Часть 2. Итерационные методы решения системы линейных алгебраических уравнений

Тема 8. Итерационные методы решения линейных уравнений.

Для решения системы $A\vec{x} = \vec{b}:\ \begin{cases} A \in R^{n\times n},\ \det A \neq 0 \\ \vec{b} \in R^n \end{cases}$ – заданы,

$\vec{x} \in R^n$ – неизвестный вектор,

итерационный метод – это способ построения $\{\vec{x}^k \in R^n\}_{k=0}^{\infty}$:

$$\vec{x}^k = \vec{x}^k(\vec{b},\ \vec{x}^0,\ \vec{x}^1,\ \ldots,\ \vec{x}^{k-1}) \xrightarrow{k\to\infty} \vec{x}, \quad \text{т.е. } \lim_{k\to\infty} \|\vec{x}^k - \vec{x}\| = 0.$$

Упражнение. Доказать, что система $A\vec{x} = \vec{b}:\ \begin{cases} A \in C^{n\times n},\ \det A \neq 0 \\ \vec{b} \in C^n \end{cases}$

эквивалентна системе $\begin{bmatrix} \mathrm{Re}\ A & -\mathrm{Im}\ A \\ \mathrm{Im}\ A & \mathrm{Re}\ A \end{bmatrix} \begin{pmatrix} \mathrm{Re}\ \vec{x} \\ \mathrm{Im}\ \vec{x} \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\ \vec{b} \\ \mathrm{Im}\ \vec{b} \end{pmatrix}$ в R^{2n}.

Итерационное уточнение “прямого метода”:

систему $A\vec{x} = \vec{b}$ решаем $\vec{x} = A^{-1}\vec{b}$, где $A^{-1} \equiv (LU)^{-1}$;

из-за ошибок округления $\left\{\begin{array}{l} A^{-1} \approx \tilde{A}^{-1} \equiv (\tilde{L}\tilde{U})^{-1} \\ \tilde{\vec{x}} = \tilde{A}^{-1}\vec{b} \neq \vec{x} \end{array}\right\} \Rightarrow A(\underbrace{\tilde{\vec{x}} - \vec{x}}_{\tilde{\vec{z}}}) = \underbrace{A\tilde{\vec{x}} - \vec{b}}_{\tilde{\vec{r}}}$

$$\Rightarrow \begin{cases} \tilde{\vec{z}} = A^{-1}\tilde{\vec{r}} \approx \tilde{A}^{-1}\tilde{\vec{r}} \\ \tilde{\vec{x}} - \vec{x} \approx \tilde{A}^{-1}\tilde{\vec{r}} \end{cases} \Rightarrow \begin{cases} \vec{x} \approx \tilde{\vec{x}} - \tilde{A}^{-1}\tilde{\vec{r}} \\ \tilde{\tilde{\vec{x}}} \approx \tilde{\vec{x}} - \tilde{A}^{-1}\tilde{\vec{r}} \end{cases} \Rightarrow \begin{cases} \vec{x}^0 = \tilde{A}^{-1}\vec{b} \\ \vec{x}^{k+1} = \vec{x}^k - \tilde{A}^{-1}(A\vec{x}^k - \vec{b}) \end{cases}$$

Оценка ошибки $\begin{cases} A\vec{z}^0 \equiv \vec{r}^0 \\ \vec{z}^k \equiv A^{-1}\vec{r}^k \end{cases} \Rightarrow \begin{cases} \|\vec{r}^0\| \le \|A\| \cdot \|\vec{z}^0\| \\ \|\vec{z}^k\| \le \|A^{-1}\| \cdot \|\vec{r}^k\| \end{cases}$

$$\Rightarrow \frac{\|\vec{z}^k\|}{\|\vec{z}^0\|} \le \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\vec{r}^k\|}{\|\vec{r}^0\|} = Cond\, A \cdot \frac{\|\vec{r}^k\|}{\|\vec{r}^0\|}$$

Одношаговый (двухслойный) итерационный метод

$$\vec{x}^0 - \textit{задан},\quad \begin{cases} \vec{x}^{k+1} = \vec{x}^k - H_k(A\vec{x}^k - \vec{b}), \\ k = 0, 1, 2, \ldots; \end{cases} \quad \begin{matrix} H_k - \textit{заданные матрицы}, \\ \det H_k \neq 0 \end{matrix}$$

1. $\vec{x}^k$ – k-тое приближение (итерация) к $\vec{x}$,
2. $\vec{z}^k = \vec{x}^k - \vec{x}$ – вектор ошибки k-той итерации (приближения),
3. $\vec{r}^k = A\vec{x}^k - \vec{b} \equiv A\vec{z}^k$ – вектор невязки (residual) k-той итерации,
4. $S_k = E - H_k A$ – матрица шага для ошибки:

$$\underbrace{\vec{x}^{k+1} - \vec{x}}_{\vec{z}^{k+1}} = \underbrace{\vec{x}^k - \vec{x}}_{\vec{z}^k} - H_k(\underbrace{A\vec{x}^k - A\vec{x}}_{A\vec{z}^k}) \Rightarrow \vec{z}^{k+1} = S_k \vec{z}^k,$$

5. $T_k = E - AH_k$ – матрица шага для невязки (residual):

$$\underbrace{A\vec{x}^{k+1} - \vec{b}}_{\vec{r}^{k+1}} = \underbrace{A\vec{x}^k - \vec{b}}_{\vec{r}^k} - AH_k(\underbrace{A\vec{x}^k - \vec{b}}_{\vec{r}^k}) \Rightarrow \vec{r}^{k+1} = T_k \vec{r}^k.$$

6. Метод называется **сходящимся**, если $\lim\limits_{k\to\infty} \| z^k \| = 0 \quad \forall\ x^0 \in R^n$.

Стационарный одношаговый итерационный метод

$$\vec{x}^{\,0} - \textit{задан}, \quad \begin{cases} \vec{x}^{\,k+1} = \vec{x}^{\,k} - H\,(A\,\vec{x}^{\,k} - \vec{b}), \\ k = 0,\,1,\,2,\,\ldots\,; \end{cases} \quad H - \textit{заданная матрица}, \quad \det H \neq 0$$

1. $\vec{z}^{\,k} = \vec{x}^{\,k} - \vec{x}$ – вектор ошибки k-той итерации,
2. $S = E - HA$ – матрица шага для ошибки (***не зависит от*** k):
   $\vec{z}^{\,k+1} = S\,\vec{z}^{\,k}$,
3. $\vec{r}^{\,k} = A\vec{x}^{\,k} - \vec{b} \equiv A\vec{z}^{\,k}$ – вектор невязки k-той итерации,
4. $T = E - AH$ – матрица шага для невязки (***не зависит от*** k):
   $\vec{r}^{\,k+1} = T\,\vec{r}^{\,k}$.
5. Метод сходится, если $\begin{cases} \forall\ \vec{z}^{\,0} \in R^n \quad \vec{z}^{\,k} = S^k \vec{z}^{\,0} \to 0 \\ \forall\ \vec{r}^{\,0} \in R^n \quad \vec{r}^{\,k} = T^k \vec{r}^{\,0} \to 0 \end{cases}$

Достаточные условия сходимости.

Теорема. Если $\| S \| = \| E - HA \| < 1$,

$$\text{то } \| \vec{z}^{\,k} \| \to 0, \text{ т.е. } \vec{x}^{\,k} \to \vec{x} \quad \forall\ \vec{x}^{\,0} \in R^n.$$

Доказательство. $\vec{z}^{\,k} = S\,\vec{z}^{\,k-1} \Rightarrow \| \vec{z}^{\,k} \| = \| S\,\vec{z}^{\,k-1} \| \le \| S \| \cdot \| \vec{z}^{\,k-1} \| \le$

$$\le \| S \| \cdot \| S \| \cdot \| \vec{z}^{\,k-2} \| \le \ldots \le \| S \|^k \cdot \| \vec{z}^{\,0} \|$$

$$\underbrace{\| S \|}_{<1}^{k} \to 0 \Rightarrow \| S \|^k \cdot \| \vec{z}^{\,0} \| \to 0 \Rightarrow \| \vec{z}^{\,k} \| \to 0 \Rightarrow \vec{z}^{\,k} = \vec{x}^{\,k} - \vec{x} \to 0.$$

Теорема. Если $\| T \| = \| E - AH \| < 1$,

$$\text{то } \| \vec{z}^{\,k} \| \to 0, \text{ т.е. } \vec{x}^{\,k} \to \vec{x} \quad \forall\ \vec{x}^{\,0} \in R^n.$$

Доказательство. 1. $\vec{r}^{\,k} = T\,\vec{r}^{\,k-1} = \ldots = T^{\,k}\vec{r}^{\,0} \to 0$

2. $\vec{z}^{\,k} = A^{-1} r^k \Rightarrow \| \vec{z}^{\,k} \| = \| A^{-1} r^k \| \le \| A^{-1} \| \cdot \| r^k \| \to 0$

$$\Rightarrow \vec{z}^{\,k} = \vec{x}^{\,k} - \vec{x} \to 0.$$

Необходимое и достаточное условие сходимости.

Теорема. Стационарный итерационный метод

$$\vec{x}^0 - \textit{задан}, \quad \begin{cases} \vec{x}^{k+1} = \vec{x}^k - H\,(A\vec{x}^k - \vec{b}), \\ k = 0, 1, 2, \ldots; \end{cases} \quad H - \textit{заданная матрица}, \; \det H \neq 0$$

сходится $\Leftrightarrow$ $\rho(S) < 1$,

где $S = E - HA$ – матрица шага для ошибки $z^k = x^k - x$,

$\rho(S) = \max\limits_{\lambda(S)} |\lambda(S)|$ – спектральный радиус.

Доказательство. 1. $\vec{z}^k \to 0 \;\; \forall\, \vec{z}^0 \in R^n \;\; \overset{?}{\Rightarrow} \;\; \rho(S) < 1$:

Докажите: $\vec{z}^k = S^k \vec{z}^0 \to 0 \;\; \forall\, \vec{z}^0 \in R^n \;\; \Rightarrow \;\; \vec{z}^k = S^k \vec{z}^0 \to 0 \;\; \forall\, \vec{z}^0 \in C^n$.

Тогда $\vec{z}^k = S^k \vec{z}^0 = [\lambda(S)]^k \cdot \vec{z}^0 \to 0 \;\; \forall\, \vec{z}^0 : \; S \cdot \vec{z}^0 = \lambda(S) \cdot \vec{z}^0$

$$\Rightarrow \; [\lambda(S)]^k \to 0 \;\; \Rightarrow \;\; |\lambda(S)| < 1 \;\; \Rightarrow \;\; \rho(S) < 1.$$

$$2.\ \rho(S)<1 \ \overset{?}{\Rightarrow}\ \vec{z}^{\,k}\to 0\ \ \forall\ \vec{z}^{\,0}\in R^n:$$

Заметим: $\forall\ \varepsilon>0\ \ \exists\ Q_\varepsilon\in C^{n\times n}:\ (Q_\varepsilon)^{-1}SQ_\varepsilon=J_\varepsilon=diag\{J_{1,\varepsilon},\ ...,\ J_{m,\varepsilon}\}$

$$J_{i,\varepsilon}=\begin{bmatrix}\lambda_i & \varepsilon & & 0\\ & \lambda_i & \ddots & \\ & & \ddots & \varepsilon\\ 0 & & & \lambda_i\end{bmatrix},\ \ \lambda_i\in Sp(S);\ \|J_\varepsilon\|_\infty=\max_{\lambda_i\in Sp(S)}|\lambda_i|+\varepsilon=\rho(S)+\varepsilon$$

Так как $\rho(S)<1$, то $\exists\ \varepsilon>0:\ \|J_\varepsilon\|_\infty=\rho(S)+\varepsilon<1\ \Rightarrow\ (\|J_\varepsilon\|_\infty)^k\overset{k\to\infty}{\to}0$

Тогда

$$\vec{z}^{\,k}=S\vec{z}^{\,k-1}\equiv Q_\varepsilon J_\varepsilon Q_\varepsilon^{-1}\vec{z}^{\,k-1}=Q_\varepsilon(J_\varepsilon)^k Q_\varepsilon^{-1}\vec{z}^{\,0}$$

$$\|\vec{z}^{\,k}\|_\infty\le\|Q_\varepsilon\|_\infty\cdot\|Q_\varepsilon^{-1}\|_\infty\cdot(\|J_\varepsilon\|_\infty)^k\cdot\|\vec{z}^{\,0}\|_\infty\overset{k\to\infty}{\to}0,$$

т.е. метод сходится.

Теорема (о количестве итераций).

Если $\| S \| = \| E - HA \| < 1$, то оценка $\dfrac{\| \vec{z}^{\,k} \|}{\| \vec{z}^{\,0} \|} \le \varepsilon$ для метода

$$\vec{x}^{\,0} - \textit{задан}, \quad \begin{cases} \vec{x}^{\,k+1} = \vec{x}^{\,k} - H\,(A\vec{x}^{\,k} - \vec{b}), \\ k = 0,\ 1,\ 2,\ \dots\ ; \end{cases} \quad \begin{matrix} H - \textit{заданная матрица}, \\ \det H \ne 0 \end{matrix}$$

достигается после $\forall\ k \ge k(\varepsilon) = \dfrac{-\ln \varepsilon}{-\ln \| S \|}$ итераций.

Доказательство. При $k \ge k(\varepsilon)$ имеем

$$\| z^k \| = \| S^k z^0 \| \le \| S \|^k \cdot \| z^0 \| \le$$

$$\le \| S \|^{k(\varepsilon)} \cdot \| z^0 \| = e^{k(\varepsilon)\cdot \ln \|S\|} \cdot \| z^0 \| =$$

$$= e^{\ln \varepsilon} \cdot \| z^0 \| = \varepsilon \cdot \| z^0 \|.$$

Асимптотическая скорость сходимости.

Для итерационного метода

$$\vec{x}^0 - \textit{задан},\ \begin{cases} \vec{x}^{k+1} = \vec{x}^k - H(A\vec{x}^k - \vec{b}), \\ k = 0, 1, 2, \dots ; \end{cases} \quad H - \textit{заданная матрица}, \ \det H \neq 0$$

определим $d_k = \sqrt[k]{\| S^k \|}$.

Тогда из $\dfrac{\| z^k \|}{\| z^0 \|} \le \| S^k \| = (d_k)^k \le \varepsilon \quad \Rightarrow \quad k(\varepsilon) = \dfrac{-\ln \varepsilon}{-\ln d_k} \equiv \dfrac{-\ln \varepsilon}{-\ln \sqrt[k]{\| S^k \|}}$

или $\underbrace{-\ln \sqrt[k]{\| S^k \|}}_{\textit{скорость}} \cdot \underbrace{k(\varepsilon)}_{\textit{время}} = \underbrace{-\ln \varepsilon}_{\textit{растояние}}$

$R_k = -\ln \sqrt[k]{\| S^k \|}$ – **средняя скорость сходимости** итерационного метода за k итераций.

Упражнение. Если $\| S \| = \| E - HA \| < 1$, то $R_k \ge -\ln \| S \|$.

$R_\infty = \lim\limits_{k\to\infty} R_k = -\ln\{\lim\limits_{k\to\infty} \sqrt[k]{\| S^k \|}\}$ – **асимптотическая скорость** сходимости стационарного итерационного метода.

Теорема. Если $\rho(S) = \rho(E - HA) < 1$, то $R_\infty = -\ln\, \rho(S)$.

Доказательство. Достаточно доказать, что $\exists \lim\limits_{k\to\infty} \sqrt[k]{\| S^k \|} = \rho(S)$,

т.е. $\varliminf\limits_{k\to\infty} \sqrt[k]{\| S^k \|} = \varlimsup\limits_{k\to\infty} \sqrt[k]{\| S^k \|} = \rho(S)$.

1. $\rho(S) \le \varliminf\limits_{k\to\infty} \sqrt[k]{\| S^k \|}$:

так как $\rho(S) \le \| S \| \quad \forall \| S \| = \max\limits_{\vec{z}\ne 0} \dfrac{\| S\vec{z} \|}{\| \vec{z} \|}$ & $\lambda(S^k) = [\lambda(S)]^k$,

то $[\rho(S)]^k \le \| S^k \| \;\Rightarrow\; \rho(S) \le \sqrt[k]{\| S^k \|} \;\Rightarrow\; \rho(S) \le \varliminf\limits_{k\to\infty} \sqrt[k]{\| S^k \|}$.

2. $\overline{\lim\limits_{k\to\infty}}\sqrt[k]{\|S^k\|_\infty} \le \rho(S)$:

знаем: $\forall\ \varepsilon>0\ \ \exists\ Q_\varepsilon$:

$$\|S^k\vec{z}^{\,0}\|_\infty \le \underbrace{\|Q_\varepsilon\|_\infty\cdot\|Q_\varepsilon^{-1}\|_\infty}_{c_\varepsilon}\cdot[\rho(S)+\varepsilon]^k\cdot\|\vec{z}^{\,0}\|_\infty \quad \forall\ \vec{z}^{\,0}\in C^n.$$

$$\Rightarrow \begin{cases} \|S^k\|_\infty = \max\limits_{\vec{z}\ne 0}\dfrac{\|S^k\,\vec{z}\|_\infty}{\|\vec{z}\|_\infty} \le c_\varepsilon\cdot[\rho(S)+\varepsilon]^k \\ \Rightarrow\ \sqrt[k]{\|S^k\|_\infty} \le \sqrt[k]{c_\varepsilon}\cdot[\rho(S)+\varepsilon] \end{cases}$$

$$\Rightarrow \overline{\lim\limits_{k\to\infty}}\sqrt[k]{\|S^k\|_\infty} \le \underbrace{\lim\limits_{k\to\infty}\sqrt[k]{c_\varepsilon}}_{\equiv 1}\cdot[\rho(S)+\varepsilon] \equiv \rho(S)+\varepsilon \quad \forall\ \varepsilon>0$$

$$\Rightarrow \overline{\lim\limits_{k\to\infty}}\sqrt[k]{\|S^k\|_\infty} \le \lim\limits_{\varepsilon\to 0}[\rho(S)+\varepsilon] = \rho(S).$$

Знаем: $\exists\, \beta:\ \forall\ A \in C^{n\times n} \quad \| A \| \le \beta \cdot \| A \|_\infty \Rightarrow \| S^k \| \le \beta \cdot \| S^k \|_\infty$

$$\Rightarrow \ \varlimsup_{k\to\infty} \sqrt[k]{\| S^k \|} \le \varlimsup_{k\to\infty} \sqrt[k]{\beta \cdot \| S^k \|_\infty} \le \underbrace{\lim_{k\to\infty} \sqrt[k]{\beta}}_{\equiv 1} \cdot \varlimsup_{k\to\infty} \sqrt[k]{\| S^k \|_\infty} \le \rho(S).$$

Итак,

$$\rho(S) \le \varliminf_{k\to\infty} \sqrt[k]{\| S^k \|} \le \varlimsup_{k\to\infty} \sqrt[k]{\| S^k \|} \le \rho(S)$$

$\Rightarrow$

$$\rho(S) = \varliminf_{k\to\infty} \sqrt[k]{\| S^k \|} = \varlimsup_{k\to\infty} \sqrt[k]{\| S^k \|} = \rho(S).$$

Замечание. Определить количество итераций, необходимое для уменьшения начальной ошибки в $1/\varepsilon$ раз: $\dfrac{\| z^k \|}{\| z^0 \|} \le \varepsilon$, в конкретной норме, зная только $R_\infty = \rho(S)$, невозможно.

Тема 9. Итерационный метод Якоби.

Система $A\vec{x} = \vec{b}$ с матрицей $A = B - C$:

$$B\vec{x} = C\vec{x} + \vec{b} \;\Rightarrow\; \begin{cases} \vec{x}^{\,0} - \text{задан}, \\ B\vec{x}^{\,k+1} = C\vec{x}^{\,k} + \vec{b}, \quad k = 0, 1, \ldots \end{cases}$$

Упражнение. Доказать $\vec{x}^{\,k+1} = \vec{x}^{\,k} - B^{-1}(A\vec{x}^{\,k} - \vec{b})$, $S = E - B^{-1}A$

Пример.

$$\begin{cases} a_{1,1}x_1^{k+1} = -a_{1,2}x_2^k - a_{1,3}x_3^k + b_1 \\ a_{2,2}x_2^{k+1} = -a_{2,1}x_1^k - a_{2,3}x_3^k + b_2 \\ a_{3,3}x_3^{k+1} = -a_{3,1}x_1^k - a_{3,2}x_2^k + b_3 \end{cases} \Rightarrow \begin{cases} B = D \equiv diag\, A \\ C = D - A \\ B\vec{x}^{\,k+1} = C\vec{x}^{\,k} + \vec{b} \end{cases}$$

Если $D = diag\, A = diag\{a_{11}, a_{22}, ..., a_{nn}\}$, $\det D \neq 0$,

то итерационный процесс

$$\begin{cases} \vec{x}^0 - \text{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - D^{-1}(A\vec{x}^k - \vec{b}), \quad k = 0, 1, ... \end{cases}$$

– **метод Якоби** для решения системы $Ax = b$.

Упражнение. Матрица шага для ошибки $\vec{z}^{k+1} = S\,\vec{z}^k$:

$$S = -\begin{bmatrix} 0 & a_{1,1}^{-1}a_{1,2} & \dots & a_{1,1}^{-1}a_{1,n} \\ a_{2,2}^{-1}a_{2,1} & 0 & \dots & a_{2,2}^{-1}a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,n}^{-1}a_{n,1} & a_{n,n}^{-1}a_{n,2} & \dots & 0 \end{bmatrix},$$

$$\| S \|_\infty = \max_{1 \le i \le n} \frac{|a_{1,i}| + ... + |a_{i-1,i}| + |a_{i+1,i}| + ... + |a_{n,n}|}{|a_{i,i}|}$$

Теорема.

Если матрица A имеет диагональное преобладание по строкам:

$$|a_{i,i}| > \sum_{j=1,\ j\neq i}^{n} |a_{i,j}| \quad \forall\, i,$$

то метод Якоби сходится к решению системы $A\vec{x} = \vec{b}$.

Доказательство. Из условий теоремы $\Rightarrow$

$$\frac{|a_{1,i}| + \ldots + |a_{i-1,i}| + |a_{i+1,i}| + \ldots + |a_{n,n}|}{|a_{i,i}|} < 1 \quad \forall\, i = 1,\, 2,\, \ldots,\, n$$

$\Rightarrow$ (см. упражнение) $\|S\|_\infty < 1$,

т.е. выполняется достаточное условие сходимости стационарного итерационного метода.

Упражнение. Матрица шага для невязки $\vec{r}^{\,k+1} = T\,\vec{r}^{\,k}$:

$$T = -\begin{bmatrix} 0 & a_{1,2}a_{2,2}^{-1} & \dots & a_{1,n}a_{n,n}^{-1} \\ a_{2,1}a_{1,1}^{-1} & 0 & \dots & a_{2,n}a_{n,n}^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1}a_{1,1}^{-1} & a_{n,2}a_{2,2}^{-1} & \dots & 0 \end{bmatrix},$$

$$\| T \|_1 = \max_{1 \le j \le n} \frac{|a_{1,j}| + \ldots + |a_{j-1,j}| + |a_{j+1,j}| + \ldots + |a_{n,j}|}{|a_{j,j}|}$$

Теорема.

Если матрица A имеет диагональное преобладание по столбцам:

$$|a_{j,j}| > \sum_{i=1,\ i \ne j}^{n} |a_{i,j}| \quad \forall\ j,$$

то метод Якоби сходится к решению системы $A\vec{x} = \vec{b}$.

Доказательство. Из условий теоремы $\Rightarrow$

$$\frac{|a_{1,j}|+\ldots+|a_{j-1,j}|+|a_{j+1,j}|+\ldots+|a_{n,j}|}{|a_{j,j}|}<1 \quad \forall\ j=1,\,2,\,\ldots,\,n$$

$\Rightarrow$ (см. упражнение) $\|T\|_1<1$,

т.е. выполняется достаточное условие сходимости стационарного итерационного метода.

Замечание. $S=E-D^{-1}A=D^{-1}(E-AD^{-1})D\equiv D^{-1}TD$

$\Rightarrow\ Sp(S)=Sp(T) \quad \Rightarrow \quad \rho(S)=\rho(T)$,

но $\rho(T)\le\|T\|_1<1 \quad \Rightarrow \quad \rho(S)<1$

– необходимое и достаточное условие сходимости стационарного итерационного метода.

Необходимое и достаточное условие сходимости метода Якоби в случае симметричной матрицы с положительной диагональю

Теорема.

Если $A = A^*$ и $D \equiv diag\, A = diag\,\{a_{1,1},\, a_{2,2},\, \dots,\, a_{n,n}\} > 0$,

то метод Якоби:

$$\begin{cases} \vec{x}^0 - \textit{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - D^{-1}(A\vec{x}^k - \vec{b}),\ k = 0, 1, \dots, \end{cases}$$

сходится к решению системы $A\vec{x} = \vec{b}$

(т.е. выполняется условие сходимости: $\forall\ |\lambda(\underbrace{E - D^{-1}A}_{S})| < 1$)

$\Leftrightarrow$

$A > 0$ и $2D - A > 0$

(т.е. выполняются условия: $\forall\ \lambda(A) > 0$ и $\forall\ \lambda(2D - A) > 0$).

Доказательство.

