24 Сентября 2024

1. Вводная

Колобов Александр Георгиевич, D947

Чтобы получить экзамен нужно:
- допуск от преподавателя по практике
- сдача теории (экзамен)

Можно сдавать частями (три коллоквиума):
- Прямые методы
- Итерационные методы
- Собственные значения

Курсовой проект. Подробно разобрать конкретный метод, который каждому будет дан. Ещё будет две теоретические задачи.

2. Первая лекция

В конечномерном пространстве все нормы эквивалентны.

Определение

$\|x\|$ — число, для которого выполняются 3 аксиомы:
1. $\|x\| \geq 0, \|x\| = 0 \Leftrightarrow x = 0$
2. $\|\lambda x\| = \lambda \|x\|$
3. $\|x + y\| \leq \|x\| + \|y\|$

Определение

Нормы вектора.

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

$$\|x\|_2 = \sqrt{(x,x)} - \text{евклидова норма}$$

$$\|x\|_\infty = \max_i |x_i|$$

СЛАУ:

$$Ax = f$$

Определение

Нормы матрицы.

Евклидова норма:

$$\|A\|_E = \sqrt{\sum_{i,j} a_{ij}^2}$$

Определение

Подчинённая матричная норма:

$$\|\mathrm{A}\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Свойства:
1. $\|I\| = 1$
2. $\|AB\| \leq \|A\|\|B\|$

Евклидова норма не является подчинённой, ведь $\|I\|_E = \sqrt{n} \neq 1$.

Из определения подчинённой матричной нормы вытекает:

$$\|\mathrm{A}\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \Rightarrow \|Ax\| \leq \|A\|\|x\|$$

Определение

Согласованная матричная норма такая, что:

$$\|Ax\| \leq \|A\|\|x\|$$

$$\|A\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$$

$$\|A\|_2 = \sqrt{\max_i \lambda_i(A^* A)},\ \text{где } A^* - \text{транспонированная матрица}$$

$$\|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$$

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

$$\|Ax\|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leq \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}||x_j| =$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} |a_{ij}||x_j| = \sum_{j=1}^{n} \left(|x_j| \sum_{i=1}^{n} |a_{ij}| \right) \leq$$

$$\leq \sum_{j=1}^{n} \left(\left(\max_j \sum_{i=1}^{n} |a_{ij}| \right) |x_j| \right) = \max_j \sum_i |a_{ij}| \cdot \sum_j |x_j| =$$

$$= \left(\max_j \sum_i |a_{ij}| \right) \|x\|_1$$

Доказали, что

$$\frac{\|Ax\|_1}{\|x\|_1} \leq \max_j \sum_{i=1}^{n} |a_{ij}|$$

k — номер столбца, где достигается максимум:

$$\sum_{i=1}^{n} |a_{ik}| = \max_{j} \sum_{i=1}^{n} |a_{ij}|$$

$$e_k = \begin{pmatrix} 0 \\ \dots \\ \underbrace{1}_{k\text{—й индекс}} \\ \dots \\ 0 \\ 0 \end{pmatrix} = \max_{j} \sum_{i=1}^{n} |a_{ij}| \|x\| ???$$

$$\|Ae_k\|_1 = \sum_{i} |a_{ik}|$$

Для второй матричной нормы $\|A\|_2 = \sqrt{\max_i \lambda_i(A^*A)}$:

λ_i должны быть действительными и положительными

A^*A — неотрицательно определена и симметрична

$$(A^*A)^* = A^*(A^*)^* = A^*A$$

Определение

Матрица C называется неотрицательно определённой, если

$$\forall (y, y) \geq 0$$

Похожим образом определяются положительно определённая матрица и другие.

Свойство скалярного произведения:

$$(Bx, y) = (x, B^*y)$$

$$(A^*Ax, x) = (Ax, Ax)) \geq 0$$

Из курса линейной алгебры известно, что все собственные значения симметрической матрицы вещественны.

Значит, $\|A\|_2 = \sqrt{\max_i \lambda_i(A^*A)}$ имеет смысл для симметрических матриц.

У симметрических матриц существует полная система ортонормированных собственных векторов.

$$(A^*A)u^{(i)} = \lambda_i u^{(i)}$$

$$\left(u^{(i)}, u^{(j)}\right) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

$$x = \sum_{i=1}^{n} \alpha_i u^{(i)}$$

$$\|x\|_2 = \sqrt{(x,x)} = \sqrt{\left(\sum_{i=1}^{n} \alpha_i u^{(i)}, \sum_{j=1}^{n} \alpha_j u^{(j)}\right)} =$$

$$= \sqrt{\sum_{i=1}^{n} \alpha_i^2}$$

$$\|Ax\|_2 = \sqrt{(Ax, Ax)} = \sqrt{(A^*Ax, x)} =$$

$$= \sqrt{\left(A^*A \sum_i \alpha_i u^{(i)}, \sum_j \alpha_j u^{(j)}\right)} = \sqrt{\left(\sum_i \alpha_i A^*A u^{(i)}, \sum_j \alpha_j u^{(j)}\right)} =$$

$$= \sqrt{\left(\sum_i \alpha_i \lambda_i u^{(i)}, \sum_j \alpha_j u^{(j)}\right)} = \sqrt{\sum_i \alpha_i^2 \lambda_i} \leq$$

$$\leq \sqrt{\sum_i \alpha_i^2 \max_j \lambda_i(A^*A)} = \sqrt{\max_i \lambda_i(A^*A)} \|x\|_2$$

$$\lambda_k(A^*A) = \max_i \lambda_i(A^*A)$$

Посчитаем вторую норму. Предположим, что матрица симметричная:

$$A^* = A$$

$$\|A\|_2 = \sqrt{\max_i \lambda_i(A^*A)} = \sqrt{\max_i \lambda_i(A^2)} = \sqrt{\max_i \lambda_i^2(A)} = \max_i |\lambda_i(A)|$$

`1 Октября 2024`

3. Обусловленность матрицы систем

Определение

Число обусловленности:

$$\mu(A) = \sup_{x,y \neq 0} \left(\frac{\|Ax\|}{\|x\|} \div \frac{\|Ay\|}{\|y\|} \right)$$

$$Ax = f$$

$$A(x + \xi) = f + \varphi$$

$$A\xi = \varphi$$

Попробуем оценить следующую величину:

$$\frac{\|\xi\|}{\|x\|} \cdot \frac{\|f\|}{\|\varphi\|} = \frac{\|\xi\|}{\|x\|} \cdot \frac{\|Ax\|}{\|A\xi\|} = \frac{\|Ax\|}{\|x\|} \div \frac{\|A\xi\|}{\|\xi\|} \leq \mu(A)$$

$$\frac{\|\xi\|}{\|x\|} \leq \mu(a) \frac{\|\varphi\|}{\|f\|}$$

Представим, что $\mu(A)$ невелико, меньше единицы оно быть не может, тогда маленькое возмущение правой части гарантирует, что возмущение в решении тоже невелико.

Как влияет погрешность матрицы на результат:

$$(A + \Sigma)(x + \xi) = f + \varphi$$

$$\frac{\|\xi\|}{\|x\|} \leq \frac{\mu(A)}{1 - \mu(A)\frac{\|\Sigma\|}{\|A\|}} \left(\frac{\|\varphi\|}{\|f\|} + \frac{\|\Sigma\|}{\|A\|} \right)$$

При условии:

$$\left\|A^{-1}\right\| \|\Sigma\| < 1$$

Нам нужна формула, по которой мы будем искать число обусловленности, выведем её.

$$\mu(A) = \sup_{x \neq 0} \frac{\frac{\|Ax\|}{\|x\|}}{\inf_{y \neq 0}\left(\frac{\|Ay\|}{\|y\|}\right)} = \frac{\|A\|}{\inf_{z \neq 0} \frac{1}{\frac{\|A^{-1}z\|}{\|z\|}}} =$$

$$= \frac{\|A\|}{\frac{1}{\sup_{z \neq 0} \frac{\|A^{-1}z\|}{\|z\|}}} = \|A\| \left\|A^{-1}\right\|$$

$$z := Ay$$

Пример

$$A = \begin{pmatrix} 1 & 1.0001 \\ 1.0001 & 1 \end{pmatrix} \qquad f = \begin{pmatrix} 2.0001 \\ 2.0001 \end{pmatrix}$$

$$x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\|A\| = \max_{i=1,2} |\lambda_i(A)|$$

$$\begin{vmatrix} 1 - \lambda & 1.0001 \\ 1.0001 & 1 - \lambda \end{vmatrix} = 0$$

$$(1 - \lambda)^2 - 1.0001^2 = 0$$

$$\lambda_1 = 0.0001 \qquad \lambda_2 = -10^{-4}$$

$$\|A\| = 2.0001 \qquad \|A^{-1}\| = 10^4$$

$$\mu(A) = 2001 \text{ большое число}$$

Рассмотрим следующую систему:

$$(A + \Sigma)y = f + \varphi$$

$$\Sigma = \begin{pmatrix} 0 & 0 \\ 0 & 0.00005 \end{pmatrix} \qquad \varphi = \begin{pmatrix} 0.0001 \\ 0 \end{pmatrix}$$

Решение системы:

$$y = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

4. Прямые методы

4.1. Метод Гаусса (метод исключения неизвестных)

Метод в основе которого лежит использование элементарных преобразований матрицы с целью применения её к треугольному или диагональному виду.

$$Ax = f$$

$$\begin{cases} a_{11}x_1 + ... + a_{1n}x_n = f_1 \\ ... \\ a_{n1}x_1 + ... + a_{nn}x_n = f_n \end{cases}$$

Исключаем элемент x_1. Пусть $a_{11} \neq 0$:

$$x_1 + \frac{a_{12}}{a_{11}}x_2 + ... + \frac{a_{1n}}{a_{11}}x_n = \frac{f_1}{a_{11}} \;\mid \cdot\, a_{21}$$

$$a_{21}x_1 + a_{21}\frac{a_{12}}{a_{11}}x_2 + ... + a_{21}\frac{a_{1n}}{a_{11}}x_n = a_{21}\frac{f_1}{a_{11}}$$

Теперь из второго уравнения системы вычитаем вот это, таким образом получим уравнение, в котором нет x_1.

$$\begin{cases} a_{11}x_1 + a_{22}x_2 + \dots + a_{1n}x_n = f_1 \\ \qquad\qquad a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n = f_2^{(1)} \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = f_n \end{cases}$$

$$\begin{cases} a_{11}x_1 + a_{22}x_2 + \dots + a_{1n}x_n = f_1 \\ \qquad\qquad a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n = f_2^{(1)} \\ \dots \\ \qquad\qquad a_{n2}^{(1)}x_2 + \dots + a_{nn}^{(1)}x_n = f_n^{(1)} \end{cases}$$

Второй шаг. $a_{22}^{(1)} \neq 0$ и делаем аналогично.

После $n-1$ шагов получим матрицу, которая имеет верхний треугольный вид.

Теперь запишем формулы всех этих шагов.

1 шаг:

$$d_{i1} = \frac{a_{i1}}{a_{11}}$$

$$a_{ij}^{(1)} = a_{ij} - d_{i1}a_{1j}, \quad i, j = 2, \dots, n$$

$$f_i^{(1)} = f_i - d_{i1}f_1$$

2 шаг:

$$d_{i2} = \frac{a_{i2}^{(1)}}{a_{22}^{(1)}}$$

$$a_{ij}^{(2)} = a_{ij}^{(1)} - d_{i2}a_{2j}^{(1)}, \quad i, j = 3, \dots, n$$

$$f_i^{(2)} = f_i^{(1)} - d_{i2}f_2^{(1)}$$

$$A_{n-1}x = f^{(n-1)}$$

$$A_{n-1} =: U \text{ верхняя треугольная матрица}$$

$$f^{(n-1)} =: g$$

$$Ux = g$$

$$u_{nn}x_n = g_n$$

$$x_n = \frac{g_n}{u_{nn}}$$

k — й шаг обратного хода:

$$u_{kk}x_k + \sum_{j=k+1}^{n} u_{kj}x_j = g_k$$

$$x_k = \frac{g_k - \sum_{j=k+1}^{n} u_{kj}x_j}{u_{kk}}$$

Условие, при котором ведущий элемент не равен нулю. Во-первых матрица должна быть невырожденной.

