Лекция 13

Ilya Yaroshevskiy

13 мая 2023 г.

Содержание

1 Число обусловленности

Пример.

$$Ax = b$$

$$A = \begin{pmatrix} 9.7 & 6.6 \\ 4.1 & 2.8 \end{pmatrix} \quad b = \begin{pmatrix} 9.7 \\ 4.1 \end{pmatrix} \quad x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\|b\| = 13.8 \quad \|x\| = 1$$

$$\operatorname{cond}(A) = 2249.4$$

$$b' = \begin{pmatrix} 9.70 \\ 4.11 \end{pmatrix} \quad x' = \begin{pmatrix} 0.34 \\ 0.97 \end{pmatrix}$$

$$\Delta b = b - b' \quad \|\Delta b\| = 0.01$$

$$\Delta x = x - x' \quad \|\Delta x\| = 1.63$$

$$\frac{\|\Delta b\|}{\|b\|} = 0.00072464$$

$$\frac{\|\Delta x\|}{\|x\|} = 1.63$$

1.1 Нормы и анализ ошибок

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

$$\|Ax\| \le \|A\| \cdot \|x\|$$

$$\tilde{x} : \|Ax\| = \|A\| \cdot \|\tilde{x}\|$$

$$\|A = M = \max_{x \ne 0} \frac{\|Ax\|}{\|x\|}$$

$$\|a\| =_j \|a_j\|$$

Результат Уилкинсона

$$x^* : (A + E)x^* = b$$

, где элементы E имеют уровень ошибок округления

$$b - Ax^* = Ex^*$$

$$\|b - Ax^*\| = \|Ex^*\| \leq \|E\| \cdot \|x^*\|$$

$$\frac{\|b - Ax^*\|}{\|A\| \cdot \|x^*\|} \leq C \cdot \varepsilon_{\text{маш.}}$$

Если A не вырождена то: $x - x^* = A^{-1} \cdot (b - Ax^*)$

$$\|x - x^*\| \leq \left\|A^{-1}\right\| \cdot \|E\| \cdot \|x^*\|$$

$$\frac{\|x - x^*\|}{\|x^*\|} \leq C \cdot \|A\| \cdot \left\|A^{-1}\right\| \cdot \varepsilon_{\text{маш.}}$$

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

$$\left\|A^{-1}\right\| = \frac{1}{m} \quad \frac{\|x - x^*\|}{\|x^*\|} \leq C \cdot \text{cond}(A) \cdot \varepsilon_{\text{маш.}}$$

- a_j — столбцы A
- $\tilde{a}_j$ — столбцы A^{-1}

$$\text{cond}(A) = \max_j \|a_j\| \cdot \max_j \|\tilde{a}_j\|$$

1.2 Оценивание числа обусловленности

$$\text{cond}(A) = \frac{\max \frac{\|Ax\|}{\|x\|}}{\min \frac{\|Ax\|}{\|x\|}}$$

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

1-норма:

- a_j — столбец

$$\|A\| = \max_j \|a_j\|$$

$$\left\|A^{-1}\right\| = \frac{1}{\min_x \frac{\|Ax\|}{\|x\|}} = \max_x \frac{\|x\|}{\|Ax\|} = \max_y \frac{\left\|A^{-1}y\right\|}{\|y\|}$$

$$y \coloneqq Ax \quad Az = y$$

$$\frac{\|z\|}{\|y\|} = \underbrace{\frac{\left\|A^{-1}y\right\|}{\|y\|}}_{\text{оценка } \|A^{-1}\|}$$

Цель — подобрать y. $A^T y = c$, где c — вектор с компонентами ± 1

Пример.

$$A = \begin{pmatrix} 9.7 & 6.6 \\ 4.1 & 2.8 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0.4227 & 1 \end{pmatrix}}_{L} \cdot \underbrace{\begin{pmatrix} 9.7000 & 6.6000 \\ 0 & 0.0103 \end{pmatrix}}_{U}$$

$y : A^T y = c \implies U^T(L^T y) = c$
$c_1 = 1; c_2 = ?$

$$(L^T y)_1 = \frac{c_1}{U_{11}} = \frac{1}{9.7} = 0.1031$$

$$(L^T y)_2 = \frac{c_2 - U_{21}(L^T y)_1}{U_{22}} = \frac{\pm 1 - 6.6 \cdot 0.10031}{0.0103}$$

$(L^T y)_2$ большее, если $c_2 = -1$, значит

$$(L^T y) = (0.1031, -163)^T \quad y = (-163, 69)^T$$

$$Az = y \implies z = (12690, -18640)^T$$

$$\left\|A^{-1}\right\| \approx \frac{\|z\|}{\|y\|} = \frac{|12690| + |-18640|}{|-163| + |69|} = \frac{31330}{232} = 135.04$$

$$\|A\| = 13.8 \quad \text{cond}(A) \approx 13.8 \cdot 135.04 = 1863.6$$

2 Дополнительно о градиентных методах

$\{x_k\}$: $x^k = x^{k-1} + \alpha_k u^k \quad k \in \mathbb{N}$
$u^k \in E_n$. $(\nabla f(x), u) < 0$ — условие спуска
Как находить α_k