Заметим:

1. $D > 0 \Rightarrow \forall\ a_{i,i} > 0 \Rightarrow D^{1/2} = diag\{\sqrt{a_{1,1}}, \sqrt{a_{2,2}}, ..., \sqrt{a_{n,n}}\} > 0$

2. $Sp(S) = Sp(D^{1/2} S D^{-1/2}) \equiv Sp(E - D^{-1/2} A D^{-1/2})$

$\Rightarrow \lambda(S) = 1 - \lambda(D^{-1/2} A D^{-1/2})$

3. $A = A^* \Rightarrow D^{-1/2} A D^{-1/2} = (D^{-1/2} A D^{-1/2})^*$

$\Rightarrow \forall\ \lambda(D^{-1/2} A D^{-1/2})$ – вещественные числа

4. Из 1. – 3. $\Rightarrow \forall\ |\lambda(S)| < 1 \Leftrightarrow \begin{cases} -1 < 1 - \lambda(D^{-1/2} A D^{-1/2}) < 1 \\ \Leftrightarrow \\ 0 < \lambda(D^{-1/2} A D^{-1/2}) < 2 \end{cases}$

Итак, нужно доказать: $0 < \lambda(D^{-1/2} A D^{-1/2}) < 2 \Leftrightarrow \begin{cases} A > 0 \\ 2D - A > 0 \end{cases}$

Необходимость: $0<\lambda(D^{-1/2}AD^{-1/2})<2 \overset{?}{\Rightarrow} \begin{cases} A>0 \\ 2D-A>0 \end{cases}$

1. $\forall\ \lambda(D^{-1/2}AD^{-1/2})>0 \Rightarrow D^{-1/2}AD^{-1/2}>0$

$\Rightarrow (D^{-1/2}AD^{-1/2}\vec{y},\, \vec{y})_2>0 \quad \forall\ \vec{y}\neq 0,$

но $(D^{-1/2}AD^{-1/2}\vec{y},\, \vec{y})_2 \equiv (A\underbrace{D^{-1/2}\vec{y}}_{\vec{z}},\, \underbrace{D^{-1/2}\vec{y}}_{\vec{z}})_2>0 \quad \forall\ \vec{z}\neq 0$

$\Rightarrow A>0.$

2. $\forall\ 2-\lambda(D^{-1/2}AD^{-1/2})>0 \Rightarrow 2E-D^{-1/2}AD^{-1/2}>0$

$\Rightarrow (D^{-1/2}(2D-A)D^{-1/2}\vec{y},\, \vec{y})_2>0 \quad \forall\ \vec{y}\neq 0,$

но $(D^{-1/2}(2D-A)D^{-1/2}\vec{y},\, \vec{y})_2 \equiv ((2D-A)\underbrace{D^{-1/2}\vec{y}}_{\vec{z}},\, \underbrace{D^{-1/2}\vec{y}}_{\vec{z}})_2>0 \quad \forall\ \vec{z}\neq 0$

$\Rightarrow 2D-A>0.$

Достаточность: $\begin{cases} A>0 \\ 2D-A>0 \end{cases} \overset{?}{\Rightarrow} \; 0<\lambda(D^{-1/2}AD^{-1/2})<2$

1. $A>0 \Rightarrow (A\underbrace{\vec{z}}_{D^{-1/2}\vec{y}},\underbrace{\vec{z}}_{D^{-1/2}\vec{y}})_2 \equiv (D^{-1/2}AD^{-1/2}\vec{y},\vec{y})_2>0 \quad \forall\ \vec{z}\neq 0$

$$\Rightarrow D^{-1/2}AD^{-1/2}>0 \Rightarrow \forall\ \lambda(D^{-1/2}AD^{-1/2})>0.$$

2. $2D-A>0 \Rightarrow \begin{cases} ((2D-A)\underbrace{\vec{z}}_{D^{-1/2}\vec{y}},\underbrace{\vec{z}}_{D^{-1/2}\vec{y}})_2 \equiv \\ \equiv([2E-D^{-1/2}AD^{-1/2}]\vec{y},\vec{y})_2>0 \quad \forall\ \vec{z}\neq 0 \end{cases}$

$$\Rightarrow 2E-D^{-1/2}AD^{-1/2}>0 \Rightarrow \forall\ 2-\lambda(D^{-1/2}AD^{-1/2})>0$$

$$\Rightarrow \forall\ \lambda(D^{-1/2}AD^{-1/2})<2.$$

Упражнение. Исследуйте сходимость метода Якоби для $A\vec{x} = \vec{b}$: вычислите $\rho(S)$, где

$$A = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix} = A^*, \quad D = E > 0$$

Проверьте и сделайте вывод о сходимости метода Якоби:

$$A = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}, \quad \lambda_1(A) = 2, \;\; \lambda_2(A) = \lambda_3(A) = 0.5 \;\; \Rightarrow \;\; A > 0,$$

$$2D - A = \begin{bmatrix} 1 & -0.5 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & -0.5 & 1 \end{bmatrix}, \quad \lambda_1(2D - A) = 0 \;\; \Rightarrow \;\; 2D - A \not> 0.$$

Тема 10. Итерационный метод Зейделя.

Матрицу системы $A\vec{x}=\vec{b}$ представим в виде $A=-L+D-R$, где

$$D=diag\,A=diag\{a_{11},\,a_{22},\,...,\,a_{nn}\},$$

$$L=-\begin{bmatrix} 0 & & & & \\ a_{21} & 0 & & 0 & \\ \ddots & \ddots & \ddots & & \\ \ddots & \ddots & \ddots & 0 & \\ a_{n1} & \ddots & \ddots & a_{n,n-1} & 0 \end{bmatrix},\quad R=-\begin{bmatrix} 0 & a_{12} & \ddots & \ddots & a_{1n} \\ & 0 & \ddots & \ddots & \ddots \\ & & \ddots & \ddots & \ddots \\ & 0 & & 0 & a_{n-1,n} \\ & & & & 0 \end{bmatrix},$$

тогда $B=D-L,\ C=R$:

$$B\vec{x}=C\vec{x}+\vec{b}\ \Rightarrow\ \begin{cases}\vec{x}^{\,0}\ -\ \textit{задан}, \\ (D-L)\vec{x}^{\,k+1}=R\vec{x}^{\,k}+\vec{b},\ \ k=0,1,...\end{cases}$$

Пример.

$$\begin{cases} a_{1,1}x_1^{k+1} & = -a_{1,2}x_2^k - a_{1,3}x_3^k + b_1 \\ a_{2,1}x_1^{k+1} + a_{2,2}x_2^{k+1} & = \qquad\qquad -a_{2,3}x_3^k + b_2 \\ a_{3,1}x_1^{k+1} + a_{3,2}x_2^{k+1} + a_{3,3}x_3^{k+1} & = \qquad\qquad\qquad\qquad b_3 \end{cases} \Rightarrow \begin{cases} A = -L + D - R \\ B = D - L,\ C = R \\ B\vec{x}^{k+1} = C\vec{x}^k + \vec{b} \end{cases}$$

Определение:

Если $D = diag\, A = diag\{a_{11},\, a_{22},\, ...,\, a_{nn}\}$, $\det D \neq 0$,

то итерационный процесс

$$\begin{cases} \vec{x}^0 - \textit{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - (D-L)^{-1}(A\vec{x}^k - \vec{b}), \quad k = 0, 1, ... \end{cases}$$

– **метод Зейделя (Гаусса-Зейделя)** для решения системы $Ax = b$.

Упражнение. $S = E - (D-L)^{-1}A = E - (A+R)^{-1}A = (A+R)^{-1}R$.

Необходимое и достаточное условие сходимости метода Зейделя в случае симметричной матрицы с положительной диагональю

Теорема.

Если $A = A^{*}$ & $D > 0$,

то метод Зейделя:

$$\begin{cases} \vec{x}^{0} - \text{задан}, \\ \vec{x}^{k+1} = \vec{x}^{k} - (D-L)^{-1}(A\vec{x}^{k} - \vec{b}), \quad k = 0, 1, \dots \end{cases}$$

сходится к решению системы $A\vec{x} = \vec{b}$

(т.е. выполняется условие сходимости: $\forall \; |\lambda(\underbrace{E-(D-L)^{-1}A}_{S})| < 1$)

$\Leftrightarrow$

$A > 0$ (т.е. выполняется условие: $\forall \; \lambda(A) > 0$).

Доказательство.

Необходимость: $\forall \ |\lambda(\underbrace{E-(D-L)^{-1}A}_{S})|<1 \overset{?}{\Rightarrow} \forall \ \lambda(A)>0).$

1. **Предположим**, что $\lambda_{\min}(A)<0$

$$\Rightarrow \exists\ \vec{z}^{\,0}\neq 0: \begin{cases} A\vec{z}^{\,0}=\lambda_{\min}(A)\cdot\vec{z}^{\,0} \\ (A\vec{z}^{\,0},\vec{z}^{\,0})_2=\lambda_{\min}(A)\cdot(\vec{z}^{\,0},\ \vec{z}^{\,0})_2<0 \end{cases}$$

2. Вычислим $\vec{z}^{\,1}=(\underbrace{E-[D-L]^{-1}A}_{S})\vec{z}^{\,0}\equiv(E-\underbrace{[A+R]^{-1}A)\vec{z}^{\,0}}_{\vec{y}^{\,0}}$ и

$$(A\vec{z}^{\,1},\ \vec{z}^{\,1})_2=(A\vec{z}^{\,0}-A\vec{y}^{\,0},\ \vec{z}^{\,0}-\vec{y}^{\,0})_2=(A\vec{z}^{\,0},\ \vec{z}^{\,0})_2-$$

$$-(A\underbrace{\vec{z}^{\,0}}_{(A+R)^{-1}A\vec{y}^{\,0}},\ \vec{y}^{\,0})_2-(A\vec{y}^{\,0},\ \underbrace{\vec{z}^{\,0}}_{(A+R)^{-1}A\vec{y}^{\,0}})_2+(A\vec{y}^{\,0},\ \vec{y}^{\,0})_2=$$

$$=(A\vec{z}^{\,0},\,\vec{z}^{\,0})_2-((A+R)\vec{y}^{\,0},\,\vec{y}^{\,0})_2-(\vec{y}^{\,0},\,(A+R)\vec{y}^{\,0})_2+(A\vec{y}^{\,0},\,\vec{y}^{\,0})_2=$$

$$=(A\vec{z}^{\,0},\,\vec{z}^{\,0})_2-((A+R)\vec{y}^{\,0},\,\vec{y}^{\,0})_2-(\vec{y}^{\,0},\,(A+R)\vec{y}^{\,0})_2+(A\vec{y}^{\,0},\,\vec{y}^{\,0})_2=$$

$$\equiv(A\vec{z}^{\,0},\,\vec{z}^{\,0})_2-([\underbrace{(A+R)+(A+R)^*-A}_{(D-L)+(D-R)-(D-L-R)}]\vec{y}^{\,0},\,\vec{y}^{\,0})_2=$$

$$=(A\vec{z}^{\,0},\,\vec{z}^{\,0})_2-(D\vec{y}^{\,0},\,\vec{y}^{\,0})_2<(A\vec{z}^{\,0},\,\vec{z}^{\,0})_2=\lambda_{\min}(A)\cdot(\vec{z}^{\,0},\,\vec{z}^{\,0})_2<0$$

3. Из 2. $\Rightarrow\ (A\vec{z}^{\,k+1},\,\vec{z}^{\,k+1})_2<(A\vec{z}^{\,k},\,\vec{z}^{\,k})_2<\ldots<(A\vec{z}^{\,0},\,\vec{z}^{\,0})_2=const<0$

4. Из 3. $\Rightarrow\ \vec{z}^{\,k}\not\to 0\ \ \Rightarrow\ \ \lambda_{\min}(A)>0\ \ \&\ \ A>0.$

Необходимость утверждения теоремы доказана.

Замечание: в методе Зейделя

$$(A\vec{z}^{\,k+1},\vec{z}^{\,k+1})<(A\vec{z}^{\,k},\vec{z}^{\,k}),\ \ \forall\ \vec{z}^{\,k}\neq 0.$$

Достаточность. $A = A^* > 0 \overset{?}{\Rightarrow} \rho(S) = \rho([A+R]^{-1}R) < 1.$

1. $\lambda(S) \in C \;\;\&\;\; \{\vec{\varphi} \in C^n \;\&\; \|\vec{\varphi}\|_2 = 1\}:\;\; S\vec{\varphi} = \lambda(S)\cdot\vec{\varphi}$

$$\Rightarrow\;\; R\vec{\varphi} = \lambda(S)\cdot(A+R)\vec{\varphi}$$

$$\Rightarrow\;\; \underbrace{(R\vec{\varphi},\, \vec{\varphi})_2}_{r+\underline{i}\cdot\mu} = \lambda(S)\cdot[\; \underbrace{(A\vec{\varphi},\, \vec{\varphi})_2}_{a\ge\lambda_{\min}(A)>0} + \underbrace{(R\vec{\varphi},\, \vec{\varphi})_2}_{r+\underline{i}\cdot\mu}]$$

2. $\Rightarrow\;\; |\lambda(S)|^2 = \left|\dfrac{r+\underline{i}\cdot\mu}{(a+r)+\underline{i}\cdot\mu}\right|^2 = \dfrac{r^2+\mu^2}{a(a+2r)+r^2+\mu^2}$

3. $a \ge \lambda_{\min}(A) > 0,\; a+2r = \begin{cases} \underbrace{(A\vec{\varphi},\, \vec{\varphi})}_{a=(D\vec{\varphi},\, \vec{\varphi})-(R^T\vec{\varphi},\, \vec{\varphi})-(R\vec{\varphi},\, \vec{\varphi})} + \underbrace{2\cdot\mathrm{Re}(R\vec{\varphi},\, \vec{\varphi})}_{(R^T\vec{\varphi},\, \vec{\varphi})+(R\vec{\varphi},\, \vec{\varphi})} = \\ = (D\vec{\varphi},\, \vec{\varphi}) \ge \lambda_{\min}(A) > 0 \end{cases}$

4. $\Rightarrow\;\; |\lambda(S)|^2 = \dfrac{r^2+\mu^2}{a(a+2r)+r^2+\mu^2} \le \dfrac{r^2+\mu^2}{\lambda_{\min}^2(A)+r^2+\mu^2} < 1.$

Замечание.

Так как $0 \le x \equiv r^2 + \mu^2 = |(R\vec{\varphi}, \vec{\varphi})|^2 \le \| R \|_2^2 \cdot \underbrace{\| \vec{\varphi} \|_2^2}_{=1} = \| R \|_2^2 \equiv \rho(R^* R)$

то

$$|\lambda(S)|^2 \le \frac{x}{\lambda_{\min}^2(A) + x} \le \max_{0 \le x \le \rho(R^* R)} \frac{x}{\lambda_{\min}^2(A) + x} = \frac{\rho(R^* R)}{\lambda_{\min}^2(A) + \rho(R^* R)},$$

т.е. $\rho(S) \le \sqrt{\dfrac{\rho(R^* R)}{\lambda_{\min}^2(A) + \rho(R^* R)}} < 1$

Тема 11. Методы релаксации.

Функционал ошибки

Теорема (о строгом убывании функционала ошибки)**.**

Если: $A\vec{x} = \vec{b}$ $(\det A \neq 0)$

$\{\vec{x}^k\}_{k=0}^{\infty}$ – приближения к $\vec{x}$:

1. $\vec{z}^{k+1} \equiv \vec{x}^{k+1} - x = S(\vec{z}^k) \quad \forall\ \vec{z}^k \neq 0$

   где $S:\ R^n \to R^n$ – оператор шага для ошибки:

$$\vec{y}^k \neq 0 \to \vec{y} \neq 0 \quad \Rightarrow \quad S(\vec{y}^k) \to S(\vec{y});$$

2. $f_{\vec{x}}(\vec{x}^{k+1}) < f_{\vec{x}}(\vec{x}^k) \quad \forall\ \vec{x}^k \neq \vec{x}$

   где $f_{\vec{x}}(\vec{x}^k) \equiv \| \vec{z}^k \|$, $\vec{z}^k = \vec{x}^k - \vec{x}$ – **функционал ошибки;**

то $\| \vec{z}^k \| \to 0$, т.е. $\vec{x}^k \to \vec{x}$.

Замечание: напомним, что в методе Зейделя

$(A\vec{z}^{k+1}, \vec{z}^{k+1}) < (A\vec{z}^k, \vec{z}^k)$, $\forall\ \vec{z}^k \neq 0$ (сравните с условием 2.).

Доказательство теоремы**.**

1. Пусть $\forall\ \vec{z}^{k} \neq 0$ (если $\vec{z}^{m} = 0 \Rightarrow \vec{x}^{m} = \vec{x}$ – систему решили).
2. Из условия 2. $\Rightarrow 0 < \ldots < \|\vec{z}\|^{k+1} < \|\vec{z}\|^{k+1} < \ldots < \|\vec{z}\|^{0}$
   $\Rightarrow \exists \lim\limits_{k\to\infty} \| z^{k} \| = \alpha \geq 0.$

   Если $\alpha = 0$, то теорема доказана.
3. **Предположим** $\lim\limits_{k\to\infty} \|\vec{z}^{k}\| = \alpha > 0.$

- $\forall\ \|\vec{z}^{k}\| < \|\vec{z}^{0}\| = const \Rightarrow \begin{cases} \exists\ \vec{z}^{k_m} \neq 0 \to \vec{z} \\ \|\vec{z}\| = \alpha > 0 \end{cases}$
- $\Rightarrow$ (и из 1.) $\Rightarrow \begin{cases} S(\vec{z}^{k_m}) \to S(\vec{z}) \\ \alpha < \|\vec{z}^{k_m+1}\| = \|S(\vec{z}^{k_m})\| \to \|S(\vec{z})\| < \alpha \end{cases}$
- $\Rightarrow \alpha < \alpha$ – **противоречие.** $\Rightarrow \lim\limits_{k\to\infty} \|\vec{z}^{k}\| = \alpha = 0 \Rightarrow \vec{x}^{k} \to \vec{x}.$

Метод полной релаксации.

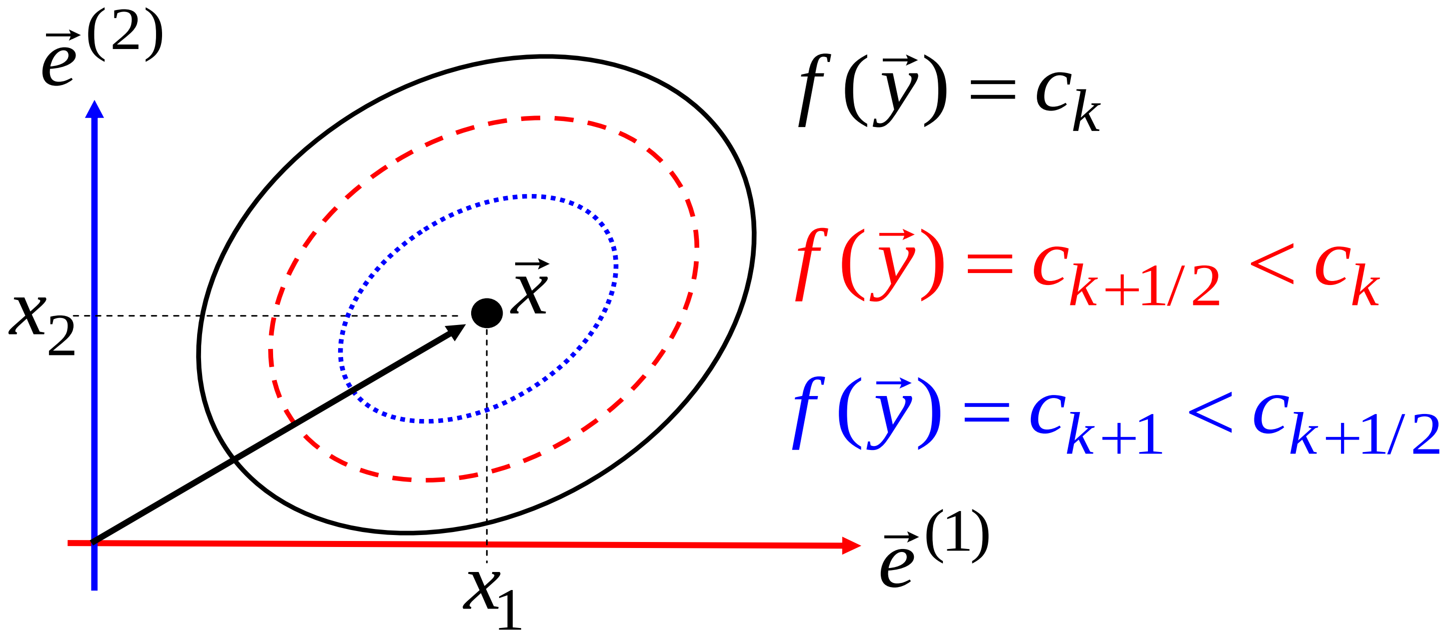
В R^n задана $A\vec{x}=\vec{b}$ с $A=A^*\equiv A^T>0$.

Зададим в R^n: 1. $(\vec{x},\vec{y})_A \equiv (A\vec{x},\vec{y}) \quad \forall\ \vec{x},\vec{y}\in R^n$ – скалярное произведение,

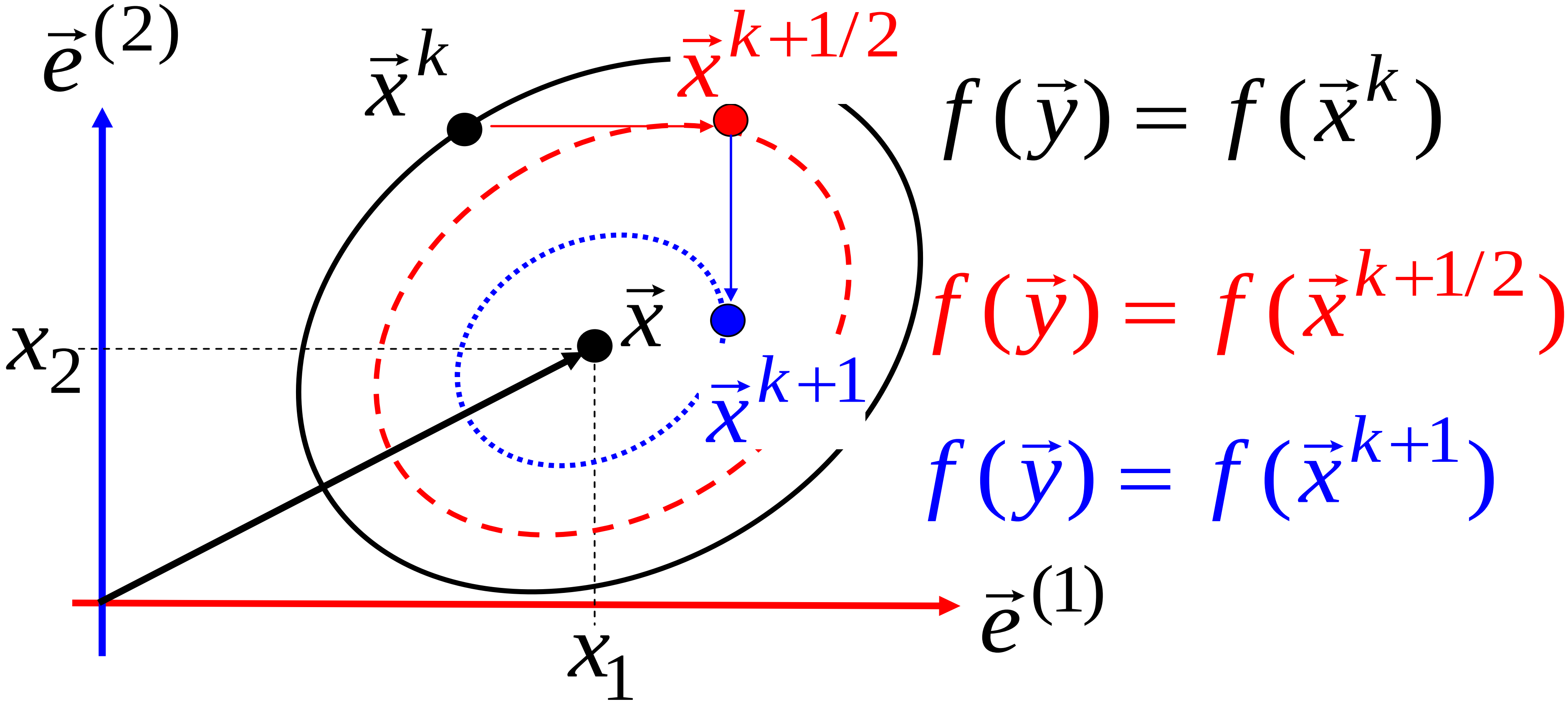
2. $\|\vec{x}\|_A = \sqrt{(\vec{x},\vec{x})_A} \quad \forall\ \vec{x}\in R^n$ – норма,

3. $f_{\vec{x}}(\vec{y}) \equiv \|\vec{y}-\vec{x}\|_A \quad \forall\ \vec{y}\in R^n$ – **функционал ошибки:**

$\vec{z}$



Для заданного $\vec{x}^k$ построим $\vec{x}^{k+1}$ за n шагов:



$$\left\{\begin{array}{l|l}
\vec{x}^{k+1/n} = \vec{x}^{k} - \alpha_{k,1} \cdot \vec{e}^{(1)} & \alpha_{k,1}: \ \| \vec{z}^{k+1/n} \|_A \le \| \vec{z}^{k} - \alpha \cdot \vec{e}^{(1)} \|_A \ \ \forall\ \alpha \\
\ldots & \ldots \\
\vec{x}^{k+i/n} = \vec{x}^{k+(i-1)/n} - \alpha_{k,i} \cdot \vec{e}^{(i)} & \alpha_{k,i}: \ \| \vec{z}^{k+i/n} \|_A \le \| \vec{z}^{k+(i-1)/n} - \alpha \cdot \vec{e}^{(i)} \|_A \ \ \forall\ \alpha \\
\ldots & \ldots \\
\vec{x}^{k+1} = \vec{x}^{k+(n-1)/n} - \alpha_{k,n} \cdot \vec{e}^{(n)} & \alpha_{k,n}: \ \| \vec{z}^{k+1} \|_A \le \| \vec{z}^{k+(n-1)/n} - \alpha \cdot \vec{e}^{(n)} \|_A \ \ \forall\ \alpha
\end{array}\right.$$

где $\vec{e}^{(i)}$ – орт в R^n, $\vec{z}^{k+i/n} = \vec{x}^{k+i/n} - \vec{x}$.

