Основы матричных вычислений. ПМИ. Семинар 1.
Нормы. Матричные нормы. Унитарные матрицы.

Весна 2025. Н.Медведь по материалам М.Рахубы, Л.Высоцкого и др.

Этот и дальнейшие конспекты не претендуют на полноту изложения, чтобы их можно было читать независимо от самого семинара. Однако, автор надеется, что его должно быть более чем достаточно, чтобы восстановить, что происходило на семинаре, поработав над некоторыми деталями самостоятельно. Количество задач, как правило, будет с некоторым запасом; не факт, что все задачи будут в итоге на семинаре разобраны.

1 Задачи для семинара

Обсуждение 1 (Обсуждение p-норм)**.** Понятие близости объектов друг к другу, к нулю – полезно, например матанализ. Метрика. В линейных пространствах норма.

Для матриц можно, вообще говоря, рассматривать их как вектор с n^2 компонентами и рассматривать векторные нормы. Так, например, определяется норма Фробениуса

$$\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$$

и норма Чебышёва

$$\|A\|_C = \max_{i,j} |a_{ij}|.$$

Но если матрица используется не просто как массив, а именно как матрица, то мы скорее всего планируем её на что-то умножать. Хотелось бы иметь норму, хорошо взаимодействующую с операцией умножения (субмультипликативную: $\|AB\| \leqslant \|A\| \cdot \|B\|$). Так удачно вышло, что фробениусова норма этому свойству, на самом деле, удовлетворяет, но это довольно неочевидно. Более естественный класс норм для этого свойства – это операторные нормы (напоминаю определение; даю частный случай только когда одна и та же норма, хотя можно сказать, что на лекции говорили про общий случай). Но определение несколько абстрактно. Хотелось бы иметь более явную конструкцию, как же такую норму для конкретной матрицы находить.

Задача 2 (Субмультипликативность нормы Фробениуса)**.** Давайте для удобства вторую матрицу обозначим не через B, а через B^T. Тогда:

$$\|AB^T\|_F = \|(a_1 a_2 \dots a_n)\begin{bmatrix} b_1^T \\ b_2^T \\ \vdots \\ b_n^T \end{bmatrix}\|_F = \|a_1 b_1^T + \ldots + a_n b_n^T\|_F \leqslant \|a_1 b_1^T\|_F + \ldots + \|a_n b_n^T\|_F =$$

$$= \sum_k \|a_k b_k^T\|_F = \sum_k \sqrt{\sum_{i,j} |a_{ki} b_{kj}|^2} = \sum_k \sqrt{\sum_i |a_{ki}|^2 \cdot \sum_j |b_{kj}|^2} = \sum_k \|a_k\|_2 \|b_k\|_2 =$$

$$= \begin{bmatrix} \|a_1\|_2 & \ldots & \|a_n\|_2 \end{bmatrix} \begin{bmatrix} \|b_1\|_2 \\ \vdots \\ \|b_n\|_2 \end{bmatrix} \leqslant \left\| \begin{bmatrix} \|a_1\|_2 & \ldots & \|a_n\|_2 \end{bmatrix} \right\|_2 \left\| \begin{bmatrix} \|b_1\|_2 \\ \vdots \\ \|b_n\|_2 \end{bmatrix} \right\|_2 = \|A\|_F \|B\|_F.$$

Изящнее равенство $\|a_i b_i^T\|_F = \|a_i\|_2 \|b_i\|_2$ можно записать бескоординатно:

$$\|a_i b_i^T\|_F = \sqrt{\operatorname{Tr}((a_i b_i^T)^*(a_i b_i^T))} = \sqrt{\operatorname{Tr}(b_i^T)^*(a_i^* a_i) b_i^T} = \sqrt{(a_i^* a_i)\operatorname{Tr}(b_i^T)^* b_i^T} = \|a_i\|_2 \|b_i\|_2.$$

Замечание: верно более сильное неравенство $\|AB\|_F \leqslant \|A\|_2 \|B\|_F$. Его можно доказать при помощи SVD-разложения и унитарной инвариантности этих норм. Если вы захотите его использовать в каких-либо задачах, его нужно будет доказать.

Упражнение 3 (Несубмультипликативность чебышевской нормы)**.** Простейший пример:

$$A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Тогда $\|A\|_C = \|B\|_C = 1, \|AB\|_C = 2$.

Преимущества чебышевской нормы: простая для вычисления, понятная интуитивно, в том числе в контексте сходимости (поэлементная сходимость).

Задача 4 (Несубмультипликативность векторной p-нормы, применённой к матрицам)**.** Можно рассмотреть матрицу, как вектор с n^2 компонентами, и для него рассмотреть ту или иную векторную норму. Интуитивно кажется, что она не имеет никакого отношения к умножению матриц, но мы видели, что для частного случая $p = 2$ (фробениусова норма) это не так. В этой задаче мы рассмотрим случай $p > 2$.

