1. ВЕКТОРНЫЕ И МАТРИЧНЫЕ НОРМЫ

1.1. Предварительные замечания

Векторные нормы. Пусть вещественному или комплексному n-вектору $\boldsymbol{x} = (x_1, \ldots, x_n)^T$ (T — операция транспонирования) поставлено в соответствие вещественное число $\|\boldsymbol{x}\|$, такое, что выполнены следующие аксиомы:

$$\|\boldsymbol{x}\| > 0\,, \quad \text{если} \quad \boldsymbol{x} \neq 0\,, \quad \|0\| = 0,$$
$$\|\alpha\boldsymbol{x}\| = |\alpha|\,\|\boldsymbol{x}\|, \quad \text{(аксиома абсолютной однородности)}$$
$$\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \quad \text{(аксиома (неравенство) треугольника)}$$

для любого числа α и любого n-вектора $\|\boldsymbol{y}\|$. Тогда число $\|\boldsymbol{x}\|$ называется *нормой* вектора $\|\boldsymbol{x}\|$.

Наиболее употребительными на практике являются следующие векторные нормы:

— абсолютная норма (норма l_1, или 1-норма)

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|;$$

— евклидова норма (норма l_2, или 2-норма)

$$\|\boldsymbol{x}\|_2 = |(\boldsymbol{x},\boldsymbol{x})| = \left|\boldsymbol{x}^T\boldsymbol{x}\right| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2},$$

где через (,) обозначено скалярное произведение векторов;

— максимальная норма (норма l_∞, или ∞-норма)

$$\|\boldsymbol{x}\|_\infty = \max_{1\leq i\leq n} |x_i|.$$

Эти нормы являются частными случаями *нормы Гельдера* (или нормы l_p) с показателем p, $p \geq 1$:

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

Две векторные нормы $\|\cdot\|_{\mathrm{I}}$ и $\|\cdot\|_{\mathrm{II}}$ называются *эквивалентными*, если существуют такие положительные вещественные числа c_1 и c_2, что для любого вектора $\boldsymbol{x}$ выполняются неравенства (*соотношения эквивалентности*)

$$c_1\,\|\boldsymbol{x}\|_{\mathrm{I}} \leq \|\boldsymbol{x}\|_{\mathrm{II}} \leq c_2\,\|\boldsymbol{x}\|_{\mathrm{I}}$$

причем c_1 и c_2 не зависят от выбора $\boldsymbol{x}$. Рассмотренные выше нормы эквивалентны.

Собственные значения матриц. Число λ называется *собственным значением* квадратной матрицы A, если существует такой ненулевой вектор $\boldsymbol{x}$, что

$$A\boldsymbol{x} = \lambda\boldsymbol{x}.$$

Любой вектор $\boldsymbol{x} \neq \mathbf{0}$, удовлетворяющий этому уравнению, называется *собственным вектором* матрицы A, соответствующим собственному значению λ. Совокупность всех собственных значений называется *спектром* матрицы, а максимальный из модулей ее собственных значений называется *спектральным радиусом*. Собственные значения λ (в общем случае комплексные) являются нулями *характеристического многочлена* $\det(A-\lambda\,E)$, т.е. нулями (корнями) *векового уравнения* $\det(A-\lambda\,E)=0$. Собственные значения матриц A и A^T совпадают. Если матрица A невырожденная и λ — любое ее собственное значение, то $\lambda\neq 0$ и λ^{-1} является собственным значением матрицы A^{-1}. Если λ — собственное значение матрицы A, то число $\lambda-\mu$ является собственным значением матрицы $A-\mu, E$, где E — единичная матрица. Число λ^2 — собственное значение матрицы A^2.

Ортогональные матрицы. Вещественная $(n\times n)$-матрица Q называется *ортогональной*, если $Q\,Q^T=E$. Матрица Q^T также ортогональная. Кроме того, $\det(Q)=\pm 1$ и $Q^{-1}=Q^T$. Произведение ортогональных матриц есть ортогональная матрица.

Симметричные положительно определенные матрицы. Это такие матрицы A, для которых квадратичная форма $\boldsymbol{x}^T A\,\boldsymbol{x}=(A\,\boldsymbol{x},\boldsymbol{x})>0$ для всех ненулевых векторов $\boldsymbol{x}$. Для этих матриц принято обозначение $A=A^T>0$. Их диагональные элементы и собственные значения положительны. Если B — прямоугольная матрица с линейно-независимыми столбцами (строками), то матрица $A=B^T B$ является симметричной положительно определенной; если у B столбцы (строки) линейно-зависимы, то $A=B^T B$ — симметричная положительно полуопределенная матрица.

