

Семинар 9

Рекуррентные соотношения

Конкретный пример Давайте в начале начнем с конкретного случая, где я все проиллюстрирую на конкретном примере – последовательность Фибоначчи. Что это такое? Это последовательность чисел $a_n \in \mathbb{R}$, где $n \in \mathbb{Z}_{\geq 0}$ удовлетворяющая следующим условиям

$$\begin{cases} a_n = a_{n-1} + a_{n-2} \\ a_0 = 1 \\ a_1 = 1 \end{cases}$$

Если мы захотим по этим правилам посчитать n -ый член последовательности, то нам понадобится $O(n)$ операций, то есть последовательно посчитать все n членов последовательности, чтобы добраться до a_n . Однако, можно несколько схитрить и сделать это быстрее за $O(\log n)$ с помощью матричных операций. Для этого введем вектор

$$x_n = \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} \text{ при этом мы знаем, что } x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Заметим, что x_n выражается через x_{n-1} следующим образом

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix} \text{ то есть } x_n = Ax_{n-1} \text{ где } A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Но тогда $x_n = Ax_{n-1} = A^2x_{n-2} = \dots = A^{n-1}x_1$. А значит для нахождения x_n нам лишь надо возвести матрицу A в степень n . Для этого подходит хорошо известный алгоритм быстрого возведения в степень для чисел, который слово в слово работает для матриц. Давайте заведем две квадратные матрицы $X, Y \in M_2(\mathbb{R})$ и число $m \in \mathbb{Z}_{\geq 0}$. В самом начале положим $X = E$, $Y = A$ и $m = n$. Будем поддерживать следующий инвариант $XY^m = A^n$. Алгоритм остановим тогда, когда $m = 0$, тогда X будет нашим ответом. Шаги алгоритма следующие. Если m четно, то $XY^{2m'} = X(Y^2)^{m'}$ поделим m на 2, а Y возведем в квадрат. Если m нечетно, то $XY^{2m'+1} = (XY)Y^{2m'}$ уменьшим m' на единицу и умножим X на Y .

На самом деле, можно проверить, что $A = CDC^{-1}$, где

$$D = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}, \quad C = \begin{pmatrix} \frac{2}{\sqrt{5}-1} & -\frac{2}{\sqrt{5}+1} \\ 1 & 1 \end{pmatrix} \text{ и } C^{-1} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{5-\sqrt{5}}{10} \\ -\frac{1}{\sqrt{5}} & \frac{5+\sqrt{5}}{10} \end{pmatrix}$$

Для простоты обозначений, будем считать $D = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$, тогда

$$A^n = (CDC^{-1})^n = CD^nC^{-1} = C \begin{pmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{pmatrix} C^{-1}$$

То есть, зная как получить из матрицы A супер крутое разложение, мы можем еще сильнее упростить вычисления степени матрицы и свести его к вычислению степеней чисел и произведения трех матриц.

Рекуррента длины 2 В случае произвольной рекурренты $a_n = aa_{n-1} + ba_{n-2}$ заметим, что если мы фиксировали a_0 и a_1 , то вся последовательность (являющаяся решением рекурренты) полностью восстанавливается однозначно.¹ Потому решать рекурренту можно следующим образом: в начале подобрать какой-то набор решений, а потом покомбинировать их и подогнать начальные условия. Давайте пройдемся по этим шагам.

В начале предлагается искать решение в виде $a_n = \lambda^n$ для ненулевого λ .² Тогда давайте пойдем, для каких $\lambda \in \mathbb{R}$ мы действительно сможем подобрать решение для рекурренты (мы пока игнорируем полностью начальные условия). Тогда уравнение для рекурренты превращается в $\lambda^n = a\lambda^{n-1} + b\lambda^{n-2}$ (это выполнено для любого $n \geq 2$). Так как $\lambda \neq 0$ по нашему выбору, то мы можем сократить на λ^{n-2} и получим, $\lambda^2 = a\lambda + b$. То есть $a_n = \lambda^n$ является решением рекурренты тогда и только тогда, когда λ является корнем квадратного уравнения $\lambda^2 - a\lambda - b = 0$.