Рассмотрим пример матрицы $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Её p-норма равна $(1^p+1^p+1^p+1^p)^{1/p} = 4^{1/p}$. Произведение двух таких норм равно $4^{2/p}$. А для квадрата норма равна $(2^p+2^p+2^p+2^p)^{1/p} = (4\cdot 2^p)^{1/p} = 4^{1/p}\cdot 2$. Для субмультипликативности на этом конкретном примере необходимо неравенство $4^{2/p} \leqslant 2 \cdot 4^{1/p}$, что равносильно $4^{1/p} \leqslant 2$, что возможно только при $p \leqslant 2$. Таким образом, рассмотренный пример показывает, что при $p > 2$ векторная норма не субмультипликативна.

Упражнение: обобщить пример с двумерного случая на произвольную размерность.

Задача 5 (Субмультипликативность операторной p-нормы)**.**

$$\|AB\|_p = \sup_{x \neq 0} \frac{\|ABx\|_p}{\|x\|_p}$$

Пусть супремум достигается на некотором векторе x_0 (он обязательно достигается по соображениям, обсуждавшимся на лекции: супремум по всему пространству кроме нуля можно заменить на супремум по единичной сфере, а она компактна; по теореме Вейерштрасса непрерывная функция на компакте достигает максимума).

Заметим, что если $\|Bx_0\|_p = 0$, то $Bx_0 = 0$, тогда $ABx_0 = 0$ и это означает, что $\|AB\|_p = 0$, откуда неравенство $\|AB\|_p \leqslant \|A\|_p \|B\|_p$ очевидно. Остаётся рассмотреть случай $Bx_0 \neq 0$.

$$\|AB\|_p = \frac{\|ABx_0\|_p}{\|x_0\|_p} = \frac{\|ABx_0\|_p}{\|Bx_0\|_p}\frac{\|Bx_0\|_p}{\|x_0\|_p} \leqslant \sup_{Bx \neq 0} \frac{\|ABx\|_p}{\|Bx\|_p} \sup_{x \neq 0} \frac{\|Bx\|_p}{\|x\|_p} \leqslant \sup_{y \neq 0} \frac{\|Ay\|_p}{\|y\|_p} \sup_{x \neq 0} \frac{\|Bx\|_p}{\|x\|_p} = \|A\|_p \|B\|_p$$

Задача 6 (Формула для матричной 1-нормы)**.** На самом деле для 1-нормы загадочный супремум равен просто максимальной сумме (модулей) по столбцу: $\|A\|_1 = \max\limits_j \sum\limits_i |a_{ij}|$. Действительно, если столбцы матрицы A обозначить через a_i, имеем:

$$\frac{\|Ax\|_1}{\|x\|_1} = \frac{\|a_1 x_1 + \ldots + a_n x_n\|_1}{\|x\|_1} \leqslant \frac{\|a_1 x_1\|_1 + \ldots + \|a_n x_n\|_1}{|x_1| + \ldots + |x_n|} = \frac{\|a_1\|_1 |x_1| + \ldots + \|a_n\|_1 |x_n|}{|x_1| + \ldots + |x_n|} \leqslant$$

$$\leqslant \frac{\max\limits_j \|a_j\|_1 |x_1| + \ldots + \max\limits_j \|a_j\|_1 |x_n|}{|x_1| + \ldots + |x_n|} = \max_j \|a_j\|_1 \frac{|x_1| + \ldots + |x_n|}{|x_1| + \ldots + |x_n|} = \max_j \|a_j\|_1 = \max_j \sum_i |a_{ij}|.$$

С другой стороны, легко понять, что для любой A при некотором x оба неравенства в равенство обратиться могут, причём даже одновременно. Действительно, второе может обратиться в равенство, если среди x_i ненулевая компонента только одна, причём именно при максимальном столбце. Тогда именно такой x и рассмотрим. Легко видеть, что первое неравенство превращается в неравенство треугольника для 1 слагаемого и тоже выполнено.

Обсуждение 7 (Формула для матричной ∞-нормы)**.** Для ∞-нормы верна аналогичная формула $\|A\|_\infty = \max\limits_i \sum\limits_j |a_{ij}|$. Мы не будем разбирать доказательство, но оно в целом аналогично предыдущему.

Задача 8 (Разложение Шура)**.** Рассмотрим разложение Шура на примере $\begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix}$. Собственные числа $1, 2$; собственные векторы $(1, -1)^T; (-1, 2)^T$ соответственно. Так как собственные числа разные, заведомо диагонализуема, вот жорданово разложение: $A = SJS^{-1} = S\Lambda S^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} S^{-1}$. Это разложение хорошо тем, что матрица диагональная вышла, а плохо тем, что S не унитарна (не ортогональна).

Возьмём только один собственный вектор и дополним его до ортогонального базиса, проделав один шаг разложения Шура (в двумерном случае всё завершится за один шаг). Мы получаем матрицу $U_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$.

Тогда $A_1 = U^*AU = \dfrac{1}{\sqrt{5}} \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix} \dfrac{1}{\sqrt{5}} \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix} = \dfrac{1}{5} \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -2 & -1 \\ 4 & 7 \end{bmatrix} = \dfrac{1}{5} \begin{bmatrix} 10 & 15 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} = T$. Мы получили с первого же шага треугольную из-за двумерности (в 2 словах напоминаю, какая блочная вышла бы).