Критерий работы метода Гаусса

Отличие от нуля всех главных миноров.

$$\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{pmatrix} \neq 0$$

$$a_{11} \neq 0$$

$$\begin{vmatrix} a_{11} & a_{12} \\ 0 & a_{22}^{(1)} \end{vmatrix} \neq 0$$

$$a_{11}a_{22}^{(1)} \neq 0$$

$$a_{22}^{(1)} \neq 0$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & 0 & a_{33}^{(2)} \end{vmatrix} \neq 0$$

$$a_{11}a_{22}^{(1)}a_{33}^{(2)} \neq 0$$

$$a_{33}^{(2)} \neq 0$$

Все главные элементы отличны от нуля.

4.2. LU-разложене

Тот же метод Гаусса, но записанный через разложение матриц.

$$D_1 := \begin{pmatrix} 1 & 0 & 0 & & \\ -d_{21} & 1 & 0 & & \\ -d_{31} & 0 & 1 & & \\ \vdots & & & & \\ -d_{n1} & 0 & 0 & \dots & 1 \end{pmatrix}$$

$$A_1 = D_1 A$$

$$D_2 := \begin{pmatrix} 1 & & & \\ 0 & 1 & \dots & 0 \\ 0 & -d_{32} & \dots & \\ 0 & -d_{42} & \dots & \\ \vdots & & & \\ 0 & -d_{n2} & \dots & 1 \end{pmatrix}$$

$$A_2 = D_2 A_1$$

$$A_{n-1}$$

$$U = A_{n-1} = D_{n-1}A_{n-2} = D_{n-1}D_{n-2}A_{n-3} = \dots$$

$$U = D_{n-1}D_{n-2}\dots D_1 A$$

$$D := D_{n-1}D_{n-2}\dots D_1$$

$$U = DA$$

$$L := D^{(-1)}$$

$$A = LU$$

Как выглядит эта матрица:

$$L = \begin{pmatrix} 1 & & & & \\ d_{21} & 1 & & & \\ d_{31} & d_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ d_{n1} & d_{n2} & d_{n3} & \dots & 1 \end{pmatrix}$$

$$LUx = f$$

$$Ux := y$$

$$Ly = f$$

Докажем необходимость в критерии работы метода Гаусса.

$$A = LU$$

$$A^{(k)} = L^{(k)}U^{(k)}$$

Подматрица k-го порядка.

$$\det A^{(k)} = \det L^{(k)} \det U^{(k)}$$

$$\det L^{(k)} = 1$$

$$\det U^{(k)} = u_{11}u_{12}\dots u_{kk}$$

$$u_{ii} \neq 0, \quad i = \overline{1,k}$$

$$\det A^{(k)} \neq 0$$

08 Октября 2024

4.3. Метод квадратного корня

Считаем, что все главные миноры не равны нулю.

$$A = LU$$

$$U = DV$$

$$D = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix} \qquad V = \begin{pmatrix} 1 & v_{12} & v_{13} & \dots & v_{1n} \\ & 1 & v_{23} & \dots & v_{2n} \\ & & & & \vdots \\ & & & & 1 \end{pmatrix}$$

$$U = \begin{pmatrix} d_1 & d_1 v_{12} & d_1 v_{13} & \dots & d_1 v_{1n} \\ & d_2 & d_2 v_{23} & \dots & d_2 v_{2n} \\ & & & & \vdots \\ & & & & d_n \end{pmatrix}$$

$$d_i = u_{ii} \quad v_{ij} = \frac{u_{ij}}{u_{ii}}$$

Есть далее понятие LDV-разложение, оно определяется единственным образом.

Пусть A - симметрическая:

$$A = A^*$$

$$A = LDV$$

$$A^* = (LDV)^* = V^* D^* L^* = (V^* D L^*) = A$$

$$\Rightarrow L = V^*, V = L^*$$

$$A = V^* D V$$

Пусть матрица A положительно определённая. Тогда по критерию Сильвестра все её угловые миноры положительны.

$$(Ax, x) > 0$$

$$(Dx, x) = (DVy, Vy) = (Ay, y) > 0$$

$$x = Vy$$

Все элементы на диагонали в D больше нуля.

$$D^{\frac{1}{2}} = \operatorname{diag}\left(\sqrt{d_1}, ..., \sqrt{d_n}\right)$$

$$A = V^* D^{\frac{1}{2}} D^{\frac{1}{2}} V$$

$$W = D^{\frac{1}{2}} V$$

$$W^* = V^* D^{\frac{1}{2}}$$

$$A = W^* W$$

$$Ax = f$$

$$W^* W x = f$$

$$y := Wx$$

$$W^* y = f$$

$$a_{ij} = \sum_{k=1}^{n} W_{ik}^* W_{kj} = \sum_{k=1}^{n} W_{ki} W_{kj} = \sum_{k=1}^{i} W_{ki} W_{kj}$$

$$a_{11} = W_{11} W_{11} \Rightarrow W_{11} = \sqrt{a_{11}}$$

$$i = 1 \quad j = 2, 3, ..., n :$$

$$a_{1j} = W_{11} W_{1j}$$

$$W_{1j} = \frac{a_{1j}}{W_{11}}$$

$$a_{ii} = \sum_{k=1}^{i} W_{ki} W_{ki} = \sum_{k=1}^{i-1} W_{ki}^2 + W_{ii}^2$$

$$W_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} W_{ki}^2}$$

$$a_{ij} = \sum_{k=1}^{i-1} W_{ki} W_{kj} + W_{ii} W_{ij}$$

$$W_{ij} = \frac{a_{ij} - \sum_{k=1}^{i-1} W_{ki} W_{kj}}{W_{ii}}$$

$$j = i + 1, ..., n$$

Это был метод квадратного корня.

Если матрица симметрическая, но не является положительно определённой, значит решением будут комплексные числа, ничего страшного).

4.4. QR разложение (Пуэр разложение)

Определение

Гиперплоскость с нормалью $p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$, проходящая через начало координат

$$(p, x) = p_1 x_1 + p_2 x_2 + ... + p_n x_n = 0$$

Ортогональное преобразование отражения относительно гиперплоскости, проходящей через начало координат:

$$y = x - 2 \frac{(p, x)}{(p, p)} p$$

$$y = x - 2 \frac{(p, x)}{(p, p)} p = x - \frac{2}{(p, p)} (p, x) p = x - \frac{2}{(p, p)} p p^* x =$$
$$= \left(I - \frac{2}{(p, p)} p p^* \right) x$$

$$P := I - \frac{2}{(p,p)} pp^*$$

$$y = Px$$

Свойства:

1. $P^2 = I$

$$P^2 x = P(Px) = P\left(x - 2\frac{(p,x)}{(p,p)}p\right) = Px - 2\frac{(p,x)}{(p,p)}Pp =$$

$$= Px - 2\frac{(p,x)}{(p,p)}\left(p - 2\frac{(p,p)}{(p,p)}p\right) = x - \cancel{2\frac{(p,x)}{(p,p)}} + \cancel{2\frac{(p,x)}{p,p}} = x$$

2. $P^* = P$

$$P^* = \left(I - \frac{2}{(p,p)}pp^*\right)^* = I - \frac{2}{(p,p)}p^{**}p^* = p$$

3. Матрица ортогональная $P^*P = P^2 = I$

4. Преобразование отражения не изменится, если вместо вектора нормали p взять вектор βp

$$P = I - \frac{2}{(\beta p, \beta p)}(\beta p)(\beta p)^* = I - \frac{2}{(p,p)}p^*$$

5. Вектор нормали p можно определить как разность между исходным и отражённым вектором: $p = x - y$

$$x - y = 2\frac{(p,x)}{(p,p)}p$$

6. $p_1 = p_2 = ... = p_k = 0$

   $\Rightarrow y_1 = x_1, y_2 = x_2, ..., y_k = x_k$

   Докажем:

$$y_i = x_i - 2\frac{p_i}{\sum_{l=1}^{n} p_l^2} \sum_{l=1}^{n} p_l x_l$$

Из этой формулы очевидно (я хз??? он так сказал на лекции)

7. $p_1 = p_2 = \ldots = p_k = 0$ и $x_{k+1} = \ldots = x_n = 0$

   $\Rightarrow y_{k+1} = \ldots = y_n = 0$

$$A = QR,$$

где Q ортогональная матрица, а R верхняя треугольная

Получим матрицу A_1 следующим образом:

$$A_1 = P_1 A$$

Как должен выглядеть первый столбец:

$$\begin{pmatrix} a_{11}^{(1)} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Введём обозначение столбцов матрицы A и A_1: $a_j, a_j^{(1)}$.

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} = a_1 \quad \begin{pmatrix} a_{11}^{(1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_1^{(1)}$$

$$P_1 a_1 = \begin{pmatrix} a_{11}^{(1)} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Вектор нормали $p^{(1)} = a_1 - a_1^{(1)}$ (разность между исходным и отражённым)

$$\begin{aligned} &p_1^{(1)} = a_{11} - a_{11}^{(1)} \\ &p_2^{(1)} = a_{21} \\ &\vdots \\ &p_n^{(1)} = a_{n1} \end{aligned}$$

При ортогональном преобразовании длина вектора не меняется:

$$\left\| a_1^{(1)} \right\|_2^2 = \|a_1\|_2^2$$

$$\left(a_{11}^{(1)} \right)^2 = \sum_{l=1}^{n} a_{l1}^2$$

$$a_{11}^{(1)} = \pm \sqrt{\sum_{l=1}^{n} a_{l1}^2}$$

$$p_1^{(1)} = a_{11} \pm \sqrt{\sum_{l=1}^{n} a_{l1}^2}$$

$$\sigma_1 := \begin{cases} 1, \ a_{11} \geq 0 \\ -1, \ a_{11} < 0 \end{cases}$$

$$p_1^{(1)} = a_{11} + \sigma_1 \sqrt{\sum_{l=1}^{n} a_{li}^2}$$

$$a_j^{(1)} = a_j - 2 \frac{\left(p^{(1)}, a_j \right)}{\left(p_1^{(1)}, p_1^{(1)} \right)} p^{(1)}, \ \text{где } j = 2, ..., n$$

По этой формуле полностью будет определена матрица A_1.

15 Октября 2024

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1,k-1}^{(1)} & a_{1k}^{(1)} & a_{1,k+1}^{(1)} & \dots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \dots & a_{2,k-1}^{(2)} & a_{2k}^{(1)} & a_{2,k+1}^{(1)} & \dots & a_{2n}^{(1)} \\ & & \ddots & & & & & \\ & & & a_{k-1,k-1}^{(k-1)} & a_{k-1,k}^{(k-1)} & a_{k-1,k+1}^{(k-1)} & \dots & a_{k-1,n}^{(k-1)} \\ & & & & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \dots & d_{k,n}^{(k)} \\ & & & & a_{k+1,k}^{(k+1)} & a_{k+1,k+1}^{(k+1)} & \dots & d_{k+1,n}^{(k+1)} \\ & & & & a_{n-1,k}^{(n-1)} & a_{n-1,k+1}^{(n-1)} & \dots & d_{n-1,n}^{(n-1)} \end{pmatrix}$$

$$A_k = P_k A_{k-1}$$

$$a_k^{(k)} = P_k a_k^{(k-1)} = \begin{pmatrix} a_{1k}^{(1)} \\ a_{2k}^{(2)} \\ \vdots \\ a_{n-1,k}^{(k-1)} \\ a_{kk}^{(k)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$p^{(k)} = a_k^{(k-1)} - a_k^{(k)}$$

$$p_l^{(k)} = 0, \quad l = 1, 2, ..., n-1$$

$$p_k^{(k)} = a_{kk}^{(k-1)} - a_{kk}^{(k)}$$

$$p_l^{(k)} = a_{lk}^{(k-1)}, \quad l = k+1, ..., n$$

$$\left\| a_k^{(k+1)} \right\|_2^2 = \left\| a_k^{(k)} \right\|_2^2$$

$$a_{kk}^{(k)} = \pm\sqrt{\sum_{l=k}^{n}\left(a_{lk}^{(k-1)}\right)^2}$$

$$\sigma_k = \begin{cases} 1,\ a_{kk}^{(k-1)} \geq 0 \\ -1,\ a_{kk}^{(k-1)} < 0 \end{cases}$$

$$p_k^{(k)} = a_{kk}^{(k-1)} + \sigma_k\sqrt{\sum_{l=k}^{n}\left(a_{lk}^{(k-1)}\right)^2}$$

$$a_j^{(k)} = a_j^{(k-1)} - 2\frac{p^{(k)}, a_j^{(k-1)}}{\left(p^{(k)}, p^{(k)}\right)}p^k$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - 2\frac{p_i^{(k)}}{\sum_{l=n}^{n}\left(p_l^{(k)}\right)^2}\sum_{l=k}^{n} p_l^{(k)} a_{lj}^{(k-1)}$$

$$j = k+1, ..., n$$

$$p_1^{(k)} = p_2^{(k)} = ... = p_{k-1}^{(k)} = 0$$

Нужно сделать $n-1$ шаг. Рассмотрим весь процесс в целом.

$$R = A_{n-1} = P_{n-1}A_{n-2} = P_{n-1}P_{n-2}A_{n-3} = ... =$$
$$= P_{n-1}P_{n-2}...P_2P_1A$$

$$Q \coloneqq P_1P_2...P_{n-1}$$

Матрица Q ортогональная (как произведение ортогональных матриц).

$$Q^* = P_{n-1}^*...P_1^* = P_{n-1}...P_1$$

$$R = Q^*A$$

$$A = QR$$

Требования для существования этого разложения:

Рассмотрим случай, когда мы делим на ноль.

$$\sum_{l=k}^{n} \left(p_l^{(k)}\right)^2 = 0 \Leftrightarrow p_l^{(k)} = 0,\ l = k, ..., n$$

Подберём вектор нормали, который будет ненулевым, но оставит вектор на месте. Чтобы формально в алгоритме шаг выполнился.