Теперь у нас две ситуации. Ситуация первая: два разных корня λ_1 и λ_2 . Тогда последовательности $a_n = \lambda_1^n$ и $a_n = \lambda_2^n$ являются решениями. Заметим, что если $a_n = aa_{n-1} + ba_{n-2}$ и $b_n = ab_{n-1} + bb_{n-2}$, тогда для последовательности $c_n = a_n + b_n$ тоже выполнено $c_n = ac_{n-1} + bc_{n-2}$. Так же для последовательности $c_n = \mu a_n$

¹Для любителей формализма тут на помощь спешит индукция.

²Почему именно в таком виде? Да, просто так, бог подсказал, во сне приснилось, догадались. Бывает.

тоже верно, что c_n – решение рекуррентности. Таким образом, если у нас есть какие-то решения, то складывая их между собой и умножая на коэффициенты, мы будем продолжать получать решения. Значит в нашем запасе есть решения вида $a_n = c_1 \lambda_1^n + c_2 \lambda_2^n$, где c_1 и c_2 – любые константы. Теперь мы хотим выбрать их так, чтобы a_0 и a_1 были любые заданные числа, то есть мы хотим решить

$$\begin{cases} c_1 \lambda_1^0 + c_2 \lambda_2^0 = a_0 \\ c_1 \lambda_1^1 + c_2 \lambda_2^1 = a_1 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$$

Мы видим, что определитель этой матрицы равен $\lambda_2 - \lambda_1$ и не равен нулю при разных λ . А значит для любой правой части существует единственное решение c_1, c_2 . А значит, мы всегда можем решить рекурренту вида $a_n = aa_{n-1} + ba_{n-2}$ с помощью выражения $c_1 \lambda_1^n + c_2 \lambda_2^n$, если уравнение $t^2 - at - b = 0$ имело разные корни.

Теперь давайте проверим, что делать, если корень получился один и кратности 2. Пусть λ – единственный корень, тогда уравнение было $t^2 - 2\lambda t + \lambda^2 = 0$. Мы уже знаем, что $a_n = \lambda^n$ является решением. Давайте проверим, что $a_n = n\lambda^n$ тоже является решением. Надо проверить

$$a_n = 2\lambda a_{n-1} - \lambda^2 a_{n-2} \Leftrightarrow n\lambda^n = 2\lambda(n-1)\lambda^{n-1} - \lambda^2(n-2)\lambda^{n-2}$$

Раскрыв все скобки, убеждаемся, что последнее равенство верное. Значит можно искать решение среди выражений вида $(c_1 + c_2 n)\lambda^n$. Посмотрим, во что превращается выбор начальных условий:

$$\begin{cases} c_1 \lambda^0 + c_2 \cdot 0 \cdot \lambda^0 = a_0 \\ c_1 \lambda^1 + c_2 \lambda^1 = a_1 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 0 \\ \lambda & \lambda \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$$

И мы видим, что и в этом случае у нас найдется единственное решение, если корень был не нулем.³

Произвольные рекурренты Теперь рассмотрим рекурренту длины k , то есть выражение

$$a_n = \alpha_1 a_{n-1} + \dots + \alpha_k a_{n-k}$$

Тогда такая рекуррента полностью определяется своими первыми k членами a_0, \dots, a_{k-1} . В начале давайте будем искать решения в виде $a_n = \lambda^n$, где $\lambda \neq 0$. Подставляя это в рекурренту, мы получим

$$\lambda^n = \alpha_1 \lambda^{n-1} + \dots + \alpha_k \lambda^{n-k}$$

Это уравнение должно быть выполнено для любого $n \geq k$. Ясно, что достаточно потребовать его выполнения для $n = k$, то есть решить уравнение

$$\lambda^k - \alpha_1 \lambda^{k-1} - \dots - \alpha_k = 0$$

Пусть комплексные корни этого уравнения $\lambda_1, \dots, \lambda_s$, а их кратности n_1, \dots, n_s .

Решения с разными корнями Теперь мы готовы записать произвольное решение. Давайте для начала запишем его в случае, когда все корни разные (то есть имеют кратность 1). В этом случае мы обозначим все корни так $\lambda_1, \dots, \lambda_k$. Тогда ясно, что выражение

$$a_n = c_1 \lambda_1^n + \dots + c_k \lambda_k^n$$

является решением для любых значений параметров c_1, \dots, c_k . Теперь надо показать, что с помощью этих параметров мы можем добиться любых начальных условий a_0, \dots, a_{k-1} . Для этого надо решить систему

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_k \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \dots & \lambda_k^{k-1} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{k-1} \end{pmatrix}$$

Слева получилась матрица с определителем Вандермонда с разными значениями λ_i , потому она не вырождена. А значит такая система имеет единственное решение для любой правой части. А это то что мы и хотели.

