

Вычислительная линейная алгебра

Константин Чепуркин, Андрей Смоленский, Андрей Лавренов

1. Итерационные методы решения систем линейных уравнений

1.1. Метод простой итерации

Пусть система уравнений записана в виде $x = Ax + b$, где $x, b \in \mathbb{R}^n$, $A \in M_n(\mathbb{R})$. Метод простой итерации заключается в том, чтобы взять произвольный вектор $x_0 \in \mathbb{R}^n$, и строить последовательность векторов по правилу

$$x_{s+1} = Ax_s + b.$$

Лемма 1. Если собственные числа A по модулю меньше 1, то последовательность x_s стремится к решению системы $x = Ax + b$.

Если система дана в виде $Ax = b$, то её можно привести к указанному виду $x = (E_n - A)x + b$. Однако, никто не гарантировал, что матрица $E_n - A$ будет иметь собственные числа, меньшие чем 1 по модулю. Отметим, что привести систему $Ax = b$ к виду $x = Ax + b$ можно разными способами. В любом случае, оценить собственные числа матрицы можно при помощи кругов Гершгорина.

Определение. Пусть A – матрица. Тогда кругами Гершгорина назовём круги с центром в a_{ii} радиуса $r_i = \sum_{j \neq i} |a_{ij}|$.

Теорема 1. Все собственные числа матрицы находятся в объединении её кругов Гершгорина.

Важным фактором, влияющим на количество необходимых итераций является выбор начального x_0 . Часто вам может быть заранее известно приближённое решение системы. Его и стоит взять в качестве x_0 .

Задача 1. Необходимо реализовать метод простых итераций, так, чтобы каждый его шаг работал за $O(n^2)$. Предполагается следующий формат ввода-вывода:

Input: $A \in M_n(\mathbb{R})$, $b \in \mathbb{R}^n$, точность $\varepsilon > 0$.

Output: x , такое что $\|x - Ax - b\| < \varepsilon$ или 0, если круги Гершгорина не лежат в единичном круге с центром в нуле и на протяжении 20 итераций подряд $\|x_s\| \geq \|x_{s-1}\| + 1$.

1.2. Метод Гаусса-Зейделя

Вариацией метода простых итераций является метод Гаусса-Зейделя. Пусть нам дана система вида $Ax = b$, где $A \in M_n(\mathbb{R})$, $b \in \mathbb{R}^n$. Пусть диагональные элементы A отличны от нуля. Представим матрицу A в виде суммы $A = L + U$, где L – нижнетреугольная матрица, а U – верхнетреугольная с нулями на диагонали. Представим систему $Ax = b$ в виде

$$Lx = -Ux + b$$

Возьмём какой-то вектор x_0 . Определим последовательность x_i : элемент x_{i+1} будет искаться, как решение системы

$$Lx_{i+1} = -Ux_i + b.$$

Заметим, что эта система имеет единственное решение в указанных предположениях. Благодаря тому, что L – нижнетреугольная, такую систему можно решить за $O(n^2)$ операций.

Задача 2. Вам необходимо реализовать этот алгоритм, так, чтобы каждый его шаг работал за $O(n^2)$. Предполагается следующий формат ввода-вывода:

Input: $A \in M_n(\mathbb{R})$, $b \in \mathbb{R}^n$, точность $\varepsilon > 0$.

Output: x , такое что $\|Ax - b\| < \varepsilon$ или 0, если на диагонали матрицы A стоят нули или на протяжении 20 итераций подряд $\|x_s\| \geq \|x_{s-1}\| + 1$.

2. QR-разложение

Нам известно, что QR -разложение напрямую связано с процедурой ортогонализации Грама-Шмидта. Однако, на практике процедура ортогонализации приводит к достаточно большим ошибкам в QR -разложении.

Для того, чтобы выяснить, как можно лучше вычислять QR -разложение, нам стоит поменять взгляд на ситуацию. В процедуре ортогонализации мы преобразовывали последовательно матрицу A к ортогональной, и все наши преобразования в итоге кодировались верхнетреугольной матрицей. Но можно и наоборот, пытаться при помощи домножения на ортогональные матрицы привести матрицу A к верхнетреугольному виду.

