



СЕМИНАРЫ ПО ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКЕ

Е. С. Тверская

МГУ им. Н.Э. Баумана
Москва

Метод Хаусхолдера (метод отражений)

Определение. Ортогональная 2×2 -матрица Q называется **матрицей отражения**, если

$$Q = \begin{bmatrix} \cos(\vartheta) & \sin(\vartheta) \\ \sin(\vartheta) & -\cos(\vartheta) \end{bmatrix}$$

Если $y = Q^T x = Qx$, то y получается отражением x относительно оси, определяемой как

$$S = \text{span} \left\{ \begin{bmatrix} \cos(\vartheta/2) \\ \sin(\vartheta/2) \end{bmatrix} \right\}$$

$x_0 \in \mathbb{R}^n$ - произвольный ненулевой вектор.

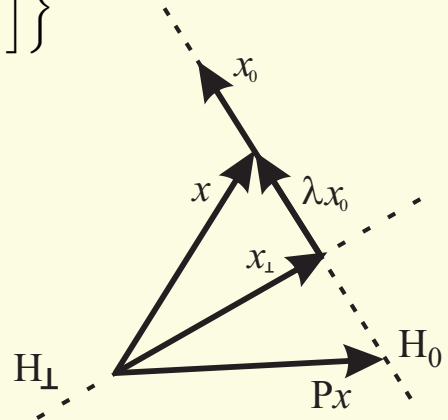
\mathcal{H}^\perp - ортогональное дополнение одномерного линейного пространства. $\mathcal{H}_0 = \text{span}\{x_0\}$.

$x = \lambda x_0 + x_\perp$ для $\forall x \in \mathbb{R}^n$,

λx_0 - ортогональная проекция вектора x на

\mathcal{H}_0 ,

x_\perp - его ортогональная составляющая.



Рассмотрим преобразование

$$Px = P(\lambda x_0 + x_{\perp}) = x_{\perp} - \lambda x_0 = x - 2\lambda x_0. \quad (1)$$

Выберем в качестве вектора x_0 вектор ω такой, что $\|\omega\| = 1$. Тогда $\lambda = (x, \omega)$ и преобразование (1) можно переписать в виде

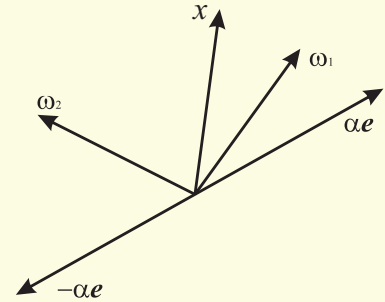
$$Px = x - 2(x, \omega)\omega = (E - 2\omega\omega^T)x.$$

Оператор P , указанного вида, называется **оператором отражений** или **оператором Хаусхолдера**. Матрицу $P = (E - 2\omega\omega^T)$, соответствующую данному оператору называют **матрицей отражений** или **матрицей Хаусхолдера**.

Свойства оператора Хаусхолдера

- Преобразование Хаусхолдера является линейным.
- Оператор Хаусхолдера - симметричный оператор.
- Оператор Хаусхолдера - ортогональный оператор.

Основная задача : подобрать вектор ω в операторе Хаусхолдера таким образом, чтобы в результате преобразования полученный вектор имел направление заданного единичного вектора e , т. е. $Px = \alpha e$ или $Px = -\alpha e$, где $\alpha > 0$. Направление, перпендикулярное плоскости отражения, будет определяться либо вектором $x - \alpha e$, либо вектором $x + \alpha e$



Тогда

$$\begin{aligned} Px &= x - 2(x, \omega)\omega = x - 2 \frac{(x, x - \alpha e)}{\|x - \alpha e\|^2} (x - \alpha e) = \\ &= x - \frac{2\|x\|^2 - 2\alpha(x, e)}{2\|x\|^2 - 2\alpha(x, e)} (x - \alpha e) = x - x + \alpha e = \alpha e. \end{aligned}$$

Пусть $v \in \mathbb{R}^n$ – ненулевой вектор. Тогда

$$P = E - \frac{2vv^T}{v^T v}, \quad (2)$$

где вектор v называют **вектором Хаусхолдера**.