Сингулярное разложение матриц. Для любой вещественной $(n\times m)$-матрица ранга r существуют ортогональная $(n\times n)$-матрица U и ортогональная $(m\times m)$-матрица V, такие, что

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

где Σ — прямоугольная диагональная $(n\times m)$-матрица с невозрастающими неотрицательными элементами σ_i $(i=1,\ldots,m)$ на диагонали. Это разложение матрицы A называется *сингулярным*, а элементы σ_i — *сингулярными числами*, причем

$$\sigma_1\geq\sigma_2\geq\ldots\geq\sigma_r>0 \quad \text{и} \quad \sigma_{r+1}=_m=0.$$

Квадраты ненулевых сингулярных чисел матрицы A совпадают с ненулевыми собственными значениями матриц $A^T A$ и $A\,A^T$, а количество ненулевых сингулярных чисел равно рангу матрицы A. Сингулярное разложение единственно с точностью до матриц U и V.

Матричные нормы. Пусть вещественной или комплексной $(n\times m)$-матрице A поставлено в соответствие некоторое вещественное число $\|A\|$, такое, что выполнены следующие аксиомы:

$$\begin{aligned}
&\|A\|>0, \quad \text{если} \quad A\neq 0\,, \quad \|0\|=0,\\
&\|\alpha\,A\|=|\alpha|\,\|A\|,\\
&\|A+B\|\leq\|A\|+\|B\|,\\
&\|A\,C\|\leq\|A\|\,\|C\| \quad \text{(аксиома мультипликативности)}
\end{aligned}$$

для любого числа α и любых матриц B и C, для которых соответствующие операции имеют смысл. Тогда число $\|A\|$ называется *нормой* матрицы A.

Если $\|A\|$ удовлетворяет только первым трем аксиомам, то такую норму называют аддитивной (обобщенной) матричной нормой. Если удовлетворяются все четыре аксиомы, то такая норма матрицы называется мультипликативной. Всякую аддитивную норму умножением на достаточно большую положительную константу можно превратить в мультипликативную. Везде, где не оговорено противное, под матричной нормой будем понимать мультипликативную матричную норму.

Наиболее употребительными на практике являются следующие матричные нормы:

— первая норма (столбцовая норма, или 1-норма)

$$\|A\|_1 = \max_{1\le j\le m} \sum_{i=1}^{n} |a_{ij}|;$$

— спектральная норма (вторая норма, или норма Гильберта)

$$\|A\|_2 = \max_{1\le i\le m} \sigma_i = \left(\max_{1\le i\le m} \lambda_i(A^TA)\right)^{1/2} = \left(\max_{1\le i\le n} \lambda_i(AA^T)\right)^{1/2},$$

где σ_i — сингулярные числа матрицы A, $\lambda_i(A^TA)$ — собственные значения матрицы A^TA и $\lambda_i(AA^T)$ — собственные значения матрицы AA^T;

— бесконечная норма (строчная норма, или ∞-норма)

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|;$$

— сферическая норма (норма Фробениуса, или евклидова норма)

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

Выписанные матричные нормы *эквивалентны*.

Матричная норма $\|\cdot\|_M$ называется *согласованной* с векторными нормами $\|\cdot\|_V$, если

$$\|A\boldsymbol{x}\|_V \le \|A\|_M \|\boldsymbol{x}\|_V$$

для любой матрицы A и всех векторов $\boldsymbol{x}$. Всякая норма матрицы согласована с какими-нибудь векторными нормами.

Пусть задана векторными норма $\|\cdot\|_V$. Тогда числовая функция

$$\|A\| = \sup_{\|\boldsymbol{x}\|_V\ne 0} \frac{\|A\,\boldsymbol{x}\|_V}{\|\boldsymbol{x}\|_V} = \sup_{\|\boldsymbol{x}\|_V=1} \|A\,\boldsymbol{x}\|_V$$

является матричной нормой и называется нормой матрицы, *подчиненной* векторной норме $\|\cdot\|_V$.

Среди всех матричных норм, согласованных с заданной векторной нормой, подчиненная норма является *минимальной* в том смысле, что в неравенстве $\|A\boldsymbol{x}\|_V \le \|A\|_M\|\boldsymbol{x}\|_V$ число $\|A\|_M$ нельзя уменьшить. Спектральная, 1- и ∞-нормы матриц являются подчиненными по отношению, соответственно, к евклидовой, 1- и ∞-нормам векторов, а значит и согласованными с ними.