Основной ссылкой для нас будет книга [GGH12] (стр. 234-256). Мы рассмотрим два основных способа нахождения QR -разложения.

2.1. Вращения Гивенса

Определение. Пусть дано число n , индексы $i, j \in \overline{1, n}$, что $i < j$, и угол φ . Возьмём $c = \cos \varphi$ и $s = \sin \varphi$. Тогда матрица вращения Гивенса имеет вид

$$G(i, j, \varphi) = \begin{matrix} & & i & & j & & \\ i & \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & s \\ & & -s & & c \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} & j & \end{matrix}.$$

Это матрица поворота по часовой стрелке на угол φ в плоскости $\langle e_i, e_j \rangle$.

Посмотрим, как домножение на матрицу $G(i, j, \varphi)$ может упростить вектор x . Все координаты $x^{new} = G(i, j, \varphi)x$ не отличаются от координат x , кроме координат с номерами i, j . В них имеем следующее:

$$x_i^{new} = cx_i + sx_j \text{ и } x_j^{new} = -sx_i + cx_j.$$

При выборе

$$c = \frac{x_j}{x_i^2 + x_j^2} \text{ и } s = \frac{-x_i}{x_i^2 + x_j^2}$$

мы обнуляем i -ую координату x_{new} .

Заметим, что домножение A на матрицу $G(i, j, \varphi)$ слева влияет только на строки i, j матрицы A . Посмотрим что происходит с этими строками. Пусть u_i, u_j соответствующие строки A , а u_i^{new}, u_j^{new} строки $G(i, j, \varphi)A$. Тогда

$$u_i^{new} = cu_i + su_j \text{ и } u_j^{new} = -su_i + cu_j$$

Задача 3. Вычислите за $O(n)$ произведение $G(i, j, \varphi)A$.

Input: матрица $A \in M_n(\mathbb{R})$, $i, j \in \overline{1, n}$, коэффициенты $c, s \in \mathbb{R}$, что $c^2 + s^2 = 1$.

Output: $G(i, j, \varphi) \cdot A$ для соответствующего φ

Как при помощи этой конструкции Найти QR-разложение матрицы? Для этого пригодна идея из метода Гаусса. Пусть на k -ом шаге изменённая матрица $A^{(k)}$ уже имеет в первых $k - 1$ столбцах нулевые элементы под диагональю. Тогда убьём элементы ниже диагонали в столбце k . Для этого пробежим все элементы ниже диагонали. Как только встретим ненулевой – возьмём его. Пусть его позиция $i \leq k$. С его помощью убьём все элементы столбца k ниже него домножая на матрицы вращений Гивенса $G(i, j, \varphi)$ для подходящих φ . Заметим, что само φ при этом вычислять не нужно.

После этого поворотом на $\pi/2$ в плоскости k, i приведём полученный элемент из позиции i на диагональ (если это нужно).

Замечание. Тут мы не следим за знаками элементов на диагонали у $A^{(k)}$, а могли бы и следить.

Замечание. Занулять элементы можно и в другом порядке. Например, можно идти снизу столбца к диагонали.

Замечание. Алгоритм работает для вырожденных матриц и вообще для матриц произвольного размера.

Замечание. На каждом шаге нужно поддерживать ортогональную матрицу $Q^{(k)}$, что $Q^{(k)\top} A = A^{(k)}$. Тогда $A = Q^{(n)} A^{(n)}$ и будет QR-разложением. Алгоритм работает за $O(n^3)$.

Задача 4. Реализуйте нахождение QR-разложения при помощи вращений Гивенса за $O(n^3)$

Input: $A \in M_n(\mathbb{R})$.

Output: Q и R , что $A = QR$, где Q – ортогональная, а R – верхнетреугольная.

2.2. Отражения Хаусхолдера

Определение. Пусть v – вектор из \mathbb{R}^n единичной длины. Тогда матрица отражения относительно гиперплоскости, перпендикулярной v , имеет вид $H_v = E_n - 2vv^\top$.