³Обратите внимание, что нулевой корень означает, что рекуррента имеет вид $a_n = 0a_{n-1} + 0a_{n-2}$, и этот случай разбирается легко.

Решения с повторяющимися корнями Пусть теперь корни $\lambda_1, \dots, \lambda_s$ с кратностями n_1, \dots, n_s . Тогда оказывается, что любое решение можно записать в виде

$$a_n = P_1(n)\lambda_1^n + \dots + P_s(n)\lambda_s^n$$

где $P_i(n) = p_0 + p_1n + \dots + p_{n_i-1}n^{n_i-1}$ – многочлен от n степени $n_i - 1$ с неопределенными коэффициентами. Заметим, что у многочлена P_i в точности n_i коэффициентов, а значит у всех многочленов $n_1 + \dots + n_s = k$ коэффициентов, то есть столько сколько надо. Можно показать, что для любых начальных условий a_0, \dots, a_{k-1} можно найти единственный набор коэффициентов для многочленов P_i .

Связь с матрицами в общем случае Как и в случае рекуренты длины 2 можно переписать все в матричной форме следующим образом

$$\begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_{n-k+1} \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & & \alpha_k \\ 1 & 0 & & 0 \\ & \ddots & & \\ 0 & & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \\ \vdots \\ a_{n-k} \end{pmatrix}$$

Давайте обозначим эту матрицу за A , а столбцы за y_n , тогда мы получили $y_n = Ay_{n-1}$ и продолжая расписывать y_i друг через друга мы получаем формулу $y_n = A^{n-k+1}y_{k-1}$. Таким образом в этом случае мы тоже можем находить n -ый член последовательности за $\log n$, только теперь константа перед логарифмом будет вида k^3 – скорость вычисления произведения матриц.⁴

Давайте я постараюсь объяснить, как глядя на это равенство получить формулы из предыдущего пункта. Это объяснение будет хорошей демонстрацией зачем нам вообще изучать какую-то общую теорию. Начнем с вычисления минимального и характеристического многочленов.

Глядя на последнюю строку степеней $E, A, A^2, \dots, A^{k-1}$ можно увидеть, что никакая их сумма с ненулевыми коэффициентами не зануляется. А в сумме $A^k - \alpha_1 A^{k-1} - \dots - \alpha_k E$ последняя строка зануляется и это единственно возможные коэффициенты. Из первого замечания следует, что минимальный многочлен имеет степень k . А так как он в частности зануляет последнюю строку, то выше написан именно он. А так как у нас степень зануляющего k , то он совпадает с характеристическим многочленом. То есть, мы получили, что

$$f_{\min}(t) = \chi_A(t) = t^k - \alpha_1 t^{k-1} - \dots - \alpha_k$$

Потому значения $\lambda_1, \dots, \lambda_s$, которые мы получили раньше – это в точности спектр матрицы A .

Теперь минутка магии. Оказывается, что в общем случае можно показать, что для такой матрицы A можно найти матрицу C такую, что верно разложение $A = CJC^{-1}$, где

$$J = \begin{pmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_s}(\lambda_s) \end{pmatrix} \quad \text{где} \quad J_m(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \quad \text{имеет размер } m$$

Тогда

$$y_{n+k-1} = A^n y_{k-1} = C J^n C^{-1} y_{k-1}$$

Обратите внимание, что C, C^{-1} и y_{k-1} не зависят от n и состоят из каких-то чисел. В то же время J^n мы с вами уже считали и можно увидеть, что там встречаются выражения вида $P(n)\lambda_i^n$, где степень P строго меньше, чем n_i . Тогда перемножив эти матрицы мы видим, что все коэффициенты вектора y_{n+k-1} являются выражениями вида $a_n = P_1(n)\lambda_1^n + \dots + P_s(n)\lambda_s^n$, где многочлены P_i степени $n_i - 1$ максимум. То есть если бы мы могли доказать, что матрица A раскладывается как описано выше, то мы могли бы получить взятые с потолка формулы из очевидного матричного равенства $y_n = Ay_{n-1}$.

⁴Ну или чуть меньше, если вы пользуетесь хитрыми алгоритмами.