Лекция 7

Ilya Yaroshevskiy

13 апреля 2021 г.

Содержание

1 Критерии Сильвестра **1**
1.1 Достаточный условия 1
1.2 Необходимые условия 1

2 Собственные значения **1**

3 Общие прицнипы многмерной оптимизации **1**
3.1 Выпуклые квадратичные функции 1
3.2 Принципы многмерной оптимизации 2
3.2.1 Скорость сходимости(минизирующих последовательностей) 3
3.2.2 Критерии окончания итерационного процесса 3

1 Критерии Сильвестра

1.1 Достаточный условия

1. $H(x^*) > 0$ и x^* — локальный минимум $\Leftrightarrow \Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0$
2. $H(x^*) < 0$ и x^* — локальный максимум $\Leftrightarrow \Delta_1 < 0, \Delta_2 > 0, \dots, (-1)^n \Delta_n > 0$

, где Δ_i — угловой минор

1.2 Необходимые условия

1. $H(x^*) \geq 0$ и x^* — может быть локальный минимум $\Leftrightarrow \Delta_1 \geq 0, \Delta_2 \geq 0, \dots, \Delta_n \geq 0$
2. $H(x^*) \leq 0$ и x^* — может быть локальный максимум $\Leftrightarrow \Delta_1 \leq 0, \Delta_2 \geq 0, \dots, (-1)^n \Delta_n \geq 0$

, где Δ_i — главный минор

2 Собственные значения

Определение. **Собственные значения** λ_i $(i = 1..n)$ $H(x^*)_{n\times n}$ находятся как корни характеристического уравнения $|H(x^*) - \lambda E| = 0$. Если $H(x)$ — вещественная, симметричная матрица, то λ_i — вещественные

3 Общие прицнипы многмерной оптимизации

3.1 Выпуклые квадратичные функции

$$f(x) = \frac{1}{2}ax^2 + bx + c$$

Определение. Функция вида

$$f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j + \sum_{j=1}^{n} b_jx_j + c \tag{1}$$

Называется **квадратичной функией** n **перменных**

Положим $a_{ij} = a_{ij} + a_{ji}$?? $\Rightarrow$ симметрия. матрица A

$$f(x) = \frac{1}{2}(Ax, x) + (b, x) + c$$

, где $b = (b_1, \dots b_n)^T \in E_n$ — вектор коэффицентов, $x = (x_1, \dots, x_n)^T$. x, y — скалярное произведение
Свойства квадратичных функций:

1. $\nabla f(x) = Ax + b$
$$\frac{\partial f}{\partial x_k} = \frac{\partial}{\partial x_k}\left(\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j + \sum_{j=1}^{n} b_j x_j + c\right) =$$
$$\frac{1}{2}\sum_{i=1}^{n}(a_{ik} + a_{ki})x_i + b_k = \sum_{i=1}^{n} a_{ki}x_i + b_k$$

2. $H(x) = A$, где $H(x)$ — Гессиан???
$$\frac{\partial^2 f}{\partial x_l \partial x_k} = \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_k}\right) = \frac{\partial}{\partial x_l}\left(\sum_{i=1}^{n} a_{ki}x_i + b_k\right)$$

3. Квадратичная функция $f(x)$ с положительно определенной матрицей A сильно выпукла
$$A = \begin{vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{vmatrix}$$
$$A - lE = \begin{vmatrix} \lambda_1 - l & 0 & \dots & 0 \\ 0 & \lambda_2 - l & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n - l \end{vmatrix}$$
В этом базисе все угловые миноры матрцы A и матрицы $A - lE$ — положительны при достаточно малом $l : 0 < l < \lambda_{\min \Rightarrow f(x)}$ — сильно выпукла

3.2 Принципы многмерной оптимизации

$$f(x) \to \min, \ x \in E_n$$
$$x^{k+1} = \Phi(x^k, x^{k+1}, \dots x)^0, \ x^0 \in E_n \tag{2}$$

— итерционная процедура(общего вида)

$\{x^k\}$:

$$\lim_{k\to\infty} f(x^k) = f^* = \min_{E_n} f(x), \ \text{если } U^* \neq \emptyset$$
$$\lim_{k\to\infty} f(x^k) = f^* = \inf_{E_n} f(x), \ \text{если } U^* = \emptyset$$