Лемма 2. Для вектора $u \in \mathbb{R}^n$ единичной длины и $e_1 = (1, 0, \dots, 0)^\top$ выполнено, что отражение, соответствующее вектору $\frac{u - e_1}{\|u - e_1\|}$, переводит u в e_1 .

Ясно, что если v – произвольный ненулевой вектор, то отражение, соответствующее $\frac{u - e_1}{\|u - e_1\|}$, где $u = \frac{v}{\|v\|}$, переводит v в $\|v\| \cdot e_1$

Лемму выше можно применять к столбцам данной матрицы A , последовательно получая нули ниже какой-то позиции. А именно, переведя первый столбец матрицы A в столбец, отличающийся от нулевого только в первой строке, можно повторить процедуру для подматрицы, полученной вычеркиванием первых столбца и строки – тогда во втором столбце будут обращены в 0 все элементы, кроме первого и второго.

Отметим, что для реализации домножения на матрицу Хаусхолдера не требуется формировать саму матрицу и применять полное матричное умножение. А именно, при домножении на матрицу отражения $I - 2vv^\top$ достаточно умножить матрицу на вектор и проделать матричное сложение исходной матрицы и матрицы ранга 1:

$$(I - 2vv^\top)A = A - (2v)(v^\top A).$$

Итого для матрицы A размера $n \times n$ потребуется $2n^2$ арифметических операций для вычисления $v^\top A$, затем еще n^2 операций для вычисления элементов матрицы $vv^\top A$, и n^2 операций для вычисления суммы $A - (2v)(v^\top A)$. Здесь мы считаем умножение на 2 быстрой операций (по сравнению с операциями с числами с плавающей запятой).

Задача 5. Реализуйте быстрое домножение на матрицу Хаусхолдера, отвечающую вектору v .

Input: $A \in M_n(\mathbb{R})$, $v \in M_n(\mathbb{R})$.

Output: $H_v \cdot A$.

Задача 6. Реализуйте QR-разложение с помощью матриц Хаусхолдера. Применяйте домножение на матрицы Хаусхолдера к подматрицам нужного размера.

Input: $A \in M_n(\mathbb{R})$.

Output: Q и R , что $A = QR$, где Q – ортогональная, а R – верхнетреугольная.

3. QR-алгоритм и другие алгоритмы для нахождения собственных чисел

3.1. Простая итерация

Простейшим алгоритмом, позволяющим что-то сказать про собственные числа и собственные вектора матрицы, является итерационный алгоритм. Напомним его описание.

Пусть дана квадратная матрица $A \in M_n(\mathbb{C})$. Возьмём начальное приближение $x_0 \in \mathbb{C}^n$ к собственному вектору матрицы A . Если начальное приближение не задано, то в качестве x_0 можно взять случайный вектор единичной длины.

Определим последовательность x_i как

$$x_{i+1} = \frac{1}{\|Ax_i\|} Ax_i.$$

Верно следующее утверждение:

Теорема 2. Пусть $A \in M_n(\mathbb{C})$ и пусть λ есть единственное максимальное по модулю собственное число. Тогда есть такое подпространство $U \leq \mathbb{C}^n$ размерности $n - 1$, что для всех начальных $x_0 \in \mathbb{C}^n \setminus U$ последовательность x_i сходится к собственному вектору матрицы A соответствующему максимальному собственному числу.

В этой ситуации, для того, чтобы оценить собственное число на i -ом шаге, надо найти произведение $x_i^\top Ax_i$. Действительно, если x_i примерно собственный вектор, то

$$x_i^\top Ax_i \simeq x_i^\top \lambda x_i = \lambda \|x_i\|^2 = \lambda.$$

Задача 7. Вам необходимо реализовать следующий алгоритм

Input: Матрица $A \in M_n(\mathbb{R})$, вектор x_0 и $\varepsilon \in \mathbb{R}$ – точность.