Упражнение 1. Решить задачу

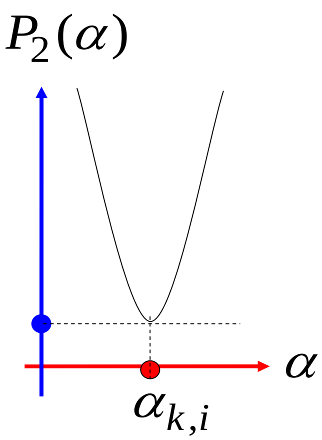
$$\alpha_{k,i}:\ \| \vec{z}^{\,k+i/n} \|_A \le \| \vec{z}^{\,k+(i-1)/n} - \alpha\cdot\vec{e}^{\,(i)} \|_A \quad \forall\ \alpha \in R$$

Решение.

$$\| \vec{z}^{\,k+(i-1)/n} - \alpha\cdot\vec{e}^{\,(i)} \|_A^2 = (A[\vec{z}^{\,k+(i-1)/n} - \alpha\cdot\vec{e}^{\,(i)}],\ \vec{z}^{\,k+(i-1)/n} - \alpha\cdot\vec{e}^{\,(i)}) =$$

$$= (A\vec{z}^{\,k+(i-1)/n},\ \vec{z}^{\,k+(i-1)/n}) - \alpha\cdot(\underbrace{A\vec{z}^{\,k+(i-1)/n}}_{\vec{r}^{\,k+(i-1)/n}},\ \vec{e}^{\,(i)}) - \alpha\cdot(A\vec{e}^{\,(i)},\ \vec{z}^{\,k+(i-1)/n}) +$$

$$+\alpha^2\cdot\underbrace{(A\vec{e}^{\,(i)},\ \vec{e}^{\,(i)})}_{a_{i,i}>0} \equiv const^{k+(i-1)/n} - 2\alpha\cdot(\vec{r}^{\,k+(i-1)/n},\ \vec{e}^{\,(i)}) + \alpha^2\cdot a_{i,i} \equiv P_2(\alpha)$$



$$\Rightarrow \begin{cases} \min\limits_{\alpha} P_2(\alpha) = P_2(\alpha_{k,i}) \\ \dfrac{d}{d\,\alpha} P_2(\alpha_{k,i}) = -2\cdot(\vec{r}^{\,k+(i-1)/n},\ \vec{e}^{\,(i)}) + 2\alpha_{k,i}\cdot a_{i,i} = 0 \\ \alpha_{k,i} = (a_{i,i})^{-1}\cdot(\vec{r}^{\,k+(i-1)/n},\ \vec{e}^{\,(i)}) \end{cases}$$

Упражнение 2. Доказать:

$$\|\vec{z}^{\,k+i/n}\|_A^2 = \|\vec{z}^{\,k+(i-1)/n}\|_A^2 - (a_{i,i})^{-1}\cdot(\vec{r}^{\,k+(i-1)/n}, \vec{e}^{\,(i)})^2 <$$

$$< \|\vec{z}^{\,k+(i-1)/n}\|_A^2, \textit{ если } (\vec{r}^{\,k+(i-1)/n})_i \neq 0$$

Упражнение 3. Доказать:

$$\|\vec{z}^{\,k+1}\|_A^2 < \|\vec{z}^{\,k}\|_A^2, \textit{ если } \vec{r}^{\,k} \equiv A\vec{z}^{\,k} \neq 0$$

Упражнение 4. Проверить: $\vec{x}^{\,k+i/n} = \vec{x}^{\,k+(i-1)/n} - (a_{i,i})^{-1}\cdot(\vec{r}^{\,k+(i-1)/n})_i \cdot \vec{e}^{\,(i)}$:

$$\vec{x}^{\,k+i/n} \equiv \begin{pmatrix} x_1^{k+1} \\ \vdots \\ x_{i-1}^{k+1} \\ x_i^{k+1} \\ x_{i+1}^{k} \\ \vdots \\ x_n^{k} \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ \vdots \\ x_{i-1}^{k+1} \\ x_i^{k} \\ x_{i+1}^{k} \\ \vdots \\ x_n^{k} \end{pmatrix} - (a_{i,i})^{-1}\cdot(\vec{r}^{\,k+(i-1)/n})_i \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Упражнение 5. Проверить:

$$(\vec{r}^{\,k+(i-1)/n})_i = a_{i,1}x_1^{k+1} + \ldots + a_{i,i-1}x_{i-1}^{k+1} + a_{i,i}x_i^k + \ldots + a_{i,n}x_n^k - b_i$$

Упражнение 6. Доказать ($A = -L + D - R$):

$$\begin{cases} x_i^{k+1} = x_i^k - (a_{ii})^{-1}[\underbrace{a_{i,1}x_1^{k+1} + \ldots + a_{i,i-1}x_{i-1}^{k+1}}_{[-L\vec{x}^{k+1}]_i} + \underbrace{a_{i,i}x_i^k + \ldots + a_{i,n}x_n^k}_{[(D-R)\vec{x}^k]_i} - b_i], \\ i = 1,\ 2,\ \ldots,\ n, \end{cases}$$

Упражнение 7. Доказать ($A = -L + D - R$):

$$\vec{x}^{k+1} = \vec{x}^k - D^{-1}(-L\vec{x}^{k+1} + (D-R)\vec{x}^k - \vec{b}) \equiv$$

$$\equiv \vec{x}^k - (D-L)^{-1}(A\vec{x}^k - \vec{b})$$

$\Rightarrow\ S = E - (D-L)^{-1}A$ – непрерывный оператор:

$$\vec{z}^{k+1} = S(\vec{z}^k)$$

Из упражнений 1. – 7. $\Rightarrow$

Теорема. Последовательность $\{\vec{x}^k\}_{k=0}^{\infty}$ метода полной релаксации для решения $A\vec{x}=\vec{b}$ $(A=-L+D-R=A^*>0)$:

$$\begin{cases} \vec{x}^0 \in R^n \text{ – } \textit{задан}, \\ \vec{x}^{k+1}=\vec{x}^k-(D-L)^{-1}(A\vec{x}^k-\vec{b}) \end{cases}$$

сходится к $\vec{x}$.

Доказательство. Из упражнений 1. – 7. $\Rightarrow$

$$\begin{cases} \vec{z}^{k+1}\equiv\vec{x}^{k+1}-\vec{x}=[\underbrace{E-(D-L)^{-1}A}_{S\text{ – } \textit{непрерывный оператор}}]\vec{z}^k\equiv S\,\vec{z}^k \\ f_{\vec{x}}(\vec{x}^{k+1})\equiv\|\vec{z}^{k+1}\|_A<\|\vec{z}^k\|_A\equiv f_{\vec{x}}(\vec{x}^k) \end{cases}$$

$\Rightarrow$ (см. теорема о строгом убывании функционала ошибки)

$$\|\vec{z}^k\|_A\to 0 \quad\Rightarrow\quad \vec{x}^k\to\vec{x}.$$

Тема 11.1. Метод неполной релаксации.

Метод неполной релаксации.

В формулы метода полной релаксации (для $A\vec{x}=\vec{b}$ с $A=A^{*}\equiv A^{T}>0$) добавим параметр ω:

$$\begin{cases}
\vec{x}^{k+1/n}=\vec{x}^{k}-\omega\cdot\alpha_{k,1}\cdot\vec{e}^{(1)} & \alpha_{k,1}=(a_{1,1})^{-1}\cdot(\vec{r}^{k},\,\vec{e}^{(1)}) \\
\ldots & \ldots \\
\vec{x}^{k+i/n}=\vec{x}^{k+(i-1)/n}-\omega\cdot\alpha_{k,i}\cdot\vec{e}^{(i)} & \alpha_{k,i}=(a_{i,i})^{-1}\cdot(\vec{r}^{k+(i-1)/n},\,\vec{e}^{(i)}) \\
\ldots & \ldots \\
\vec{x}^{k+1}=\vec{x}^{k+(n-1)/n}-\omega\cdot\alpha_{k,n}\cdot\vec{e}^{(n)} & \alpha_{k,n}=(a_{n,n})^{-1}\cdot(\vec{r}^{k+(n-1)/n},\,\vec{e}^{(n)})
\end{cases}$$

где $\vec{e}^{(i)}$ – орт в R^{n}, $\quad\vec{z}^{k+i/n}=\vec{x}^{k+i/n}-\vec{x}$.

Упражнение 1. $\| \vec{z}^{\,k+i/n} \|_A^2 = \| \vec{z}^{\,k+(i-1)/n} \|_A^2 - \omega(2-\omega)\dfrac{(\vec{r}^{\,k+(i-1)/n}, \vec{e}^{\,(i)})^2}{a_{i,i}}$

Доказательство. $\vec{x}^{\,k+i/n} = \vec{x}^{\,k+(i-1)/n} - \omega \cdot \alpha_{k,i} \cdot \vec{e}^{\,(i)}$; $\alpha_{k,i} = \dfrac{(\vec{r}^{\,k+(i-1)/n}, \vec{e}^{\,(i)})}{a_{i,i}}$

$$\Rightarrow \left\{ \begin{aligned} \| \vec{z}^{\,k+i/n} \|_A^2 &= (A[\vec{z}^{\,k+(i-1)/n} - \omega \cdot \alpha_{k,i} \cdot \vec{e}^{\,(i)}], \vec{z}^{\,k+(i-1)/n} - \omega \cdot \alpha_{k,i} \cdot \vec{e}^{\,(i)}) = \\ &= \underbrace{(A\vec{z}^{\,k+(i-1)/n}, \vec{z}^{\,k+(i-1)/n})}_{\|\vec{z}^{\,k+(i-1)/n}\|_A^2} - \omega \cdot \alpha_{k,i} \cdot (\underbrace{A\vec{z}^{\,k+(i-1)/n}}_{\vec{r}^{\,k+(i-1)/n}}, \vec{e}^{\,(i)}) - \\ &- \omega \cdot \alpha_{k,i} \cdot \underbrace{(A\vec{e}^{\,(i)}, \vec{z}^{\,k+(i-1)/n})}_{(\vec{e}^{\,(i)}, A\vec{z}^{\,k+(i-1)/n})} + [\omega \cdot \alpha_{k,i}]^2 \cdot \underbrace{(A\vec{e}^{\,(i)}, \vec{e}^{\,(i)})}_{a_{i,i}} = \\ &= \| \vec{z}^{\,k+(i-1)/n} \|_A^2 - 2\omega \frac{(\vec{r}^{\,k+(i-1)/n}, \vec{e}^{\,(i)})^2}{a_{i,i}} + \omega^2 \frac{(\vec{r}^{\,k+(i-1)/n}, \vec{e}^{\,(i)})^2}{a_{i,i}} \end{aligned} \right.$$

Упражнение 2. Доказать:

$$\forall \quad \omega \in (0,2) \quad \| \vec{z}^{\,k+1} \|_A^2 < \| \vec{z}^{\,k} \|_A^2, \text{ } если \text{ } \vec{r}^{\,k} \equiv A\vec{z}^{\,k} \neq 0$$

Упражнение 3. Проверить:

$$\vec{x}^{\,k+i/n} = \vec{x}^{\,k+(i-1)/n} - \omega \cdot (a_{i,i})^{-1} \cdot (\vec{r}^{\,k+(i-1)/n})_i \cdot \vec{e}^{\,(i)}:$$

$$\vec{x}^{\,k+i/n} \equiv \begin{pmatrix} x_1^{k+1} \\ \vdots \\ x_{i-1}^{k+1} \\ x_i^{k+1} \\ x_{i+1}^{k} \\ \vdots \\ x_n^{k} \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ \vdots \\ x_{i-1}^{k+1} \\ x_i^{k} \\ x_{i+1}^{k} \\ \vdots \\ x_n^{k} \end{pmatrix} - \omega \cdot (a_{i,i})^{-1} \cdot (\vec{r}^{\,k+(i-1)/n})_i \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Упражнение 4. Проверить:

$$(\vec{r}^{\,k+(i-1)/n})_i = a_{i,1}x_1^{k+1} + \ldots + a_{i,i-1}x_{i-1}^{k+1} + a_{i,i}x_i^k + \ldots + a_{i,n}x_n^k - b_i$$

Упражнение 5. Доказать ($A = -L + D - R$):

$$\begin{cases} x_i^{k+1} = x_i^k - \omega(a_{ii})^{-1}[\underbrace{a_{i,1}x_1^{k+1} + \ldots + a_{i,i-1}x_{i-1}^{k+1}}_{[-L\vec{x}^{k+1}]_i} + \underbrace{a_{i,i}x_i^k + \ldots + a_{i,n}x_n^k}_{[(D-R)\vec{x}^k]_i} - b_i], \\ i = 1, 2, \ldots, n, \end{cases}$$

Упражнение 6. Доказать ($A = -L + D - R$):

$$\vec{x}^{k+1} = \vec{x}^k - \omega D^{-1}(-L\vec{x}^{k+1} + (D-R)\vec{x}^k - \vec{b}) \equiv$$

$$\equiv \vec{x}^k - \omega(D - \omega L)^{-1}(A\vec{x}^k - \vec{b})$$

$\Rightarrow\ S = E - \omega(D - \omega L)^{-1}A$ – непрерывный оператор:

$$\vec{z}^{\,k+1} = S(\vec{z}^{\,k})$$

Из упражнений 1. – 6. $\Rightarrow$

Теорема. Последовательность $\{\vec{x}^k\}_{k=0}^{\infty}$ метода неполной релаксации для решения $A\vec{x} = \vec{b}$ $(A = -L + D - R = A^* > 0)$:

$$\begin{cases} \vec{x}^0 \in R^n \ - \textit{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - \omega(D - \omega L)^{-1}(A\vec{x}^k - \vec{b}) \end{cases}$$

сходится к $\vec{x}$ при $\forall \ \omega \in (0, 2)$.

Доказательство. Из упражнений 1. – 6. $\Rightarrow$

$$\begin{cases} \vec{z}^{k+1} \equiv \vec{x}^{k+1} - \vec{x} = [\ \underbrace{E - \omega(D - \omega L)^{-1} A}_{S\ -\ \textit{непрерывный оператор}}\ ]\vec{z}^k \equiv S\,\vec{z}^k \\ f_{\vec{x}}(\vec{x}^{k+1}) \equiv \| \vec{z}^{k+1} \|_A < \| \vec{z}^k \|_A \equiv f_{\vec{x}}(\vec{x}^k) \end{cases}$$

$\Rightarrow$ (см. теорема о строгом убывании функционала ошибки)

$$\| \vec{z}^k \|_A \to 0 \quad \Rightarrow \quad \vec{x}^k \to \vec{x}.$$

Оценка сходимости методов релаксации

Предварительные замечания

Итак, $A = -L + D - R = A^T \equiv A^* > 0$,

$$\left.\begin{array}{l} \vec{x}^0 \in R^n, \ \ \omega \in (0,\, 2) \ \textit{заданы}, \\ \vec{x}^{k+1} = \vec{x}^k - \omega(D - \omega L)^{-1}(A\vec{x}^k - \vec{b}) \end{array}\right\} \ \& \ \left\{\begin{array}{l} \| \vec{z}^{k+1} \|_A^2 < \| \vec{z}^k \|_A^2 \neq 0 \\ k = 0, 1, \ldots \end{array}\right.$$

$$A = (-L + 0.5 \cdot D) + (0.5 \cdot D - R) \equiv R_1 + R_1^*:$$

$$(A\vec{y},\ \vec{y}) = 2 \cdot (R_1\vec{y},\ \vec{y}) > 0 \quad \forall\ \vec{y} \in R^n \ \Rightarrow \ R_1 > 0 \text{ в } R^n.$$

$$D - \omega L = (1 - 0.5\omega)D + \omega(0.5 \cdot D - L) = (1 - 0.5\omega)D + \omega R_1 =$$

$$= (1 - 0.5\omega)(D + \frac{\omega}{(1 - 0.5\omega)} R_1) \equiv \frac{\omega}{\tau}(D + \tau \cdot R_1) \equiv \frac{\omega}{\tau} B > 0 \text{ в } R^n,$$

$$\text{где } \tau = \frac{\omega}{1 - 0.5\omega} \in (0,\ \infty) \text{ при } \omega \in (0,\ 2).$$

Тогда метод релаксации может быть переписан в виде

$$\vec{x}^0 \in R^n \ \textit{задан},$$

$$\vec{x}^{k+1} = \vec{x}^k - \tau \cdot B^{-1}(A\vec{x}^k - \vec{b}), \ \ B = D + \tau \cdot R_1 > 0 \ \ \forall \ \tau > 0,$$

$$k = 0, 1, \ldots$$

и он сходится при любом $\tau > 0$.

Очевидно, что

$$\vec{z}^{k+1} = [E - \tau \cdot B^{-1}A]\vec{z}^k \equiv S\vec{z}^k$$

Замечание. Мы знаем, что для любого вектора $z \neq 0$

$$\| S\vec{z} \|_A < \| \vec{z} \|_A \Rightarrow \| S \|_A = \sup_{\|\vec{z}\|_A = 1} \| S\vec{z} \|_A = \| S\vec{z}_{\sup} \|_A < 1.$$

Здесь мы применили теорему о достижении максимума непрерывной функции на сфере конечномерного пространства.

Оценим $\| S \|_A^2 = \max\limits_{\vec{z} \neq 0} \dfrac{(AS\vec{z}\,, S\vec{z}\,)}{(A\vec{z}\,, \vec{z}\,)}$: $S = E - \tau\, B^{-1} A \Rightarrow$

$$\| S \|_A^2 = \max_{\vec{z} \neq 0} \left[1 - 2\tau \frac{(B^{-1} A\vec{z}\,, A\vec{z}\,)}{(A\vec{z}\,, \vec{z}\,)} + \tau^2 \frac{(AB^{-1} A\vec{z}\,, B^{-1} A\vec{z}\,)}{(A\vec{z}\,, \vec{z}\,)} \right].$$

Так как

$$(A \underbrace{B^{-1} A\vec{z}}_{\vec{y}}, \underbrace{B^{-1} A\vec{z}}_{\vec{y}}) = 2(R_1\vec{y},\ \vec{y}) \equiv 2\tau^{-1}([B - D]\vec{y},\ \vec{y}) =$$

$$= 2\tau^{-1} \underbrace{(B\vec{y},\ \vec{y})}_{=\overline{(\vec{y},\ B\vec{y})}=(\vec{y},\ B\vec{y})} - 2\tau^{-1}(D\vec{y},\ \vec{y}) =$$

$$= 2\tau^{-1}(B^{-1} A\vec{z},\ A\vec{z}) - 2\tau^{-1}(DB^{-1} A\vec{z},\ B^{-1} A\vec{z})$$

то $\| S \|_A^2 \equiv \max\limits_{\vec{z} \neq 0} \left[1 - 2\tau \dfrac{(DB^{-1} A\vec{z}\,, B^{-1} A\vec{z}\,)}{(A\vec{z}\,, \vec{z}\,)} \right] = 1 - 2\tau \cdot \min\limits_{\vec{z} \neq 0} \dfrac{(DB^{-1} A\vec{z}\,, B^{-1} A\vec{z}\,)}{(A\vec{z}\,, \vec{z}\,)}$

Итак, $\| S \|_A^2 = 1 - 2\tau \cdot \underbrace{\min_{\vec{z}\in R^n:\ \vec{z}\neq 0} \frac{(DB^{-1}A\vec{z}, B^{-1}A\vec{z})}{(A\vec{z}, \vec{z})}}_{\gamma(\tau)}$, $\begin{cases} A = A^T > 0 \quad in\ R^n \\ D = diag\ A > 0 \\ B = D + \tau \cdot R_1 > 0 \end{cases}$

Упражнение 1.

$$\gamma(\tau) \equiv \min_{\vec{z}\in R^n:\ \vec{z}\neq 0} \frac{(DB^{-1}A\vec{z}, B^{-1}A\vec{z})}{(A\vec{z}, \vec{z})} = \lambda_{\min}(A^{1/2}(B^{-1})^T DB^{-1}A^{1/2})$$

Доказательство.

1. $A = A^T > 0 \quad in\ R^n \ \Rightarrow \ A = Q\Lambda Q^T,\ \Lambda = diag\{\lambda_1(A), \ldots, \lambda_n(A)\} > 0 \Rightarrow$

$\exists \ \ A^{1/2} = Q\Lambda^{1/2}Q^T = (A^{1/2})^T > 0,\ \Lambda^{1/2} = diag\{\sqrt{\lambda_1(A)}, \ldots, \sqrt{\lambda_n(A)}\} > 0$

$$\Rightarrow \gamma(\tau) = \min_{\vec{z}\neq 0} \frac{(DB^{-1}A\vec{z}, B^{-1}A\vec{z})}{(A\vec{z}, \vec{z})} = \min_{\vec{y}=A^{1/2}\vec{z}\neq 0} \frac{(DB^{-1}A^{1/2}\vec{y}, B^{-1}A^{1/2}\vec{y})}{(\vec{y}, \vec{y})}$$

2. $\gamma(\tau) = \lambda_{\min}(C) = \min\limits_{\vec{y} \neq 0} \dfrac{(C\vec{y}, \vec{y})}{(\vec{y}, \vec{y})}, \quad C = A^{1/2}(B^{-1})^T DB^{-1}A^{1/2}$:

$$C = A^{1/2}(B^{-1})^T DB^{-1}A^{1/2} = C^T > 0$$

$$\Rightarrow C = Q_C \Lambda_C Q_C^T,\ \Lambda_C = diag\{\lambda_1(C), ..., \lambda_n(C)\} > 0$$

$$\frac{(C\vec{y}, \vec{y})}{(\vec{y}, \vec{y})} = \frac{(Q_C \Lambda_C Q_C^T \vec{y}, \vec{y})}{(Q_C Q_C^T \vec{y}, \vec{y})} = \frac{(\Lambda_C [Q_C^T \vec{y}], [Q_C^T \vec{y}])}{([Q_C^T \vec{y}], [Q_C^T \vec{y}])} \geq$$

$$\geq \frac{\lambda_{\min}(C) \cdot ([Q_C^T \vec{y}], [Q_C^T \vec{y}])}{([Q_C^T \vec{y}], [Q_C^T \vec{y}])} = \lambda_{\min}(C)$$

$$\frac{(C\vec{y}, \vec{y})}{(\vec{y}, \vec{y})} = \frac{(\lambda_{\min}(C) \cdot \vec{y},\ \vec{y})}{(\vec{y}, \vec{y})} = \lambda_{\min}(C), \quad C\vec{y} = \lambda_{\min}(C) \cdot \vec{y}$$

Упражнение 2. $\gamma(\tau)=\lambda_{\min}(C)\geq[\rho(A^{-1}D)+\tau+\tau^2\rho(A^{-1}R_1D^{-1}R_1^T)]^{-1}$

Доказательство. Пусть

$$A^{1/2}(B^{-1})^T DB^{-1}A^{1/2}\vec{y}=\gamma\cdot\vec{y} \quad \Rightarrow \begin{cases} A^{1/2}\vec{y}=\gamma\cdot[(B^{-1})^T DB^{-1}]^{-1}A^{-1/2}\vec{y} \\ A\underbrace{A^{-1/2}\vec{y}}_{\vec{v}}=\gamma\cdot[BD^{-1}B^T]\underbrace{A^{-1/2}\vec{y}}_{\vec{v}} \end{cases}$$

Т.к. $BD^{-1}B^*=(D+\tau R_1)D^{-1}(D+\tau R_1^*)=D+\tau A+\tau^2R_1D^{-1}R_1^*$,

то $\gamma(\tau)=\dfrac{(A\vec{v},\vec{v})}{(D\vec{v},\vec{v})+\tau(A\vec{v},\vec{v})+\tau^2([R_1D^{-1}R_1^T]\vec{v},\vec{v})}\geq$

$$\geq\frac{1}{\max\limits_{\vec{v}\neq 0}\dfrac{(D\vec{v},\vec{v})}{(A\vec{v},\vec{v})}+\tau+\tau^2\max\limits_{\vec{v}\neq 0}\dfrac{([R_1D^{-1}R_1^T]\vec{v},\vec{v})}{(A\vec{v},\vec{v})}}\geq$$

$$\geq\frac{1}{\rho(A^{-1}D)+\tau+\tau^2\rho(A^{-1}R_1D^{-1}R_1^T)}.$$

Упражнение 3. Если известны оценки $\begin{cases} 1/\delta \geq \rho(A^{-1}D) \\ \Delta \geq \rho(A^{-1}R_1 D^{-1} R_1^*) \end{cases}$,

то проверить, что

$$\| S \|_A^2 = 1 - 2\tau\,\gamma(\tau) \leq 1 - 2\tau \frac{1}{\delta^{-1} + \tau + \tau^2 \Delta} = \frac{1 - \tau\delta + \tau^2 \delta\Delta}{1 + \tau\delta + \tau^2 \delta\Delta} \equiv g(\tau) < 1 \;\; \forall\, \tau > 0$$

Упражнение 4. Постоянная $0 < \delta \leq \lambda_{\min}(D^{-1}A)$:

$$\delta \cdot (D\vec{v}, \vec{v}) \leq (A\vec{v}, \vec{v}) \quad \forall\, \vec{v} \in R^n$$

Упражнение 5. Постоянная $\Delta \geq \lambda_{\max}(A^{-1}R_1 D^{-1} R_1^*) > 0$:

$$(R_1 D^{-1} R_1^T \vec{v}, \vec{v}) \leq \Delta \cdot (A\vec{v}, \vec{v}) \quad \forall\, \vec{v} \in R^n$$

Упражнение 6. Найти минимум функции $g(\tau)$:

$$g(\tau_*) = \min_{\tau>0} \frac{1-\tau\,\delta+\tau^2\delta\Delta}{1+\tau\,\delta+\tau^2\delta\Delta}$$

и параметры ω_*: $\tau_* = \dfrac{\omega_*}{1-0.5\omega_*}$

Ответ. $\tau_* = \dfrac{1}{\sqrt{\delta\Delta}}$; $\omega_* = \dfrac{2}{1+2\sqrt{\delta\Delta}}$; $g(\tau_*) = \dfrac{1-\sqrt{\delta/(4\Delta)}}{1+\sqrt{\delta/(4\Delta)}}$.