Например,

$$p^{(k)} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sqrt{2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Теперь можем говорить, что QR разложение существует для любых матриц.

Рассмотрим вопрос решения уравнений с помощью QR разложений.

$$Ax = f$$

$$QRx = f$$

$$Rx = Q^* f$$

Распишем Q^*:

$$Q^* = P_{n-1} ... P_1$$

$$Q^* f := g$$

$$f^{(1)} = P_1 f$$

$$f^{(2)} = P_2 f^{(1)}$$

$$g = f^{(n-1)} = P_{n-1} f^{(n-2)}$$

$$f^{(1)} = f - 2\frac{\left(p^{(1)}, f\right)}{\left(p^{(1)}, p^{(1)}\right)} p^{(1)}$$

$$f^{(k)} = f^{(k-1)} - 2\frac{p^{(k)}}{\sum_{l=k}^{n}\left(p_l^{(k)}\right)^2} \sum_{l=k}^{n} p_l^{(k)} f_l^{(k-1)}, \quad i = k, k+1, ..., n$$

4.5. Метод окаймления

Способ отыскания обратной матрицы.

$$A = A_n = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n-1} & a_{2n} \\ \vdots & & & & \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} & a_{n-1,n} \\ \hline a_{n,1} & a_{n,2} & \dots & a_{n,n-1} & a_{n,n} \end{array}\right)$$

$$A_{n-1}$$

$$v_n = \left(a_{n1}, a_{n2}, ..., a_{n,n-1}\right)$$

$$u_n = \left(a_{1n}, a_{2,n}, ..., a_{n-1,n}\right)^*$$

$$A_n = \begin{pmatrix} A_n - 1 & u_n \\ v_n & a_{nn} \end{pmatrix}$$

$$A_n^{-1} = \begin{pmatrix} P_{n-1} & r_n \\ q_n & \frac{1}{\alpha_n} \end{pmatrix} = D_n$$

Предполагаем, что A_{n-1}^{-1} известна

$$\begin{pmatrix} A_n - 1 & u_n \\ v_n & a_{nn} \end{pmatrix} \begin{pmatrix} P_{n-1} & r_n \\ q_n & \frac{1}{\alpha_n} \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{cases} A_{n-1}P_{n-1} + u_n q_n = E \\ v_n P_{n-1} + a_{nn} q_n = 0 \\ A_{n-1} r_n + \frac{u_n}{\alpha_n} = 0 \\ v_n r_n + \frac{a_{nn}}{\alpha_n} = 1 \end{cases}$$

Возмьём третье и выразим:

$$\boxed{r_n = -A_{n-1}^{-1} \frac{u_n}{\alpha_n}}$$

Подставляем в четвёртое:

$$-\frac{v_n A_{n-1}^{-1} u_n}{\alpha_n} + \frac{a_{nn}}{\alpha_n} = 1$$

$$\boxed{\alpha_n = a_{nn} - v_n A_{n-1}^{-1} u_n}$$

$$\boxed{P_{n-1} = A_{n-1}^{-1} - A_{n-1}^{-1} u_n q_n}$$

$$v_n \left(A_{n-1}^{-1} - A_{n-1}^{-1} u_n q_n\right) + a_{nn} q_n = 0$$

$$v_n A_{n-1}^{-1} + (a_{nn} - v_n A_{n-1}^{-1} u_n) q_n = 0$$

$$v_n A_{n-1}^{-1} + \alpha_n q_n = 0$$

$$\boxed{q_n = -\frac{v_n A_{n-1}^{-1}}{\alpha_n}}$$

1. $$-A_{n-1}^{-1} u_n \qquad \left(\beta_{1n}, ..., \beta_{n-1,n}\right)^*$$
2. $$-v_n A_{n-1}^{-1} \qquad \left(\gamma_{n1}, ..., \gamma_{n,n-1}\right)$$
3. $$\alpha_n = a_{nn} + \sum_{i=1}^{n-1} a_{ni} \beta_{in} = a_{nn} + \sum_{i=1}^{n-1} a_{in} \gamma_{ni}$$

4. $$d_{ik} = d_{ik} + \frac{\beta_{ik}\gamma_{nk}}{\alpha_n}, \quad i, k \leq n - 1$$

$$d_{in} = \frac{\beta_{in}}{\alpha_n} \quad d_{nk} = \frac{\gamma_{nk}}{\alpha_n}$$

Работа в первом блоке завершена. Можно сдавать первый коллоквиум по прямым методам. Д947.

`22 Октября 2024`

5. Итерационные методы

$$Ax = f$$

$$\left\{x^{(k)}\right\}$$

$$\lim_{k \to \infty} x^{(k)} = x^*$$

$$\lim_{k \to \infty} \left\|x^{(k)} - x^*\right\| = 0$$

Каноническая форма записи:

$$B \frac{x^{(k+1)} - x^{(k)}}{\tau_k} + Ax^{(k)} = f, \ k = 0, 1, ...$$

$x^{(0)}$ начальное приближение

B – невырожденная матрица

τ – итерационный параметр (может меняться на разных шагах)

Если τ не зависит от k, то метод называется стационарным.

Если методы сходятся, то они сходятся к точному решению. Поэтому нужно доказывать только сходимость.

Скорость сходимости. Надо добиться наивысшей скорости сходимости. Зависит от параметра и от матрицы B.

Сейчас нас будут интересовать только стационарные методы:

$$B\frac{x^{(k+1)} - x^{(k)}}{\tau} + Ax^{(k)} = f,\ k = 0, 1, \dots$$

$$Bx^{(k+1)} = Bx^{(k)} - \tau Ax^{(k)} + \tau f$$

$$Bx^{(k+1)} = (B - \tau A)x^{(k)} + \tau f$$

Матрица B должна быть простая, чтобы систему можно было просто решать, а то игра не стоит свеч.

Показатели:

1. Вектор погрешности или вектор ошибки $z^{(k)} = x^{(k)} - x^*$.
2. Вектор невязки: $r^{(k)} = Ax^{(k)} - f$

$$\lim_{k\to\infty} \left\|z^{(k)}\right\| = 0 - \text{сходимость к точному решению}$$

Добавим и отнимем точное решение:

$$B\frac{x^{(k+1)} - x^* - x^{(k)} + x^*}{\tau} + Ax^{(k)} - Ax^* = 0$$

$$B\frac{z^{(k+1)} - z^k}{\tau} + Az^{(k)} = 0$$

$$Bz^{(k+1)} = Bz^{(k)} - \tau Az^{(k)}$$

$$z^{(k+1)} = \left(I - \tau B^{-1}A\right)z^{(k)}$$

Матрица перехода $S = \left(I - \tau B^{-1}A\right)$:

$$z^{(k+1)} = Sz^{(k)}$$

Для того, чтобы можно было сравнивать скорости сходимости для разных методом существует асимптотическая скорость сходимости.

$$\left\|z^{(k)}\right\| \le \frac{1}{e}\left\|z^{(0)}\right\|$$

Сколько шагов нужно сделать, чтобы начальная ошибка уменьшилась в e раз.

$$z^{(k+1)} = Sz^{(k)} \Rightarrow z^{(k)} = S^k z^{(0)}$$

$$\left\|z^{(k)}\right\| \leq \|S\|^k \left\|z^{(0)}\right\|$$

Самый грубый подход. Потребуем, чтобы:

$$\|S\|^k \leq \frac{1}{e}$$

Вытащим отсюда k:

$$k \ln\|S\| \leq -1$$

Для сходящихся методов норма матрицы всегда меньше единицы. Мы это докажем.

$$k \geq \frac{1}{-\ln\|S\|}$$

$$R := -\ln\|N\| - \text{ассимптотическая скорость сходимости}$$

Теперь рассмотрим основополагающую теорему о сходимости стационарных методов.

Критерий сходимости

Для того, чтобы стационарный метод сходился при любом начальном приближении, небходимо и достаточно, чтобы все собственные значения матрицы перехода были по модулю меньше единицы.

$$|(\lambda(S))| < 1$$

Критерий плох тем, что надо постоянно искать собственные значения. Поэтому это теорема базовая.

Докажем необходимость. Пусть метод сходится при любом начальном приближении. Докажем, что все собственные значения по модулю меньше единицы. От противного. Пусть метод сходится, но существует хотя бы одно $|\lambda(S)| \geq 1$. Обозначим за u собственный вектор, который соответствует этому собственному значению. $Su = \lambda u$. В качестве начального приближения возьмём $x^{(0)} := x^* + u$. Тогда $z^{(0)} = u$.

$$z^{(k)} = S^k z^{(0)} = S^k u = \lambda^k u$$

$$\left\|z^{(k)}\right\| = \left\|\lambda^k u\right\| = |\lambda|^k \|u\|$$

$$u \neq 0, \quad k \to \infty$$

$$\left\|z^{(k)}\right\| \underset{k \to \infty}{\not\to} 0$$

Получили противоречие.

Сформулируем лемму, которая понадобится для доказательства достаточности.

Лемма

Пусть все собственные значения матрицы S:

$$|\lambda_i(s)| < q, \ i = 1, ..., n$$

Тогда существует такая невырожденная матрица T, что матрица $\Lambda = TST^{-1}$ удовлетворяет такому условию: $\|\Lambda\|_\infty \leq q$

Напоминание:

$$\Lambda \sim S \Rightarrow \text{собственные значения у них совпадают}$$

Теперь докажем достаточность.

$$|\lambda(S)| < 1$$

$$q < 1$$

$$\max_i |\lambda_i(s)| < q$$

Надо показать, что $\lim_{k\to\infty} \left\| z^{(k)} \right\| = 0$.

Теперь привлекаем лемму. Существует такая невырожденная матрица T, что $\|\Lambda\|_\infty \le q$.

$$S = T^{-1}\Lambda T$$

Подставим S в определение $z^{(k)}$.

$$S^k = T^{-1}\Lambda^k T$$

$$z^{(k)} = S^k z^{(0)} = T^{-1}\Lambda^k T z^{(0)}$$

$$\left\| z^{(k)} \right\|_\infty = \left\| S^k z^{(0)} = T^{-1}\Lambda^k T z^{(0)} \right\|_\infty \le$$

$$\le \left\| T^{-1} \right\|_\infty \|\Lambda\|_\infty^k \|T\|_\infty \left\| z^{(0)} \right\|_\infty \le$$

$$\le \mu_{\infty(T)} \left\| \Lambda^k \right\|_\infty \left\| z^{(0)} \right\|_\infty \le \mu_\infty(T) q^k \left\| z^{(0)} \right\|_\infty \xrightarrow[k\to\infty]{} 0$$

Теорема

Для сходимости двуслойного стационарного метода при любом начальном приближении достаточно, чтобы хотя бы одна из норм матрицы перехода S, согласованная с какой-нибудь векторной, была меньше единицы.

$$\|S\| < 1$$

Собственное значение и вектор:

$$Su = \lambda u$$

$$|\lambda| \|u\| = \|\lambda u\| = \|Su\| \le \|S\| \|u\|$$

$$|\lambda| \le \|S\|$$

Собственное значение всегда ограничено нормой.

$$\|S\| < 1 \Rightarrow |\lambda| < 1$$

`29 Октября 2024`

5.1. Метод простой итерации

Когда матрица $B = I$

Тогда матрица перехода:

$$S = I - \tau A$$

$$x^{(k+1)} = x^{(k)} - \tau\left(Ax^{(k)} - f\right)$$

Оптимизация скорости сходимости по параметру τ, то есть найдём такой τ оптимальный.