$$f(x^{k-1} + \alpha_k u^k) \leq (1 - \lambda_k) f(x^{k-1}) + \lambda_k \min_{\alpha \in E} f(x^{k-1} + \alpha u^k) \quad \lambda_k \in [0, 1]$$

$$f(x^{k-1} + \alpha_k u^k) \leq f(x^{k-1}) \tag{1}$$

— если это выполнено, то $\{x^k\}$ — релаксационная $\lambda_k = 0$, тогда 1, если же $\lambda_k = 1$, то для нахождения α_k^* решаем задачу одномерной оптимизации. Для случая $\lambda_k \in (0, 1)$:

$$f(x^{k-1}) - f(x^k) \geq \lambda_k (f(x^{k-1}) - f(x^{k-1} + \alpha_k^* u_k))$$

$$\omega(x) = -\nabla f(x)$$

2.1 Градиентный спуск

$$x^k = x^{k-1} + \beta_k \underbrace{\frac{\omega^k}{|\omega^k|}}_{u^k}$$

$$\beta_k = \underbrace{\alpha}_{const} |\omega^k|$$

$$x^k = x^{k-1} + \alpha \omega^k$$

В окрестности точки $\tilde{x}$, может быть $f(x^k) > f(x^{k-1})$

Теорема 2.1. Пусть $f(x)$ ограничена снизу и дифференцируема в E_n, а ее градиент удовлетворяет условию Липшица, т.е. $\forall x, y \in E_n$

$$|\nabla f(x) - \nabla f(y)| \leq L|x - y|$$

$L > 0$ — константа. Тогда $\{x^k\}$, определяемое рекурентным соотношением

$$x^k = x^{k-1} + \alpha \omega^k$$

с $\alpha \in (0, \frac{2}{L})$ является релаксационной. При этом справедлива оценка

$$f(x^k) \leq f(x^{k-1}) - \alpha \left(1 - \frac{\alpha L}{2}\right) \cdot |\nabla f(x^{k-1})|^2$$

и $|\nabla f(x^k)| \to 0$, при $k \to +\infty$

Если $f(x)$ удовлетворяет теореме, то при $\alpha = \frac{1}{L}$, $\{x^k\}$ — релаксационная последовательность и не происходит "проскакивания" стационарной точки

$$\left\|\omega^k\right\| = \left\|\nabla f(x^{k-1})\right\| \leq L|x^{k-1} - \tilde{x}|$$

$\tilde{x}$ — стационарная точка $\implies \nabla f(\tilde{x}) = 0$
При $\alpha \leq \frac{1}{L}$

$$|x^k - x^{k-1}| = \alpha|\omega^k| \leq |x^{k-1} - \tilde{x}|$$

Пусть $f(x)$ ограничена снизу, а $\{x^k\}$, при $\gamma_0 > 0$:

$$f(x^{k-1}) - f(x^k) \geq \gamma_0 |\omega^k|^2 \quad k = \overline{1, m} \tag{2}$$

$$f(x^0) - f(x^m) \geq \gamma_0 \sum_{k=1}^{m} |\nabla f(x^{k-1})|^2$$

$$\sum_{k=1}^{\infty} |\nabla f(x^{k-1})|^2$$

— знакоположительный ряд, т.е. $\nabla f(x^k) \xrightarrow[k \to \infty]{} 0$

2.2 Модификация

$$x^k = x^{k-1} + \alpha_k \omega^k \quad k \in \mathbb{N}$$

$\alpha_k > 0$

$$\varphi_k(\alpha) = f(x^{k-1} + \alpha\omega^k)$$

$$\varphi_k'(0) = (\nabla f(x), \omega^k) = -|\omega^k|^2 < 0$$

φ_k — убывает в окрестности $\alpha = 0$ до $\tilde{\alpha}_k$. $\alpha_k = \alpha_k^*$ — исчерпывающий спуск. Если $f(x)$ удовлетворяет теореме, то $\{x^k\}$ по исчерпывающему спуску удовлетворяет условиям 2:

$$\varphi_k'(\alpha) = (\nabla f(x^{k-1} + \alpha\omega^k), \omega^k)$$

$$x^k = x^{k-1} + \alpha_k^* \omega^k$$

$$(\omega^{k+1}, \omega^k) = (-\nabla f(x^k), \omega^k) = -\varphi_k'(\alpha_k^*) = 0$$

$$\left\|\omega^k\right\|^2 = (\omega^k, \omega^k - \omega^{k+1}) \leq |\omega^k| \cdot |\omega^k - \omega^{k+1}| =$$

$$= |\omega^k| \cdot |\nabla f(x^{k-1}) - \nabla f(x^k)| \leq L|\omega^k| \cdot |x^{k-1} - x^k| = L \cdot \alpha_k^* |\omega^k|^2$$

$\alpha_k^* \geq \frac{1}{L}$:

$$f(x^{k-1}) - f(x^k) \geq f(x^{k-1}) - f(\tilde{x}^k) \geq \frac{1}{2L}|\omega^k|^2$$

$$\tilde{x}^k = x^{k-1} + \underbrace{\frac{1}{L}}_{\alpha} \omega^k$$