Ортогонально подобные матрицы. Матрица A называется ортогонально подобной матрице B, если существует такая ортогональная матрица Q, что $A = Q^TBQ$. Ортогонально подобные матрицы имеют одинаковые собственные значения.

Задачи и решения

1. Найти константы эквивалентности, связывающие нормы $\|\boldsymbol{x}\|_\infty$, $\|\boldsymbol{x}\|_1$, $\|\boldsymbol{x}\|_2$, а также векторы, на которых они достигаются.

Решение. Из неравенств $\max_{1\le i\le n}|x_i| \le \sum_{i=1}^n |x_i| \le n \max_{1\le i\le n}|x_i|$ следует

$$||\boldsymbol{x}||_\infty \le ||\boldsymbol{x}||_1 \le n\,||\boldsymbol{x}||_\infty\,.$$

Поскольку $\sum_{i=1}^n x_i^2 \le \left(\sum_{i=1}^n |x_i|\right)^2$, то $||\boldsymbol{x}||_2 \le ||\boldsymbol{x}||_1$. Из неравенства Коши–Буняковского имеем

$$\sum_{i=1}^n |x_i| \le \left(\sum_{i=1}^n 1\right)^{1/2} \left(\sum_{i=1}^n x_i^2\right)^{1/2} = n^{1/2}\left(\sum_{i=1}^n x_i^2\right)^{1/2}.$$

Следовательно,

$$n^{-1/2}||\boldsymbol{x}||_1 \le ||\boldsymbol{x}||_2 \le ||\boldsymbol{x}||_1\,.$$

Из неравенств $\max_{1\le i\le n} x_i^2 \le \sum_{i=1}^n x_i^2 \le n \max_{1\le i\le n} x_i^2$ следует

$$||\boldsymbol{x}||_\infty \le ||\boldsymbol{x}||_2 \le n^{1/2}\,||\boldsymbol{x}||_\infty\,.$$

2. Показать, что функционал из определения подчиненной нормы действительно является матричной нормой.

Решение. Первые два условия легко следуют из условий для векторной нормы. Покажем, что удовлетворяется третье условие. Действительно, в силу неравенства треугольника для векторной нормы имеем

$$||A+B|| = \sup_{||\boldsymbol{x}||=1} ||(A+B)\,\boldsymbol{x}|| \le \sup_{||\boldsymbol{x}||=1} ||A\,\boldsymbol{x}|| + \sup_{||\boldsymbol{x}||=1} ||B\,\boldsymbol{x}|| = ||A|| + ||B||\,.$$

Из определения подчиненной нормы следует, что

$$\frac{||A\,\boldsymbol{x}||}{||\boldsymbol{x}||} \le ||A|| \quad \text{для всех} \quad \boldsymbol{x} \ne 0.$$

Таким образом, справедливо неравенство

$$||AB|| = \sup_{||\boldsymbol{x}||=1} ||AB\,\boldsymbol{x}|| \le \sup_{||\boldsymbol{x}||=1} \big(||A||\,||B\,\boldsymbol{x}||\big) = ||A|| \sup_{||\boldsymbol{x}||=1} ||B\,\boldsymbol{x}|| = ||A||\,||B||\,,$$

из которого следует, что удовлетворяется и четвертое условие.

3. Найти матричные нормы, подчиненные векторным нормам $||\cdot||_\infty$ и $||\cdot||_2$.

Решение. Получим оценку сверху для величины $||A\boldsymbol{x}||_\infty$:

$$||A\boldsymbol{x}||_\infty = \max_i \left|\sum_j a_{ij}x_j\right| \le \max_i \left(\sum_j |a_{ij}| \max_j |x_j|\right) \le \max_i \left(\sum_j |a_{ij}|\right) ||\boldsymbol{x}||_\infty$$

Покажем, что эта оценка достигается. Пусть максимум по i имеет место при $i = l$; тогда возьмем $\boldsymbol{x} = \big(\operatorname{sign}(a_{l1}), \operatorname{sign}(a_{l2}), \ldots, \operatorname{sign}(a_{ln})\big)$. Имеем $||\boldsymbol{x}||_\infty = 1$ и точные равенства во всей цепочке выше. Таким образом, $||A||_\infty = \max_i \left(\sum_{j=1}^n |a_{ij}|\right)$.