Асимптотическая скорость сходимости $R = -\ln\|S\|$.

Нужно максимизировать R, то есть, минимизировать $\|S\|$.

Предположим, что матрица A симметрическая и положительно определённая.

$$A = A^* > 0$$

Все собственные значения вещественные и положительные. Обозначим минимальное и максимальное собственное значение за α и β:

$$0 < \alpha \leq \lambda_i(A) \leq \beta$$

.

$$\|S\|_2 = \max|\lambda_i(S)|$$

$$\lambda_i(S) = \lambda_i(I - \tau A) = \text{свойства линейности} = 1 - \tau\lambda_i(A)$$

$$\|S\|_2 = \max_{1 \leq i \leq n}|1 - \tau\lambda_i(A)|$$

Рассмотрим функцию на отрезке $[\alpha, \beta]$:

$$f(\tau, \lambda) = |1 - \tau\lambda|$$

Принцип максимума (на веру):

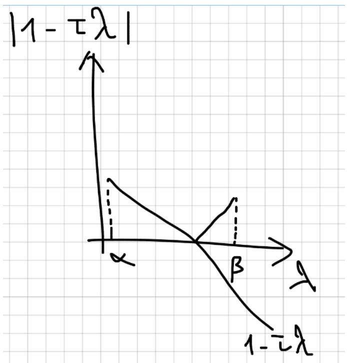
$$\max_{\alpha \le \lambda < \beta} |1 - \lambda\tau| = \max(f(\tau, \alpha), f(\tau, \beta))$$

Рассмотрим три случая (на самом деле один, остальные дома рассмотрим).

1. $\frac{1}{\beta} \le \tau \le \frac{1}{\alpha}$

$$1 - \tau\alpha \ge 1 - \frac{1}{\alpha}\alpha = 0$$

$$1 - \tau\beta \le 1 - \frac{1}{\beta}\beta = 0$$



Следующее очев, потому что слева n значений, а справа отрезок $[\alpha, \beta]$:

$$\max_{1 \le i \le n} |1 - \tau\lambda_i(A)| \le \max_{\alpha \le \lambda \le \beta} f(\tau, \lambda)$$

$$\max_{1 \le i \le n} |1 - \tau\lambda_i(A)| \le \max_{\alpha \le \lambda \le \beta} f(\tau, \lambda) = \max(f(\tau, \alpha), f(\tau, \beta))$$

Следующее очевидно по аналогичным причинам: слева n значений, справа — два:

$$\max_{1 \le i \le n} |1 - \tau\lambda_i(A)| \ge \max(|1 - \tau\alpha|, |1 - \tau\beta|)$$

$$\max_{1\le i\le n}|1-\tau\lambda_i(A)| \ge \max(|1-\tau\alpha|, |1-\tau\beta|) = \max(f(\tau,\alpha), f(\tau,\beta))$$

Получаем, что:

$$\max_{1\le i\le n}|1-\tau\lambda_i(A)| = \|S\|_2 = \max(f(\tau,\alpha), f(\tau,\beta))$$

Вот оно оптимальное значение:

$$\tau_0 = \frac{2}{\alpha+\beta}$$

Проведём доказательство следующим образом:

$$1-\tau_0\alpha = 1-\frac{2}{\alpha+\beta}\alpha = \frac{\beta-\alpha}{\alpha+\beta} > 0$$

$$1-\tau_0\beta = 1-\frac{2}{\alpha+\beta}\beta = \frac{\alpha-\beta}{\alpha+\beta} < 0$$

$$f(\tau_0,\alpha) = \frac{\beta-\alpha}{\beta+\alpha}$$

$$f(\tau_0,\beta) = \frac{\beta-\alpha}{\beta+\alpha}$$

$$\|S\|_2 = \frac{\beta-\alpha}{\beta+\alpha} < 1$$

Рассмотрим первый случай $\tau < \tau_0$:

$$1-\tau\alpha > 1-\tau_0\alpha = \frac{\beta-\alpha}{\beta+\alpha}$$

$$f(\tau,\alpha) = |1-\tau\alpha| > \frac{\beta-\alpha}{\beta+\alpha}$$

$$\max(f(\tau,\alpha), f(\tau,\beta)) \ge f(\tau,\alpha) > \frac{\beta-\alpha}{\beta+\alpha}$$

Получили, что $\|S\|_2 > \frac{\beta-\alpha}{\beta+\alpha}$

Аналогично во втором случае $\tau > \tau_0$.

$$R = -\ln \|S\|_2 = -\ln \frac{\beta - \alpha}{\beta + \alpha}$$

К чему можно преобразовать эту формулу асимптотической скорости сходимости? Свяжем её с определением качества сходимости, которое мы назвали числом обусловленности.

$$\mu_2(A) = \|A\|_2 \left\|A^{-1}\right\|_2 = \max_i \lambda_i(A) \cdot \max_i \lambda_i(A^{-1}) =$$

$$= \max_i \lambda_i(A) \cdot \max_i \frac{1}{\lambda_i}(A) = \frac{\max_i \lambda_i(A)}{\min_i \lambda_i(A)} = \frac{\beta}{\alpha}$$

$$R = -\ln \frac{\beta - \alpha}{\beta + \alpha} = \ln \frac{\beta + \alpha}{\beta - \alpha} = \ln \frac{\frac{\beta}{\alpha} + 1}{\frac{\beta}{\alpha} - 1} = \ln \frac{\mu_2(A) + 1}{\mu_2(A) - 1}$$

5.2. Методы Якоби (якобы Метод)

Тут матрица B будет диагональной.

$$A \coloneqq A_l + D + A_v$$

A_l — нижнетреугольная, A_v — верхнетреугольная, $D = \operatorname{diag}(a_{11}, ..., a_{nn})$.

Пусть все диагональные элементы отличны от нуля $\alpha_{ii} \neq 0$.

$$B \coloneqq D,\ \tau \coloneqq 1$$

Подставим в каноническую форму записи:

$$D\left(x^{(k+1)} - x^{(k)}\right) + (A_l + D + A_v)x^{(k)} = f$$

$$Dx^{(k+1)} - Dx^{(k)} + A_l x^{(k)} + Dx^{(k)} + A_v x^{(k)} = f$$

$$Dx^{(k+1)} = f - (A_l + A_v)x^{(k)}$$

Эту матричную запись легко переписать в скалярный за счёт диагональной матрицы.

$$a_{ii}x_i^{(k+1)} = f_i - \sum_{j=1, j\neq i}^{n} a_{ij}x_j^{(k)}$$

$$x_i^{(k+1)} = \frac{f_i - \sum_{j=1, j\neq i}^{n} a_{ij}x_j^{(k)}}{a_{ii}}$$

Потрясём матрицу S:

$$S = I - \tau B^{-1}A = I - D^{-1}(A_l + D + A_v) = -D^{-1}(A_l + A_v)$$

Для того, чтобы метод Якобы сходился, необходимо и достаточно, чтобы собственные значения этой матрицы были меньше единички по модулю.

Но нас такое не вдохновляет, мы хотим по виду A сразу понять.

Определение

Строгое диагональное преобладание или строгое условие Адамара, это когда $\forall i$:

$$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}} |a_{ij}|$$

Теорема Адамара

Матрица со строгим диагональным преобладанием невырождена.

Пусть матрица имеет строгое диагональное обладание и вырождена (противное). Тогда $Ax = 0$ имеет ненулевое решение. Выделим максимальную по модулю компоненту этого решения, $|x_k| = \max_i |x_i| > 0$. Поработаем с ним:

$$\sum_{j=1}^{n} a_{nj}x_j = 0$$

$$a_{kk}x_k + \sum_{\substack{j=1 \\ j\neq n}}^{n} a_{kj}x_k = 0$$

$$a_{kk}x_k = -\sum_{\substack{j=1 \\ j\neq k}}^{n} a_{kj}x_j$$

$$|a_{kk}||x_k| \leq \sum_{j=1}^{n}|a_{kj}||x_j| \leq |x_k|\sum_{j=1}^{n}|a_{kj}| =$$

$$= |kk| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n}|a_{kj}|$$

Следствие из теоремы Адамара

У вырожденной матрица нарушено ходя бы одно строгое условие Адамара.

Теорема Гершгорина

Каждое собственное значение матрицы A принадлежит по крайней мере одному из кругов Гершгорина этой матрицы.

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n}|a_{ij}|,\ i = 1, ..., n$$

Пусть λ — собственное значение.

$$\det(A - \lambda I) = 0$$

$$|a_{kk} - \lambda| \leq \sum_{\substack{j=1 \\ j\neq n}}^{n}|a_{kj}|$$

Если матрица имеет строгое диагональное преобладание, то матрица Якоби сходится при любом начальном приближении.

$$S = -D^{-1}(A_l + A_v)$$

$$S_{ii} = 0$$

$$S_{ij} = -\frac{a_{ij}}{a_{ii}}$$

$$|\lambda_i(S) - 0| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} \left| -\frac{a_{ij}}{a_{ii}} \right| = \sum_{\substack{j=1 \\ j\neq i}}^{n} \left| \frac{a_{ij}}{a_{ii}} \right| < 1$$

Тогда по критерию этот метод сходится. $|\lambda_i(S)| < 1$.

`5 Ноября 2024`

5.3. Метод Зейделя

$$B = A_l + D, \quad \tau = 1$$

Подставляем в каноническую форму:

$$(A_l + D)\left(x^{(k+1)} - x^{(k)}\right) + (A_l + D + A_v)x^{(k)} = f$$

$$(A_l + D)x^{(k+1)} = -A_v x^{(k)} + f$$

Запишем в скалярном виде:

$$\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} + a_{ii}x_i^{(k+1)} = -\sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + f_i$$

$$x_i^{(k+1)} = \frac{f_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}}{a_{ii}} \quad (1)$$

Вот так вот крутимся-крутимся.

Вопрос со сходимостью. Сейчас будем работать с матрицей S:

$$S = I - \tau B^{-1} A = I - (A_l + D)^{-1}(A_l + D + A_v) = -(A_l + D)^{-1} A_v$$

Для того, чтобы сходился метод, необходимо и достаточно, чтобы собственные числа этой матрицы были по модулю меньше единицы.

$$\det\Big(-(A_l + D)^{-1} A_v - \lambda I\Big) = 0$$

$$\det\Big(-(A_l + D)^{-1}(A_v + \lambda(A_l + D))\Big) = 0$$

$$\det\Big(-(A_l + D)^{-1}\Big) \cdot \det(A_v + \lambda(A_l + D))) = 0$$

$$\text{т.к. } \det\Big(-(A_l + D)^{-1}\Big) \neq 0$$

$$\det(A_v + \lambda(A_l + D)) = 0$$

$$|\lambda| < 1$$

Теорема

Если в матрице A диагональное преобладание, то метод Зейделя сходится при любом начальном приближении.

Это очень мощная, хорошая теорема. Сейчас мы её докажем.

Пусть λ произвольный корень уравнения:

$$\det(A_v + \lambda(A_l + D)) = 0$$

У вырожденной матрицы нарушено хотя бы одно строгое условие Адамара.

Вот матрица:

$$A_v + \lambda(A_l + D)$$

Выпишем k — ю строчку, в которой нарушено диагональное преобладание:

$$|\lambda a_{kk}| \le \sum_{j=1}^{k-1} |\lambda a_{kj}| + \sum_{j=k+1}^{n} |a_{kj}|$$

$$|\lambda| \left(|a_{kk}| - \sum_{j=1}^{k-1} |a_{kj}| \right) \le \sum_{j=k+1}^{n} |a_{kj}|$$

$$|\lambda| \le \frac{\sum_{j=k+1}^{n} |a_{kj}|}{|a_{kk}| - \sum_{j=1}^{k-1} |a_{kj}|}$$

$$|a_{kk}| > \sum_{j=1}^{k-1} |a_{kj}| + \sum_{j=k+1}^{n} |a_{kj}|$$

$$\sum_{j=k+1}^{n} |a_{kj}| < |a_{kk}| - \sum_{j=1}^{k-1} |a_{kj}|$$

$$|\lambda| < 1$$

Вот это доказателство вот этой теоремы.

Есть ещё одна теорема о сходимости, но мы её доказывать не будем, только посмотрим. Давайте сформулируем.

Теорема

Если матрица системы A симметрическая и положительно определённая, то метод Зейделя сходится при любом начальнмо приближении.