Замечание. При решении задач математической физики:

$$\delta\Delta << 1 \;\Rightarrow\; 1 < \omega_* = \frac{2}{1+2\sqrt{\delta\Delta}} < 2$$

– метод верхней релаксации.

При $0<\omega<1$ – метод верхней релаксации.

При $\omega=1$ – метод полной релаксации.

Из упражнений 1. – 6. $\Rightarrow$

Теорема (оценка сходимости методов релаксации)**.**

Для ошибки $\vec{z}^k = \vec{x}^k - \vec{x}$ метода релаксации ($\tau = \dfrac{\omega}{1-0.5\omega}$):

$$\begin{cases} \vec{x}^0 \in R^n \; \textit{задан}, \quad k = 0, 1, ..., \\ \vec{x}^{k+1} = \vec{x}^k - \tau \cdot B^{-1}(A\vec{x}^k - \vec{b}), \; B = D + \tau \cdot R_1 > 0 \;\; \forall \; \tau > 0, \end{cases}$$

имеет место оценки:

$$\forall \; \tau > 0: \; \| \vec{z}^{k+1} \|_A^2 \le g(\tau) \cdot \| \vec{z}^k \|_A^2, \;\; g(\tau) = \frac{1 - \tau\delta + \tau^2 \delta\Delta}{1 + \tau\delta + \tau^2 \delta\Delta} < 1$$

$$\tau_* = \frac{1}{\sqrt{\delta\Delta}}: \; \| \vec{z}^{k+1} \|_A^2 \le g(\tau_*) \cdot \| \vec{z}^k \|_A^2, \;\; g(\tau_*) = \frac{1 - \sqrt{\delta / (4\Delta)}}{1 + \sqrt{\delta / (4\Delta)}} < 1$$

где $0 < \delta \le \lambda_{\min}(D^{-1}A)$ & $\Delta \ge \lambda_{\max}(A^{-1}R_1 D^{-1} R_1^T) > 0$.

Тема 11.2. Пример оценок сходимости методов релаксации.

Система

$$\begin{cases} u_0 = 0 \\ u_i + h^{-2}(-u_{i-1} + 2u_i - u_{i+1}) = f_i, \quad h^{-2} = (n+1)^2, \quad i = 1, \ldots, n; \\ u_{n+1} = 0 \end{cases}$$

аппроксимирует краевую задачу для дифференциального уравнения:

$$\begin{cases} u(x) - u''(x) = f(x), \ x \in (0, 1); \\ u(0) = u(1) = 0 \end{cases}$$

на сетке $x_0 = 0 < x_1 = h < x_2 = 2h < \ldots < x_n = nh < x_{n+1} = (n+1)h \equiv 1$:

$$u(x_i) \approx u_i; \ \ u''(x_i) \approx \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2}; \quad f_i = f(x_i).$$

– это модельная задача:

$$A = \begin{bmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & & & 0 \\ -1 & 1 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 1 & 0 \\ 0 & & & -1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & -1 & & & 0 \\ 0 & 1 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 1 & -1 \\ 0 & & & 0 & 1 \end{bmatrix}$$

Собственные значения матрицы A известны:

$$\lambda_k = 4 \cdot \sin^2 \frac{k \cdot \pi}{2(n+1)}, \; k = 1, \ldots, n, \qquad \lambda_{\min}(A) = 4 \cdot \sin^2 \frac{\pi}{2(n+1)} \approx \frac{\pi^2}{(n+1)^2},$$

$$\lambda_{\max}(A) = 4 \cdot \sin^2 \frac{n \cdot \pi}{2(n+1)} \approx 4.$$

Оценим постоянную $\delta \cdot D \le A$:

$$\delta = \lambda_{\min}(D^{-1}A) = 0.5 \cdot \lambda_{\min}(A) \approx \frac{\pi^2}{2(n+1)^2} << 1.$$

Оценим постоянную $R_1 D^{-1} R_1^* \le \Delta \cdot A$:

т.к. $A = 2 \cdot R_1 D^{-1} R_1^* + diag\{1, 0, ..., 0\} \ge 2 \cdot R_1 D^{-1} R_1^*$, то

$$0.5 \cdot E \ge A^{-1} \cdot R_1 D^{-1} R_1^* \quad \Rightarrow \quad \Delta = 0.5.$$

Для метода полной релаксации ($\omega = 1,\ \tau = \dfrac{\omega}{1 - 0.5\omega} = 2$):

$$x^{k+1} = x^k - (D - L)^{-1}(Ax^k - b)$$

$$\| z^{k+1} \|_A \le \| S \|_A \cdot \| z^k \|_A \le \sqrt{\frac{1 - 2\delta + 4\delta\Delta}{1 + 2\delta + 4\delta\Delta}} \cdot \| z^k \|_A =$$

$$= \sqrt{\frac{1}{1 + 4\delta}} \cdot \| z^k \|_A \approx (1 - 2\delta) \cdot \| z^k \|_A \approx (1 - \frac{\pi^2}{(n+1)^2}) \cdot \| z^k \|_A$$

Для метода неполной релаксации с “оптимальным” параметром

$$\omega_* = \frac{2}{1+2\sqrt{\delta\Delta}},\ \tau_* = \omega_* / (1-0.5\omega_*) = 1/\sqrt{\delta\Delta}:$$

$$\| z^{k+1} \|_A \le \| S \|_A \cdot \| z^k \|_A \le \sqrt{\frac{1-\tau_*\delta+\tau_*^2\,\delta\Delta}{1+\tau_*\delta+\tau_*^2\,\delta\Delta}} \cdot \| z^k \|_A \le \sqrt{\frac{1-\sqrt{\delta/(4\Delta)}}{1+\sqrt{\delta/(4\Delta)}}} \cdot \| z^k \|_A$$

$$= \sqrt{\frac{1-\sqrt{\delta/2}}{1+\sqrt{\delta/2}}} \cdot \| z^k \|_A \approx (1-\sqrt{\delta/2}) \cdot \| z^k \|_A \approx (1-\frac{\pi}{2(n+1)}) \cdot \| z^k \|_A$$

Для метода Якоби:

$$x^{k+1} = x^k - D^{-1}(Ax^k - b)$$

$$\| z^{k+1} \|_A \le \| S \|_A \cdot \| z^k \|_A =$$

$$= \max |1 - 0.5 \cdot \lambda(A)| \cdot \| z^k \|_A =$$

$$\approx (1 - 0.5 \cdot \frac{\pi^2}{(n+1)^2}) \cdot \| z^k \|_A =$$

Действительно,

$$\| S \|_A^2 = \sup \frac{(A(E - D^{-1}A)z, (E - D^{-1}A)z)}{(Az, z)} = \sup \frac{\| (E - A^{1/2}D^{-1}A^{1/2})A^{1/2}z \|^2}{(A^{1/2}z, A^{1/2}z)} =$$

$$= \rho^2(E - A^{1/2}D^{-1}A^{1/2}) = \rho^2(E - D^{-1}A) = \rho^2(E - 0.5 \cdot A) \approx [1 - \frac{\pi^2}{2(n+1)^2}]^2$$

следовательно, $\| z^{k+1} \|_A \le [1 - \frac{\pi^2}{2(n+1)^2}] \cdot \| z^k \|_A$.

Тема 12. Методы наискорейшего спуска и минимальных невязок. Градиент функции

$\vec{y} = ?$**:** $\quad \vec{x}^{k+1} = \vec{x}^k + \alpha \cdot \vec{y}$ *и* $\| \underbrace{\vec{x}^{k+1} - \vec{x}}_{\vec{z}^{k+1}} \|^2 = \min\limits_{\alpha} \| \underbrace{\vec{x}^k - \vec{x}}_{\vec{z}^k} + \alpha \cdot \vec{y} \|^2$**?**

Градиент функции $f(\vec{z}): R^n \to R$:

$$f(\vec{z} + \alpha \cdot \vec{y}) = f(\vec{z}) + \left. \frac{d\, f(\vec{z} + \alpha \cdot \vec{y})}{d\, \alpha} \right|_{\alpha = 0} \cdot \alpha + O(\alpha^2) =$$

$$= f(\vec{z}) + \left[\frac{\partial f(\vec{z})}{\partial z_1} y_1 + \ldots + \frac{\partial f(\vec{z})}{\partial z_n} y_n \right] \cdot \alpha + O(\alpha^2) =$$

$$= f(\vec{z}) + (\nabla f, \vec{y}) \cdot \alpha + O(\alpha^2) \approx f(\vec{z}) + (\nabla f, \vec{y}) \cdot \alpha.$$

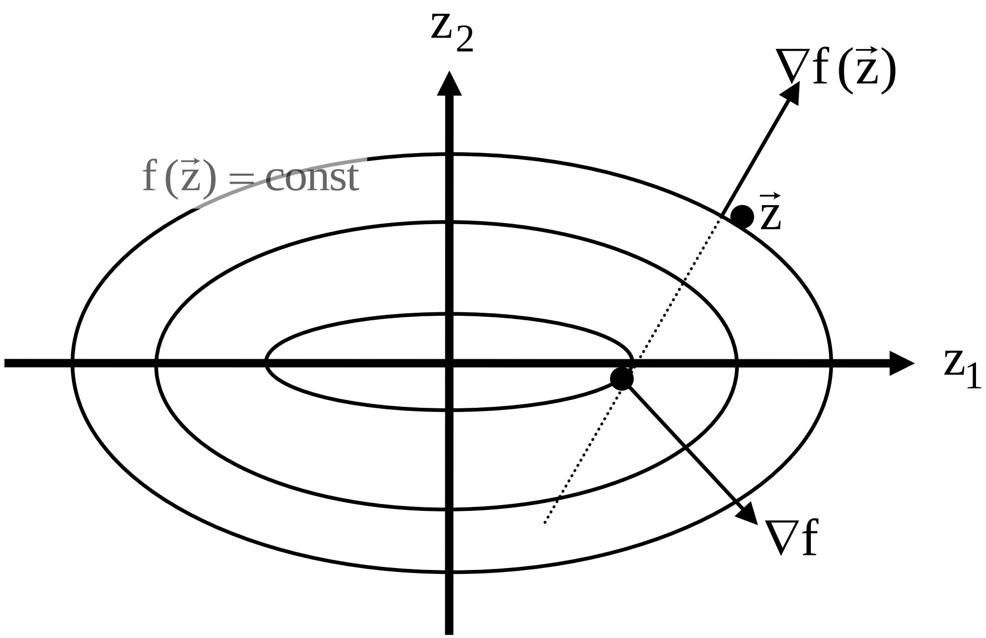
Так как $(\nabla f, \vec{y}) = \cos(\nabla f, \vec{y}) \cdot \| \nabla f \| \cdot \| \vec{y} \|$

принимает минимальное или максимальное значения,

когда вектора ∇f, $\vec{y}$ параллельны, то

$\vec{y} = \nabla f$ **– направление максимального возрастания** $f(\vec{z})$**,**

$\vec{y} = -\nabla f$ **– направление максимального убывания** $f(\vec{z})$**.**



Пример спуска по антиградиенту.

Градиент функции $f(\vec{z}) = \|\vec{z}\|_C^2 \; : R^n \to R$

Если $f(\vec{z}) \equiv \|\vec{z}\|_C^2 = (C\vec{z}, \vec{z}),\quad C = C^* > 0$**, то**

$$\left.\begin{aligned} f(\vec{z} + \alpha \cdot y) &= f(\vec{z}) + (\nabla f, \vec{y}) \cdot \alpha + O(\alpha^2) = \\ &= f(\vec{z}) + (C\vec{z}, \vec{y}) \cdot 2\alpha + \alpha^2 \cdot (C\vec{y}, \vec{y}) \end{aligned}\right\} \Rightarrow \nabla f(\vec{z}) = 2C\vec{z}$$

Следовательно, итерационный процесс имеет вид ($\vec{y} = -0.5 \cdot C\vec{z}$)

$$\vec{x}^{k+1} = \vec{x}^k - \alpha_k \cdot C\vec{z}^k, \quad \vec{z}^{k+1} = \vec{z}^k - \alpha_k \cdot C\vec{z}^k,$$

$$\alpha_k : \|\vec{z}^{k+1}\|_C^2 = \min_{\alpha} [\|\vec{z}^k\|_C^2 - 2\alpha \cdot (C\vec{z}^k, C\vec{z}^k) + \alpha^2 \cdot (CC\vec{z}^k, C\vec{z}^k)],$$

Так как

$$\alpha(\vec{z}^k) \equiv \alpha_k = \frac{(C\vec{z}^k, \vec{z}^k)}{(CC\vec{z}^k, C\vec{z}^k)} \Rightarrow \|\vec{z}^{k+1}\|_C^2 = \|\vec{z}^k\|_C^2 - \frac{(C\vec{z}^k, C\vec{z}^k)^2}{(CC\vec{z}^k, C\vec{z}^k)} < \|\vec{z}^k\|_C^2$$

и отображение $S(\vec{z}) = \vec{z} - \alpha(\vec{z}) \cdot C\vec{z}$ непрерывно при $\vec{z} \neq 0$, то по теореме о строгом убывании функционала итерационный процесс сходится.

Метод наискорейшего спуска

Если матрица системы $A\vec{x} = \vec{b}$

симметрична и положительно определена,

то, выбирая $C = A$, получаем метод наискорейшего спуска:

$$\begin{cases} \vec{x}^0 - \textit{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - \tau_k (A\vec{x}^k - \vec{b}), \;\; \tau_k = \dfrac{(\vec{r}^k, \vec{r}^k)}{(A\vec{r}^k, \vec{r}^k)}, \\ \quad k = 0, 1, \dots \end{cases}$$

энергетическая норма $\| \vec{z}^k \|_A$ ошибки которого строго убывает.

Метод минимальных невязок

В итерационном процессе $\vec{x}^{k+1} = \vec{x}^k - \tau_k (A\vec{x}^k - \vec{b})$

параметр τ_k выбираем из условия минимизации невязки:

$$(\vec{r}^{k+1}, \vec{r}^{k+1}) = \min_{\tau} (\vec{r}^k - \tau A\vec{r}^k,\ \vec{r}^k - \tau A\vec{r}^k).$$

Теорема. Если матрица вещественной системы $Ax = b$ положительно определена, то метод минимальных невязок:

$$\begin{cases} \vec{x}^0 - \textit{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - \tau_k (A\vec{x}^k - \vec{b}),\ \tau_k = \dfrac{(A\vec{r}^k, \vec{r}^k)}{(A\vec{r}^k, A\vec{r}^k)}, \\ k = 0, 1, \ldots \end{cases}$$

сходится, а норма $\| \vec{r}^k \|_2 = \sqrt{(\vec{r}^k, \vec{r}^k)}$ его невязок строго убывает.

Доказательство.

Во-первых; $\min\limits_{\tau}(\vec{r}^{\,k}-\tau A\vec{r}^{\,k},\ \vec{r}^{\,k}-\tau A\vec{r}^{\,k})=$

$$=\min_{\tau}[(\vec{r}^{\,k},\vec{r}^{\,k})-\tau\cdot(\vec{r}^{\,k},A\vec{r}^{\,k})-\tau\cdot(A\vec{r}^{\,k},\vec{r}^{\,k})+\tau^2\cdot(A\vec{r}^{\,k},A\vec{r}^{\,k})]$$

$$=\min_{\tau}[(\vec{r}^{\,k},\vec{r}^{\,k})-2\tau\cdot(A\vec{r}^{\,k},\vec{r}^{\,k})+\tau^2\cdot(A\vec{r}^{\,k},A\vec{r}^{\,k})]$$

достигается при $\tau_k=(A\vec{r}^{\,k},\vec{r}^{\,k})/(A\vec{r}^{\,k},A\vec{r}^{\,k})$ и

$$\|\vec{r}^{\,k+1}\|_2^2=\|\vec{r}^{\,k}\|_2^2-\frac{|(A\vec{r}^{\,k},\vec{r}^{\,k})|^2}{(A\vec{r}^{\,k},A\vec{r}^{\,k})}<\|\vec{r}^{\,k}\|_2^2,$$

если $\vec{r}^{\,k}\neq 0$ & $(A\vec{r}^{\,k},\vec{r}^{\,k})\neq 0$,

то невязка строго убывает при $A>0$.

Во-вторых; из строгого убывания нормы невязки $\vec{r}^{\,k} = A\vec{z}^{\,k} \Rightarrow$

$$\|\vec{z}^{\,k+1}\|^2_{A^*A} \equiv (\vec{r}^{\,k+1}, \vec{r}^{\,k+1}) < (\vec{r}^{\,k}, \vec{r}^{\,k}) \equiv \|\vec{z}^{\,k}\|^2_{A^*A},$$

т.е. A^*A-норма ошибки строго убывает.

Кроме того, очевидно, что оператор S:

$$\vec{z}^{\,k+1} = S(\vec{z}^{\,k}) = \vec{z}^{\,k} - \tau_k(\vec{z}^{\,k}) \cdot A\vec{z}^{\,k}$$

непрерывен всюду за (исключением, быть может, 0) и по теореме о строгом убывании функционала итерационный процесс сходится.

Тема 13. Метод простой итерации

В методах наискорейшего спуска и минимальных невязок:

параметр τ_k = два скалярных произведения + умножение $A\vec{r}^k$;

постоянный параметра $\tau_k \equiv \tau \quad \Rightarrow$ меньше вычислений.

Предварительные замечания

Вопрос: $\tau = ?$, чтобы стационарный итерационный метод

$$\left.\begin{aligned} &\vec{x}^0 - \textit{задан}, \\ &\vec{x}^{k+1} = \vec{x}^k - \tau \cdot (A\vec{x}^k - \vec{b}), \\ &\quad k = 0, 1, \ldots \end{aligned}\right\}$$ **сходился к решению системы** $A\vec{x} = \vec{b}$**.**

Ответ: $\Leftrightarrow \quad \rho(S) = \rho(E - \tau \cdot A) < 1.$

Упражнение 1. $\rho(\underbrace{E-\tau\cdot A}_{S}) = \max\limits_{\lambda(A)\in Sp(A)} |1-\tau\cdot\lambda(A)|$

Упражнение 2. Если $\lambda(A) = \operatorname{Re}\lambda(A) + \underline{i}\cdot\operatorname{Im}\lambda(A)$,

то $|\lambda(S)|^2 = 1 - 2\tau\cdot\operatorname{Re}\lambda(A) + \tau^2\cdot|\lambda(A)|^2$

Упражнение 3. Если $\tau\cdot\operatorname{Re}\lambda(A) \le 0$, то $|\lambda(S)| \ge 1 \Rightarrow$ метод не сходится.

Упражнение 4. Если $\exists \quad \lambda_1(A) \;\; \& \;\; \lambda_2(A): \;\; \operatorname{Re}\lambda_1(A) < 0 \;\; \& \;\; \operatorname{Re}\lambda_2(A) > 0$,

то $\rho(S) = \rho(E-\tau\cdot A) \ge 1 \;\Rightarrow$ метод не сходится $\forall\, \tau \in R$.

Упражнение 5. Если $\operatorname{Re}\lambda(A) > 0 \;\; \forall\, \lambda \in Sp\,A$,

то $\forall\;\; 0 < \tau < \min\limits_{\lambda(A)\in Sp\,A} \dfrac{2\cdot\operatorname{Re}\lambda(A)}{|\lambda(A)|^2}$

$$\max_{\lambda(S)\in Sp\,S} |\lambda(S)|^2 = \max_{\lambda(A)\in Sp\,A} |1-\tau\cdot\lambda(A)|^2 < 1,$$

т.е. метод сходится.

Метод простой итерации

Теорема. Если матрица $A = A^* > 0$ системы $A\vec{x} = \vec{b}$, то метод простой итерации (метод Ричардсона):

$$\begin{cases} \vec{x}^0 - \textit{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - \tau \cdot (A\vec{x}^k - \vec{b}), \\ \quad k = 0, 1, \ldots \end{cases}$$

сходится при $\forall\ \tau \in (0, \dfrac{2}{\rho(A)})$.

Доказательство. Так как $A = A^* > 0$, то $\lambda(A) = \operatorname{Re}\lambda(A) > 0$ и (см. Упражнение 5.) метод сходится, если

$$0 < \tau < \min_{\lambda(A) \in Sp\,A} \frac{2 \cdot \operatorname{Re}\lambda(A)}{|\lambda(A)|^2} = \min_{\lambda(A) \in Sp\,A} \frac{2}{|\lambda(A)|} = \frac{2}{\rho(A)}.$$

Оптимальный выбор параметра τ метода Ричардсона

Пусть границы спектра матрицы $A = A^* > 0$ известны:

$$0 < \lambda_{\min}(A) \le \lambda(A) \le \lambda_{\max}(A) = \rho(A)$$

Определим $\tau_{опт}$:

$$\rho(S_{\tau_{опт}}) = \min_{\tau} \rho(S_\tau) \,| = \min_{\tau} \max_{\lambda(A)\in Sp\,A} |1 - \tau\cdot\lambda(A)| \le$$

$$\le \min_{\tau} \max_{\lambda_{\min}\le\lambda\le\lambda_{\max}} |1 - \tau\cdot\lambda| =$$

$$= \min_{\tau} \max\{|1-\tau\cdot\lambda_{\min}(A)|, |1-\tau\cdot\lambda_{\max}(A)|\} =$$

$$= |1-\tau_{опт}\cdot\lambda_{\min}(A)| = |1-\tau_{опт}\cdot\lambda_{\max}(A)|$$

$$\Rightarrow \tau_{опт} = \frac{2}{\lambda_{\min}+\lambda_{\max}}, \quad \rho(S_{\tau_{опт}}) = \frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}} < 1.$$

Теорема. Матрица $A = A^* > 0$, метод $\vec{x}^{k+1} = \vec{x}^k - \tau \cdot (A\vec{x}^k - \vec{b})$,

с оптимальным параметром $\tau_{опт} = 2 / [\lambda_{\max}(A) + \lambda_{\min}(A)]$,

тогда справедливы оценки:

$$\begin{cases} \|\vec{z}^k\|_2 \le \rho^k(S_{\tau_{опт}}) \cdot \|\vec{z}^0\|_2, \\ \|\vec{z}^k\|_A \le \rho^k(S_{\tau_{опт}}) \cdot \|\vec{z}^0\|_A, \; \text{где } \rho(S_{\tau_{опт}}) = \dfrac{\lambda_{\max}(A) - \lambda_{\min}(A)}{\lambda_{\max}(A) + \lambda_{\min}(A)} < 1. \end{cases}$$

Доказательство.

Так как $A = A^*$, то $S_\tau = S_\tau^*$ и, следовательно, $\|S_\tau\|_2 = \rho(S_\tau) \;\; \forall \tau$,

то $\|\vec{z}^k\|_2 \le \rho(S_{\tau_{опт}}) \cdot \|\vec{z}^{k-1}\|_2 \le \ldots \le \rho^k(S_{\tau_{опт}}) \cdot \|\vec{z}^0\|_2$.

Далее,

Упражнение 6: $\|S_\tau\|_A = \rho(S_\tau) \;\; \forall \tau$.

$\Rightarrow \|\vec{z}^k\|_A \le \rho(S_{\tau_{опт}}) \cdot \|\vec{z}^{k-1}\|_A \le \ldots \le \rho^k(S_{\tau_{опт}}) \cdot \|\vec{z}^0\|_A$.

Оценки сходимости МНС и ММН

Теорема. Если $A = A^* > 0$, то для ошибки $\vec{z}^k = \vec{x}^k - \vec{x}$ метода наискорейшего спуска:

$$\vec{x}^{k+1} = \vec{x}^k - \tau_k (A\vec{x}^k - b), \quad \tau_k = \frac{(\vec{r}^k, \vec{r}^k)}{(A\vec{r}^k, \vec{r}^k)}, \quad k = 0, 1, \ldots,$$

справедливы оценки:

$$\| \vec{z}^k \|_A \le \left[\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \right]^k \| \vec{z}^0 \|_A, \quad \| \vec{z}^k \|_2 \le \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \left[\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \right]^k \| \vec{z}^0 \|_2$$

Доказательство. Т.к. $\begin{cases} \| \vec{z}^{k+1} \|_A = \inf\limits_{\tau} \| \vec{z}^k - \tau A\vec{z}^k \|_A \le \| S_{\tau_{опт}} \vec{z}^k \|_A \le \\ \qquad \le \| S_{\tau_{опт}} \|_A \cdot \| \vec{z}^k \|_A = \rho(S_{\tau_{опт}}) \| \vec{z}^k \|_A \end{cases}$

то $\| \vec{z}^k \|_A \le \left[\rho_{опт} \right]^k \| \vec{z}^0 \|_A$, где $\rho(S_{\tau_{опт}}) = \dfrac{\lambda_{\max}(A) - \lambda_{\min}(A)}{\lambda_{\max}(A) + \lambda_{\min}(A)}$.

