Доказательство теоремы Фробениуса - Перрона

Мария Школьник

Определение 1. Матрица называется положительной (неотрицательной), если все ее элементы положительны (неотрицательны). Обозначение: $M > 0$ $(M \geq 0)$

Определение 2. Спектральным радиусом матрицы называется наибольший из модулей ее собственных значений.

Определение 3. Пусть $A = (a_{ij})$ *и* $B = (b_{ij})$ *- две матрицы* $n \times r$*. Тогда* $A \geq B$ $(A > B)$*, если* $a_{ij} \geq b_{ij}$ $(a_{ij} > b_{ij})$ $\forall$ i, j.

Обозначение. Пусть - матрица с комплексными элементами. Тогда C^+ *- матрица, у которой* $(C^+)_{ij} = |C_{ij}|$.

Графы и неприводимость.

Каждой неотрицательной $n \times n$ *матрице* A *соответствует граф* G_A *с множеством вершин (1, ..., n), где ребро (i, j) существует, если* $A_{ij} > 0$.

Неотрицательная матрица A *называется неприводимой, если* G_A *сильно связан. То есть для любых* $i, j \in (1, .., n)$ *существует* k *такое ,что* $(A^k)_{ij} > 0$*. Если матрица не является неприводимой, то ее называют приводимой.*

Пусть π *- произвольная перестановка 1, ..., n. Тогда* $\pi(A)$ *- матрица, элементы которой определяются равенством* $A_{ij} = (\pi(A))_{(\pi(i))(\pi(j))}$*. Другими словами, можно написать* $\pi(A) = P^{-1}AP$*, где* P *- матрица перестановки* π*. Тогда можем получить фурмулу:*

$$(\pi(A))_{\pi(i),\pi(j)} = (P^{-1})_{\pi(i),i} \cdot A_{ij} \cdot P_{j,\pi(j)}$$

Таким образом, $\pi(A)$ *обозначает тот же линейный опрератор, что и матрица* A*, только в другом базисе.(Что касается графа, такая перестановка просто переименовывает вершины.)*

Замечание 1

Матрица A является приводимой, если для какой-либо перестановки π матрица $\pi(A)$ имеет вид $\pi(A) = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$, где B и D - квадратные матрицы. Простой способ увидеть это - взглянуть на соответствующий граф: граф не является сильно связанным, если его вершины могут быть разбиты на два множества таким образом, что из одного в другое не идет ни одно ребро. Соответствующая нумерация вершин дает матрицу смежности вида $\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$.

Теорема Фробениуса - Перрона

Пусть A - неприводимая квадратная матрица с неотрицательными вещественными элементами и r $\in R_{>0}$ - ее спектральный радиус. Тогда справедливы следующие утверждения:

(1) r собственное значение матрицы A.
(2) соответствующий r собственный вектор имеет строго положительные координаты
(3) r как собственное значение матрицы имеет (геометрическую и алгебраическую) кратность, равную 1
(4) не существует другого (отличного от x) положительного собственного вектора матрицы A.
(5) r увеличивается, если какой-либо элемент матрицы увеличивается
(6) все собственные значения матрицы A с абсолютным значением r имеют кратность 1. Если их h, то они являются (комплексными) корнями уравнения $\lambda^h = r^h$.
(7) спектр матрицы A, рассматриваемый как мультимножество, переходит сам в себя при вращении комплексной плоскости на угол $\frac{2\pi}{h}$

Доказательство

Лемма 1. Пусть A $\in M_n$, A $\geq$ 0 - неприводимая матрица. Тогда $(E + A)^{n-1} > 0$.

Доказательство.