Метод Зейделя — это частный случай метода релаксации, но это мы попозже разберём.

Перейдём к следующему методу, но сначала проделаем некоторую работу. Давайте под другим углом посмотрим на этот метод, посмотрим переход:

$$y = \begin{pmatrix} x_1^{(k+1)} \\ \vdots \\ x_{i-1}^{(k+1)} \\ x_i^{(k)} \\ x_{i+1}^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix} \qquad z = \begin{pmatrix} x_1^{(k+1)} \\ \vdots \\ x_{i-1}^{(k+1)} \\ x_i^{(k+1)} \\ x_{i+1}^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix}$$

$$t = Ay - f \qquad r = Az - f$$

$$t_i = \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} + a_{ii} x_i^{(k)} + \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} - f_i$$

$$r_i = \textstyle\sum_{j=1}^{i-1} a_{ij} x_{j((k+1))} + a_{ii} x_i^{(k+1)} + \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} - f_i$$

$$r_i \equiv 0 \text{ соответсвтует методу Зейделя}$$

То есть обнуление невязки это и есть метод. Очев, подставьте в (1), чтоб было очев.

Вычтем из одного уравнения другое:

$$t_i - r_i = a_{ii}\left(x_i^{(k)} - x_i^{(k+1)}\right) := a_{ii}\alpha$$

$$r_i = t_i - \alpha a_{ii} \equiv 0$$

Метод Зейделя — это метод полной релаксации.

5.4. Метод частичной релаксации

Тут мы хотим, чтобы r был не ноль, но поменьше t.

$$|t_i - \alpha a_{ii}| < |t_i|$$

$$\left|\alpha - \frac{t_i}{a_{ii}}\right| < \left|\frac{t_i}{a_{ii}}\right|$$

1)

$$\frac{t_i}{a_{ii}} > 0 \qquad \left|\alpha - \frac{t_i}{a_{ii}}\right| < \frac{t_i}{a_{ii}}$$

$$-\frac{t_i}{a_{ii}} < \alpha - \frac{t_i}{a_{ii}} < \frac{t_i}{a_{ii}} \qquad 0 < \alpha < 2\frac{t_i}{a_{ii}}$$

2)

$$\frac{t_i}{a_{ii}} < 0 \qquad 2\frac{t_i}{a_{ii}} < \alpha < 0$$

$$\alpha = \omega\frac{t_i}{a_{ii}} \qquad \omega \in (0, 2)$$

Тогда будет уменьшение компоненты невязки.

$$|r_i| < |t_i|$$

$$r_i = t_i - \alpha a_{ii} = t_i - \omega\frac{t_i}{a_{ii}}a_{ii} = (1-\omega)t_i$$

$$r_i = (1-\omega)t_i$$

$$x_i^{(k+1)} = \frac{\omega f_i - \omega\sum_{j=1}^{i-1} a_{ij}x_j^{(k)} + (1-\omega)a_{ii}x_j^{(k)} - \omega\sum_{j=i+1}^{n} a_{ij}x_j^{(k)}}{a_{ii}}$$

Теперь восстановим каноническую форму записи, ведь мы не знаем ни τ, ни B. В общем, надо из скалярной формы записи получить матричную.

$$S = I - \tau B^{-1}A$$

Переработаем формулу для $x_i^{(k+1)}$. К этой пижне добавим и вычтем $\sum_{j=1}^{i-1} a_{ij}x_j^{(k)}\omega$. Это дома проделать обязательно, на лекции нет времени:

$$a_{ii}\left(x_i^{(k+1)} - x_i^{(k)}\right) + \omega \sum_{j=1}^{i-1} a_{ij}\left(x_j^{(k+1)} - x_j^{(k)}\right) + \omega \sum_{j=1}^{n} a_{ij} x_j^{(i)} = \omega f_i$$

Нужно будет заполучить:

$$B\frac{x^{(k+1)} - x^{(k)}}{\tau} + Ax^{(k)} = f$$

Переходим теперь к матричной форме записи:

$$D\left(x^{(k+1)} - x^{(k)}\right) + \omega A_l\left(x^{(k+1)} - x^{(k)}\right) + \omega A x^{(k)} = \omega f$$

Возьмём:

$$\begin{aligned} \tau &:= \omega \\ B &:= D + \omega A_l \end{aligned}$$

$$(D + \omega A_l)\frac{x^{(k+1)} - x^{(k)}}{\omega} + Ax^{(k)} = f$$

Вот она — каноническая формула записи для этого метода. А $\omega = 1$ — это метод Зейделя.

$$S = I - \tau B^{(-1)} A = I - \omega (D + \omega A_l)^{-1} A$$

Первый шаг, который здесь делается, это убрать эту некрасивую обратную матрицу.

$$\det\left(I - \omega (D + \omega A_l)^{-1} - \lambda I\right) = 0$$

$$\det\left((D + \omega A_l)^{-1}\right) \cdot \det((1 - \lambda)(D + \omega A_l) - \omega(A_l + D + A_v)) = 0$$

$$\det((1 - \lambda)(D + \omega A_l) - \omega(A_l + D + A_v)) = 0$$

$$\det((1 - \omega)D - \lambda(D + \omega A_l) - \omega A_v) = 0$$

$$|\lambda| < 1$$

Оказывается теорема с достаточными условиями сходимости (симметрическая + положительно определённая) работает для всех релаксационных методов.

Теперь мы идём в комплексную плоскость. Рассмотрим конформное отображение круга в следующую пижню:

$$|\lambda| < 1 \Leftrightarrow \lambda = \frac{\mu + 1}{\mu - 1}, \mathrm{Re}\ \mu < 0$$

Для сходимости достаточно, чтобы все реальные части μ были отрицательны.

`12 Ноября 2024`

5.5. Метод Ричардсона

5.5.1. Многочлены Чебышёва

Определение

$$T_n(x),\ \ n \geq 0$$

$$T_0(x) = 1 \quad T_1(x) = x$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

$$T_2(x) = 2x^2 - 1$$

$$T_3(x) = 4x^3 - 3x$$

Вспомним тригонометрическую формулу:

$$\cos((n+1)\theta) = 2\cos\theta\cos n\theta - \cos((n-1)\theta)$$

$$\theta := \arccos x$$

$$T_n(x) = \cos(n \arccos x)$$

$$\cos((n+1)\theta) = T_{n+1}(x)$$

$$2\cos\theta\cos n\theta = 2x\cos(n\arccos x) = 2xT_n(x)$$

$$\cos((n-1)\theta) = T_{n-1}(x)$$

Ограничение, $|x| \leq 1$.

Рекуррентная запись полинома Чебышёва соответствует характеристическому уравнению:

$$\mu^2 - 2x\mu + 1 = 0$$

$$\mu_{1,2} = x \pm \sqrt{x^2 - 1}$$

$$T_n(x) = C_1\mu_1^n + C_2\mu_2^n - \text{на веру, доказывать не будем}$$

$$T_n(x) = \frac{\left(x + \sqrt{x^2-1}\right)^n + \left(x - \sqrt{x^2-1}\right)^n}{2}$$

Нули полинома:

$$T_n(x) = \cos(n\arccos x) = 0$$

$$n\arccos x = \frac{\pi}{2} + m\pi = \frac{\pi(2m+1)}{2}$$

$$x_m = \cos\frac{\pi(2m+1)}{2n}, \text{ где } m = 0, 1, ..., n-1$$

$$|T_n(x)| \leq 1$$

Точки экстремума:

$$x_{(m)} = \cos\frac{\pi m}{n}, \text{ где } m = 0, ..., n$$

Если построить немного другой полином (убирает коэффициент при старшем члене):

$$\overline{T}_n(x) = 2^{1-n}T_n(x) = x^n + ...$$

Полиномы называются *наименее уклоняющиеся от нуля.*

Лемма

Если $P_n(x)$ это многочлен степени n со старшим коэффициентом 1, то справебыдло следующее:

$$\max_{[-1,1]} |P_n(x)| \geq \max_{[-1,1]} \left| \overline{T}_n(x) \right| = 2^{1-n}$$

Преобразование:

$$x' := \frac{b+a}{2} + \frac{b-a}{2} x$$

Перегоняет отрезок:

$$[-1, 1] \longrightarrow [a, b]$$

Переформулируем на отрезок $[a, b]$:

$$\overline{T}_n \left(\frac{2x - (b+a)}{b-a} \right)$$

$$\max_{[a,b]} |P_n(x)| \geq \max_{[a,b]} \left| \overline{T}_n^{[a,b]}(x) \right| = (b-a)^n 2^{1-2n}$$

$$x_m = \frac{b+a}{2} + \frac{b-a}{2} \cos \frac{\pi(2m+1)}{2n}$$

Теперь можем переходить к методу Ричардсона.

5.5.2. Метод Ричардсона

Нестанционарный, но напоминает метод простой итерации.

$$B = I$$

$$\frac{x^{(k+1)} - x^{(k)}}{\tau_{k+1}} + Ax^{(k)} = f$$

$$x^{(k+1)} = x^{(k)} - \tau_{k+1} A x^{(k)} + \tau_{k+1} f$$

По аналогии с методом простой итерации мы получим:

$$z^{(k+1)} = S_{k+1} z^{(k)}$$

Здесь матрица перехода на каждом шаге своя, так как метод нестационарный.

$$S_{k+1} = I - \tau_{k+1} A$$

$$z^{(k)} = S_k z^{(k-1)} = S_k S_{k-1} z^{(k-1)} = S_k S_{k-2} ... S_1 z^{(0)} := T_k z^{(0)}$$

Это не Полином Чебышёва! Это *разрешающая матрица k-го шага.*

$$z^{(k)} = T_k z^{(0)}$$

$$z^{(k)} = S^k z^{(0)}$$

Будем минимизировать резрешающую матрицу.

$$A := A^* > 0, \quad 0 < \alpha \leq \lambda_i(A) \leq \beta$$

$$T_k = (I - \tau_k A)(I - \tau_{k-1} A)...(I - \tau_1 A)$$

$$T_k - \text{симметрическая}$$

$$T_k^* = (I - \tau_1 A)^* (I - \tau_2 A)^* ... (I - \tau_k A)^* = (I - \tau_1 A)(I - \tau_2 A)...(I - \tau_k A)$$

Любые два полинома от матрицы перестановочны: $P(A)Q(A) = Q(A)P(A)$

$$\|T_k\|_2 = \max |\lambda_i(T_k)|$$

$$T_k = (I - \tau_k A)(I - \tau_{k-1} A)...(I - \tau_1 A)$$

$$A x^{(i)} = \lambda_i(A) x^{(i)}, \quad x^{(i)} \neq 0$$

$$\left(I - \tau_j A\right) x^{(i)} = \left(1 - \tau_j \lambda_i(A)\right) x^{(i)}$$

Теперь проводим цепочку:

$$T_k x^{(i)} = (I - \tau_k A)(I - \tau_{k-1} A)...(I - \tau_2 A)(I - \tau_1 A) x^{(i)}$$

$$(I - \tau_1 A)x^{(i)} = (1 - \tau_1 \lambda_i)x^{(i)} - \text{число}$$

$$T_k x^{(i)} = (1 - \tau_1 \lambda_i)(1 - \tau_2 \lambda_i)... = \prod_{j=1}^{k} \big(1 - \tau_j \lambda_i\big) x^{(i)}$$

$$\|T_k\|_2 = \max_{1 \le i \le n} \left| \prod_{j=1}^{k} \big(1 - \tau_j \lambda_i(A)\big) \right|$$

$$\|T_k\|_2 \le \max_{[\alpha, \beta]} \left| \prod_{j=1}^{k} \big(1 - \tau_j \mu\big) \right|$$

Введём полином и опишем его свойства:

$$P_k(\mu) := \prod_{j=1}^{k} \big(1 - \tau_j \mu\big)$$

$$P_k(0) = 1$$

Надо подобрать такой полином по набору параметров $\tau_1, ..., \tau_n$, чтобы он наименее отклонялся от нуля.