Output: λ – приближение к максимальному собственному числу A , вектор $v \in \mathbb{C}^n$, такой, что $\|v\| = 1$ и $\|Av - \lambda v\| < \varepsilon$, либо 0, если таких λ и v не удаётся найти за разумное время.

3.2. QR-алгоритм

Во многих задачах нам нужно найти только максимальное собственное число. Но бывают и такие задачи, где нужно найти спектр матрицы целиком. Например, если мы смотрим на задачу об изоморфизме графов.

Одним из наиболее употребимых алгоритмов для нахождения всех собственных чисел матрицы является QR-алгоритм. Несмотря на то, что QR-алгоритм применим для произвольных квадратных матриц над \mathbb{C} мы ограничимся только симметричными вещественными матрицами в виду того, что в этой ситуации проще описать результат работы QR-алгоритма. Основной ссылкой для нас будет [GGH12] (стр. 385-392 и 458-465).

Точнее, в нашей ситуации QR-алгоритм должен сойтись к диагональной матрице, на диагонали которой стоят собственные числа исходной матрицы. Понятно, что завершиться надо будет на каком-то конечном шаге, то есть на необязательно диагональной матрице. В качестве меры недиагональности матрицы выберем радиусы кругов Гершгорина.

Опишем теперь сам QR-алгоритм. Пусть $A_0 = A$ – симметричная вещественная матрица. Построим последовательность матриц A_k . На k -ом шаге для матрицы A_{k-1} строится QR -разложение $A_{k-1} = Q_k R_k$, а затем строится матрица $A_k = R_k Q_k$. Имеет место следующий результат:

Теорема 3. Пусть A – симметричная матрица с различными по модулю собственными числами. Тогда последовательность матриц A_k сходится к диагональной матрице, на диагонали которой стоят собственные числа A . При этом на каждом шаге выполнено $A_k = Q^{(k)\top} A Q^{(k)}$, где $Q^{(k)} = Q_1 \dots Q_k$.

Замечание. Таким образом, матрица $Q^{(k)}$ – это матрица составленная примерно из собственных векторов A .

Задача 8. Вам необходимо реализовать этот алгоритм. Предполагается следующий формат ввода-вывода:

Input: Симметричная матрица $A \in M_n(\mathbb{R})$, $\varepsilon \in \mathbb{R}$ – точность.

Output: Диагональные элементы матрицы A_k и матрица $Q^{(k)}$, для такого k , что радиусы кругов Гершгорина у A_k меньше ε .

3.3. Трёхдиагональные матрицы

В том виде, в котором QR-алгоритм описан выше, он крайне затратен по времени. Поэтому были придуманы модификации, ускоряющие его работу. Первая из них – тридиагонализация исходной матрицы.

Определение. Матрица A называется трёхдиагональной, если $a_{ij} = 0$ при $|i - j| \geq 2$. Иными словами, это матрицы вида:

$$\begin{pmatrix} a_1 & b_1 & & \\ c_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & c_{n-1} & a_n \end{pmatrix}.$$

Покажем, что для любой симметричной матрицы A существует такая матрица ортогональная матрица Q , что $Q^\top A Q$ – трёхдиагональна. Для этого нам понадобятся матрицы отражений Хаусхолдера. Опишем алгоритм тридиагонализации матрицы A .

Разберёмся с первым столбцом. Пусть $u_1 \in \mathbb{R}^{n-1}$ – это первый столбец A без верхнего элемента. Тогда, есть такой $u \in \mathbb{R}^{n-1}$, что для $H_u \in M_{n-1}(\mathbb{R})$ выполнено

$$H_u u_1 = c e_1 \in \mathbb{R}^{n-1}.$$

Сделаем из u вектор $v \in \mathbb{R}^n$ добавив сверху нулевой элемент. Тогда матрица H_v как матрица размера n будет иметь вид

$$H_v = \begin{pmatrix} 1 & 0 \\ 0 & H_u \end{pmatrix}.$$

Тогда у матрицы AH_v первая строка имеет вид

$$(*, c, 0 \dots, 0).$$

Когда мы домножим AH_v слева на матрицу H_v , то в силу структуры H_v первая строка останется такой же, как и у AH_v . По симметричности, первый столбец $H_v A H_v$ будет иметь такой же вид. Значит мы можем тридиагонализировать первые столбец и строчку A .