На практике отражения Хаусхолдера будем использовать для обнуления выбранных компонент вектора.

В частности, пусть задан ненулевой вектор $x \in \mathbb{R}^n$ и мы хотим, чтобы вектор Px был коллинеарен вектору e_1 – первому столбцу матрицы E_n . Тогда

$$Px = \left(E - \frac{2vv^T}{v^T v} \right) x = x - 2 \frac{v^T x}{v^T v} v.$$

Положим $v = x + \alpha_1 e_1$. Тогда

$$v^T x = x^T x + \alpha_1 x_1,$$

$$v^T v = x^T x + 2\alpha_1 x_1 + \alpha_1^2,$$

и

$$Px = \left(1 - 2 \frac{x^T x + \alpha_1 x_1}{x^T x + 2\alpha_1 x_1 + \alpha_1^2} \right) x - 2\alpha_1 \frac{v^T x}{v^T v} e_1.$$

Для коллинеарности Px и e_1 нужно, чтобы коэффициент при $x \in \mathbb{R}^n$ был нулевым. Тогда

$$v = x \pm \sqrt{x^T x} e_1 \tag{3}$$

и

$$Px = \left(E - \frac{2vv^T}{v^T v} \right) x = \mp \sqrt{x^T x} e_1.$$

Вычисление вектора Хаусхолдера

Рассмотрим некоторые нюансы. Если x почти коллинеарен вектору e_1 , то вектор

$$v = x - \text{sign}x_1 \sqrt{x^T x} e_1$$

имеет малую норму. Поэтому возможно появление большой относительной ошибки при вычислении множителя $\frac{2}{v^T v}$.

Эту трудность можно обойти, если взять α_1 с тем же знаком, что и знак первой компоненты вектора x , т. е.

$$v = x + \text{sign}x_1 \sqrt{x^T x} e_1.$$

заметим, что при таком выборе $\|v\|_\infty = |v_1|$.

Полезно также придерживаться такой нормировки вектора v , что $v_1 = 1$. Данное условие можно выполнить, если первую компоненту взять равной единице, т. е. $v_1 = 1$, а все остальные компоненты получить как

$$v_i = \frac{v_i}{x_1 + \text{sign} x_1 \sqrt{x^T x}}.$$

Данное представление иногда упрощает алгоритмы, в которых нужно хранить вектор Хаусхолдера.

Пример. Дан вектор $x = (3, 1, 5, 1)^T$. Найдем матрицу оператора Хаусхолдера, который отражает вектор x относительно гиперплоскости перпендикулярной вектору v таким образом, что $Px = \alpha_1 e_1$, где $e_1 = (1, 0, 0, 0)^T$.

С учетом нормировки вектор Хаусхолдера $v = (1, 1/9, 5/9, 1/9)^T$. Тогда матрица Хаусхолдера

$$P = \frac{1}{54} \begin{pmatrix} -27 & -9 & -45 & -9 \\ -9 & 53 & -5 & -1 \\ -45 & -5 & 29 & -5 \\ -9 & -1 & -5 & 53 \end{pmatrix}$$

Замечание. Преобразование Хаусхолдера не требует явного формирования матрицы Хаусхолдера, так как

$$PA = \left(E - \frac{2vv^T}{v^T v} \right) A = A + v\varphi^T,$$

где $\varphi = \beta A^T v$ и $\beta = -\frac{2}{v^T v}$.

Схема алгоритма. Пусть матрицы $A \in \mathbf{M}_{n \times n}(\mathbb{R})$ и $n = 5$. Пусть матрица Хаусхолдера P_1 и P_2 таковы, что

$$P_2 P_1 A = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \otimes & \times & \times \\ 0 & 0 & \otimes & \times & \times \\ 0 & 0 & \otimes & \times & \times \end{pmatrix}$$

Найдем матрицу Хаусхолдера $\tilde{P}_3 \in \mathbf{M}_{3 \times 3}(\mathbb{R})$ такую, что

$$\tilde{P}_3 \begin{pmatrix} \otimes \\ \otimes \\ \otimes \end{pmatrix} = \begin{pmatrix} \times \\ 0 \\ 0 \end{pmatrix}.$$

Если $P_3 = \mathbf{diag}(E_2, \tilde{P}_3)$, то

$$P_3 P_2 P_1 A = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times & \times \end{pmatrix}$$

QR разложение матрицы.