Оказывается любой полином со свойством $P_k(0) = 1$ представим в виде:

$$P_k(\mu) := \prod_{j=1}^{k} \big(1 - \tau_j \mu\big)$$

Любой полином раскладывается:

$$Q_k = a(\mu - \mu_1)(\mu - \mu_2)...(\mu - \mu_k)$$

$$Q_{k(0)} = 1 - \text{сохраняем условие}$$

$$Q_k = a \prod_{j=1}^{k} (\mu - \mu_j)$$

$$Q_k(0) = a\prod_{j=1}^{k}\left(-\mu_j\right) = (-1)^k)a\prod_{j=1}^{k}\mu_j = 1$$

$$a = \frac{(-1)^k}{\prod_{j=1}^{k}\mu_j}$$

$$Q_k(\mu) = \frac{(-1)^k}{\prod_{j=1}^{k}\mu_j}\prod_{j=1}^{k}\left(\mu - \mu_j\right) = (-1)^k\prod_{j=1}^{k}\frac{\mu - \mu_j}{\mu_j} = \prod_{j=1}^{k}\frac{\mu_j - \mu}{\mu_j}$$

$$\tau_j = \frac{1}{\mu_j}$$

Лемма

Среди всех многочленов с вещественными корнями и удовлетворяющих условию $P_k(0) = 1$, наименее отклоняется от нуля:

$$P_k^0(\mu) = \frac{T_k\left(\frac{2\mu-(\beta+\alpha)}{\beta-\alpha}\right)}{t_k}$$

Здесь уже T_k это полином Чебышёва.

$$t_k = T_k\left(-\frac{\alpha+\beta}{\beta-\alpha}\right)$$

Ещё в этой лемме утверждается, что:

$$\max_{\alpha\leq\mu\leq\beta}\left|P_k^0(\mu)\right| = \frac{1}{|t_k|}$$

19 Ноября 2024

$$P_k^0(\mu) = \prod_{j=1}^{n}\left(1 - \tau_j^0\mu\right),\ \mu_j = \frac{1}{\tau_j^0}$$

$$T_k(\frac{\frac{2\mu-(\alpha+\beta)}{\beta-\alpha}}{t_n}$$

$$\mu_j = \frac{\alpha+\beta}{2} + \frac{\beta-\alpha}{2}\cos\frac{(2j-1)\pi}{2k}$$

$$\tau_j^0 = \frac{2}{\left(\alpha+\beta+(\beta-\alpha)\cos\frac{(2j-1)\pi}{2k}\right)},\ j = 1, ..., k$$

$$\|T_k\|_2 \le \frac{1}{|t_k|}$$

Найдём количество шагов, чтобы начальную погрешность уменьшить в e раз.

$$\left\|z^{(0)}\right\|_2$$

$$\frac{1}{|t_k|} \le \frac{1}{e}$$

$$\left\|z^{(k)}\right\|_2 \le \|T_k\|_2 \left\|z^{(0)}\right\|_2$$

$$\left\|z^{(k)}\right\|_2 \le \frac{1}{|t_k|}\left\|z^{(0)}\right\|_2 \le \frac{1}{e}\left\|z^{(0)}\right\|_2$$

$$|t_k| \ge e$$

$$t_k = T_k\left(-\frac{\alpha+\beta}{\beta-\alpha}\right)$$

$-\frac{\alpha+\beta}{\beta-\alpha}$ выпадает за отрезок $[-1, 1]$. Поэтому используем другое представление полинома Чебышёва:

$$T_k(x) = \frac{\left(x-\sqrt{x^2-1}\right)^k + \left(x+\sqrt{x^2-1}\right)^k}{2}$$

Подставляем сюда $x = -\frac{\alpha+\beta}{\beta-\alpha}$.

$$x - \sqrt{x^2 - 1} = -\frac{\sqrt{\beta} + \sqrt{\alpha}}{\sqrt{\beta} - \sqrt{\alpha}}$$

$$x + \sqrt{x^2 - 1} = -\frac{\sqrt{\beta} - \sqrt{\alpha}}{\sqrt{\beta} + \sqrt{\alpha}}$$

$$\rho := \frac{\sqrt{\beta} + \sqrt{\alpha}}{\sqrt{\beta} - \sqrt{\alpha}} > 1$$

$$t_k = T_{k(x)} = \frac{1}{2}\left[(-1)^k \rho^k + \frac{(-1)^k}{\rho^k}\right] = \frac{(-1)^k}{2} \frac{\rho^{2k} + 1}{\rho^k}$$

Смотрим по модулю:

$$\frac{\rho^2 k + 1}{2\rho^k} \geq e$$

Теперь порешаем это неравенство.

$$\rho^{2k} + 1 \geq 2\rho^k e$$

$$\rho^{2k} - 2\rho^k e + e^2 \geq e^2 - 1$$

$$\left(\rho^k - e\right)^2 \geq e^2 - 1$$

$$\rho^k \geq e + \sqrt{e^2 - 1}$$

$$k \ln \rho \geq \ln\left(e + \sqrt{e^2 - 1}\right)$$

$$k \geq \frac{\ln\left(e + \sqrt{e^2 - 1}\right)}{\ln \frac{\sqrt{\beta} + \sqrt{\alpha}}{\sqrt{\beta} - \sqrt{\alpha}}}$$

$$k \geq \frac{1}{\ln \frac{\beta + \alpha}{\beta - \alpha}}$$

Рассмотрим пример, когда у нас не очень хорошая матрица:

$$\alpha = 10^{-3} \quad \beta = 1$$

$$\mu_2(A) = \frac{1}{10^{-3}} = 10^3$$

Если мы сейчас посчитаем число шагов в методе Ричардсона и методе простой итерации, то:

- Простая итерация — порядка 500 шагов.
- Метод итерация — порядок 26 шагов.

Метод применяют в так называемом циклическом виде. А именно. Фиксируют k. Например $k = 10$. Считаем t_j^0. Пока не получим решение с нужной нам точностью ε.

Но это ещё не всё. Обычно такие сложные нестационарные методы в себе содержат вычислительную неустойчивость. Всё сделаем, начинаем считать и накапливается вычислительная ошибка. Обычно в процессе вычисления ошибка съедается, но здесь ситуация друга. Почему этот метод может быть неустойчивым?

Вот возьмём один шажок:

$$z^{(j)} = \left(I - \tau_j^0 A\right) z^{(j-1)},\ j = 1, ..., k$$

Но это если мы не учитываем погрешности машинной арифметики. Реально происходит:

$$\tilde{z}^{(j)} = \left(I - \tau_j^0 A\right) \tilde{z}^{(j-1)} + \rho^j$$

$$z^{(k)} = \left(I - \tau_k^0\right) z^{(k-1)} + \rho^{(k)} =$$

$$= \left(I - \tau_k^0 A\right)\left(I - \tau_{k-1}^0 A\right) z^{(k-2)} + \left(I - \tau_k^0 A\right) \rho^{(k-1)} + \rho^k =$$

$$= T_k z^{(0)} + \sum_{i=1}^{k} \left(\prod_{j=i+1}^{n} \left(I - \tau_j^0 A\right) \right) \rho^{(i)}$$

$$\left\|T_k\right\|_2 \leq \frac{1}{|t_k|} = \frac{2\rho^k}{1 + \rho^{2k}} < 1$$

5.6. Метод наибыстрейшего градиентного спуска.

Идея метода. Есть некая функция $F(x)$. Где x — это вектор. Общая формула:

$$x^{(k+1)} = x^{(k)} - \delta_k \nabla F\left(x^{(k)}\right)$$

Градиент смотрит в сторону возрастания, поэтому мы берём со знаком минус.

δ_k находится из условия минимума:

$$F\left(x^{(k+1)}\right) = F\left(x^{(k)} - \delta_k \nabla F\left(x^{(k)}\right)\right)$$

Как это связано с решением СЛАУ?

$$Ay = f$$

Введём функцию многих переменных:

$$F(y) = (Ay, y) - 2(f, y)$$

Ограничения:

$$A = A^* > 0,\ \ A - \text{вещественная}$$

$$Ax^* = f$$

$$F(y) = (A(y - x^*), y - x^*) - (Ax^*, x^*)$$

Если, $y = x^*$:

$$F(x^*) = -(Ax^*, x^*)$$

В случае $y \neq x^*$:

$$F(y) > F(x^*)$$

$$\nabla F = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \ldots, \frac{\partial F}{\partial x_n}\right)$$

$$F = \sum_{i=1}^{n} (Ax_i)x_i - 2\sum_{i=1}^{n} f_i x_i$$

$$\sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_j x_i - 2\sum_{i=1}^{n} f_i x_i$$

$$\frac{\partial F}{\partial x_i} = 2(Ax_i - f_i)$$

$$\nabla F = 2(Ax - f)$$

$$x^{(k+1)} = x^{(k)} - 2\delta_k\big(Ax^{(k)} - f\big)$$

$$2\delta_k \coloneqq \Delta_k$$

$$x^{(k-1)} = x^{(k)} - \Delta_k\big(Ax^{(k)} - f\big) \coloneqq \varphi(\Delta_k)$$

$$\varphi'(\Delta_k) = 0$$

$$\varphi'(\Delta_k) = \Big[\big(Ax^{(k+1)}, x^{(k+1)}\big) - 2\big(f, x^{(k+1)}\big)\Big]' =$$

$$= \left(A\frac{\mathrm{d}x^{(k+1)}}{\mathrm{d}\Delta_k}, x^{(k+1)}\right) + \left(Ax^{(k+1)}, \frac{\mathrm{d}x^{(k+1)}}{\mathrm{d}\Delta_k}\right) - 2\left(f, \frac{\mathrm{d}x^{(k+1)}}{\mathrm{d}\Delta_k}\right) =$$

$$= 2\left(Ax^{(k+1)} - f, \frac{\mathrm{d}x^{(k+1)}}{\mathrm{d}\Delta_k}\right) = 0$$

$$\frac{\mathrm{d}x^{(k+1)}}{\mathrm{d}\Delta_k} = -\big(Ax^{(k)} - f\big)$$

$$-2\big(Ax^{(k+1)} - f, Ax^{(k)} - f\big) = 0$$

$$\big(Ax^{(k)} - f - \Delta_k A\big(Ax^{(k)} - f\big), Ax^{(k)} - f\big) = 0$$

$$\Delta_k = \frac{\left(Ax^{(k)}, Ax^{(k)} - f\right)}{\left(A\left(Ax^{(k)} - f\right), Ax^{(k)} - f\right)}$$

На самом деле метод похож на метод простой итерации.

Вопросы, связанные со сходимостью. Теорема такая:

$$F_0(y) \coloneqq F(y) + (Ax^*, x^*) = (A(y - x^*), y - x^*)$$

$$z^{(k)} = x^{(k)} - x^*$$

$$F_0\left(x^{(k)}\right) = \left(Ax^{(k)}, z^{(k)}\right)$$

Теорема:

$$F_0\left(x^{(k)}\right) \leq \left(\frac{\beta - \alpha}{\beta + \alpha}\right)^{2n} F_0\left(x^{(0)}\right)$$

$$\alpha = \min_i \lambda_i(A) \qquad \beta = \max_i \lambda_i(A)$$

$$\left\|z^{(k)}\right\|_2 \leq \left(\frac{\beta - \alpha}{\beta + \alpha}\right)^n \sqrt{\frac{\beta}{\alpha}} \cdot \left\|z^{(0)}\right\|_2$$

$$z^{(k)} = \sum_{i=1}^n c_i u^{(i)}$$

$$\|z_k\|_2^2 = \sum_{i=1}^n c_i^2$$

$$\left(Az^{(k)}, z^{(k)}\right) = \sum_{i=1}^n \lambda_i c_i^2 \leq \beta \sum_{i=1}^n c_i^2 = \beta \left\|z^{(k)}\right\|_2$$

$$\alpha \left\|z^{(k)}\right\|_2^2 \leq F_0\left(x^{(k)}\right) \leq \beta \left\|z^{(k)}\right\|_2^2$$

$$\left\|z^{(k)}\right\|_2 \le \sqrt{\frac{F_0(x^k)}{\alpha}} \le \sqrt{\frac{1}{\alpha}\left(\frac{\beta-\alpha}{\beta+\alpha}\right)^{2n} F_0\left(x^{(0)}\right)} \le$$

На этом вторая часть окончена. Можно сдавать коллоквиум по второй части. Далее будет третья часть, это проблема собственных значений.

`3 Декабря 2024`

6. Проблема собственных значений

Рассмторим плохой пример:

$$A = \begin{pmatrix} 1 & 10 & & & \\ & 1 & 10 & & \\ & & \ddots & \ddots & \\ & & & 1 & 10 \\ & & & & 1 \end{pmatrix} \quad n = 25$$

$$\lambda_i = 1 \quad \forall i = 1, ..., 25$$

$$\tilde{A}(\delta) = \begin{pmatrix} 1 & 10 & & & \\ & 1 & 10 & & \\ & & \ddots & \ddots & \\ & & & 1 & 10 \\ \delta & & & & 1 \end{pmatrix} \quad n = 25$$

$$\det \tilde{A}(\delta) = 1 + \delta \cdot 10^{24}$$

$$\delta := -10^{-24}$$

Маленькая ошибка, но пиздец, ведь тогда $\det \tilde{A}(\delta) = 0$.