Так как $\lambda_{\min}(A)\cdot(\vec{z},\vec{z}) \le (A\vec{z},\vec{z}) \le \lambda_{\max}(A)\cdot(\vec{z},\vec{z})$ и, следовательно,

$$\sqrt{\lambda_{\min}(A)}\cdot\|\vec{z}\|_2 \le \|\vec{z}\|_A \le \sqrt{\lambda_{\max}(A)}\cdot\|\vec{z}\|_2,$$

то $\|\vec{z}^k\|_A \le \sqrt{\lambda_{\max}(A)/\lambda_{\min}(A)}\cdot\left[\rho_{опт}\right]^k\cdot\|\vec{z}^0\|_A$.

Теорема. Если $A = A^* > 0$,

то для ошибки $\vec{z}^k = \vec{x}^k - \vec{x}$ метода минимальных невязок:

$$\vec{x}^{k+1} = \vec{x}^k - \tau_k(A\vec{x}^k - \vec{b}), \quad \tau_k = \frac{(A\vec{r}^k,\vec{r}^k)}{(A\vec{r}^k,A\vec{r}^k)}, \quad k = 0, 1, \ldots,$$

справедливы оценки:

$$\|\vec{r}^k\|_2 \le \left[\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}}\right]^k\|\vec{r}^0\|_2, \quad \|\vec{z}^k\|_2 \le \frac{\lambda_{\max}}{\lambda_{\min}}\left[\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}}\right]^k\|\vec{z}^0\|_2$$

Доказательство. Так как

$$\| \vec{r}^{\,k+1} \|_2 = \inf_{\tau} \| \vec{r}^{\,k} - \tau A\vec{r}^{\,k} \|_2 \le \| \vec{r}^{\,k} - \tau_{опт} A\vec{r}^{\,k} \|_2 \le$$

$$\le \| S_{\tau_{опт}} \|_2 \cdot \| \vec{r}^{\,k} \|_2 = \rho(S_{\tau_{опт}}) \cdot \| \vec{z}^{\,k} \|_2$$

то $\| \vec{r}^{\,k} \|_2 \le \left[\rho_{опт}\right]^k \cdot \| \vec{r}^{\,0} \|_2$, где

$$\rho_{опт} = \rho(S_{\tau_{опт}}) = \frac{\lambda_{\max}(A) - \lambda_{\min}(A)}{\lambda_{\max}(A) + \lambda_{\min}(A)}.$$

Так как $\| \vec{r}^{\,k} \|_2 = \| A\vec{z}^{\,k} \|_2$ и

$$|\lambda_{\min}|^2 \cdot (\vec{z}, \vec{z}) \le (A\vec{z}, A\vec{z}) \le |\lambda_{\max}|^2 \cdot (\vec{z}, \vec{z}),$$

то из неравенств $\lambda_{\min}(A) \cdot \| \vec{z}^{\,k} \|_2 \le \| \vec{r}^{\,k} \|_2 \le \lambda_{\max}(A) \cdot \| \vec{z}^{\,k} \|_2$

следует оценка $\| \vec{z}^{\,k} \|_2 \le \dfrac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \rho_{опт}^k \cdot \| \vec{z}^{\,0} \|_2$.

Тема 14. Метод Ричардсона с чебышевскими параметрами.

Предварительные замечания

В предыдущем разделе для $A\vec{x} = \vec{b}$ с матрицей $A = A^* > 0$ для метода Ричардсона (простой итерации):

$$\begin{cases} \vec{x}^0 - \textit{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - \tau \cdot (A\vec{x}^k - \vec{b}), \\ \quad k = 0, 1, \ldots \end{cases}$$

определили параметр $\tau = \tau_{опт} = 2/(\lambda_{min} + \lambda_{\max})$:

$$\rho(S_{\tau_{опт}}) = \inf_{\tau} \rho(S_\tau) = (\lambda_{max} - \lambda_{\min})/(\lambda_{\max} + \lambda_{\min})$$

где $S_\tau = E - \tau \cdot A,\ 0 < \lambda_{min} = \lambda_{min}(A) \le \lambda_{\max} = \lambda_{\max}(A) = \rho(A)$.

Если $0<\alpha\le\beta$: $0<\alpha\le\lambda_{min}<\lambda_{\max}\le\beta$,

то оптимальный параметр $\tau\equiv\tau_{опт}(\alpha,\beta)=2/(\alpha+\beta)$:

$$\inf_{\tau}\rho(S_\tau)\equiv\inf_{\tau}[\max_{\lambda\in Sp(A)}|1-\tau\cdot\lambda|]\le\inf_{\tau}[\max_{\alpha\le\lambda\le\beta}|1-\tau\cdot\lambda|]\equiv$$

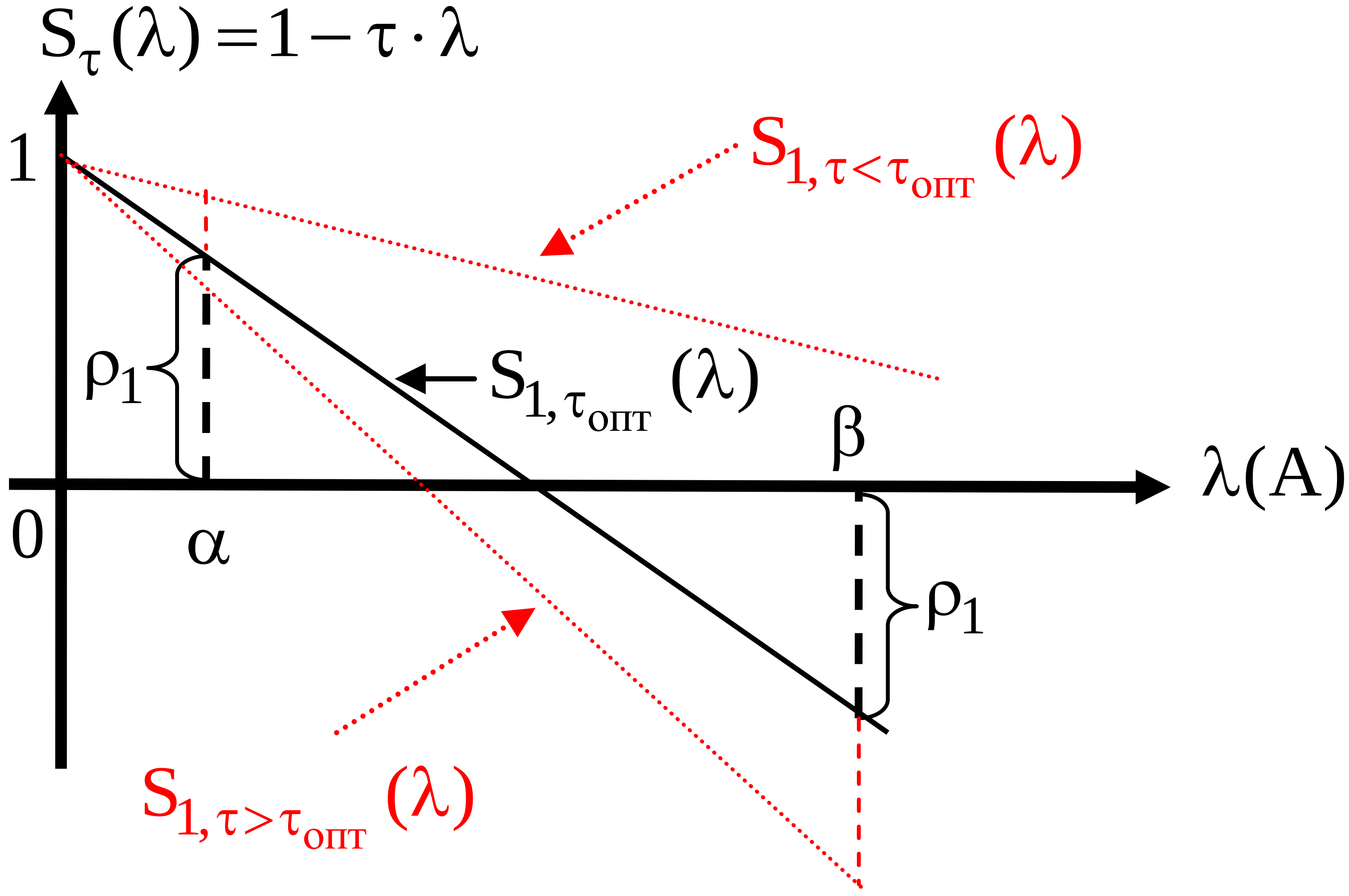
$$\equiv\inf_{\tau}\|\underbrace{S_{1,\tau}(\lambda)}_{1-\tau\cdot\lambda}\|_{C[\alpha,\beta]}=$$

$$=\inf_{\tau}[\max\{|1-\tau\cdot\alpha|,|1-\tau\cdot\beta|\}]=\underbrace{(\beta-\alpha)/(\beta+\alpha)}_{\approx1-2\alpha/\beta}\equiv\rho_1,$$

$$\tau_{опт}=2/(\alpha+\beta),$$

где $S_\tau(A)=E-\tau\cdot A \quad\Leftrightarrow\quad S_{1,\tau}(\lambda)=1-\tau\cdot\lambda$.

Решение этой минимаксной задачи на следующем графике:



Двуциклический метод Ричардсона.

Теперь сделаем две итерации (2-циклический метод Ричардсона), но с разными параметрами:

$$\left.\begin{aligned}\vec{x}^{2k+1} &= \vec{x}^{2k} - \tau_1 \cdot (A\vec{x}^{2k} - \vec{b}),\\ \vec{x}^{2k+2} &= \vec{x}^{2k+1} - \tau_2 \cdot (A\vec{x}^{2k+1} - \vec{b})\end{aligned}\right\} \Rightarrow \quad \vec{z}^{2k+2} = \underbrace{(E - \tau_2 A)(E - \tau_1 A)}_{S_{2,\tau=[\tau_1,\tau_2]}(A)} \vec{z}^{2k}$$

Упражнение 1. $\lambda(A) \in Sp(A) \;\Rightarrow\; \lambda(S_{2,\tau}(A)) = [1 - \tau_2 \lambda(A)] \cdot [E - \tau_1 \lambda(A)]$.

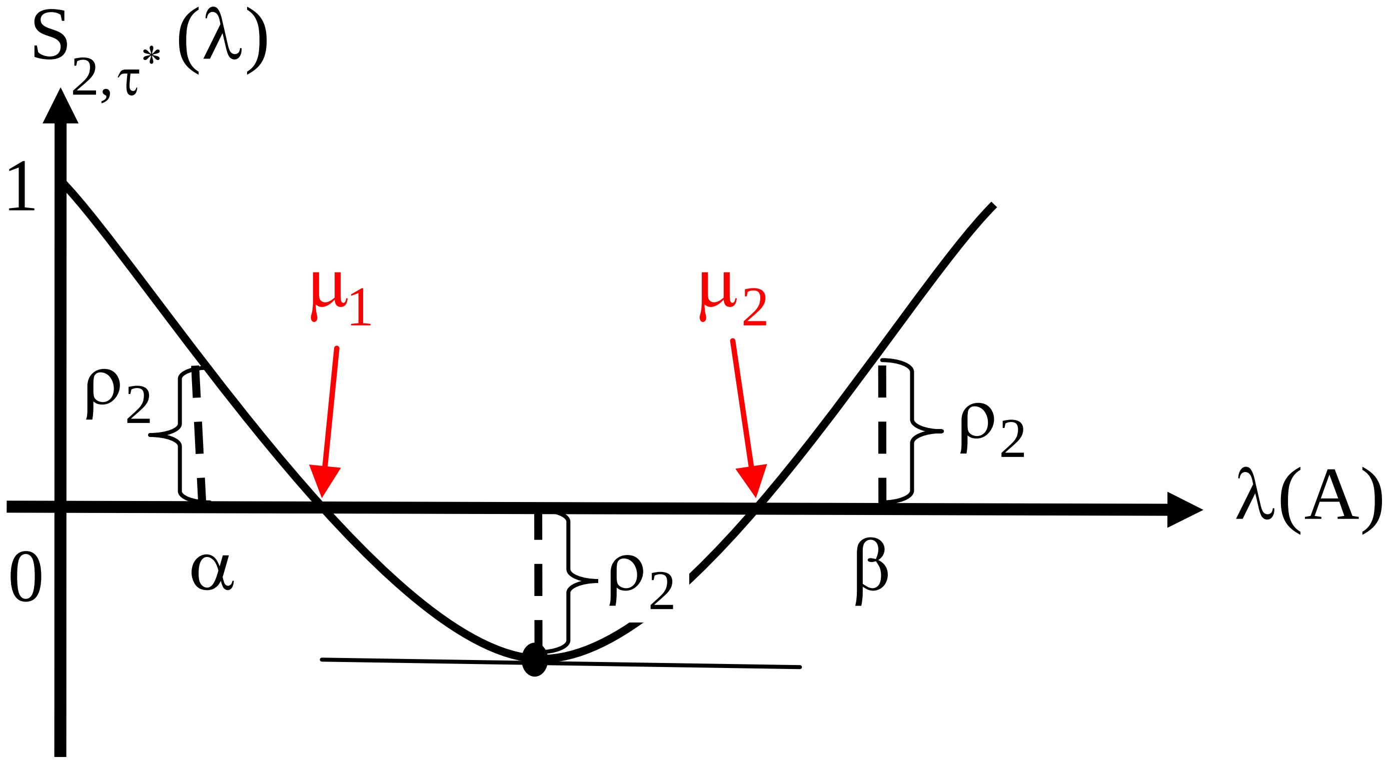
Упражнение 2. Определить параметры τ_1 и τ_2:

$$\inf_{\tau=(\tau_1,\tau_2)} \rho(S_{2,\tau}) \equiv \inf_{\tau} \{ \max_{\lambda \in Sp(A)} | \underbrace{[1 - \tau_2 \lambda(A)] \cdot [E - \tau_1 \lambda(A)]}_{S(\lambda)_{2,\tau=[\tau_1,\tau_2]}(\lambda)} | \}$$

$$\le \inf_{\tau} \{ \| S_{2,\tau}(\lambda) \|_{C[\alpha,\beta]} \} \equiv \rho_2.$$

Замечание. Значения параметров $\tau^* = (\tau_1^* = \mu_1^{-1},\ \tau_2^* = \mu_2^{-1})$

определяются из условий:



Упражнение 3. Доказать: $\rho_2 < [\rho_1]^2$,

т.е. 2-циклический метод Ричардсона сходится “быстрее” метода простой итерации.

m-циклический метод Ричардсона.

Параметры $\tau^* = \{\tau_i = \tau_i^*\}_{i=1}^{m}$ m-циклического метода Ричардсона:

$$\begin{cases} \vec{x}^{mk+1} = \vec{x}^{mk} - \tau_1 \cdot (A\vec{x}^{mk} - \vec{b}), \\ \ldots \\ \vec{x}^{mk+m} = \vec{x}^{mk+m-1} - \tau_m \cdot (A\vec{x}^{mk+m-1} - \vec{b}), \end{cases}$$

$$k = 0, 1, \ldots$$

$$\vec{z}^{mk+1} = (E - \tau_1 A) \cdot \vec{z}^{mk},$$

$$\ldots$$

$$\vec{z}^{mk+m} = (E - \tau_m A) \cdot \vec{z}^{mk+m-1}$$

$$S_{m,\tau} = (E - \tau_m A) \ldots (E - \tau_1 A)$$

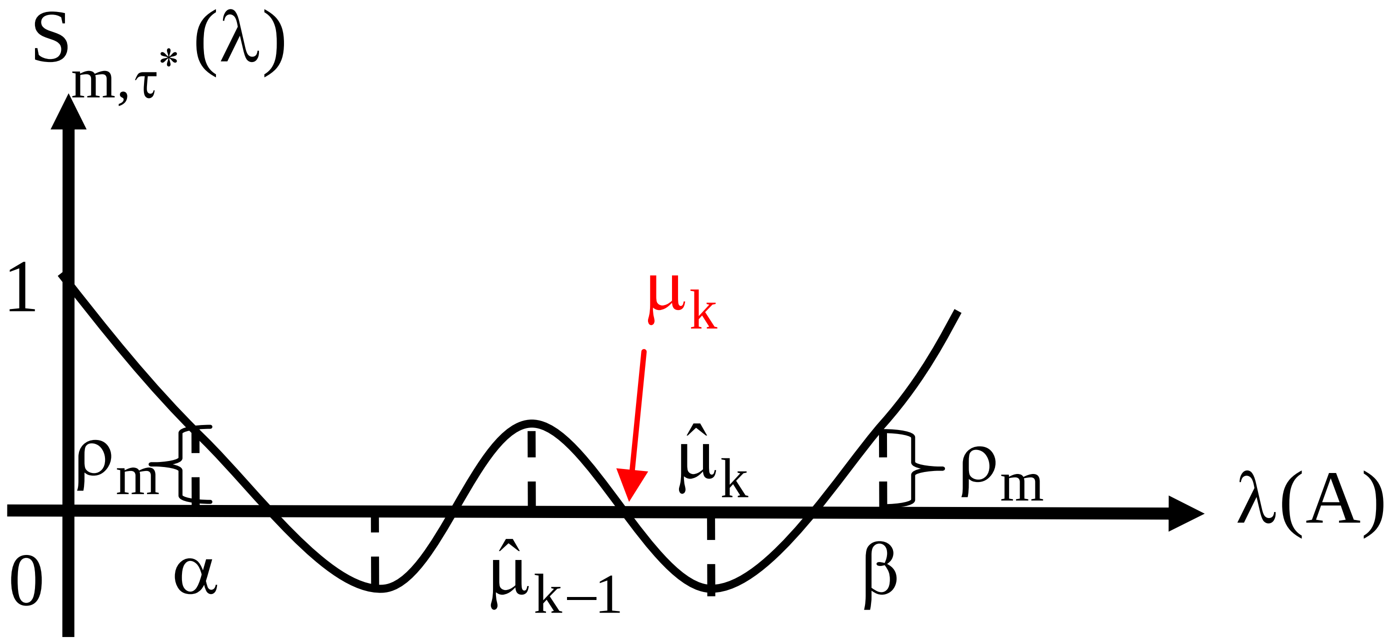
строим из условия минимизации оценки для $\rho(S_{m,\tau=[\tau_1, ,..., \tau_m]})$:

$$\inf_{\tau=(\tau_1, \ldots, \tau_m)} \rho(S_{m,\tau}) \equiv \inf_{\tau} [\max_{\lambda \in Sp(A)} |(1 - \tau_1 \cdot \lambda) \cdot \ldots \cdot (1 - \tau_m \cdot \lambda)|] \le$$

$$\le \inf_{\tau} [\max_{\alpha \le \lambda \le \beta} |\underbrace{(1 - \tau_1 \cdot \lambda) \cdot \ldots \cdot (1 - \tau_m \cdot \lambda)}_{S_{m,\tau}(\lambda)}|] \equiv$$

$$\equiv \inf_{\tau} \| S_{m,\tau}(\lambda) \|_{C[\alpha,\beta]} = \rho_m.$$

Решение этой задачи изображено на графике:



- $0 < \alpha \le \hat{\mu}_{k-1} < \mu_k < \hat{\mu}_k \le \beta, \quad k = 1, 2, \ldots, m;$
- $S_{m,\tau^*}(0) = 1;$ $\quad \begin{cases} S_{m,\tau^*}(0) = 1, \quad S_{m,\tau^*}(\mu_k) = 0 \\ S_{m,\tau^*}(\hat{\mu}_k) = (-1)^k \rho_m \end{cases}$

– “чебышевский альтернанс”.

Итак, мы должны:

1. Доказать: если $S_{m,\tau^*}(\lambda)=(1-\tau_1^*\cdot\lambda)\cdot\ldots\cdot(1-\tau_m^*\cdot\lambda)$

   имеет “чебышевский альтернанс” на интервале $[\alpha,\beta]$,

   то $S_{m,\tau^*}(\lambda)$ наименее уклоняется от нуля на интервале $[\alpha,\ \beta]$

   во множестве полиномов $S_m(\lambda),\ S_m(0)=1$:

   $$\| S_{m,\tau^*}(\lambda)\|_{C[\alpha,\beta]} = \inf_{S_m(\lambda),\ S_m(0)=1} \| S_m(\lambda)\|_{C[\alpha,\beta]}.$$

2. Построить $S_{m,\tau^*}(\lambda)$,

   найти его корни $0<\alpha\le\mu_1<\mu_2<\ldots<\mu_m\le\beta$,

   вычислить $\rho_m=\| S_{m,\tau^*}(\lambda)\|_{C[\alpha,\beta]}$.

3. Задать $\tau_1^*=1/\mu_1,\ \ \tau_2^*=1/\mu_2,\ \ \ldots,\ \ \tau_m^*=1/\mu_m$.

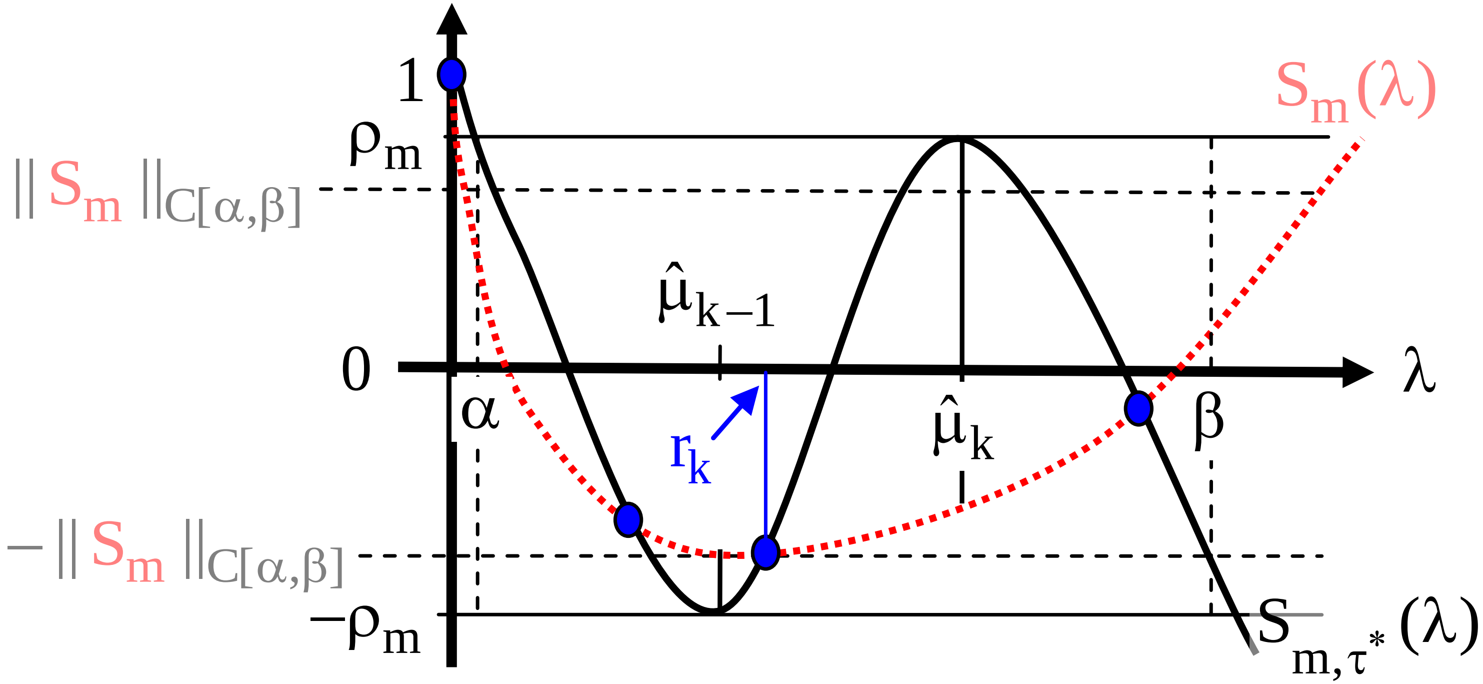
4. Построить итерационный метод:

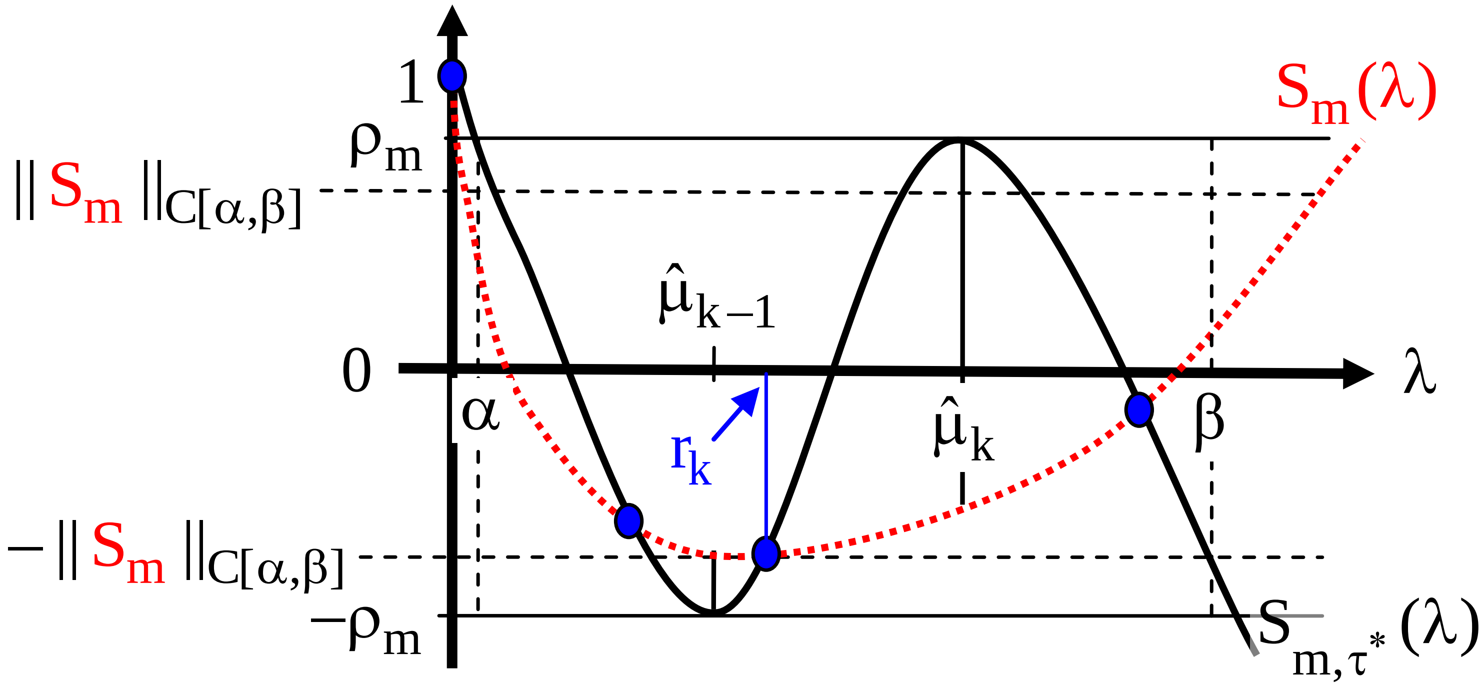
   $\vec{x}^{k+1}=\vec{x}^k-\tau_{k+1}(A\vec{x}^k-\vec{b});\ \ \tau_{k+m}=\tau_k:\ \ \tau_j=\tau_j^*,\ j=1,\ldots,m.$

Теорема.

