k} \leq L + \varepsilon$. Учитывая произвольность ε, получаем $\limsup_{k\to\infty} r_k^{1/k} \leq L$.

Итоги

Итоги

Определения

1. Положительно определённая матрица.
2. Евклидова норма вектора.
3. Неравенство треугольника для нормы.
4. p-норма вектора.
5. Как выглядит единичный шар в p - норме на плоскости для $p = 1, 2, \infty$?
6. Норма Фробениуса для матрицы.
7. Спектральная норма матрицы.
8. Скалярное произведение двух векторов.
9. Скалярное произведение двух матриц, согласованное с нормой Фробениуса.
10. Собственные значения матрицы. Спектр матрицы.
11. Связь спектра матрицы и её определенности.
12. Спектральное разложение матрицы.
13. Сингулярное разложение матрицы.
14. Связь определителя и собственных чисел для квадратной матрицы.
15. Связь следа и собственных чисел для квадратной матрицы.
16. Линейная сходимость последовательности.
17. Сублинейная сходимость последовательности.
18. Сверхлинейная сходимость последовательности.
19. Квадратичная сходимость последовательности.
20. Тест корней для определения скорости сходимости последовательности.
21. Тест отношений для определения скорости сходимости последовательности.

Теоремы

1. Критерий положительной определенности матрицы через знаки собственных значений матрицы.
2. Тест корней
3. Тест отношений