Выполнив n таких шагов, мы получаем верхнюю треугольную матрицу $P_n P_{n-1} \dots P_2 P_1 A = R$.

Таким образом, положив $Q = P_1 P_2 \dots P_{n-1} P_n$ имеем $A = QR$, где Q - ортогональная матрица, а R - верхняя треугольная матрица.

Метод Гивенса (метод вращений)

Определение. Ортогональная 2×2 -матрица Q называется **матрицей вращения**, если

$$Q = \begin{bmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{bmatrix}$$

Если $y = Q^T x$, то y получается поворотом x на угол ϑ против часовой стрелки.

Отражения Хаусхолдера полезны при крупномасштабных обнулениях. Однако, когда необходимо избирательное зануление элементов предпочтительнее использовать **вращения Гивенса**.

Матрица Гивенса имеет следующий вид

$$G = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & c & \dots & s & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & -s & \dots & c & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{matrix} i \\ \\ k \\ \\ k \\ \\ i \quad k \end{matrix}$$

где $c = \cos(\vartheta)$, а $s = \sin(\vartheta)$.

$$\text{Если } y = G^T x, \text{ то } y_j = \begin{cases} cx_i - sx_k, & j = i, \\ sx_i + cx_k, & j = k, \\ x_j, & j \neq k, i. \end{cases}$$

Тогда y_k можно обратить в ноль, положив $c = \frac{x_i}{\sqrt{x_i^2 + x_k^2}}$, $s = \frac{-x_k}{\sqrt{x_i^2 + x_k^2}}$.

Данное преобразование равносильно повороту вектора x на угол ϑ против часовой стрелки в координатной плоскости (i, k) .

в которой

$$\begin{aligned} a_{1j}^{(1)} &= c_{12}a_{1j} - s_{12}a_{2j}, & a_{2j}^{(1)} &= s_{12}a_{1j} + c_{12}a_{2j}, & j &= \overline{1, m} \\ f_1^{(1)} &= c_{12}f_1 - s_{12}f_2, & f_2^{(1)} &= s_{12}f_1 + c_{12}f_2. \end{aligned}$$

Заметим, что $a_{21}^{(1)} = 0$ в силу специального выбора коэффициентов c_{12} и s_{12} .

Если в исходной системе коэффициент $a_{21} = 0$, т. е. система уже приведена к нужному виду, то полагают $c_{12} = 1$, а $s_{12} = 0$.

Преобразование исходной системы равносильно умножению слева матрицы A и столбца f на матрицу T_{12} , которая имеет вид

$$T_{12} = \begin{pmatrix} c_{12} & -s_{12} & 0 & 0 & \dots & 0 \\ s_{12} & c_{12} & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Для исключения неизвестного x_1 из 3-го уравнения вычислим числа

$$c_{13} = \frac{a_{11}^{(1)}}{\sqrt{\left(a_{11}^{(1)}\right)^2 + a_{31}^2}}, \quad s_{13} = \frac{-a_{31}}{\sqrt{\left(a_{11}^{(1)}\right)^2 + a_{31}^2}},$$

которые обладают следующими свойствами:

$$c_{13}^2 + s_{13}^2 = 1, \quad s_{13}a_{11}^{(1)} + c_{13}a_{31} = 0.$$

Затем 1-е уравнение системы заменяют линейной комбинацией 1-го и 3-го уравнений с коэффициентами c_{13} и $-s_{13}$, т. е.

$$\begin{aligned} \left(a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \dots + a_{1m}^{(1)}x_m\right)c_{13} + \left(a_{31}x_1 + a_{32}x_2 + \dots + a_{3m}x_m\right)(-s_{13}) = \\ = f_1^{(1)}c_{13} + f_3(-s_{13}). \end{aligned}$$

А 3-е уравнение – линейной комбинацией 1-го и 3-го уравнений с коэффициентами s_{13} и c_{13} , т. е.