Если $S_{m,\tau^*}(\lambda)$ имеет “чебышевский альтернанс” на интервале $[\alpha,\beta]$, то $\| S_{m,\tau^*}(\lambda) \|_{C[\alpha,\beta]} = \inf\limits_{S_m(\lambda),\, S_m(0)=1} \| S_m(\lambda) \|_{C[\alpha,\beta]}$.

Доказательство. Пусть $\exists\ S_m(\lambda)$: $\| S_{m,\tau^*}(\lambda) \|_{C[\alpha,\beta]} > \| S_m(\lambda) \|_{C[\alpha,\beta]}$:





$\Rightarrow R_m(\lambda) \equiv S_{m,\tau^*}(\lambda) - S_m(\lambda)$ имеет $m+1$ попарно различных корней:

$0 = r_0 < (\alpha \le r_1 \le \hat{\mu}_1) < \ldots < (\hat{\mu}_{k-1} \le r_k \le \hat{\mu}_k) < \ldots < (\hat{\mu}_{m-1} \le r_m \le \hat{\mu}_m = \beta)$.

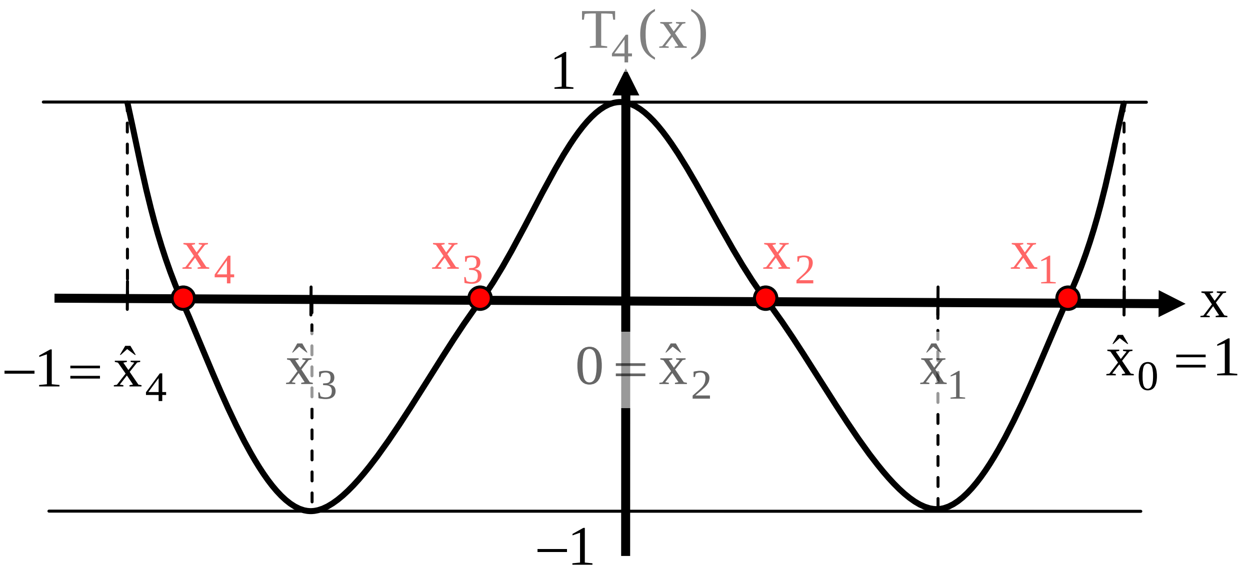
Но тогда $R_m(\lambda) \equiv 0$ (полином степени m имеет не больше m корней),

т.е. $\neg\exists\, S_m(\lambda)$: $\quad \| S_{m,\tau^*}(\lambda) \|_{C[\alpha,\beta]} > \| S_m(\lambda) \|_{C[\alpha,\beta]}$

$\Rightarrow \| S_{m,\tau^*}(\lambda) \|_{C[\alpha,\beta]} \le \| S_m(\lambda) \|_{C[\alpha,\beta]} \;\; \forall \;\; S_m(\lambda),\ S_m(0) = 1.$

Полином Чебышева

На интервале [–1, 1] рассмотрим функцию $T_m(x) = \cos(m \cdot \arccos(x))$.



Очевидно, что $(k = m,\ m-1,\ \dots,\ 0)$:

1. $\hat{x}_k = \cos(\dfrac{k\pi}{m}) \Rightarrow m \cdot \arccos(\hat{x}_k) = k\pi \Rightarrow T_m(\hat{x}_k) = (-1)^k$;

2. $x_k = \cos\dfrac{(2k-1)\pi}{2m} \Rightarrow m \cdot \arccos(x_k) = \dfrac{(2k-1)\pi}{2} \Rightarrow T_m(x_k) = 0$

Лемма.

Функция $T_m(x) = \cos(m \cdot \arccos(x))$ – полином степени m.

Доказательство.

$$T_0(x) \equiv 1, \quad T_1(x) \equiv x \text{ – полиномы.}$$

$$\cos((k+1)\varphi) + \cos((k-1)\varphi) = 2\cos(\varphi) \cdot \cos(k\varphi),$$

$$\varphi = \arccos x \quad \Rightarrow$$

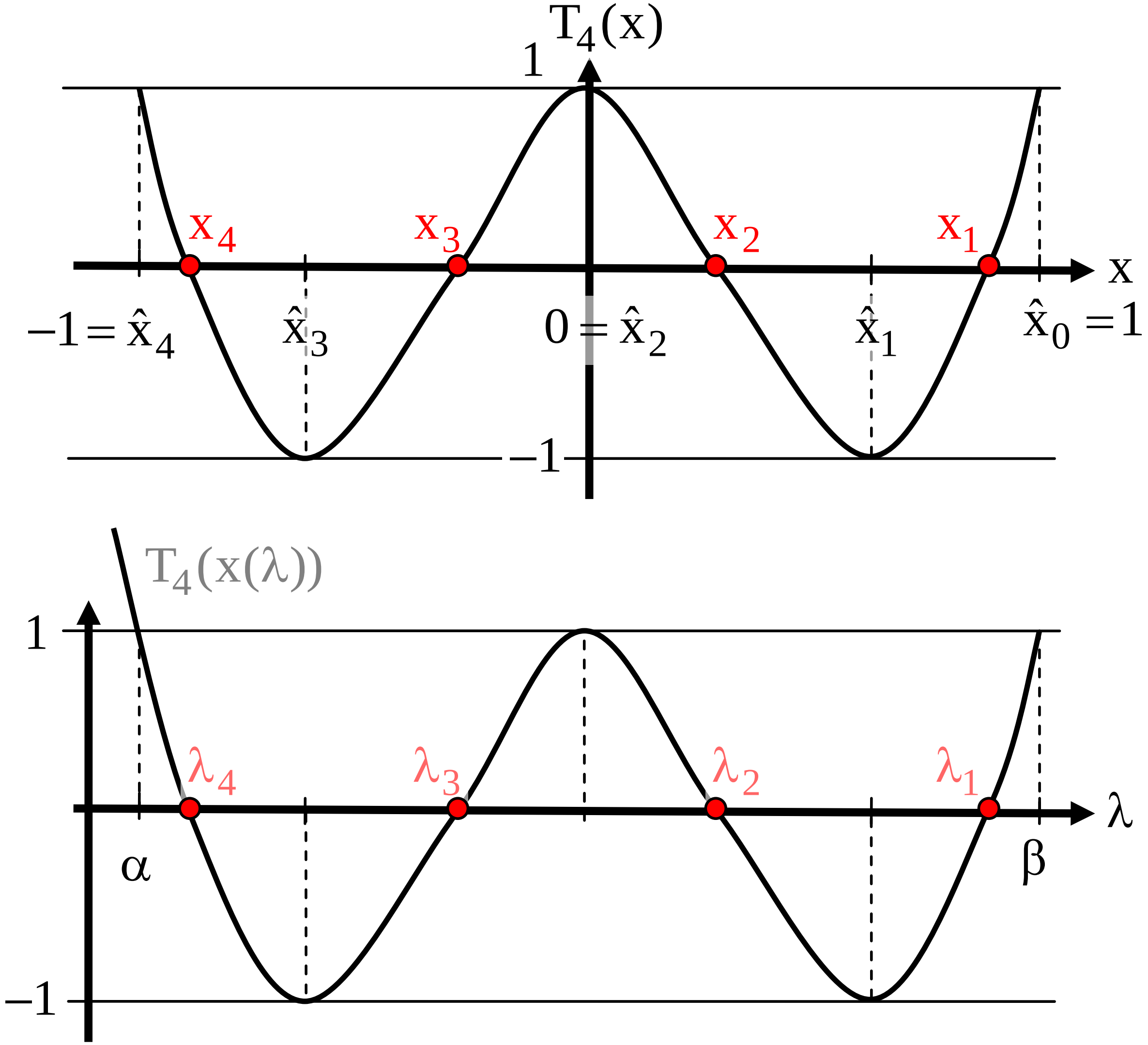
$$T_{k+1}(x) = 2 \cdot T_1(x) \cdot T_k(x) - T_{k-1}(x) \text{ – полином степени } k+1 \ \forall\ k > 1.$$

Замечание. Полином Чебышева $T_m(x)$ имеет альтернанс:

$$-1 = \hat{x}_m < \hat{x}_{m-1} < \ldots < \hat{x}_1 < \hat{x}_0 = 1,$$

$$\| T_m(x) \|_{C[-1,\,1]} = 1, \quad T_m(\hat{x}_k) = (-1)^k.$$

Построение $S_{m,\tau^*}(\lambda)$, $S_{m,\tau^*}(0)=1$, с чебышевским альтернансом:

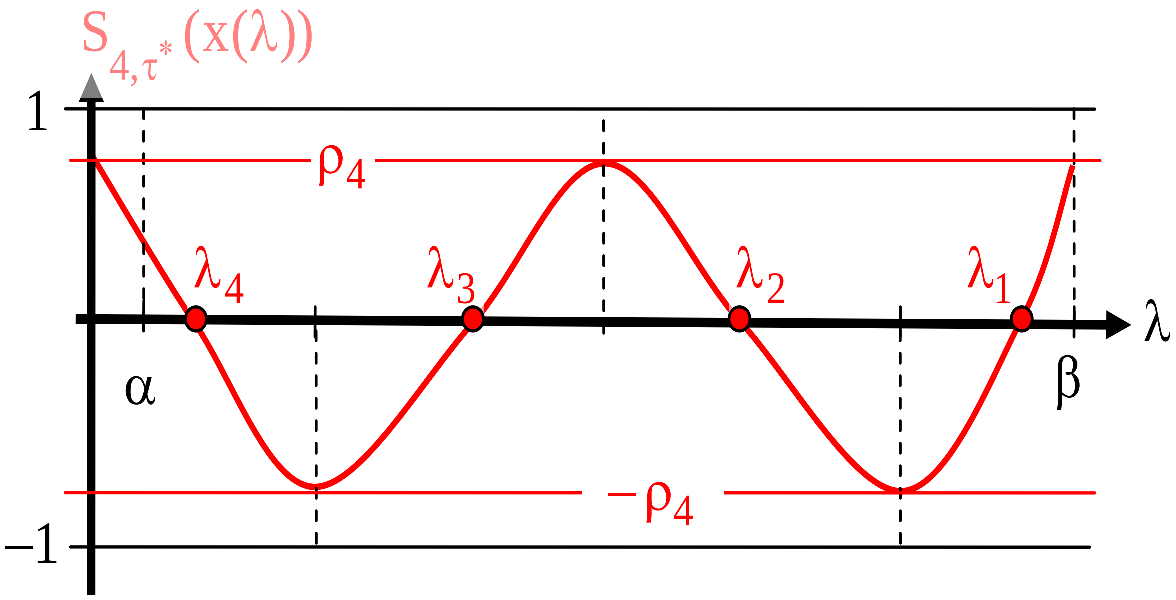


$$x \in [-1, 1] \;\rightarrow\; \lambda \in [\alpha, \beta]$$

$$\lambda(x) = \frac{\beta-\alpha}{2}x + \frac{\beta+\alpha}{2}$$

$$x(\lambda) = \frac{2}{\beta-\alpha}\lambda - \frac{\beta+\alpha}{\beta-\alpha}$$

Отнормируем полином $T_m(x(\lambda))$:



получим:

$$S_{m,\tau^*}(\lambda) = \frac{1}{T_m(x(0))} T_m(x(\lambda)),$$

т.е. $S_{m,\tau^*}(0) = 1$

с корнями

$$\mu_k = \lambda(x_k) = \frac{\beta-\alpha}{2} x_k + \frac{\beta+\alpha}{2}, \;\; x_k = \cos\left(\frac{(2k-1)\pi}{2m}\right), \;\; k = 1,\, 2,\, \dots,\, m.$$

Замечание. $\rho_m = \| S_{m,\tau^*}(\lambda) \|_{C[\alpha,\beta]} = | T_m(x(0)) |^{-1}$.

Теорема.

$$\rho_m = \| S_{m,\tau^*}(\lambda) \|_{C[\alpha,\beta]} = \frac{2\gamma^m}{1+\gamma^{2m}} < 1, \qquad \gamma = \frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}}.$$

Доказательство.

$$\rho_m = \| S_{m,\tau^*}(\lambda) \|_{C[\alpha,\beta]} = | T_m(x(0)) |^{-1}, \qquad x(0) = -\frac{\beta+\alpha}{\beta-\alpha} < -1.$$

Для вычисления $T_m(x(0))$ воспользуемся формулой

$$T_m(x) = \frac{(x+\sqrt{x^2-1})^m + (x-\sqrt{x^2-1})^m}{2} \text{ при } |x| > 1.$$

Заметим, что

$$x(0)+\sqrt{x^2(0)-1}=-\frac{\beta+\alpha}{\beta-\alpha}+\sqrt{\frac{(\beta+\alpha)^2}{(\beta-\alpha)^2}-1}=\frac{-\beta-\alpha+2\sqrt{\beta}\sqrt{\alpha}}{\beta-\alpha}=$$

$$=\frac{-(\sqrt{\beta})^2+2\sqrt{\beta}\sqrt{\alpha}-(\sqrt{\alpha})^2}{(\sqrt{\beta}+\sqrt{\alpha})\cdot(\sqrt{\beta}-\sqrt{\alpha})}=-\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}}\equiv-\gamma,$$

$$x(0)-\sqrt{x^2(0)-1}=\frac{1}{x(0)+\sqrt{x^2(0)-1}}=-\frac{1}{\gamma}.$$

Тогда

$$T_m(x(0))=\frac{(-\gamma)^m+(-\gamma^{-1})^m}{2}=(-1)^m\frac{1+\gamma^{2m}}{2\gamma^m}\quad\Rightarrow\quad\rho_m=\frac{2\gamma^m}{1+\gamma^{2m}}<1.$$

Доказательство формулы

$$T_k(x) = \frac{(x+\sqrt{x^2-1})^k + (x-\sqrt{x^2-1})^k}{2} \text{ при } |x|>1.$$

$T_0(x)=1$ и $T_1(x)=x$ – полиномы.

Осталось проверить, что $T_{k+1}(x) = 2\cdot T_1(x)\cdot T_k(x) - T_{k-1}(x)$ или

$$\underbrace{(x+\sqrt{x^2-1})^{k+1} + (x-\sqrt{x^2-1})^{k+1}}_{2T_{k+1}(x)} =$$

$$= 2x[\underbrace{(x+\sqrt{x^2-1})^k + (x-\sqrt{x^2-1})^k}_{2T_k(x)}] - [\underbrace{(x+\sqrt{x^2-1})^{k-1} + (x-\sqrt{x^2-1})^{k-1}}_{2T_{k-1}(x)}]$$

Пусть $y=\sqrt{x^2-1}$, тогда

$$(x+y)^{k+1}+(x-y)^{k+1}=x\cdot[(x+y)^k+(x-y)^k]+y\cdot[(x+y)^k-(x-y)^k]=$$

$$=x\cdot[(x+y)^k+(x-y)^k]+$$

$$+y\cdot[x\cdot\{(x+y)^{k-1}-(x-y)^{k-1}\}+y\cdot\{(x+y)^{k-1}+(x-y)^{k-1}\}]=$$

$$=x\cdot[(x+y)^k+(x-y)^k]+$$

$$+y\cdot x\cdot\{(x+y)^{k-1}-(x-y)^{k-1}\}+(x^2-1)\cdot\{(x+y)^{k-1}+(x-y)^{k-1}\}=$$

$$=2x\cdot[(x+y)^k+(x-y)^k]-\{(x+y)^{k-1}+(x-y)^{k-1}\}+$$

$$+(-x^2-xy+yx+x^2)\cdot(x+y)^{k-1}+(-x^2+xy-yx+x^2)\cdot(x-y)^{k-1}=$$

$$=2x\cdot[(x+y)^k+(x-y)^k]-\{(x+y)^{k-1}+(x-y)^{k-1}\},$$ что и тр. док.

Формулы m-циклического метода Ричардсона с чебышевскими параметрами

Итак, для решения системы $A\vec{x}=\vec{b}$, $A=A^*>0$, $Sp(A)\subset[\alpha,\beta]$, $\alpha>0$, двучленные формулы m-циклического метода Ричардсона:

$$\begin{cases} \vec{x}^{k+1}=\vec{x}^k-\tau_{k+1}(A\vec{x}^k-\vec{b}), \quad k=0,1,\dots, \\ \{\tau_1,\dots,\tau_m\};\{\tau_{m+1}=\tau_1,\dots,\tau_{2m}=\tau_m\};\dots\ (\tau_{m+j}=\tau_j), \end{cases}$$

$$\tau_j\equiv\tau_j^*=2/[(\beta+\alpha)+(\beta-\alpha)\cdot\cos\frac{(2j-1)\pi}{2m}],\ j=1,2,\dots,m,$$

а матрица перехода $S_{m,\tau^*}(A)=(E-\tau_m^*A)\cdot\ldots\cdot(E-\tau_1^*A)$ для его ошибки за полный цикл ($\vec{z}^{(t+1)\cdot m}=S_{m,\tau^*}\cdot\vec{z}^{t\cdot m}$) симметрична, и, следовательно,

$$\|S_{m,\tau^*}(A)\|_A=\|S_{m,\tau^*}(A)\|_2\le$$
$$\le\|S_{m,\tau^*}(\lambda)\|_{C[\alpha,\beta]}=\rho_m\approx 2\gamma^m,\ \gamma=\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}}.$$

Численная неустойчивость двучленных формул метода Ричардсона.

Из-за ошибок округления

реализация формул $x^{k+1} = x^k - \tau_{k+1}(Ax^k - b)$ неустойчива,

т.к. норма оператора шага $\| E - \tau_{k+1}A \|_2$ для ошибки может быть значительно больше 1 (в методе простой итерации эта норма меньше 1), в то время как $\| (E - \tau_1 A) \cdot \ldots \cdot (E - \tau_m A) \|_2 \le \rho_m < 1$.

Поэтому, если несколько итераций выполняется с параметрами τ_{k+1}, для которых нормы операторов шага

$$\| E - \tau_{k+1}A \|_2 = \max\{ |1 - \tau_{k+1} \underbrace{\lambda_{\min}(A)}_{\alpha>0}|, |1 - \tau_{k+1} \underbrace{\lambda_{\max}(A)}_{\beta}| \} >> 1,$$

то процесс вычислений “прервется из-за переполнения”.

Трехчленные формулы реализации метода Ричардсона.

Цель: для решения системы $A\vec{x} = \vec{b}$, $A = A^* > 0$, $Sp(A) \subset [\alpha, \beta]$, $\alpha > 0$,

построить приближения $\vec{x}^0$, $\vec{x}^1$, ..., $\vec{x}^k$, ... :

$$\begin{cases} \|\vec{z}^1\|_2 = \|S_{1,\tau^*}(A)\vec{z}^0\|_2 \le \underbrace{\|S_{1,\tau^*}(A)\|_2}_{\rho_1} \cdot \|\vec{z}^0\|_2 \\ \dots \\ \|\vec{z}^k\|_2 = \|S_{k,\tau^*}(A)\vec{z}^0\|_2 \le \underbrace{\|S_{k,\tau^*}(A)\|_2}_{\rho_k} \cdot \|\vec{z}^0\|_2\ , \ \dots; \ \vec{z}^k = \vec{x}^k - \vec{x} \end{cases}$$

где $S_{k,\tau^*}(\lambda) = \dfrac{1}{T_k(x(0))} T_k(x(\lambda)) \equiv \dfrac{1}{t_k} T_k(x(\lambda)) \quad \Leftrightarrow$

$$\Leftrightarrow \ S_{k,\tau^*}(A) = \frac{1}{T_k(x(0))} T_k(x(A)) \equiv \frac{1}{t_k} T_k(x(A)).$$

Упражнение 1. Из $\begin{cases} T_0(x)=1, \quad T_1(x)=x, \\ T_{k+1}(x)=2\cdot T_1(x)\cdot T_k(x)-T_{k-1}(x), \end{cases}$

$$\Rightarrow \; t_k \equiv T_m\left(-\frac{\beta+\alpha}{\beta-\alpha}\right): \; \begin{cases} t_0=1, \quad t_1=-\dfrac{\beta+\alpha}{\beta-\alpha}, \\ t_{k+1}=2t_1t_k-t_{k-1}. \end{cases}$$

$$\Rightarrow \; \begin{cases} S_{0,\tau^*}(\lambda)=1, \quad S_{1,\tau^*}(\lambda)=1-\dfrac{2}{\beta+\alpha}\lambda, \\ S_{k+1,\tau^*}(\lambda)=\left(1+\dfrac{t_{k-1}}{t_{k+1}}\right)\cdot\left[1-\dfrac{2}{\beta+\alpha}\lambda\right]\cdot S_{k,\tau^*}(\lambda)-\dfrac{t_{k-1}}{t_{k+1}}\cdot S_{k-1,\tau^*}(\lambda)= \\ =S_{k,\tau^*}(\lambda)-\left(1+\dfrac{t_{k-1}}{t_{k+1}}\right)\dfrac{2}{\beta+\alpha}\lambda\cdot S_{k,\tau^*}(\lambda)+\dfrac{t_{k-1}}{t_{k+1}}\left[S_{k,\tau^*}(\lambda)-S_{k-1,\tau^*}(\lambda)\right] \end{cases}$$

Упражнение 2. $S_{0,\tau^*}(A)=E,\ \ S_{1,\tau^*}(A)=E-\dfrac{2}{\beta+\alpha}A$

$$S_{k+1,\tau^*}(A)=S_{k,\tau^*}(A)-(1+\frac{t_{k-1}}{t_{k+1}})\frac{2}{\beta+\alpha}A\cdot S_{k,\tau^*}(A)+\frac{t_{k-1}}{t_{k+1}}[S_{k,\tau^*}(A)-S_{k-1,\tau^*}(A)]$$

Упражнение 3. $\underbrace{S_{0,\tau^*}(A)\vec{z}^{\,0}}_{\vec{z}^{\,0}}=\vec{z}^{\,0},\ \ \underbrace{S_{1,\tau^*}(A)\vec{z}^{\,0}}_{\vec{z}^{\,1}}=\vec{z}^{\,0}-\dfrac{2}{\beta+\alpha}A\vec{z}^{\,0}$

$$\underbrace{S_{k+1,\tau^*}(A)\vec{z}^{\,0}}_{\vec{z}^{\,k+1}}=\underbrace{S_{k,\tau^*}(A)\vec{z}^{\,0}}_{\vec{z}^{\,k}}-(1+\frac{t_{k-1}}{t_{k+1}})\frac{2}{\beta+\alpha}A\cdot\underbrace{S_{k,\tau^*}(A)\vec{z}^{\,0}}_{\vec{z}^{\,k}}+$$

$$+\frac{t_{k-1}}{t_{k+1}}[\underbrace{S_{k,\tau^*}(A)\vec{z}^{\,0}}_{\vec{z}^{\,k}}-\underbrace{S_{k-1,\tau^*}(A)\vec{z}^{\,0}}_{\vec{z}^{\,k-1}}]$$

Упражнение 4. $\vec{x}^0$ – *задан*, $\vec{x}^1 = \vec{x}^0 - \dfrac{2}{\beta+\alpha}\underbrace{(A\vec{x}^0 - \vec{b})}_{\vec{r}^0}$

$$\vec{x}^{k+1} = \vec{x}^k - \left(1 + \frac{t_{k-1}}{t_{k+1}}\right)\frac{2}{\beta+\alpha}\vec{r}^k + \frac{t_{k-1}}{t_{k+1}}[\vec{x}^k - \vec{x}^{k-1}]$$

Упражнение 5. $\dfrac{t_{k-1}}{t_{k+1}} = \omega_k \omega_{k+1}$:

$$\omega_1 = \frac{t_0}{t_1} \equiv -\frac{\beta-\alpha}{\beta+\alpha}, \quad \omega_{k+1} \equiv \frac{t_k}{t_{k+1}} = \frac{1}{2t_1 - \omega_k}, \quad k = 1, 2, \dots . \qquad \Rightarrow$$

$$x^{k+1} = x^k + \omega_k \omega_{k+1}(x^k - x^{k-1}) - \frac{2}{\beta+\alpha}(1 + \omega_k \omega_{k+1}) \cdot (Ax^k - b), \;\; k = 1, 2, \dots ,$$

– трехслойный метод Ричардсона:

$$\| \vec{z}^k \|_2 = \| S_{k,\tau^*}(A)\vec{z}^0 \|_2 \le \underbrace{\| S_{k,\tau^*}(A) \|_2}_{\rho_k} \cdot \| \vec{z}^0 \|_2 \;\; \forall\, k$$

Тема 15. Метод сопряженных градиентов. Предварительные замечания

Циклический метод Ричардсона

$$\vec{x}^{k+1} = \vec{x}^k - \tau_{k+1}^*(A\vec{x}^k - \vec{b}),\ \ \vec{z}^{k+1} = (E - \tau_{k+1}^* A)\vec{z}^k,\ \ \vec{z}^{k\cdot t+m} = S_{m,\tau^*}(A)\vec{z}^{k\cdot t}$$

с чебышевскими параметрами $\tau_{k+m}^* = \tau_k^*$

для решения системы $A\vec{x}^k = \vec{b}$: $A = A^* > 0,\ \lambda(A) \in [\alpha,\ \beta],\ \alpha > 0,$

за m итераций уменьшает ошибку в $1/\rho_m$ раз:

$$\|\vec{z}^{t\cdot m}\|_A \le \rho_m \cdot \|\vec{z}^{(t-1)\cdot m}\|_A \le [\rho_m]^t \cdot \|\vec{z}^0\|_2 \quad \forall\ \vec{z}^0,$$

где

$$\rho_m = \| S_{m,\tau^*}(\lambda) \|_{C[\alpha,\beta]} \equiv 2\gamma^m / (1+\gamma^{2m}) < 1,\ \ \gamma = (\sqrt{\beta} - \sqrt{\alpha}) / (\sqrt{\beta} + \sqrt{\alpha}),$$

а параметры $\tau_{k+m}^* = \tau_k^*(\alpha, \beta)$ одинаковые $\forall\ \vec{z}^0$.