Мы ограничимся собственными значениями для симметрических матриц.

$$|A - \delta I| = 0$$

$$p_n \lambda^n + p_{n-1}\lambda^{n-1} + ... + p_1 \lambda + p_0 = 0$$

Нужно найти коэффициенты и затем решить уравнение само.

Рассмотрим другой подход, преобразование матриц.

$$B = T^{-1}AT$$

чтобы B была простой.

$$v := \text{собственный вектор } B$$

$$Bv = \lambda v$$

$$T^{-1}ATv = \lambda v$$

$$A(Tv) = \lambda(Tv)$$

$$u := Tv$$

$$Au = \lambda u$$

$$A := A^*$$

$$U^*AU = \Lambda - \text{диагональная}$$

$$U^*U = I \qquad U^* = U^{-1}$$

$$\Lambda = \begin{pmatrix} \lambda_1(A) & & 0 \\ & \ddots & \\ 0 & & \lambda_n(A) \end{pmatrix}$$

Собственные вектора это столбцы U:

$$U^*AU = \Lambda \quad | \cdot U$$

$$AU = U\Lambda$$

$$u^{(i)} - i\text{-й стобец } U$$

$$Au^{(i)} = \lambda_i(A)u^{(i)}$$

Для симметрических матриц ничего не развалится. Проблем для них не существует.

Вариационное описание собственных чисел:

$$\lambda_1(A) \leq \ldots \leq \lambda_n(A)$$

Определение

Отношение Рэлея:

$$\frac{(Ax, x)}{(x, x)}$$

$$A = U \Lambda U^*$$

$$\frac{(Ax, x)}{(x, x)} = \frac{(U \Lambda U^* x, x)}{(U U^* x, x)} = \frac{(\Lambda U^* x, U^* x)}{(U^* x, U^* x)} = [y := U^* x] = \frac{(\Lambda y, y)}{(y, y)}$$

$$\min_{x \neq 0} \frac{(Ax, x)}{(x, x)} = \min_{y \neq 0} \frac{(\Lambda y, y)}{(y, y)} \stackrel{?}{=} \lambda_1(A)$$

Докажем этот факт, что это и есть наименьшее собственное значение A.

$$\frac{(\Lambda y, y)}{(y, y)} = \frac{\sum_{i=1}^{n} \lambda_i(A) \cdot y_i^2}{\sum_{i=1}^{n} y_i^2} \geq \frac{\sum \lambda_1(A) y_i^2}{\sum y_i^2} = \lambda_1(A)$$

$$y = e_1 = (1, 0, \ldots, 0)^*$$

$$\frac{(\Lambda e_1, e_1)}{e_1, e_1} = \frac{\lambda_1(A)}{1}$$

$$\min_{y \neq 0} \frac{(\Lambda y, y)}{(y, y)} = \lambda_1(A) = \min_{x \neq 0} \frac{(Ax, x)}{(x, x)} =$$

вариационное описание первого собственного значения

$$= \frac{\left(Au^{(1)}, u^{(1)}\right)}{\left(u^{(1)}, u^{(1)}\right)}$$

$u^{(1)}$ – первый собственный вектор

$$x = Ue_1$$

Рассмотрим подпространство векторов, ортогональных первому собственному вектору:

$$\left(x, u^{(1)}\right) = 0$$

$$\left(x, u^{(1)}\right) = [y \coloneqq U^* x] = (Uy, Ue_1) = (U^* Uy, e_1) = (y, e_1)$$

$$\min_{x \neq 0} \frac{(Ax, x)}{(x, x)} = \min_{y \neq 0} \frac{(\Lambda y, y)}{(y, y)}$$

$$\left(x, u^{(1)}\right) = 0 \qquad (y, e_1) = 0$$

Тогда:

$$y_1 = 0$$

$$\frac{(\Lambda y, y)}{(y, y)} = \frac{\sum_{i=2}^{n} \lambda_i y_i^2}{\sum_{i=2}^{n} y_i^2} \geq \frac{\sum_{i=2}^{n} \lambda_2 y_i^2}{\sum_{i=2}^{n} y_i^2} = \lambda_2(A)$$

$$\min_{x \neq 0} \frac{(Ax, x)}{(x, x)} = \min_{y \neq 0} \frac{(\Lambda y, y)}{(y, y)} = \lambda_2(A)$$

$$e_2 = (0 \;\; 1 \;\; 0 \;\; \dots \;\; 0)^*$$

$$\min_{\substack{x \neq 0 \\ (x, u^{(1)}) = 0}} \frac{(Ax, x)}{(x, x)} = \min_{\substack{y \neq 0 \\ (y, e_1) = 0}} \frac{(\Lambda y, y)}{(y, y)} = \lambda_2(A) = \frac{(\Lambda e_2, e_2)}{(e_2, e_2)} = \frac{\left(Au^{(2)}, u^{(2)}\right)}{\left(u^{(2)}, u^{(2)}\right)}$$

По аналогии выпишем λ_i:

$$\min_{\substack{x \neq 0 \\ (x,u^{(1)})=0 \\ \vdots \\ (x,u^{(i-1)})=0}} \frac{(Ax,x)}{(x,x)} = \lambda_i(A), \quad \forall i = 1, ..., n$$

$$S_i \qquad u^{(1)}, ..., u^{(i)}$$

$$x \neq 0 \qquad x \in S_i$$

$$\frac{(Ax,x)}{(x,x)} \leq \lambda_i(A)$$

$$x := \sum_{j=1}^{i} c_j u^j$$

$$(x,x) = \sum_{j=1}^{i} c_j^2$$

$$(Ax,x) = \sum_{j=1}^{i} \lambda_j c_j^2$$

$$\frac{(Ax,x)}{(x,x)} = \frac{\sum_{j=1}^{i} \lambda_j c_j^2}{\sum_{j=1}^{i} c_j^2} = \lambda_i$$

Рассмотрим совершенно произвольное подпространство размерностью $n - i + 1$:

$$R_{n-i+1}$$

$$S_i$$

Сумма размерностей этих пространств $(n + 1)$.

$$R_{n-i+1} \wedge S_i \neq 0$$

$$x_0 \neq 0 \wedge x_0 \in R_{n-i+1} \wedge x_0 \in S_i$$

Будет выполняться:

$$\frac{(Ax_0, x_0)}{(x_0, x_0)} \leq \lambda_i(A)$$

Так как $x_0 \in S_i$.

$$\min_{\substack{x \neq 0 \\ x \in R_{n-i+1}}} \frac{(Ax, x)}{(x, x)} \leq \lambda_i(A)$$

$$R_{n-i+1} = R_{n-(i-1)}$$

$$u^{(1)}, ..., u^{(i-1)}$$

$$\left(x, u^{(1)}\right) = 0$$
$$\vdots$$
$$\left(x, u^{(i-1)}\right) = 0$$

$$\max_{R_{n-i+1}} \min_{\substack{x \neq 0 \\ x \in R_{n-i+1}}} \frac{(Ax, x)}{(x, x)} = \lambda_i(A)$$

Это вариационный принцип Куранта-Фишера.

$$A = A^*, \ B = B^*$$

$$\lambda_1(A) \leq ... \leq \lambda_n(A)$$

$$\lambda_i(B) \leq ... \leq \lambda_n(B)$$

И пусть:

$$(Ax, x) \leq (Bx, x)$$

Тогда:

$$\lambda_i(A) \leq \lambda_i(B)$$

Доказательство очень интересно звучит:

$$\frac{(Ax,x)}{(x,x)} \leq \frac{(Bx,x)}{(x,x)} \quad \forall x$$

$$\min_{\substack{x\neq 0 \\ x\in R_{n-i+1}}} \frac{(Ax,x)}{(x,x)} \leq \min_{\substack{x\neq 0 \\ x\in R_{n-i+1}}} \frac{(Bx,x)}{(x,x)}$$

$$\max_{R_{n-i+1}} \min_{\substack{x\neq 0 \\ x\in R_{n-i+1}}} \frac{(Ax,x)}{(x,x)} \leq \max_{R_{n-i+1}} \min_{\substack{x\neq 0 \\ x\in R_{n-i+1}}} \frac{(Bx,x)}{(x,x)}$$

$$\lambda_i(A) \leq \lambda_i(B)$$

Какое отношение это имеет к практике? Рассмотрим второе следствие:

$$A = A^*, \ B = B^*$$

$$\lambda_1(A) \leq ... \leq \lambda_n(A)$$

$$\lambda_i(B) \leq ... \leq \lambda_n(B)$$

Тогда:

$$|\lambda_i(B) - \lambda_i(A)| \leq \|B - A\|_2 \quad \forall i = 1, ..., n$$

Докажем:

$$C := B - A$$

$$(Bx,x) = (Ax,x) + (Cx,x)$$

Вспоминаем неравенство Коши-Буняковского:

$$|(Cx,x)| \leq \|Cx\|_2 \cdot \|x\|_2 \leq \|C\|_2 \cdot \|x\|_2^2 = (\|C\|_2 Ix, x)$$

$$-(\|C\|_2 Ix, x) \leq (Cx,x) \leq (\|C\|_2 Ix, x)$$

Прибавим к обеим частям (Ax,x):

$$((A - \|C\|_2 I)x, x) \leq (Bx,x) \leq ((A + \|C\|_2 I)x, x)$$

$$\lambda_i(A - \|C\|_2 I) \leq \lambda_i(B) \leq \lambda_i(A + \|C\|_2 I)$$

$$\lambda_i(A) - \|C\|_2 \leq \lambda_i(B) \leq \lambda_i(A) + \|C\|_2$$

Ни что иное как другая форма записи того утверждения, что нам нужно доказать:

$$|\lambda_i(B) - \lambda_i(A)| \leq \|B - A\|_2$$

Теперь главное: нафига мы это делали? Что это значит? Если две матрицы близки, то тогда и собственные значения них близки. Это ни что иное, как устойчивость.

`10 Декабря 2024`

6.1. Метод вращения

$$\Lambda = U^* A U$$

$$U^* = U^{-1}$$

Как строятся итерации?

$$A_k = U_k^* A U_k \underset{k \to \infty}{\longrightarrow} \Lambda$$

Матрица вращения:

$$C_{lm} = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & c & \dots & -s & & \\ & & & \vdots & \ddots & \vdots & & \\ & & & s & \dots & c & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}$$

$$C_{lm}[m, l] = s$$

$$c^2 + s^2 = 1$$

$$c = \cos\varphi \quad s = \sin\varphi$$

$$C_{lm}^* C_{lm} = C_{lm} C_{lm}^* = I$$

$$|a_{lm}| = \max_{i<j} |a_{ij}|$$

Матрицу вращения строим опираясь на этот максимальный наддиагональный элемент.

$$B = C_{lm}^* A C_{lm}$$

Мы построим преобразование так, чтобы обнулить этот максимальный элемент.

Мы докажем, что все элементы, которые выше диагонали, будут постепенно стремиться к нулю.

$$B^* = (C_{lm}^* A C_{lm})^* = C_{lm}^* A C_{lm} = B$$

Теперь распишем скалярные формулы.

$$D := A C_{lm}$$

$$B = C_{lm}^* D$$

$$d_{il} = c a_{il} + s a_{im}, \quad i = 1, ..., n$$

$$d_{im} = -sa_{il} + ca_{im}, \quad i = 1, ..., n$$

$$b_{li} = cd_{li} + sd_{mi}, \quad i = 1, ..., n$$

$$b_{mi} = -sd_{li} + cd_{mi}, \quad i = 1, ..., n$$

Рассмотрим случай $i \neq l,\ i \neq m$:

$$d_{li} = a_{li} \quad d_{mi} = a_{mi}$$

$$b_{li} = ca_{li} + sa_{mi}$$

$$b_{mi} = -sa_{li} + ca_{mi}$$

Теперь, если $i = l$:

$$a_{lm} = a_{ml}$$

$$\begin{gathered} b_{ll} = cd_{ll} + sd_{ml} = c(ca_{ll} + sa_{lm}) + s(ca_{ml} + sa_{mm}) = \\ = c^2 a_{ll} + 2csa_{lm} + s^2 a_{mm} \end{gathered}$$

Если $i = m$:

$$\begin{gathered} b_{lm} = cd_{lm} + sd_{mm} = c(-sa_{ll} + ca_{lm}) + s(-sa_{ml} + ca_{mm}) = \\ = cs(a_{mm} - a_{ll}) + (c^2 - s^2)a_{lm} \end{gathered}$$

$$\begin{gathered} b_{mm} = -sd_{lm} + cd_{mm} = -s(-sa_{ll} + ca_{lm}) + c(-sa_{ml} + ca_{mm}) = \\ = s^2 a_{ll} + c^2 a_{mm} - 2csa_{lm} \end{gathered}$$

Наша задача:

$$b_{lm} = 0$$

$$cs(a_{mm} - a_{ll}) + (c^2 - s^2)a_{lm} = 0$$

$$cs = \cos\varphi \sin\varphi = \frac{1}{2}\sin 2\varphi$$

$$c^2 - s^2 = \cos^2\varphi - \sin^2\varphi = \cos 2\varphi$$

$$\frac{1}{2}\sin 2\varphi \cdot (a_{mm} - a_{ll}) + \cos 2\varphi \cdot a_{lm} = 0$$

$$\operatorname{tg} 2\varphi = \frac{2a_{lm}}{a_{ll} - a_{mm}}$$

Будем считать, что $-\frac{\pi}{2} \le 2\varphi \le \frac{\pi}{2}$, то есть $-\frac{\pi}{4} \le \varphi \le \frac{\pi}{4}$. Так определяем c и s.