, где U^* – множестве точек глобального минимума функции $f(x)$
$\{x^k\}$+ условие 2 = минимизирующая последовательность для $f(x)$
Если для $U^* \neq \emptyset$ выполняется условие

$$\lim_{k\to\infty} \rho(x^k, U^*) = 0$$

, то x^k сходится к множеству U^*. Если U^* содежит единственную точку x^*, то для $\{x^k\}$ сходящейся к U^* будет справедливо $\lim_{k\to\infty} x^k = x^*$

Определение. $\rho(x, U) = \inf_{y\in U} \rho(x, y)$ — растояние от точки x до множества U

Примечание. Минимизирующая последовательность $\{x^k\}$ может и не сходится к точке минимума

Теорема 3.1 (Вейерштрасса)**.** Если $f(x)$ непрерывна в E_n и множество $U^\alpha = x : f(x) \leq \alpha$ для некоторого α непусто и ограничено, то $f(x)$ достигает глобального минимума в E_n

3.2.1 Скорость сходимости(минизирующих последовательностей)

Определение. $\{x^k\}$ сходится к точке x^* **линейно** (со скоростью геометрической последовательности), если $\exists q \in (0,1)$:

$$\rho(x^k, x^*) \le q\rho(x^{k-1}, x^*) \tag{3}$$
$$\rho(x^k, x^*) \le q^k\rho(x^0, x^*)$$

Определение. Сходимость называется **сверхлинейной** если

$$\rho(x^k, x^*) \le q_k\rho(x^{k-1}, x^*)$$

, и $q_k \xrightarrow[k\to\infty]{} +0$

Определение. Квадратичная сходимость:

$$\rho(x^k, x^*) \le \left[c\rho(x^{k-1}, x^*)\right]^2,\ c > 0$$

3.2.2 Критерии окончания итерационного процесса

$$\begin{gathered}\rho(x^{k+1}, x^*) < \varepsilon_1 \\ |f(x^{k+1}) - f(x^k)| < \varepsilon_2 \\ \|\nabla f(x^k)\| < \varepsilon_3\end{gathered} \tag{4}$$

, где ε_i — заранее заданные точности

$$x^{k+1} = x^k + \alpha_k p^k,\ k = 0, 1, \ldots \tag{5}$$

, где p^k — направление поиска из x^k в x^{k+1}, α_k — величина шага

$$f(x^{k+1}) < f(x^k)$$

— условие выбора α_k

Определение. В итерационном процессе 5 производится **исчерпывающий спуск**, если величина шага α_k находится из решения одномерной задачи минизации:

$$\Phi_k(\alpha) \to \min_\alpha,\ \Phi_k(\alpha) = f(x^k + \alpha p^k) \tag{6}$$

Теорема 3.2. Если функция $f(x)$ дифференцируема в пространстве E_n, то в итерационном процессе 5 с выбором шага с ичерпывающим спуском для любого $k \ge 1$:

$$(\nabla f(x^{k+1}), p^k) = 0 \tag{7}$$

— это значит что эти два вектора ортогональны

для $\Phi_k(\alpha)$ необходимое условие минимума функции:

$$\frac{d\Phi_k(\alpha)}{d\alpha} = \sum_{j=1}^{n} \frac{\partial f(x^{k+1})}{\partial x_j} \cdot \frac{dx_j^{k+1}}{d\alpha} = 0$$

учитывая $x_j^{k+1} = x_j^k + \alpha p_j^k \Rightarrow \frac{dx_j^k}{d\alpha} = p_j^k$

Теорема 3.3. Для квадратичной функции $f(x) = \frac{1}{2}(Ax, x) + (b, x) + c$ величина α_k исчерпывающего спуска в итерационном процессе

$$x^{k+1} = x^k + \alpha_k p^k, =0, 1, \ldots$$

равна

$$\alpha_k = -\frac{(\nabla f(x^k), p^k)}{(Ap^k, p^k)} = -\frac{(Ax^k + b, p^k)}{(Ap^k, p^k)} \tag{8}$$