Достаточно доказать, что $(E+A)^{n-1}\cdot$x $> 0 \ \forall$ x ≥ 0, x $\neq 0$. Сравним количество ненулевых координат произвольного ненулевого вектора x и вектора y = (A + E)x. Если $x_i > 0$, то $y_i = (Ax)_i + x_i > 0$. Таким образом, множество ненулевых координат y является подмножеством такого множества вектора x. Предположим, что два эти набора совпадают. Выберем при помощи перестановки π такой базис, чтобы координаты векторов x и y в данном базисе имели вид: $\begin{bmatrix} u \\ 0 \end{bmatrix}$ и $\begin{bmatrix} v \\ 0 \end{bmatrix}$ соответственно, где u и v - векторы одинакового размера; u, v $>$ 0. Представвим π(A) в виде блочной матрицы $\begin{bmatrix} B & C \\ D & F \end{bmatrix}$, где B и F - квадратные матрицы, и размер B равен размеру вектора u. Теперь равенство

$$\begin{bmatrix} B & C \\ D & F \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix}$$

подразумевает D = 0, поскольку u $>$ 0. Тогда матрица A приводима по замечанию 1, что противоречит условию леммы. Следовательно, y имеет строго меньше нулевых координат, чем x. Применяя этот факт для векторов x, (A + E)x, ..., $(A + E)^{n-2}$ x и учитывая что число нулевых координат x не превосходит n - 1, получаем: $(A + E)^{n-1} > 0$.

Шаг 1.

Введем функцию r: $R^n \to$ R такую, что:

$$r(x) = \min(\frac{(Ax)_i}{x_i})\ (1)$$

Заметим, что r(x) непрерывна в любой точке x > 0, но может иметь точки разрыва на границе множества X = (x $\in R^n$ | x $\geq$ 0, x $\neq$ 0). Очевидным свойством r(x) является то, что r(x)x $\leq$ Ax и, более того:

$$r(x) = \max(p \mid px \geq Ax)\ (2)$$

Мы собираемся доказать существование такого вектора z, что

$$r(z) = \sup(\ r(x) \mid x \in X)\ (3)$$

По теореме Вейерштрасса непрерывная функция из компакта в подмножество R достигает своего минимума (и максимума) на этом компакте. Множество X не является ни ограниченным, ни замкнутым. Однако r(x) = r(αx) $\forall\ \alpha > 0$ и, поэтому, sup(r(x) | x $\in$ X) = sup(r(x) | x $\in X^1$), где X^1 = (x $\in$ X | |x| = 1). Заметим, что множество X^1 является компактом, но r(x) может иметь разрыв в некоторых точках данного множества, которые не являются положительными векторами. Тогда вместо X^1 рассмотрим множество

$$Y = (y \mid y = (A+E)^{n-1}x, \text{ где } x \in X^1)$$

Данное множество компактно. Более того, все векторы множества Y положительны по Лемме 1. Следовательно, функция r непрерывна на Y и по теореме Вейерштрасса достигает своего минимума и максимума на данном множестве. Пусть z - такой вектор, что r(z) = max(r(y) | y $\in$ Y). Также y = $(A+E)^{n-1}$x для какого-то x и px $\leq$ Ax, а, следовательно, py $\leq$ Ay. Применяя (2) это означает, что r(x) $\leq$ r(y). Тогда мы нашли вектор необходимый для (3):

$$\sup(r(x) \mid x \in X) = \sup(r(x) \mid x \in X^1) \leq \sup(r(y) \mid y \in Y) = \max(r(y) \mid y \in Y) = r(z).$$

Шаг 2 Пусть t = r(z) и u - вектор из X, удовлетворяющий равенству r(u) = t. Опишем некоторые свойства r и u:

$(1')$ Если x > 0, тогда Ax > 0 и, следовательно, r(x) > 0 и t > 0.

$(2')$ Пусть y = $(A+E)^{n-1}$u. Используя (2), ru $\leq$ Au. Если это не равенство, то $(A+E)^{n-1}$ (Au - tu) > 0 по Лемме 1, подразумевая ry < Au. Тогда существует $\varepsilon > 0$ такой, что (r + ε)y < Ay. Следовательно, r(y) $\leq$ t + ε что противоречит условию (3) Следовательно, tu = Au. Получаем, что t - собственное значение матрицы A с собственным вектором u.