Тригонометрия долго считается на компьютере, поэтому в конце мы рассмотрим модификацию.

Утверждение при таком шаге:

$$\sum_{i \ne j} b_{ij}^2 \le \left(1 - \frac{2}{n(n-1)}\right) \sum_{i \ne j} a_{ij}^2$$

При умножении на ортогональную матрицу евклидова норма матрицы не изменяется (!):

$$\|B\|_E = \|C_{lm}^* A C_{lm}\|_E = \|A\|_E$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2$$

$$\sum_{i} b_{ii}^2 + \sum_{i \ne j} b_{ij}^2 = \sum_{i} a_{ii}^2 + \sum_{i \ne j} a_{ij}^2$$

$$\sum_{i \ne j} b_{ij}^2 - \sum_{i \ne j} a_{ij}^2 = \sum_{i} a_{ii}^2 - \sum_{i} b_{ii}^2 = a_{ll}^2 + a_{mm}^2 - b_{ll}^2 - b_{mm}^2$$

Следующий приёмчик. Рассмотрим три квадратные матрицы с потолка, чтобы доказать теорему.

$$\overline{A} = \begin{pmatrix} a_{ll} & a_{lm} \\ a_{ml} & a_{mm} \end{pmatrix}$$

$$\overline{C} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

$$\overline{B} = \begin{pmatrix} b_{ll} & 0 \\ 0 & b_{mm} \end{pmatrix}$$

$$\overline{B} = \overline{C}^* \overline{AC}$$

Это проверятся непосредственно умножением.

$$\left\|\overline{B}\right\|_E = \left\|\overline{A}\right\|_E$$

$$b_{ll}^2 + b_{mm}^2 = a_{ll}^2 + a_{mm}^2 + 2a_{lm}^2$$

Получается:

$$\sum_{i \neq j} b_{ij}^2 - \sum_{i \neq j} a_{ij}^2 = -2a_{lm}^2$$

$$\sum_{i \neq j} a_{ij}^2 \leq \sum_{i \neq j} a_{lm}^2 = a_{lm}^2 \sum_{i \neq j} 1 = a_{lm}^2 n(n-1)$$

$$\sum_{i \neq j} a_{ij}^2 \leq a_{lm}^2 n(n-1)$$

$$-2a_{lm}^2 \leq -\frac{2}{n(n-1)} \sum_{i \neq j} a_{ij}^2$$

$$\sum_{i \neq j} b_{ij}^2 - \sum_{i \neq j} a_{ij}^2 \leq -\frac{1}{n(n-1)} \sum_{i \neq j} a_{ij}^2$$

$$\sum_{i \neq j} b_{ij}^2 \leq \left(1 - \frac{2}{n(n-1)}\right) \sum_{i \neq j} a_{ij}^2$$

Перекур.

$$A_1 = C_1^* A C_1$$

$$C_1 = C_{l_1 m_1}$$

$$l_1 m_1$$

$$A_2 = C_2^* A_1 C_2$$

$$A_k = C_k^* A_{k-1} C_k$$

$$A_k = C_k^* C_{k-1}^* A_{k-2} C_{k-1} C_k = \ldots = C_k^* \ldots C_1^* \cdot A \cdot C_1 \ldots C_k \coloneqq$$
$$\coloneqq (C_1 \ldots C_k)^* \cdot A(C_1 \ldots C_k)$$

$$U_k \coloneqq C_1 \ldots C_k$$

$$A_k = U^* A U$$

Устремляем k в бесконечность. Покажем, что $A_k \to \Lambda$ (к диагональной матрице):

$$\sum_{i \neq j} \left[a_{ij}^{(k)}\right]^2 \leq \left(1 - \frac{2}{n(n-1)}\right) \sum_{i \neq j} \left[a_{ij}^{(k-1)}\right]^2 \leq \ldots \leq$$
$$\leq \left(1 - \frac{2}{n(n-1)}\right)^k \sum_{i \neq j} a_{ij}^2 \longrightarrow 0$$

Когда мы останавливаем итерационный процесс? Когда сумма квадратов внедиагональных элементов меньше ε.

Введём в рассмотрение матрицу (убрали все внедиагональные элементы у A_k:

$$\Lambda_k$$

$$\varepsilon_k = A_k - \Lambda_k$$

$$\lambda_i(\Lambda_k) = a_{ii}^{(k)}$$

$$\lambda_i(A) = \lambda_i(A_k)$$

$$|\lambda_i(A_k) - \lambda_i(\Lambda_k)| \leq \|\varepsilon_k\|_2 \leq \|\varepsilon_k\|_E = \sqrt{\sum \left[a_{ij}^{(k)}\right]^2} < \varepsilon$$

$$\left|\lambda_i(A) - a_{ii}^{(k)}\right| < \varepsilon$$

`17 Декабря 2024`

Вернёмся к методу вращения. Вопросик обкашляем. Как выбирать опорный элемент?

Напоминаем:

$$\operatorname{tg} 2\varphi = \frac{2a_{lm}}{a_{ll} - a_{mm}} \qquad -\frac{\pi}{4} \le \varphi \le \frac{\pi}{4}$$

$$c = \cos\varphi \qquad s = \sin\varphi$$

$$x := 2a_{lm} \qquad y := a_{ll} - a_{mm}$$

$$\operatorname{tg} 2\varphi = \frac{x}{y}$$

Если $y = 0$:

$$2\varphi = \frac{\pi}{2} \Leftrightarrow \varphi = \frac{\pi}{4}$$

$$c = s = \frac{\sqrt{2}}{2}$$

Если $y \ne 0$:

$$\operatorname{tg} 2\varphi = \frac{x}{y} = \frac{\frac{\operatorname{sign}(xy)|x|}{\sqrt{x^2+y^2}}}{\frac{|y|}{\sqrt{x^2+y^2}}} = \frac{\overline{s}}{\overline{c}}$$

$$\overline{s}^2 + \overline{c}^2 = 1$$

$$\overline{c} = \cos\psi \qquad \overline{s} = \sin\psi$$

$$\operatorname{tg} 2\varphi = \operatorname{tg}\psi$$

$$\psi = 2\varphi + k\pi$$

Так как:

$$-\frac{\pi}{2} \le 2\varphi \le \frac{\pi}{2}$$

$$\cos 2\varphi > 0$$

$$\cos \psi > 0$$

То:

$$\psi = 2\varphi + 2k\pi$$

$$\cos \psi = \cos 2\varphi \qquad \sin \psi = \sin 2\varphi$$

$$\cos 2\varphi = 2\cos^2 \varphi - 1 = 2c^2 - 1$$

$$\sin 2\varphi = 2 \sin \varphi \cos \varphi = 2cs$$

Получаем формулы для нахождения c и s:

$$2c^2 - 1 = \frac{|y|}{\sqrt{x^2 + y^2}}$$

$$2cs = \frac{\operatorname{sign}(xy)|x|}{\sqrt{x^2 + y^2}}$$

Теперь хотим не для симметрических, а для произвольных матриц чтоб работало.

Используем QR-метод (метод Пуэра):

$$A = QR$$

$$A = Q_1 R_1 \qquad A_1 = R_1 Q_1 \qquad A_1 = Q_1^* A Q_1$$

$$A_1 = Q_2 R_2 \qquad A_2 = R_2 Q_2 \qquad A_2 = Q_2^* A_1 Q_2$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$A_{k-1} = Q_k R_k \qquad A_k = R_k Q_k \qquad \ldots$$

Для всякой матрицы есть перевод в нормальную Жорданову форму:

$$\Lambda = T^{-1} A T$$

$$\Lambda = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_m \end{pmatrix} \qquad J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}$$

Теорема

Пусть все собственные значения невырожденной матрицы вещественны и различны по модулю.

Пусть все главные миноры T^{-1} отличны от нуля:

$$\Lambda = T^{-1}AT$$

Тогда последовательность A_k сходится к верхней треугольной матрице.

Мы на каждом шаге делаем QR (Пуэр) разложение, его вычислительная сложность $O(n^3)$. Это дорого.

Верхняя почтитреугольная матрица Хессенберга

Матрица, у которой все элементы лежащие ниже первой поддиагонали равны нулю.

6.2. Итерационный метод (прямых, обратных).

Этим методом можно найти минимальное и максимальное собственные значения.

Рассмотрим случай, когда собственные значения различны и вещественны по модулю.

На самом деле, алгоритм работает и на комплексных числах, но мы просто рассмотрим на простом случае, чтобы было наглядно.

Занумеруем собственные числа в порядке убывания модуля:

$$|\lambda_1(A)| > \ldots > |\lambda_n(A)|$$

Полная система собственных векторов, их можно рассмтривать как базис:

$$u^{(1)}, ..., u^{(n)}$$

$$Au^{(i)} = \lambda_i u^{(i)}$$

Итерационный процесс:

$$x^{(0)},\ x^{(1)} = Ax^{(0)},\ ...,\ x^{(k)} = Ax^{(k-1)}$$

$$x^{(k)} = A^k x^{(0)}$$

$$x^{(0)} = \sum_{i=1}^{n} c_i u^{(i)}$$

$$x^{(k)} = A^k x^{(0)} = A^k \sum_{i=1}^{n} c_i u^{(i)} = \sum_{i=1}^{n} c_i A^k u^{(i)} = \sum_{i=1}^{n} c_k \lambda_i^k u^{(i)} =$$

$$= c_1 \lambda_1^k u^{(1)} + \sum_{i=2}^{n} c_k \lambda_i^k u^{(i)} = \lambda_1^k \left(c_1 u^{(1)} + \sum_{i=2}^{n} c_k \frac{\lambda_i^k}{\lambda_1} u^{(i)} \right) :=$$

$$:= \lambda_1^k v^{(k)}$$

$$k \to \infty \qquad v^{(k)} \to c_1 u^{(1)}$$

То есть $v^{(k)}$ стремится к первому собственному вектору.

$$x^{(k)} = \lambda_1^k v^{(k)}$$

Хуйня полная, если $|\lambda_1| > 1$, то расходится, если $|\lambda_1| < 1$, то ноль будет.

Помогает нормировочка. Таким вот образом, на каждом шаге:

У $x^{(k)}$ выделяем максимальну компоненту:

$$|\alpha_k| := \max_{1 \leq i \leq n} \left| x_i^{(k)} \right|$$

$$x^{(k)} = A\frac{x^{(k-1)}}{\alpha_{k-1}}$$

$$\frac{x^{(k)}}{\alpha_k} \to \frac{u^{(1)}}{\alpha} \qquad \alpha = \max_i |u_i|^1$$

$$\frac{x^{(k-1)}}{\alpha_{k-1}} \xrightarrow[k\to\infty]{} \frac{u^{(1)}}{\alpha}$$

$$A\frac{x^{(k-1)}}{\alpha_{k-1}} \to A\frac{u^{(1)}}{\alpha} = \lambda_1\frac{u^{(1)}}{\alpha}$$

Это был метод прямых итераций. Метод обратных итераций делается с A^{-1}:

$$x^{(k)} = A^{-1}\frac{x^{(k-1)}}{\alpha_{k-1}}$$

$$Ax^{(k)} = \frac{x^{(k-1)}}{\alpha_{k-1}}$$

Тогда:

$$\frac{1}{\alpha_k} \longrightarrow \lambda_n(A)$$