Сделаем теперь более общее замечание. Пусть вектор $v \in \mathbb{R}^n$ имеет нулями первые $k+1$ координат. Тогда матрица H_v имеет вид

$$H_v = E_n + H,$$

где у матрицы H первые $k+1$ строк и столбцов нулевые. Пусть так же у матрицы A первые k столбцов уже тридиагонализированы (а из симметричности и первые k строк). Тогда первые k строк и столбцов матрицы

$$H_v A H_v = A + H A + A H + H A H$$

будут такими же, как у матрицы A , то есть останутся тридиагонализированными.

Отсюда следует, что если мы на шаге $k+1$ уже тридиагонализировали первые k строк и столбцов A и дальше будем использовать для тридиагонализации только вектора с $k+1$ первыми нулями, то мы не будем менять уже тридиагонализированную часть матрицы A . Но тогда давайте выкинем первые k строк и столбцов, коль скоро они не будут меняться. Тогда мы получим матрицу размера $n-k$ у которой надо тридиагонализировать первый столбец используя вектор с первой нулевой координатой. А это мы уже умеем.

Замечание. Каждый шаг этого алгоритма занимает $O(n^2)$, а всего алгоритм тридиагонализации требует $O(n^3)$ операций.

Задача 9. Вам необходимо реализовать этот алгоритм, так, чтобы он работал за $O(n^3)$. Предполагается следующий формат ввода-вывода:

Input: Симметричная матрица $A \in M_n(\mathbb{R})$

Output: A' – тридиагонализация матрицы A , ортогональная матрица Q , что $Q^T A Q = A'$.

3.4. Применение к QR-алгоритму

Вернёмся к QR-алгоритму. Улучшения в работе этого алгоритма для трёхдиагональной матрицы, связаны со следующей серией замечаний:

Замечание. QR-разложение для трёхдиагональной матрицы можно сделать за $O(n^2)$ зануляя поддиагональные элементы матрицы A при помощи вращений Гивенса.

Замечание. Верхнетреугольная матрица R в QR-разложении трёхдиагональной матрицы имеет только две ненулевых диагонали над главной диагональю.

Замечание. Если в QR-алгоритме матрица A_{k-1} была трёхдиагональной, то и матрица A_k – тоже трёхдиагональна.

Задача 10. Вам необходимо реализовать QR-алгоритм для трёхдиагональных матриц, так, чтобы каждый шаг QR-алгоритма требовал $O(n^2)$ операций. Предполагается следующий формат ввода-вывода:

Input: Трёхдиагональная матрица $A \in M_n(\mathbb{R})$, ε – точность

Output: Диагональные элементы матрицы A_k и матрица $Q^{(k)}$, для такого k , что радиусы кругов Гершгорина у A_k меньше ε .

3.5. Сдвиг

Заметим, что в пределе диагональные элементы матрицы $A^{(k)}$ соответствуют собственным числам матрицы A , упорядоченным по убыванию их модуля. ближе всего к корректному значению оказываются последние элементы.

То, что мы в какой-то момент знаем приближение к собственным числам позволяет нам ускорить сходимость нашего метода. Точнее, пусть s – приближение к минимальному по модулю собственному

числу A . Рассмотрим QR-разложение матрицы $A - sE_n = QR$. Тогда $Q^T A Q = RQ + sE_n$ – матрица, сопряжённая исходной. Оказывается, что такой шаг QR-алгоритма со сдвигом приводит к очень быстрой сходимости нижнего элемента к минимальному по модулю собственному числу. Когда внедиагональные элементы нижней строки и столбца матрицы станут близкими к 0, то можно сказать, что минимальное собственное число матрицы A найдено. После этого можно выкинуть последний столбец и строку матрицы, перейдя к матрице меньшего размера.