Но, если $\tau_k = \tau_k(\vec{z}^0)$:

$$\| \vec{z}^m \|_A = \| \underbrace{(E - \tau_m(\vec{z}^0)A) \cdot \ldots \cdot (E - \tau_1(\vec{z}^0)A)}_{S_{m,\tau(\vec{z}^0)}(A)} \vec{z}^0 \|_A =$$

$$= \inf_{\tau} \| \underbrace{(E - \tau_m A) \cdot \ldots \cdot (E - \tau_1 A)}_{S_{m,\tau}(A)} \vec{z}^0 \|_A$$

то $\| \vec{z}^m \|_A \le \rho_m(\vec{z}^0) \cdot \| \vec{z}^0 \|_A$, $\rho_m(\vec{z}^0) \le \rho_m$

Упражнение 1.

$$\vec{z}^m = \vec{z}^0 - q_1 A\vec{z}^0 - \ldots - q_m A^m \vec{z}^0 \in \vec{z}^0 + L\{A\vec{z}^0, \ldots, A^m \vec{z}^0\}$$

Упражнение 2. Если $L_m\{\underbrace{\underset{\vec{r}^0 = A\vec{x}^0 - \vec{b}}{A\vec{z}^0}, \ldots, \underset{A^{m-1}\vec{r}^0}{A^m \vec{z}^0}}_{\text{известные векторы}}\} = L_m\{\underbrace{\vec{g}_1, \ldots, \vec{g}_m}_{\text{новый базис}}\}$,

$$\text{то } \vec{z}^m = \vec{z}^0 - \alpha_1 \vec{g}_1 - \ldots - \alpha_m \vec{g}_m.$$

$\Rightarrow$ **Упражнение 4.** $\begin{cases} \vec{x}^{\,0} - \textit{задан} \\ \vec{x}^{\,k} = \vec{x}^{\,k-1} - \alpha_k \vec{g}_k, \; k = 1, 2, \ldots \end{cases}$

$$\| \vec{z}^{\,m} \|_A = \min_{\tau} \| (E - \tau_m A) \ldots (E - \tau_1 A) \vec{z}^{\,0} \|_A = \min_{\alpha} \| \vec{z}^{\,0} - \alpha_1 \vec{g}_1 - \ldots - \alpha_m \vec{g}_m \|_A$$

Параметры $\alpha_1, \ldots, \alpha_m$ удовлетворяют системе уравнений

$$\frac{1}{2} \frac{\partial (A\vec{z}^{\,m}, \vec{z}^{\,m})}{\partial \alpha_i} = (A \frac{\partial \vec{z}^{\,m}}{\partial \alpha_i}, \vec{z}^{\,m}) \equiv (A\vec{g}_i, \vec{z}^{\,0} - \alpha_1 \vec{g}_1 - \ldots - \alpha_m \vec{g}_m) = 0, \quad i = 1, \ldots, m$$

$\Rightarrow$

$$\begin{bmatrix} (A\vec{g}_1, \vec{g}_1) & (A\vec{g}_1, \vec{g}_2) & \ldots & (A\vec{g}_1, \vec{g}_m) \\ (A\vec{g}_2, \vec{g}_1) & (A\vec{g}_2, \vec{g}_2) & \ldots & (A\vec{g}_2, \vec{g}_m) \\ \vdots & \vdots & \ldots & \vdots \\ (A\vec{g}_m, \vec{g}_1) & (A\vec{g}_m, \vec{g}_2) & \ldots & (A\vec{g}_m, \vec{g}_m) \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} (Az^0, \vec{g}_1) \\ (Az^0, \vec{g}_2) \\ \vdots \\ (Az^0, \vec{g}_m) \end{pmatrix} = \begin{pmatrix} (\vec{r}^{\,0}, \vec{g}_1) \\ (\vec{r}^{\,0}, \vec{g}_2) \\ \vdots \\ (\vec{r}^{\,0}, \vec{g}_m) \end{pmatrix}$$

Если базис $\{\vec{g}_i\}_{i=1}^{m}$ является A–ортогональным, т.е. $(A\vec{g}_i, \vec{g}_j) = 0, \quad i \neq j$, то

$$\alpha_k = \frac{(A\vec{z}^0, \vec{g}_k)}{(A\vec{g}_k, \vec{g}_k)} \equiv \frac{(\vec{r}^0, \vec{g}_k)}{(A\vec{g}_k, \vec{g}_k)}, \qquad \begin{cases} \vec{z}^m = \vec{z}^0 - \alpha_1\vec{g}_1 - \alpha_2\vec{g}_2 - \ldots - \alpha_m\vec{g}_m, \\ \vec{x}^m = \vec{x}^0 - \alpha_1\vec{g}_1 - \alpha_2\vec{g}_2 - \ldots - \alpha_m\vec{g}_m. \end{cases}$$

Важное замечание. В случае A–ортогонального базиса $\{\vec{g}_i\}_{i=1}^{k}$ в L_k, $k = \overline{1, m}$,

вектор $\vec{z}^k = \vec{z}^0 - \alpha_1\vec{g}_1 - \ldots - \alpha_k\vec{g}_k$ является решением задачи

$$\| \vec{z}^k \|_A = \min_{\tau} \| (E - \tau_k A)\ldots(E - \tau_1 A)\vec{z}^0 \|_A = \min_{\tau} \| S_{k,\tau}(A)\vec{z}^0 \|_A$$

т.е., в отличие от метода Ричардсона, минимизация нормы ошибки осуществляется на каждом внутреннем шаге цикла, и, следовательно,

$$\| \vec{z}^k \|_A \le \rho_k \cdot \| \vec{z}^0 \|_A \quad \forall\, k = 1, \ldots, m.$$

Построение A–ортогонального базиса

A–ортогонализуем систему векторов $\{Az^0, \ldots, A^m z^0\}$.

$$\vec{g}_1 = A\vec{z}^{\,0} \equiv \vec{r}^{\,0} \in L_1 \quad \Rightarrow \quad \alpha_1 = \frac{(\vec{r}^{\,0}, \vec{g}_1)}{(A\vec{g}_1, \vec{g}_1)} \quad \Rightarrow \quad \begin{cases} \vec{z}^{\,1} = \vec{z}^{\,0} - \alpha_1 \vec{g}_1 \\ \vec{r}^{\,1} = A\vec{z}^{\,1} = A\vec{z}^{\,0} - \alpha_1 A\vec{g}_1 \end{cases}$$

Утверждение 1. $\vec{r}^{\,1} \in L_2$, так как $\vec{r}^{\,1} = A\vec{z}^{\,0} - \alpha_1 A^2 \vec{z}^{\,0} \in L_2 = L\{A\vec{z}^{\,0}, A^2\vec{z}^{\,0}\}$

Утверждение 2. $(\vec{r}^{\,1}, \vec{g}_1) = 0$, т.е. $r^1 \perp L_1$.

$$(\vec{r}^{\,1}, \vec{g}_1) \equiv (A\vec{z}^{\,0} - \alpha_1 A\vec{g}_1,\ \vec{g}_1) \equiv (\vec{r}^{\,0} - \alpha_1 A\vec{g}_1,\ \vec{g}_1) =$$

$$= (\vec{r}^{\,0},\ \vec{g}_1) - \frac{(\vec{r}^{\,0}, \vec{g}_1)}{(A\vec{g}_1, \vec{g}_1)} (A\vec{g}_1,\ \vec{g}_1) = 0$$

Далее применяем метод математической индукции.

Пусть построены:

$$\begin{cases} \vec{g}_1 & A-\textit{ортогональный базис в} & L_1 = L\{A\vec{z}^{\,0}\} \\ \{\vec{g}_1, \vec{g}_2\} & \ldots & L_2 = L\{A\vec{z}^{\,0}, A^2\vec{z}^{\,0}\} \\ \ldots & \ldots & \ldots \\ \{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_k\} & \ldots & L_k = L\{A\vec{z}^{\,0}, A^2\vec{z}^{\,0}, \ldots, A^k\vec{z}^{\,0}\} \end{cases}$$

и вычислены $\alpha_1 = \dfrac{(\vec{r}^{\,0}, \vec{g}_1)}{(A\vec{g}_1, \vec{g}_1)}$, $\alpha_2 = \dfrac{(\vec{r}^{\,0}, \vec{g}_2)}{(A\vec{g}_2, \vec{g}_2)}$, …, $\alpha_k = \dfrac{(\vec{r}^{\,0}, \vec{g}_k)}{(A\vec{g}_k, \vec{g}_k)} \quad \Rightarrow$

$$
\left\{
\begin{array}{l}
\left.\begin{array}{l}\vec{z}^{\,1} = \vec{z}^{\,0} - \alpha_1 \vec{g}_1 \\ \vec{r}^{\,1} = A\vec{z}^{\,1}\end{array}\right\} \\
\left.\begin{array}{l}\vec{z}^{\,2} = \vec{z}^{\,0} - \alpha_1 \vec{g}_1 - \alpha_2 \vec{g}_2 \\ \vec{r}^{\,2} = A\vec{z}^{\,2}\end{array}\right\} \\
\ldots \\
\left.\begin{array}{l}\vec{z}^{\,k} = \vec{z}^{\,0} - \alpha_1 \vec{g}_1 - \alpha_2 \vec{g}_2 - \ldots - \alpha_k \vec{g}_k \\ \vec{r}^{\,k} = A\vec{z}^{\,k}\end{array}\right\}
\end{array}
\right.
$$

$$
\begin{cases}\vec{r}^{\,1} \in L_2, \quad \vec{r}^{\,1} \perp L_1: \\ (\vec{r}^{\,1}, \vec{g}_1) = 0\end{cases}
$$

$$
\begin{cases}\vec{r}^{\,2} \in L_3, \quad \vec{r}^{\,2} \perp L_2: \\ (\vec{r}^{\,2}, \vec{g}_1) = 0, \quad (\vec{r}^{\,2}, \vec{g}_2) = 0\end{cases}
$$

...

$$
\begin{cases}\vec{r}^{\,k} \in L_{k+1}, \quad \vec{r}^{\,k} \perp L_k: \\ (\vec{r}^{\,k}, \vec{g}_1) = 0, \quad (\vec{r}^{\,k}, \vec{g}_2) = 0, \\ \qquad \ldots, (\vec{r}^{\,k}, \vec{g}_k) = 0\end{cases}
$$

Наша задача. Построить $\vec{g}_{k+1} = \vec{r}^{\,k} - \gamma_1 \vec{g}_1 - \ldots - \gamma_k \vec{g}_k \in L_{k+1}$: $A\vec{g}_{k+1} \perp L_k$.

Доказать: $\vec{r}^{\,k+1} \in L_{k+2}, \quad \vec{r}^{\,k+1} \perp L_{k+1}$.

Вычислим $\vec{g}_{k+1} = \underbrace{\vec{r}^{\,k}}_{\in L_{k+1}} - \underbrace{\gamma_1 \vec{g}_1 - \ldots - \gamma_k \vec{g}_k}_{\in L_k} \in L_{k+1},\ A\vec{g}_{k+1} \perp L_k$:

$$\begin{cases} (A\vec{g}_{k+1}, \vec{g}_i) \equiv (A\vec{r}^{\,k}, \vec{g}_i) - \gamma_1 \cdot (A\vec{g}_1, \vec{g}_i) - \ldots - \gamma_k \cdot (A\vec{g}_k, \vec{g}_i) = 0, \\ \qquad i = 1, \ldots, k, \end{cases}$$

т.е.

$$\begin{bmatrix} (A\vec{g}_1, \vec{g}_1) & \underbrace{(A\vec{g}_2, \vec{g}_1)}_{0} & \ldots & \underbrace{(A\vec{g}_k, \vec{g}_1)}_{0} \\ \underbrace{(A\vec{g}_1, \vec{g}_2)}_{0} & (A\vec{g}_2, \vec{g}_2) & \ldots & \underbrace{(A\vec{g}_k, \vec{g}_2)}_{0} \\ \vdots & \vdots & \ldots & \vdots \\ \underbrace{(A\vec{g}_1, \vec{g}_k)}_{0} & \underbrace{(A\vec{g}_2, \vec{g}_k)}_{0} & \ldots & (A\vec{g}_k, \vec{g}_k) \end{bmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_m \end{pmatrix} = \begin{pmatrix} (A\vec{r}^{\,k}, \vec{g}_1) \\ (A\vec{r}^{\,k}, \vec{g}_2) \\ \vdots \\ (A\vec{r}^{\,k}, \vec{g}_k) \end{pmatrix} = \begin{pmatrix} (\vec{r}^{\,k}, A\vec{g}_1) \\ (\vec{r}^{\,k}, A\vec{g}_2) \\ \vdots \\ (\vec{r}^{\,k}, A\vec{g}_k) \end{pmatrix}$$

$$\Rightarrow \quad \gamma_i = (\vec{r}^{\,k}, A\vec{g}_i) / (A\vec{g}_i, \vec{g}_i).$$

Уточним формулу $\gamma_i = (\vec{r}^{\,k}, A\vec{g}_i)/(A\vec{g}_i, \vec{g}_i)$:

так как $\vec{r}^{\,k} \perp L_k$ и $A\vec{g}_i \in L_{i+1}$

то $(A\vec{r}^{\,k}, \vec{g}_i) = (\vec{r}^{\,k}, A\vec{g}_i) = 0 \quad \Rightarrow \quad \begin{cases} \gamma_i = 0, \; i = \overline{1, k-1} \\ \gamma_k = (A\vec{r}^{\,k}, \vec{g}_k)/(A\vec{g}_k, \vec{g}_k) \end{cases}$

$\Rightarrow \quad \vec{g}_{k+1} = \vec{r}^{\,k} - \gamma_k \vec{g}_k, \; \gamma_k = (A\vec{r}^{\,k}, \vec{g}_k)/(A\vec{g}_k, \vec{g}_k),$

Далее,

$$\alpha_{k+1} = \frac{(\vec{r}^{\,0}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})} \stackrel{?}{=} \frac{(\vec{r}^{\,k}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})},$$

$$\begin{cases} \vec{z}^{\,k+1} = \vec{z}^{\,0} - \alpha_1 \cdot \vec{g}_1 - \ldots - \alpha_{k+1} \cdot \vec{g}_{k+1} = \vec{z}^{\,k} - \alpha_{k+1} \cdot \vec{g}_{k+1} \\ \vec{r}^{\,k+1} = A\vec{z}^{\,k+1} = \vec{r}^{\,0} - \alpha_1 \cdot A\vec{g}_1 - \ldots - \alpha_{k+1} \cdot A\vec{g}_{k+1} \end{cases}$$

Утверждение 3. $\vec{r}^{\,k+1} \in L_{k+2}, \quad \vec{r}^{\,k+1} \perp L_{k+1}.$

1. $$\vec{r}^{\,k+1} = A\vec{z}^{\,k+1} = \underbrace{\vec{r}^{\,0} - \alpha_1 \cdot A\vec{g}_1 - \ldots - \alpha_k \cdot A\vec{g}_k}_{\vec{r}^{\,k} \in L_{k+1}} - \alpha_{k+1} \cdot \underbrace{A\vec{g}_{k+1}}_{\in AL_{k+1} = L_{k+2}} \in L_{k+2}$$

2. $$\begin{cases} (\vec{r}^{\,k+1}, \vec{g}_j) = \underbrace{\vec{g}_1 - \alpha_1 \cdot A\vec{g}_1 - \ldots - \alpha_k \cdot A\vec{g}_k}_{\vec{r}^{\,k} \in L_{k+1}} - \alpha_{k+1} \cdot \underbrace{A\vec{g}_{k+1}}_{\in AL_{k+1} = L_{k+2}} = \\ = \underbrace{(\vec{r}^{\,k}, \vec{g}_j)}_{=0\ \forall\ j \le k} - \alpha_{k+1} \cdot \underbrace{(A\vec{g}_{k+1}, \vec{g}_j)}_{=0\ j \le k} \end{cases}$$

$$\begin{cases} (\vec{r}^{\,k+1}, \vec{g}_{k+1}) = (\vec{r}^{\,k}, \vec{g}_{k+1}) - \alpha_{k+1} \cdot (A\vec{g}_{k+1}, \vec{g}_{k+1}) \equiv \\ \equiv (\vec{r}^{\,k}, \vec{g}_{k+1}) - \dfrac{(\vec{r}^{\,k}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})} \cdot (A\vec{g}_{k+1}, \vec{g}_{k+1}) = 0 \end{cases}$$

$\Rightarrow \ \vec{g}_{k+1}$ построен.

Замечание. Мы использовали формулу:

$$\alpha_{k+1} = \frac{(\vec{r}^{\,0}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})} \equiv \frac{(\vec{r}^{\,k}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})}.$$

Доказательство.

$$\alpha_{k+1} = \frac{(\vec{r}^{\,0}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})} = \frac{(\vec{r}^{\,1} + \alpha_1 A\vec{g}_1, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, g_{\vec{k}+1})}$$

$$= \frac{(\vec{r}^{\,1}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})} = \ldots = \frac{(\vec{r}^{\,k}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})}$$

Итак,

$$\gamma_k = (A\vec{r}^{\,k}, \vec{g}_k) / (A\vec{g}_k, \vec{g}_k) \Rightarrow \begin{cases} \vec{g}_{k+1} = \vec{r}^{\,k} - \gamma_k \vec{g}_k \in L_{k+1} \\ \alpha_{k+1} = \dfrac{(\vec{r}^{\,k}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})} \end{cases}$$

Формулы метода сопряженных градиентов

Итак, для решения системы $A\vec{x}=\vec{b}$ $(A=A^*>0)$ построен метод:

$$\vec{x}^0 \text{ задан}, \quad \vec{r}^0 = A\vec{x}^0-\vec{b}$$

$$\vec{g}_1=\vec{r}^0 \;\Rightarrow\; \alpha_1=\frac{(\vec{r}^0,\vec{g}_1)}{(A\vec{g}_1,\vec{g}_1)} \;\Rightarrow\; \begin{cases}\vec{x}^1=\vec{x}^0-\alpha_1\vec{g}_1\\ \vec{r}^1=\vec{r}^0-\alpha_1 A\vec{g}_1\end{cases}$$

$$\begin{cases}\gamma_k=\dfrac{(A\vec{r}^k,\vec{g}_k)}{(A\vec{g}_k,\vec{g}_k)} \;\Rightarrow\; \vec{g}_{k+1}=\vec{r}^k-\gamma_k\vec{g}_k \;\Rightarrow\; \alpha_{k+1}=\dfrac{(\vec{r}^k,\vec{g}_{k+1})}{(A\vec{g}_{k+1},\vec{g}_{k+1})}\\ \Rightarrow\; \begin{cases}\vec{x}^{k+1}=\vec{x}^k-\alpha_{k+1}\vec{g}_{k+1}\\ \vec{r}^{k+1}=\vec{r}^k-\alpha_{k+1}A\vec{g}_{k+1}\end{cases}\end{cases}$$

$$k=1,\ 2,\ \ldots$$

– метод сопряженных градиентов.

Теорема.

Если $A = A^* > 0$, то метод сопряженных градиентов продолжается до получения решения системы $A\vec{x} = \vec{b}$ за $m = k(\vec{b}) \le n$ итераций (пока $\vec{r}^{\,k} \ne 0$) и $\forall\ k$

$$\| \vec{z}^{\,k} \|_A \le \frac{2\gamma^k}{1+\gamma^{2k}} \| \vec{z}^{\,0} \|_A, \qquad \gamma = \frac{\sqrt{\lambda_{\max}} - \sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}} + \sqrt{\lambda_{\min}}}.$$

$\exists\ k(\vec{b}) \le n\colon\ \vec{z}^{\,k(\vec{b})} = 0$

Доказать теорему в качестве упражнения.

Замечание. В формулах метода сопряженных градиентов нет констант $\lambda_{\min}(A) \le \alpha < \beta \le \lambda_{\max}(A)$, но есть скалярные произведения: $(A\vec{r}^{\,k}, \vec{g}_k)$, $(\vec{r}^{\,k}, \vec{g}_{k+1})$, $(A\vec{g}_{k+1}, \vec{g}_{k+1})$ – это цена незнания констант: $0 < \alpha < \beta$.

Трехслойные формулы метода сопряженных градиентов

Утверждение 4. Двухслойные формулы метода сопряженных градиентов:

$$\vec{x}^1 = \vec{x}^0 - \alpha_1 \cdot \vec{r}^{\,0},\ \alpha_1 = (\vec{r}^{\,0}, \vec{r}^{\,0}) / (A\vec{r}^{\,0}, \vec{r}^{\,0});$$

$$\begin{cases} \vec{g}_{k+1} = \vec{r}^{\,k} - \gamma_k \vec{g}_k \\ \vec{x}^{k+1} = \vec{x}^k - \alpha_{k+1} \vec{g}_{k+1} \end{cases}$$

можно преобразовать:

$$\vec{x}^1 = \vec{x}^0 - \alpha_1 \cdot \vec{r}^{\,0},\ \alpha_1 = (\vec{r}^{\,0}, \vec{r}^{\,0}) / (A\vec{r}^{\,0}, \vec{r}^{\,0});$$

$$\begin{cases} \vec{x}^{k+1} = \vec{x}^k - \alpha_{k+1} \vec{g}_{k+1} \equiv \vec{x}^k - \alpha_{k+1}(\vec{r}^{\,k} - \gamma_k \underbrace{\vec{g}_k}_{-(\vec{x}^k - \vec{x}^{k-1})/\alpha_k}) = \\ \qquad = \vec{x}^k - \alpha_{k+1}\vec{r}^{\,k} - \beta_{k+1}(\vec{x}^k - \vec{x}^{k-1}) \end{cases}$$

– трехслойные формулы метода сопряженных градиентов.