$$\begin{aligned} \left(a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \dots + a_{1m}^{(1)}x_m\right)s_{13} + \left(a_{31}x_1 + a_{32}x_2 + \dots + a_{3m}x_m\right)c_{13} = \\ = f_1^{(1)}(-s_{13}) + f_3c_{13}. \end{aligned}$$

Действие матрицы T_{kl} на вектор x эквивалентно его повороту вокруг оси, перпендикулярной плоскости Ox_kx_l на угол φ_{lk} такой, что $c_{kl} = \cos \varphi_{kl}$ и $s_{kl} = \sin \varphi_{kl}$. Заметим, что

$$T_{kl}^T = T_{kl}^{-1}.$$

Следовательно матрица T_{kl} – ортогональная матрица.

На 2-м шаге метода вращений, из уравнений системы (4) с номерами $i = \overline{3, m}$ исключают неизвестное x_2 . Для этого каждое i -е уравнение комбинируют со вторым уравнением. В результате приходи к системе

$$\begin{aligned} a_{11}^{(m-1)} x_1 + a_{12}^{(m-1)} x_2 + a_{12}^{(m-1)} x_3 + \dots + a_{1m}^{(m-1)} x_m &= f_1^{(m-1)}, \\ a_{22}^{(m-1)} x_2 + a_{23}^{(m-1)} x_3 + \dots + a_{2m}^{(m-1)} x_m &= f_2^{(m-1)}, \\ a_{33}^{(2)} x_3 + \dots + a_{3m}^{(2)} x_m &= f_3^{(2)}, \\ &\dots\dots\dots \\ a_{3m}^{(2)} x_3 + \dots + a_{mm}^{(2)} x_m &= f_m^{(2)}. \end{aligned}$$

В матричном виде получаем

$$A^{(2)} x = f^{(2)},$$

$$\text{где } A^{(2)} = T_{2m} T_{2(m-1)} \dots T_{24} T_{23} A^{(1)}, \quad f^{(2)} = T_{2m} T_{2(m-1)} \dots T_{24} T_{23} f^{(1)}.$$

После завершения $(m - 1)$ -го шага система принимает вид

$$\begin{aligned} a_{11}^{(m-1)} x_1 + a_{12}^{(m-1)} x_2 + a_{13}^{(m-1)} x_3 + \dots + a_{1m}^{(m-1)} x_m &= f_1^{(m-1)}, \\ a_{22}^{(m-1)} x_2 + a_{23}^{(m-1)} x_3 + \dots + a_{2m}^{(m-1)} x_m &= f_2^{(m-1)}, \\ a_{33}^{(m-1)} x_3 + \dots + a_{3m}^{(m-1)} x_m &= f_3^{(m-1)}, \\ &\dots \\ &\dots \\ a_{mm}^{(m-1)} x_m &= f_m^{(m-1)} \end{aligned}$$

или $A^{(m-1)}x = f^{(m-1)}$, где

$$A^{(m-1)} = T_{m-1,m}A^{(m-2)}, \quad f^{(m-1)} = T_{m-1,m}f^{(m-2)}$$

Введем обозначение R для полученной верхней треугольной матрицы $A^{(m-1)}$. Она связана с исходной матрицей равенством

$$R = TA$$

где $T = T_{m-1,m} \dots T_{2m}T_{2(m-1)} \dots T_{24}T_{23} \dots T_{1m}T_{1(m-1)} \dots T_{13}T_{12}$ – матрица результирующего вращения. Матрица T является ортогональной, так как является произведением ортогональных матриц.

Введем обозначение $Q = T^{-1} = T^T$. Таким образом получаем QR разложение матрицы A .

Обратный ход метода вращения проводится точно также как и для метода Гаусса.

Метод вращения обладает хорошей числовой устойчивостью, однако он более трудоемок по сравнению с методом Гаусса.

Иллюстрация метода вращений

$$\begin{pmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{pmatrix} \longrightarrow \begin{pmatrix} \times & \times & \times \\ 0 & \times & \times \\ \times & \times & \times \end{pmatrix} \longrightarrow \begin{pmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{pmatrix} \longrightarrow \begin{pmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{pmatrix}$$