Возьмём в качестве приближения к минимальному собственному числу матрицы A ближайшее к a_{nn} собственное число матрицы

$$\begin{pmatrix} a_{n-1,n-1} & a_{n-1,n} \\ a_{n,n-1} & a_{n,n} \end{pmatrix}$$

после нескольких итераций. Такой сдвиг называется сдвигом Уилкинсона.

Задача 11. Вам необходимо реализовать этот алгоритм, так, чтобы каждый шаг требовал $O(n^2)$ операций. Предполагается следующий формат ввода-вывода:

Input: Трёхдиагональная матрица $A \in M_n(\mathbb{R})$, ε – точность

Output: Диагональные элементы матрицы A_k и матрица $Q^{(k)}$, для такого k , что радиусы кругов Гершгорина у A_k меньше ε .

4. Спектры графов

В этом разделе мы обсудим простейшие свойства спектров графов. Сформулируем творческую задачу: реализуйте алгоритм, тестирующий изоморфизм графов, опирающийся на описанные ранее алгоритмы нахождения собственных чисел и собственных векторов. Попробуйте выжать максимум из информации о собственных числах и собственных векторах матрицы смежности.

Задача 12. Реализуйте тест графов на неизоморфность.

Input: Матрицы смежности $A(G_1)$ и $A(G_2)$.

Output: 0 – если графы не изоморфны, 1 – если это неясно.

Определение. Граф G называется (n, d, α) -экспандером (алгебраическим), если G – d -регулярный граф на n вершинах с $\lambda = \max(|\lambda_2|, |\lambda_n|) \leq \alpha d$.

В построении графов-расширителей используют следующий приём – строят маленький граф-расширитель, а потом применяют комбинаторные конструкции, чтобы сделать его больше. Мы не будем обсуждать комбинаторные конструкции. Мы хотим чтобы вы нашли с точностью $\varepsilon = 10^{-5}$ для данных небольших регулярных графов наилучшие показатели α .

- 1) В качестве вершин графа G рассмотрим множество $\mathbb{Z}/n \times \mathbb{Z}/n$ и проведём для пары (x, y) ребра в вершины

$$(x \pm 2y, y), (x \pm (2y + 1), y), (x, y \pm 2x), (x, y \pm (2x + 1)).$$

- 2) Пусть p – простое число. В качестве вершин графа G возьмём $\mathbb{Z}/p \cup \{\infty\}$. Вершина x соединена с $x + 1$, $x - 1$ и x^{-1} . При $x = 0$ считаем $0^{-1} = \infty$ и обратно.

Задача 13. Нужно реализовать две программы, которые считают оптимальное α для данных типов графов

Input: n в первом случае и p – во втором.

Output: $\alpha = \max(|\lambda_2|, |\lambda_n|)/d$

В качестве дополнения отметим, что любой регулярный граф с чётной степенью вершины d является графом Шрайера [Gross 77](#) (для нечётной степени тоже есть результат, но там не все графы). Осталось понять, что такое графы Шрайера. Пусть дано конечное множество X . В качестве вершин графа Шрайера возьмём элементы X . А откуда взять рёбра?

Пусть на X действовала группа G . А у группы G было зафиксировано множество образующих $S = \{g_1, \dots, g_d\}$, такое что $S^{-1} = S$ (любой обратный к элементу S снова лежит в S). Соединим два элемента $x, y \in X$ ребром, если есть такое $g \in S$, что $gx = y$. Понятно, что из условия $S^{-1} = S$ получается неориентированный граф, возможно с петлями и кратными рёбрами. Степень вершины в таком графе равна d – это и есть граф Шрайера. Заметим, что у такого графа может быть нечётная степень только, если среди g_i есть элемент порядка 2.

Посмотрите на конструкции выше – это два графа Шрайера. Именно графы Шрайера некоммутативных групп дают базовые примеры семейств экспандеров (да, абелевы не подходят). Подробнее про это смотри [Expander graphs and their applications](#).

Список литературы

[GGH12] Van Loan Ch. F. Golub G. H. *Matrix Computations*. Johns Hopkins University Press, 2012.