Упражнение. Постройте формулы для невязок метода сопряженных градиентов:

$$\begin{cases} \vec{x}^1 = \vec{x}^0 - \alpha_1 \cdot \vec{r}^0, \;\; \alpha_1 = (\vec{r}^0, \vec{r}^0) / (A\vec{r}^0, \vec{r}^0); \\ \vec{x}^{k+1} = \vec{x}^k - \alpha_{k+1}\vec{r}^k - \beta_{k+1}(\vec{x}^k - \vec{x}^{k-1}), \;\; k \geq 1 \end{cases}$$

Ответ:

$$\begin{cases} \vec{r}^0 = A\vec{x}^0 - \vec{b} \\ \vec{r}^1 = \vec{r}^0 - \alpha_1 \cdot A\vec{r}^0, \;\; \alpha_1 = (\vec{r}^0, \vec{r}^0) / (A\vec{r}^0, \vec{r}^0); \\ \vec{r}^{k+1} = \vec{r}^k - \alpha_{k+1}A\vec{r}^k - \beta_{k+1}(\vec{r}^k - \vec{r}^{k-1}), \;\; k \geq 1 \end{cases}$$

Утверждение 5. Параметры α_{k+1} и β_{k+1} можно определять из условий ортогональности невязки $\vec{r}^{k+1}$ невязкам $\vec{r}^{k-1}$ и $\vec{r}^k$:

$$\begin{cases} (\vec{r}^{k+1}, \vec{r}^{k-1}) \equiv (\vec{r}^k, \vec{r}^{k-1}) - \beta_{k+1} \cdot (\vec{r}^k - \vec{r}^{k-1}, \vec{r}^{k-1}) - \alpha_{k+1} \cdot (A\vec{r}^k, \vec{r}^{k-1}) = 0 \\ (\vec{r}^{k+1}, \vec{r}^k) \equiv (\vec{r}^k, \vec{r}^k) - \beta_{k+1} \cdot (\vec{r}^k - \vec{r}^{k-1}, \vec{r}^k) - \alpha_{k+1} \cdot (A\vec{r}^k, \vec{r}^k) = 0 \end{cases}$$

Упражнение. Постройте формулы для ошибок метода сопряженных градиентов:

$$\begin{cases} \vec{x}^1 = \vec{x}^0 - \alpha_1 \cdot \vec{r}^{\,0}, \;\; \alpha_1 = (\vec{r}^{\,0}, \vec{r}^{\,0}) / (A\vec{r}^{\,0}, \vec{r}^{\,0}); \\ \vec{x}^{k+1} = \vec{x}^k - \alpha_{k+1}\vec{r}^{\,k} - \beta_{k+1}(\vec{x}^k - \vec{x}^{k-1}), \;\; k \geq 1 \end{cases}$$

Ответ:

$$\begin{cases} \vec{z}^{\,0} = \vec{x}^0 - \vec{x} \\ \vec{z}^{\,1} = \vec{z}^{\,0} - \alpha_1 \cdot A\vec{z}^{\,0}, \;\; \alpha_1 = (\vec{r}^{\,0}, \vec{r}^{\,0}) / (A\vec{r}^{\,0}, \vec{r}^{\,0}); \\ \vec{z}^{\,k+1} = \vec{z}^{\,k} - \alpha_{k+1}A\vec{z}^{\,k} - \beta_{k+1}(\vec{z}^{\,k} - \vec{z}^{\,k-1}), \;\; k \geq 1 \end{cases}$$

Утверждение* 6. Параметры α_{k+1} и β_{k+1} можно определять из условия минимизации A-нормы ошибки:

$$\| \vec{z}^{\,k+1} \|_A^2 = \min_{\alpha,\beta} (A[\vec{z}^{\,k} - \beta \cdot (\vec{z}^{\,k} - \vec{z}^{\,k-1}) - \alpha \cdot \vec{r}^{\,k}], \vec{z}^{\,k} - \beta \cdot (\vec{z}^{\,k} - \vec{z}^{\,k-1}) - \alpha \cdot \vec{r}^{\,k}).$$

Упражнение. Решить задачу:

$$\vec{x}^{k-1},\ \vec{x}^{k},\ \vec{r}^{k} = A\vec{x}^{k} - \vec{b} \quad - \quad \textit{заданы}$$

вычислить $\alpha_{k+1},\ \beta_{k+1}$:

$$\begin{cases} (A[\underbrace{\vec{z}^{k} - \beta_{k+1}\cdot(\vec{z}^{k} - \vec{z}^{k-1}) - \alpha_{k+1}\cdot\vec{r}^{k}}_{\vec{z}^{k+1}}], \underbrace{\vec{z}^{k} - \beta_{k+1}\cdot(\vec{z}^{k} - \vec{z}^{k-1}) - \alpha_{k+1}\cdot\vec{r}^{k}}_{\vec{z}^{k+1}}) = \\ \qquad = \min\limits_{\alpha,\beta}(A[\vec{z}^{k} - \beta\cdot(\vec{z}^{k} - \vec{z}^{k-1}) - \alpha\cdot\vec{r}^{k}], \vec{z}^{k} - \beta\cdot(\vec{z}^{k} - \vec{z}^{k-1}) - \alpha\cdot\vec{r}^{k}) \end{cases}$$

Ответ: $\alpha_{k+1},\ \beta_{k+1}$ – решение системы:

$$\begin{cases} \beta_{k+1}\cdot(r^{k} - r^{k-1}, x^{k} - x^{k-1}) + \alpha_{k+1}\cdot(r^{k} - r^{k-1}, r^{k}) = (r^{k}, x^{k} - x^{k-1}), \\ \beta_{k+1}\cdot(r^{k} - r^{k-1}, r^{k}) \qquad + \alpha_{k+1}\cdot(Ar^{k}, r^{k}) \qquad = (r^{k}, r^{k}). \end{cases}$$

Тема 16. Переобусловливатель (preconditionary).

Если систему уравнений

$$A\vec{x} = \vec{b}$$

преобразовать к системе уравнений

$$B^{-1}A\vec{x} = B^{-1}\vec{b},$$

то обусловленность

$cond(B^{-1}A)$ матрицы $B^{-1}A$ новой системы

может оказаться значительно меньше

$cond(A)$ – обусловленности матрицы исходной системы

и влияние ошибок округления на решение системы уменьшится.

Матрица B называется переобусловливателем для матрицы A.

Определение. Положительно определенные матрицы A и B называются эквивалентными по спектру с постоянными $\gamma_1 \geq \gamma_0 > 0$, если

$$\gamma_0(B\vec{v},\vec{v}) \leq (A\vec{v},\vec{v}) \leq \gamma_1(B\vec{v},\vec{v}) \quad \forall\ \vec{v} \in R^n.$$

Теорема. **Если** $A = A^* > 0$ **и** $B = B^* > 0$ **и** $\gamma_0 B \leq A \leq \gamma_1 B$**:**

$$\gamma_0(B\vec{v},\vec{v}) \leq (A\vec{v},\vec{v}) \leq \gamma_1(B\vec{v},\vec{v}) \quad \forall\ \vec{v} \in R^n,$$

то $\forall\ \lambda(B^{-1}A) \in [\gamma_0, \gamma_1]$**.**

Доказательство. $B = B^* > 0 \Rightarrow \exists\ B^{1/2} = (B^{1/2})^* > 0 \Rightarrow$

$$Sp(B^{-1}A) = Sp(B^{+1/2}[B^{-1}A]B^{-1/2}) = Sp(\underbrace{B^{-1/2}AB^{-1/2}}_{C=C^*>0 \Rightarrow \forall\ \lambda(C)\equiv\lambda(B^{-1}A)>0})$$

$$B^{-1}A\vec{y} = \lambda(B^{-1}A)\cdot\vec{y} \Rightarrow$$

$$\Rightarrow\ \lambda(B^{-1}A) = \frac{(A\vec{y},\vec{y})}{(B\vec{y},\vec{y})} \ \Rightarrow\ \gamma_0 \leq \lambda(B^{-1}A) = \frac{(A\vec{y},\vec{y})}{(B\vec{y},\vec{y})} \leq \gamma_1.$$

Метод простой итерации с переобусловливателем

Теорема. **Если** $A = A^* > 0 \;\; \& \;\; B = B^* > 0$**, то метод простой итерации**

$$\begin{cases} \vec{x}^0 - \text{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - \tau \cdot B^{-1}(A\vec{x}^k - \vec{b}), \quad k = 0, 1, ..., \end{cases}$$

сходится $\forall \;\; \tau \in (0,\ 2/\gamma_1)$**,**

и справедливы оценки ($S_\tau = E - \tau \cdot A$**):**

$$\| \vec{z}^{k+1} \|_A \le \| S_\tau \|_A \cdot \| \vec{z}^k \|_A \le \rho_\tau \cdot \| \vec{z}^k \|_A,$$

$$\| \vec{z}^{k+1} \|_B \le \| S_\tau \|_B \cdot \| \vec{z}^k \|_B \le \rho_\tau \cdot \| \vec{z}^k \|_B,$$

$$\rho_\tau = \max\{|1 - \tau \cdot \gamma_0|, |1 - \tau \cdot \gamma_1|\} < 1$$

при оптимальном параметре $\tau = \tau_{опт} = 2/(\gamma_1 + \gamma_0)$ имеем $\rho_{\tau_{опт}} = \dfrac{\gamma_1 - \gamma_0}{\gamma_1 + \gamma_0}$**.**

Доказательство. Мы знаем, что, если все собственные значения матрицы $B^{-1}A$ вещественны и положительны, то условием сходимости метода простой итерации является выбор параметра τ из интервала $(0,\ 2/\rho(B^{-1}A))$, а спектральный радиус $\rho(S_\tau)$ оценивается сверху величиной $\rho_\tau = \max\{|1-\tau\cdot\gamma_0|, |1-\tau\cdot\gamma_1|\}$, где γ_0 и γ_1 оценки спектра матрицы $B^{-1}A$.

Следовательно, метод простой итерации сходится при любом $\tau \in (0,\ 2/\rho(B^{-1}A))$

и ρ_τ минимально, если $\tau = \tau_{опт} = 2/(\gamma_1+\gamma_0)$.

Для доказательства оценок для ошибки $z^{k+1} = S_\tau z^k$ достаточно установить равенство энергетических A- и B-норм матрицы перехода $S_\tau \equiv E - \tau B^{-1}A$ ее спектральному радиусу:

1. $\| S_\tau \|_A = \rho(S_\tau)$:

$$\| S_\tau \|_A^2 = \inf_{\vec{z}\neq 0} \frac{(A(E-\tau B^{-1}A)\vec{z}, (E-\tau B^{-1}A)\vec{z})}{(A\vec{z}, \vec{z})} =$$

$$= \inf_{\vec{v}=A^{1/2}\vec{z}\neq 0} \frac{((E-\tau A^{1/2}B^{-1}A^{1/2})\vec{v}, (E-\tau A^{1/2}B^{-1}A^{1/2})\vec{v})}{(\vec{v}, \vec{v})} =$$

$$= \| E-\tau A^{1/2}B^{-1}A^{1/2} \|_2^2 = \rho^2(E-\tau A^{1/2}B^{-1}A^{1/2}) =$$

$$= \rho^2(A^{1/2}S_\tau A^{-1/2}) = \rho^2(S_\tau)$$

2. $\| S_\tau \|_A = \rho(S_\tau)$:

$$\| S_\tau \|_B^2 = \inf_{z \neq 0} \frac{(B(E-\tau B^{-1}A)\vec{z}, (E-\tau B^{-1}A)\vec{z})}{(B\vec{z}, \vec{z})} =$$

$$= \inf_{\vec{v}=B^{1/2}\vec{z} \neq 0} \frac{((E-\tau B^{-1/2}AB^{-1/2})\vec{v}, (E-\tau B^{-1/2}AB^{-1/2})\vec{v})}{(\vec{v}, \vec{v})} =$$

$$= \| E-\tau B^{-1/2}AB^{-1/2} \|_2^2 = \rho^2(E-\tau B^{-1/2}AB^{-1/2}) =$$

$$= \rho^2(B^{1/2}S_\tau B^{-1/2}) = \rho^2(S_\tau),$$

Теорема Самарского.

Теорема. Если $A = A^* > 0 \;\;\&\;\; B > 0.5\tau A \quad (\tau > 0)$,

то метод простой итерации

$$\begin{cases} \vec{x}^0 - \textit{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - \tau \cdot B^{-1}(A\vec{x}^k - \vec{b}), \quad k = 0, 1, \ldots, \end{cases}$$

сходится.

Доказательство. Энергетическая норма ошибки $\vec{z}^{k+1} = S_\tau \vec{z}^k$ убывает:

$$\| \vec{z}^{k+1} \|_A^2 \equiv (A\vec{z}^{k+1}, \vec{z}^{k+1}) = (A\vec{z}^k, \vec{z}^k) - 2\tau(A \underbrace{B^{-1}A\vec{z}^k}_{\vec{w}^k}, \vec{z}^k) +$$

$$+ \tau^2 (A \underbrace{B^{-1}A\vec{z}^k}_{\vec{w}^k}, \underbrace{B^{-1}A\vec{z}^k}_{\vec{w}^k}) =$$

$$= (A\vec{z}^k, \vec{z}^k) - 2\tau([B - 0.5\tau A]\vec{w}^k, \vec{w}^k) < \| z^k \|_A^2$$

и, т.к. оператор S_τ непрерывен, то итерационный процесс сходится.

Метод наискорейшего спуска с переобусловливателем

На каждом шаге метода простой итерации параметр τ можно выбирать из условия минимизации энергетической нормы ошибки $\vec{z}^{k+1} = S_\tau \vec{z}^k$.

Теорема. Если $A = A^* > 0 \;\; \& \;\; B > 0$, то итерационный метод

$$\begin{cases} \vec{x}^0 - \textit{задан}, \\ \vec{x}^{k+1} = \vec{x}^k - \tau_{k+1} \cdot B^{-1}(A\vec{x}^k - \vec{b}), \quad \tau_{k+1} = \dfrac{(B^{-1}\vec{r}^k, \vec{r}^k)}{(AB^{-1}\vec{r}^k, B^{-1}\vec{r}^k)}, \\ \qquad k = 0, 1, ..., \end{cases}$$

сходится.

Если, кроме всего прочего, $B = B^*$ и известны границы γ_0 и γ_1 спектра матрицы $B^{-1}A$, то

$$\| \vec{z}^{k+1} \|_A \le \rho_{\tau_{опт}} \cdot \| \vec{z}^k \|_A \le (\rho_{\tau_{опт}})^{k+1} \cdot \| \vec{z}^0 \|_A, \; \rho_{\tau_{опт}} = \frac{\gamma_1 - \gamma_0}{\gamma_1 + \gamma_0}.$$

Доказательство. Очевидно, что минимум энергетической нормы

$$\| \vec{z}^{k+1} \|_A^2 = \inf_{\tau} [(A\vec{z}^k, \vec{z}^k) - 2\tau \cdot (AB^{-1}A\vec{z}^k, \vec{z}^k) + \tau^2 (AB^{-1}A\vec{z}^k, B^{-1}A\vec{z}^k)] =$$

$$= (A\vec{z}^k, \vec{z}^k) - \frac{(B^{-1}\vec{r}^k, \vec{r}^k)^2}{(AB^{-1}\vec{r}^k, B^{-1}\vec{r}^k)} \ \text{при} \ \tau = \tau_{k+1} = \frac{(B^{-1}\vec{r}^k, \vec{r}^k)}{(AB^{-1}\vec{r}^k, B^{-1}\vec{r}^k)}$$

и норма ошибки строго убывает, так как

$$(B^{-1}\vec{r}^k, \vec{r}^k) \equiv ([B^{-1}\vec{r}^k], B[B^{-1}\vec{r}^k]) > 0.$$

Так как оператор $S_{\tau_{k+1}}\vec{z}^k = \vec{z}^k - \tau_{k+1}(\vec{z}^k) \cdot B^{-1}A\vec{z}^k$ непрерывен всюду (кроме $\vec{z} = 0$), то итерационный процесс сходится.

Оценка нормы ошибки при $B = B^* > 0$ устанавливается элементарно:

$$\| \vec{z}^{k+1} \|_A = \inf_{\tau} \| S_\tau \vec{z}^k \|_A \le \| S_{\tau_{опт}} \vec{z}^k \|_A \le \| S_{\tau_{опт}} \|_A \cdot \| \vec{z}^k \|_A \le \rho_{\tau_{опт}} \cdot \| \vec{z}^k \|_A,$$

что и требовалось доказать.

Метод сопряженных градиентов с переобусловливателем

Напомним, что метод сопряженных градиентов для решения системы

$$A\vec{x} = \vec{b},\ A = A^* > 0,$$

был построен из условия минимизации энергетической нормы ошибки за m шагов.

Построим метод сопряженных градиентов для решения системы

$$B^{-1}A\vec{x} = B^{-1}\vec{b},\ A = A^* > 0,\ B = B^* > 0,$$

т.е. решим задачу построения ошибки $\vec{z}^m$ такой, что

$$\begin{gathered} \| \vec{z}^m \|_A = \min_{\tau} \| (E - \tau_m B^{-1}A)\ldots(E - \tau_1 B^{-1}A)\vec{z}^0 \|_A = \\ = \min_{\alpha} \| z^0 - \alpha_1 \vec{g}_1 - \ldots - \alpha_m \vec{g}_m \|_A \end{gathered},$$

где $L_m = L\{B^{-1}A\vec{z}^0, \ldots, (B^{-1}A)^m z^{-0}\} \equiv L\{\vec{g}_1, \ldots, \vec{g}_m\}$, а систему векторов $\{\vec{g}_i\}_{i=1}^m$ нужно построить.

Как и ранее, параметры α_1, ..., α_m удовлетворяют системе уравнений

$$\frac{1}{2}\frac{\partial(A\vec{z}^m,\vec{z}^m)}{\partial\alpha_i}=(A\frac{\partial\vec{z}^m}{\partial\alpha_i},z^m)\equiv(A\vec{g}_i,\ \vec{z}^0-\alpha_1\vec{g}_1-\ldots-\alpha_m\vec{g}_m)=0,\quad i=1,\ldots,m$$

,

или

$$\begin{bmatrix} (A\vec{g}_1,\vec{g}_1) & (A\vec{g}_1,\vec{g}_2) & \ldots & (A\vec{g}_1,\vec{g}_m) \\ (A\vec{g}_2,\vec{g}_1) & (A\vec{g}_2,\vec{g}_2) & \ldots & (A\vec{g}_2,\vec{g}_m) \\ \vdots & \vdots & \ldots & \vdots \\ (A\vec{g}_m,\vec{g}_1) & (A\vec{g}_m,\vec{g}_2) & \ldots & (A\vec{g}_m,\vec{g}_m) \end{bmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_m\end{pmatrix}=\begin{pmatrix}(A\vec{z}^0,\vec{g}_1)\\ (A\vec{z}^0,\vec{g}_2)\\ \vdots\\ (A\vec{z}^0,\vec{g}_m)\end{pmatrix}.$$

Матрица этой системы – матрица Грамма в скалярном произведении $(\vec{x},\vec{y})_A\equiv(A\vec{x},\vec{y})$ базиса $\{\vec{g}_i\}_{i=1}^m$ в L_m, ее определитель не равен нулю, следовательно, решение α существует и единственно.

Если базис $\{\vec{g}_i\}_{i=1}^m$ является A–ортогональным, т.е. $(A\vec{g}_i, \vec{g}_j) = 0, \quad i \neq j$, то

$$\alpha_k = \frac{(A\vec{z}^{\,0}, \vec{g}_k)}{(A\vec{g}_k, \vec{g}_k)} \equiv \frac{(\vec{r}^{\,0}, \vec{g}_k)}{(A\vec{g}_k, \vec{g}_k)}, \qquad \begin{cases} \vec{z}^{\,m} = \vec{z}^{\,0} - \alpha_1\vec{g}_1 - \alpha_2\vec{g}_2 - \ldots - \alpha_m\vec{g}_m, \\ \vec{x}^{\,m} = \vec{x}^{\,0} - \alpha_1\vec{g}_1 - \alpha_2\vec{g}_2 - \ldots - \alpha_m\vec{g}_m. \end{cases}$$

Построение A–ортогонального базиса

A–ортогонализуем систему векторов $\{B^{-1}A\vec{z}^{\,0}, \ldots, (B^{-1}A)^m \vec{z}^{\,0}\}$.

Зададим $\vec{g}_1 = B^{-1}A\vec{z}^{\,0} \equiv B^{-1}\vec{r}^{\,0} \in L_1$, тогда $\alpha_1 = \dfrac{(\vec{r}^{\,0}, \vec{g}_1)}{(A\vec{g}_1, \vec{g}_1)}$, $\vec{z}^{\,1} = \vec{z}^{\,0} - \alpha_1\vec{g}_1$.

Вычислим вектор $\vec{r}^{\,1} = A\vec{z}^{\,1} = A\vec{z}^{\,0} - \alpha_1 A\vec{g}_1 = A\vec{z}^{\,0} - \alpha_1 A(B^{-1}A\vec{z}^{\,0})$.

Легко проверить, что $B^{-1}\vec{r}^{\,1} = B^{-1}A\vec{z}^{\,0} - \alpha_1(B^{-1}A)^2 \vec{z}^{\,0} \in L_2$ и $(\vec{r}^{\,1}, \vec{g}_1) = 0$, т.е. $\vec{r}^{\,1} \perp L_1$ относительно обычного скалярного произведения.

Далее применяем метод математической индукции.

Предположим, что мы построили A-ортогональный базис $\{\vec{g}_i\}_{i=1}^{k}$ в

$$L_k = L\{(B^{-1}A)\vec{z}^{\,0}, \ldots, (B^{-1}A)^k\,\vec{z}^{\,0}\}$$ такой, что

1. $\{\vec{g}_i\}_{i=1}^{j}$ – A-ортогональный базис в $L_j = L\{(B^{-1}A)\vec{z}^{\,0}, \ldots, (B^{-1}A)^j\,\vec{z}^{\,0}\}$, $j \le k$,

2. $$\begin{cases} \vec{z}^{\,j} = \vec{z}^{\,0} - \alpha_1 \cdot \vec{g}_1 - \ldots - \alpha_j \cdot \vec{g}_j \\ \vec{x}^{\,j} = \vec{x}^{\,0} - \alpha_1 \cdot \vec{g}_1 - \ldots - \alpha_j \cdot \vec{g}_j \end{cases}, \ \alpha_i = \frac{(\vec{r}^{\,0}, \vec{g}_i)}{(A\vec{g}_i, \vec{g}_i)},$$

3. $r^j = Az^j$: $B^{-1}r^j \in L_{j+1}$ и $r^j \perp L_j$, т.е. $(r^j, g_i) = 0 \ \ \forall\, i = \overline{1,\ j}$.

Определим $\vec{g}_{k+1} = B^{-1}\vec{r}^{\,k} - \gamma_1\vec{g}_1 - \ldots - \gamma_k\vec{g}_k \in L_{k+1}$ из условий

$$(A\vec{g}_{k+1}, \vec{g}_i) \equiv (AB^{-1}\vec{r}^{\,k}, \vec{g}_i) - \gamma_1 \cdot (A\vec{g}_1, \vec{g}_i) - \ldots - \gamma_k \cdot (A\vec{g}_k, \vec{g}_i) = 0,$$

$$i = 1, \ldots, k.$$

Так как $(A\vec{g}_j, \vec{g}_i) = 0,\ i \neq j$, то $\gamma_i = (AB^{-1}\vec{r}^{\,k}, \vec{g}_i) / (A\vec{g}_i, \vec{g}_i)$,

так как $\vec{r}^{\,k} \perp L_k$ и $B^{-1}A\vec{g}_i \in L_{i+1} \subset L_k$ при $i+1 \leq k$,

то $(AB^{-1}\vec{r}^{\,k}, \vec{g}_i) = (\vec{r}^{\,k}, B^{-1}A\vec{g}_i) = 0$ и $\gamma_i = 0,\ i = \overline{1, k-1}$,

(здесь потребовалась симметричность матрицы B^{-1}).

Тогда

1. $\vec{g}_{k+1} = B^{-1}\vec{r}^{\,k} - \gamma_k \vec{g}_k,\ \gamma_k = (B^{-1}\vec{r}^{\,k}, A\vec{g}_k) / (A\vec{g}_k, \vec{g}_k)$

2. $\vec{z}^{\,k+1} = \vec{z}^{\,k} - \alpha_{k+1} \cdot \vec{g}_{k+1}, \quad \alpha_{k+1} = \dfrac{(\vec{r}^{\,0}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})}$,

3. $\vec{r}^{\,k+1} = A\vec{z}^{\,k+1} = \vec{r}^{\,0} - \alpha_1 \cdot A\vec{g}_1 - \ldots - \alpha_{k+1} \cdot A\vec{g}_{k+1} \perp L_{k+1}$,

   т.к. легко проверяются равенства $(\vec{r}^{\,k+1}, \vec{g}_i) = 0,\ i = \overline{1, k+1}$;

   очевидно, что

$$B^{-1}\vec{r}^{\,k+1} = B^{-1}A\vec{z}^{\,k+1} \in L\{\ B^{-1}\vec{r}^{\,0},\ L\{(B^{-1}A)^2\,\vec{z}^{\,0},\ \ldots,\ (B^{-1}A)^{k+2}\,\vec{z}^{\,0}\}\ \} \equiv L_{k+2}.$$

Базисный вектор $\vec{g}_{k+1}$ построен и все предположения метода математической индукции для него выполняются, из которых следует:

$$\alpha_{k+1} = \frac{(\vec{r}^{\,0}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})} = \frac{(\vec{r}^{\,1} + \alpha_1 A\vec{g}_1, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})} = \frac{(\vec{r}^{\,1}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})} =$$

$$= \ldots = \frac{(\vec{r}^{\,k}, \vec{g}_{k+1})}{(A\vec{g}_{k+1}, \vec{g}_{k+1})}$$

Формулы метода сопряженных градиентов с переобусловливателем

$$\vec{g}_1 = B^{-1}\vec{r}^{\,0} \equiv B^{-1}(A\vec{x}^{\,0} - \vec{b},)$$

$$\vec{x}^{\,1} = \vec{x}^{\,0} - \alpha_1 \cdot \vec{g}_1, \qquad \alpha_1 = (\vec{r}^{\,0}, \vec{g}_1) / (A\vec{g}_1, \vec{g}_1)$$

$$\begin{cases} \vec{g}_{k+1} = B^{-1}\vec{r}^{\,k} - \gamma_k \vec{g}_k, & \gamma_k = (B^{-1}\vec{r}^{\,k}, A\vec{g}_k) / (A\vec{g}_k, \vec{g}_k) \\ \vec{x}^{\,k+1} = \vec{x}^{\,k} - \alpha_{k+1} \cdot \vec{g}_{k+1}, & \alpha_{k+1} = (\vec{r}^{\,k}, \vec{g}_{k+1}) / (A\vec{g}_{k+1}, \vec{g}_{k+1}), \quad k > 0 \end{cases}$$