По определению матричной нормы, подчиненной евклидовой векторной норме, имеем

$$||A||_2 = \sup_{\boldsymbol{x}\ne 0} \frac{||A\boldsymbol{x}||_2}{||\boldsymbol{x}||_2} = \sup_{\boldsymbol{x}\ne 0} \sqrt{\frac{(A\boldsymbol{x}, A\boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})}} = \sup_{\boldsymbol{x}\ne 0} \sqrt{\frac{(A^TA\boldsymbol{x}, \boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})}}.$$

Отметим, что $(A^TA)^T = A^T(A^T)^T = A^TA$, т.е. матрица $B = A^TA$ — симметричная, и $(A^TA\boldsymbol{x},\boldsymbol{x}) = (A\boldsymbol{x}, A\boldsymbol{x}) \ge 0$; следовательно, все $\lambda(B) \ge 0$. Рассуждая далее, как и в задаче 4, получим

$$\sup_{x\ne 0} \frac{(B\boldsymbol{x},\boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})} = \max_i \lambda_i(B),$$

а равенство достигается на соответствующем собственном векторе. Поэтому

$$||A||_2 = \sqrt{\max_i \lambda(A^T A)}\,.$$

Следует отметить важный частный случай симметричной матрицы: $A = A^T$. Здесь

$$||A||_2 = \max_i |\lambda(A)|.$$

Если же эта матрица к тому же положительно определена, т.е. $A = A^T > 0$, то

$$||A||_2 = \max_i \lambda(A).$$

4. Пусть A — вещественная $(n \times m)$-матрица, $\boldsymbol{x}$ — вещественный m-вектор и $\boldsymbol{y}$ — вещественный n-вектор. Доказать следующие три свойства спектральной нормы $||\cdot||$:

$$||A||_2 = \max_{\substack{||\boldsymbol{x}||_2=1 \\ ||\boldsymbol{y}||_2=1}} |\boldsymbol{y}^T A\boldsymbol{x}|, \quad ||A^T||_2 = ||A||_2, \quad ||A^T A||_2 = ||AA^T||_2 = ||A||_2^2.$$

Решение. Для доказательства первого свойства спектральной нормы надо показать, что существуют такие векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ единичной длины, на которых максимум достигается. В силу неравенства Коши–Буняковского и учитывая, что спектральная норма подчинена евклидовой векторной норме, получим неравенство

$$\left|\boldsymbol{y}^T A\boldsymbol{x}\right| = (\boldsymbol{y}, A\boldsymbol{x}) \leq ||\boldsymbol{y}||_2\, ||A\boldsymbol{x}||_2 \leq ||\boldsymbol{y}||_2\, ||\boldsymbol{x}||_2\, ||A||_2 = ||A||_2\,.$$

Пусть вектор $\boldsymbol{x}$ такой, что $||A\boldsymbol{x}||_2 = ||A||_2$, т.е. на нем достигается максимум в определении подчиненной нормы, и возьмем $\boldsymbol{y} = A\boldsymbol{x}/||A\boldsymbol{x}||_2$. Тогда $||\boldsymbol{y}||_2 = 1$ и

$$\left|\boldsymbol{y}^T A\boldsymbol{x}\right| = \frac{(A\boldsymbol{x})^T}{||A\boldsymbol{x}||_2}\, A\boldsymbol{x} = \frac{(A\boldsymbol{x}, A\boldsymbol{x})}{||A\boldsymbol{x}||_2} = ||A\boldsymbol{x}||_2 = ||A||_2\,.$$

Следовательно, искомые векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ построены и первое свойство спектральной нормы доказано.

Из первого свойства спектральной нормы и из равенства

$$||A^T||_2 = \max_{\substack{||\boldsymbol{x}||_2=1 \\ ||\boldsymbol{y}||_2=1}} \left|\boldsymbol{y}^T A^T\boldsymbol{x}\right| = \max_{\substack{||\boldsymbol{x}||_2=1 \\ ||\boldsymbol{y}||_2=1}} (\boldsymbol{y}, A^T\boldsymbol{x}) = \max_{\substack{||\boldsymbol{x}||_2=1 \\ ||\boldsymbol{y}||_2=1}} (A\boldsymbol{y}, \boldsymbol{x}) = \max_{\substack{||\boldsymbol{y}||_2=1 \\ ||\boldsymbol{x}||_2=1}} \left|\boldsymbol{x}^T A\boldsymbol{y}\right| = ||A||_2$$

следует ее второе свойство. Заметим, что поскольку здесь мы применяем первое свойство к матрице A^T, то в обозначениях, принятых в этом равенстве, вектор $\boldsymbol{y}$ имеет размерность m, а вектор $\boldsymbol{x}$ — размерность n.