$(3')$ Вектор u строого положительный, поскольку
y = $(A+E)^{n-1}$u = $(r+1)^{n-1}$u и y $>$ 0 по Лемме 1.

$(4')$ Пусть α - комплексное собственное значение матрицы A с собственным вектором v. Равенство Av = αv означает $(Av)^+ = (\alpha v)^+$. Тогда для любого i:

$$((\alpha v)^+)_j = |\alpha v_j| = |\alpha|\ |v_j| = (|\alpha| v^+)_j \ (4)$$
$$((Av)^+)_j = |A_{j1}v_1 + ... + A_{jn}v_n| \leq A_{j1}|v_1| + ... + A_{jn}|v_n| = (Av^+)_j \ (5)$$

и тогда $|\alpha| v^+ \leq Av^+$. Поскольку $v^+ \geq 0$, получаем $|\alpha| \leq$ r$(v^+) \leq$ t. Следовательно, t - такое собственное значение, которое $\geq$ любого другого собственного значения матрицы A. Следовательно, t = r - спектральный радиус матрицы A.

$(5')$ Имеем r(v^+) = r, тогда v^+ - положительный собственный вектор матрицы A соответствующий r. Но $Av^+ = rv^+$ подразумевает равенство в (5). Поэтому, комплексные числа $v_1, ..., v_n$ должны иметь один и тот же аргумент φ. Тогда v = $e^{i\varphi}v^+$. Таким образом, любой собственный вектор, соответствующий r является коллинеарным положительному вектору. Остается заметить, что r не могут соответствовать два линейно независимых собственных вектора, так как два таких вектора будут иметь ненулевую линейную комбинацию u, что u - неотрицательный вектор с хотя бы одной нулевой координатой, что невозможно по (3). То есть r соответствует единственный положительный собственный вектор.
Таким образом мы доказали, что геометрическая кратность t как собственного значения матрицы A равна 1. Докажем, что алгебраическая кратность t также равна 1. Для этого достаточно показать, что жорданова форма матрицы A имеет клетку размера 1 с числом t. Рассмотрим матрицу A^T: она неотрицательна, неприводима и имеет тот же самый характеристический многочлен, что и матрица A. Следовательно, существует такой вектор y, что $A^Ty = ry$.
Рассмотрим A как матрицу линейного оператора в некотором базисе. Ортогональное дополнение вектора y: $y' = (x|y^Tx = 0)$ инвариантно относительно A. Заметим, что z не принадлежит y' поскольку y и z - положительные векторы. Таким образом, R^n раскладывается в прямую сумму двух инвариантных подпространств: y' и <z>. Тогда матрица A в базисе (z, y_1, ..., y_{n-1}) имеет блочно-диагональный вид A' = $\begin{bmatrix} r & 0 \\ 0 & Y \end{bmatrix}$ с матрицей Y размера n-1. Приводя матрицу Y к жордановой форме некоторой заменой базиса, мы также приводим матрицу A' к этой же самой форме. (Поскольку, матрицы A и A', будучи подобными, имеют одинаковую жорданову форму). В верхнем углу у этой матрицы есть клетка размера 1 с номером r. Таким образом, мы доказали пункт (3) теоремы.

Лемма 2. Пусть A $\geq$ 0 - неприводимая n*n матрица и C - такая комплексная матрица, что $C^+ \leq$ A. Пусть α - какое-то собственное значение матрицы C. Тогда $|\alpha| \leq$ r, и равенство достигается, если C $= e^{i\varphi}DAD^{-1}$, где $\alpha = re^{i\varphi}$ и D - такая диагональная матрица, что $P^+ =$ E.