Покажем теперь справедливость третьего свойства спектральной нормы. Из второго свойства следует неравенство

$$||A^T A||_2 \leq ||A^T||_2 ||A||_2 = ||A||_2^2.$$

Возьмем такой вектор $\boldsymbol{x}$, что $||\boldsymbol{x}||_2 = 1$ и $||A\,\boldsymbol{x}||_2 = ||A||_2$, и применим первое свойство к матрице $A^T A$, положив $\boldsymbol{y} = \boldsymbol{x}$. Тогда получим неравенство

$$||A^T A||_2 \geq \left|\boldsymbol{x}^T A^T A\boldsymbol{x}\right| = (A\boldsymbol{x}, A\boldsymbol{x}) = ||A\,\boldsymbol{x}||_2^2 = ||A||_2^2.$$

Из этих двух неравенств следует $||A^T A||_2 = ||A||_2^2$. Аналогично показывается, что $||AA^T||_2 = ||A||_2^2$. Таким образом, третье свойство спектральной нормы доказано.

5. Пусть A — вещественная прямоугольная матрица. Показать, что умножение ее справа или слева на ортогональную матрицу Q соответствующих порядков не меняет ее спектральную норму.

Решение. Из третьего свойства спектральной нормы следует

$$||QA||_2^2 = ||(QA)^T QA||_2 = ||A^T Q^T QA||_2 = ||A^T A||_2 = ||A||_2^2.$$

Из второго свойства спектральной нормы и полученного равенства следует

$$||AQ||_2 = ||(AQ)^T||_2 = ||Q^T A^T||_2 = ||A^T||_2 = ||A||_2.$$

В частности, из

$$||Q\boldsymbol{x}||_2^2 = (Q\boldsymbol{x}, Q\boldsymbol{x}) = (Q\boldsymbol{x})^T Q\boldsymbol{x} = \boldsymbol{x}^T Q^T Q\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = (\boldsymbol{x}, \boldsymbol{x}) = ||\boldsymbol{x}||_2^2$$

следует, что умножение вектора $\boldsymbol{x}$ на ортогональную матрицу сохраняет его длину.

6. Показать справедливость неравенства

$$\|A\|_2^2 \le \|A\|_1 \|A\|_\infty.$$

Решение. Для согласованных векторных и матричных норм из

$$A\boldsymbol{x} = \lambda\boldsymbol{x}$$

имеем

$$|\lambda|\,||\boldsymbol{x}|| \le ||A||\,||\boldsymbol{x}||, \quad |\lambda| \le ||A||,$$

т.е. модуль любого собственного значения матрицы не больше любой ее нормы. Из этого заключения и из третьего свойства спектральной нормы получим решение задачи:

$$||A||_2^2 = \max\lambda(A^T A) \le ||A^T A||_1 \le ||A||_1 ||A^T||_1 = ||A||_1 ||A||_\infty.$$

7. Доказать, что числовая функция $M(A) = n\,\max_{ij}|a_{ij}|$ от элементов матрицы A является матричной нормой.

Решение. Заметим, что

$$\eta(A) = \max_{i,j}|a_{ij}| = \sup_{\boldsymbol{x}\ne 0}\frac{||A\boldsymbol{x}||_\infty}{||\boldsymbol{x}||_1},$$

поскольку

$$||A\boldsymbol{x}||_\infty = \max_i\left|\sum_j a_{ij}x_j\right| \le \max_{i,j}|a_{ij}|\sum_j|x_j| \quad \text{и} \quad \sum_j|x_j| = ||\boldsymbol{x}||_1,$$

причем равенство достигается, когда все элементы матрицы A одинаковы.

Получим оценку снизу:

$$n\,\max_{i,j}|a_{ij}| = n\,\sup_{\boldsymbol{x}\ne 0}\frac{||A\boldsymbol{x}||_\infty}{||\boldsymbol{x}||_1} \ge \sup_{\boldsymbol{x}\ne 0}\frac{n\,||A\boldsymbol{x}||_\infty}{n\,||\boldsymbol{x}||_\infty} \ge ||A||_\infty,$$

т.е. $||A||_\infty \le M(A)$. Здесь мы использовали неравенство

$$||\boldsymbol{x}||_\infty \le ||\boldsymbol{x}||_1 \le n\,||\boldsymbol{x}||_\infty.$$

Оценка сверху получается аналогично:

$$n\,\max_{ij}|a_{ij}| \le n\,\sup_{\boldsymbol{x}\ne 0}\frac{||A\boldsymbol{x}||_\infty}{||\boldsymbol{x}||_\infty} \le n\,||A||_\infty,$$

т.е. $M(A) \le n\,||A||_\infty$. Полученные неравенства определяют соотношение эквивалентности для $M(A)$ и $||A||_\infty$.