Доказательство.
Пусть $Cy = \alpha y$. Ранее показали, что: $|\alpha| y^+ \leq C^+y^+ \leq Ay^+$.
Следовательно, $|\alpha| \leq C^+y^+ \leq Ay^+$ по (2) и (3). Теперь пусть $|\alpha|$ = r, тогда r(y^+) = r, и поэтому y - положительный собственный вектор A, соответствующий r. Получаем равенство:

$$ry^+ = C^+y^+ = Ay^+$$

Тогда $(A - C^+)y^+ = 0$, $A - C^+ \geq 0$, $y^+ > 0$, значит, $C^+ = A$. Пусть

$$
\text{y} = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{bmatrix} = \begin{bmatrix} |y_1|e^{i\varphi_1} \\ |y_2|e^{i\varphi_2} \\ ... \\ |y_n|e^{i\varphi_n} \end{bmatrix},
$$

$$
\text{D} = \begin{bmatrix} e^{i\varphi_1} & ... & 0 \\ ... & ... & ... \\ 0 & ... & e^{i\varphi_n} \end{bmatrix}
$$

Тогда $y = Dy^+$. Подставляя $\alpha = re^{i\varphi}$, получим:

$$
Cy = CDy^+ = re^{i\varphi}Dy^+ \text{ и } e^{-i\varphi}D^{-1}CDy^+ = ry^+.
$$

Пусть $e^{-i\varphi}D^{-1}CD =$ F. $Fy^+ = Ay^+$ и $F^+ = C^+$ ($=$ A), подразумевая $Fy^+ = F^+y^+$. Поскольку y > 0 и действительный части элементов матрицы $(F^+ - F)$ неотрицательны, значит, они должны быть равны нулю, а это возможно только в том случае, когда $F^+ = F$. Следовательно, F $=$ A и C $= e^{i\varphi}PAP^{-1}$.

Шаг 3.

Предположим, что мы увеличили некоторые элементы A, получив при этом новую неприводимую матрицу B такую, что B $\geq$ A, B $\neq$ A. Применив Лемму 2, получим, что модуль любого собственного значения A не превосходит спектральный радиус матрицы B. Отсюда следует пункт (5) теоремы.

Шаг 4.

Пусть $\lambda_0 =$ r, $\lambda_1 = re^{i\varphi}$, ..., $\lambda_{h-1} = re^{i\varphi_{h-1}}$ - все собственный значения матрицы A с абсолютным значением r и $0 = \varphi_0 \leq \varphi_1 \leq ... \leq \varphi_{h-1} < 2\pi$. Применим Лемму 2 для C $=$ A, $\alpha = \lambda_k$, получим равенство:
A $= e^{i\varphi_k D_k A(D_k)^{-1}}$ (k $=$ 0, ..., h-1, $(D_k)^+ =$ E). Значит, матрица $e^{-i\varphi_k}$A $= D_kA(D_k)^{-1}$ подобна матрице A и имеет тот же самый спектр. С другой стороны, умножая матрицу на скаляр, мы умножаем все ее собственные значения на этот скаляр. Таким образом, спектр матрицы A сохраняется при вращении комплексной плоскости на угол φ_k. Поскольку поворот на угол φ_k отображает r в λ_k, их кратности совпадают, значит, каждое λ_k имеет кратность 1. Далее, спектр сохраняется для любых k, l $\in$ 1, ..., h-1 при вращении на $\varphi_k + \varphi_l$, -φ_k, и конечно на 0. Каждый из этих поворотов переводит r в некоторое λ_j, значит, суммы и разности, взятые по модулю 2π углов из $\Phi = (\varphi_0, ..., \varphi_{h-1})$ также принадлежат этому множеству. Значит, Φ - группа по сложению по модулю 2π.
Заметим, что $\varphi_1 + \varphi_1 = \varphi_2$, $\varphi_1 + \varphi_2 = \varphi_3$, ..., $\varphi_1 + \varphi_{h-1} = 0 \pmod{2\pi}$. Получаем $\varphi_k =$ kφ_1. Следовательно, $\varphi_k = \frac{2\pi k}{h}$. Таким образом, доказали утверждения (6) и (7) теоремы.