11.10.19 ИО I

$$f(x) \to \min_{x \in \mathbb{R}^n}, \quad x, \nabla f(x)$$

$$x^+ = x - \alpha \nabla f(x)$$ Градиентный метод

Необх. условие 1го порядка: $\nabla f(x) = 0$

— " — 2го порядка: $\nabla^2 f(x) \geq 0$

$$a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

$h^T \nabla^2 f(x) h \geq 0 \quad \forall h \in \mathbb{R}^n \quad (\langle \nabla^2 f(x) h, h \rangle \geq 0)$

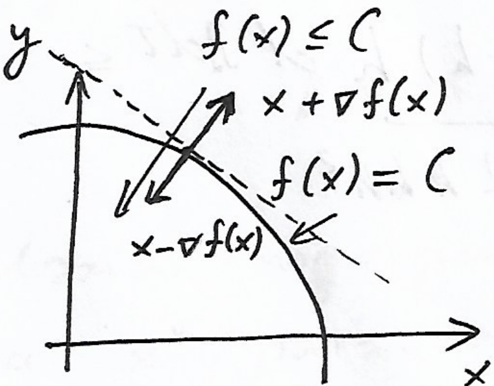
$\sim \quad \lambda_i \geq 0 \quad \forall i$ с.зн. $(\min \lambda \geq 0)$

Условия не яв-ся достаточными!

Пример: $f(x) = x^3$

Дост. условие 2го порядка: $\min \lambda(x) > 0$

Неединственность в окрестности: $f(x,y) = x^2$



Направление $h \in \mathbb{R}^n$, $\alpha \in \mathbb{R}_+$

$$f(x + \alpha h) - f(x) = \int_0^{\alpha} \nabla^T f(x + \tau h) h \, d\tau$$

Скаляр: $f(x)$

x, y: $f(y) - f(x) = \int_x^y f'(t) \, dt$

$$f(x+\alpha h) = f(x) + \underbrace{\alpha}_{\text{коэф-т разложения в ряд}} \underbrace{\nabla^T f(x) h}_{\text{пр-я по направлению}} + \frac{\alpha^2}{2} h^T \nabla^2 f(x) h + \dots$$

коэф-т пр-я по направлению

разложения в ряд

$$f(y) - f(x) = \int_x^y f'(\tau)\, d\tau =$$

$$= \int_x^y f'(0)\, d\tau + \int_x^y (f'(\tau) - f'(0))\, d\tau =$$

$$= \underbrace{f'(0)(y-x) + \int_0^\tau f''(s)\,\tau\, ds}_{\text{ряд Тейлора}}$$

ряд Тейлора

$$f(x+\alpha h) - f(x) = \int_0^\alpha \nabla f^T(x+\tau h) h\, d\tau = \quad \pm \nabla f^T(x) h$$

$$= \alpha \nabla f^T(x) h + \int_0^\alpha (\underbrace{\nabla f^T(x+\tau h) - \nabla f^T(x)) h}_{=\int_0^\tau h^T \nabla^2 f(x+sh) h\, ds}\, d\tau =$$

$$= \alpha \nabla f^T(x) h + \int_0^\alpha \int_0^\tau \underbrace{h^T \nabla^2 f(x+sh) h}_{\le L\|h\|^2}\, ds\, d\tau \le$$

Ограниченность Гессиана,

Липшицевость градиента:

$$\exists L : \forall y \quad h^T \nabla^2 f(y) h \le L \|h\|^2$$

$$\le \alpha \nabla f^T(x) h + L \|h\|^2 \underbrace{\int_0^\alpha \tau \, d\tau}_{\frac{\alpha^2}{2}} \ominus$$

Пусть $h = -\nabla f(x)$

$$\ominus -\alpha \|\nabla f(x)\|^2 + \frac{\alpha^2 L}{2} \|\nabla f(x)\|^2 =$$

$$= \left(\frac{\alpha^2 L}{2} - \alpha\right) \|\nabla f(x)\|^2 \Big|_{\alpha^*} = -\frac{1}{2L} \|\nabla f(x)\|^2$$

$$\boxed{\alpha^* = \frac{1}{L}}$$

$$f(x + \alpha h) - f(x) \le -\frac{1}{2L} \|\nabla f(x)\|^2, \quad h = -\nabla f(x)$$

градиентный метод: $0, 1 \dots M$

$$\underbrace{\sum_{k=0}^{M} \left(f(x^{k+1}) - f(x^k)\right)}_{f(x^M) - f(x^0)} \le -\frac{1}{2L} \sum_{k=0}^{M} \|\nabla f(x^k)\|^2$$

$$\frac{1}{2L} \sum_{k=0}^{M} \|\nabla f(x^k)\|^2 \le f(x^0) - f(x^M) \le C$$

$$\|\nabla f(x^k)\| \to 0, \quad k \to \infty$$

Можно построить пример $\|\nabla f(x^k)\|$ сх-ся, а x^k расх-ся. Но на практике x^k сх-ся вместе с $\|\nabla f(x^k)\|$

$$\|\nabla f(x^k)\|^2 \sim O\left(\frac{1}{k}\right) \quad \left\{\|\nabla f(x^k)\| \sim O\left(\frac{1}{\sqrt{k}}\right)\right\}$$

Пусть f – выпукла:

$\forall x, y \in \mathbb{R}^n: \quad f(y) - f(x) \geq \nabla f^T(x)(y-x)$

График выпуклой ф-ии лежит выше её касательной пл-ти.

$x: \nabla f(x) = 0 \quad \Rightarrow \quad f(y) \geq f(x) \quad \forall y$

~~Ф-ия обладает~~

Выпуклая ф-ия обладает минимумом x^*:

$\| x^{k+1} - x^* \|^2 = \| x^k - \alpha \nabla f(x^k) - x^* \|^2 =$

$= \| x^k - x^* \|^2 - 2\alpha \nabla f^T(x^k)(x^k - x^*) + \alpha^2 \| \nabla f(x^k) \|^2 \leq$

$\| \{ f(x^*) - f(x^k) \geq \nabla f^T(x^k)(x^* - x^k) \}$

выпуклость

$\leq \| x^k - x^* \|^2 + 2\alpha \left(f(x^*) - f(x^k) \right) + \alpha^2 \| \nabla f(x^k) \|^2$

$\sum_{k=0}^{M}: \quad \| x^M - x^* \|^2 \leq \| x^0 - x^* \|^2 + 2\alpha \sum_{k=0}^{M} (f(x^*) - f(x^k)) +$

$+ \sum_{k=0}^{M} \| \nabla f(x^k) \|^2$

$\sum_{k=0}^{M} f(x^k) - f(x^*) \leq \infty$

$f(x^k) - f(x^*) \sim O(\frac{1}{k})$

Сх-ть ф-ий

Сильная выпуклость ф-ии

$$\exists \lambda > 0 \quad \forall x: \quad h^T \nabla^2 f(x)\, h \geq \lambda \|h\|^2$$

$$f(y) - f(x) \geq \nabla f^T(x)(y-x) + \frac{\lambda}{2}\|y-x\|^2$$

$$f^* - f(x) \geq \min_{y \in \mathbb{R}^n} \left(\nabla f^T (y-x) + \frac{\lambda}{2}\|y-x\|^2\right)$$

$$\nabla f(x) + \lambda(y-x) = 0, \quad y - x = -\frac{\nabla f(x)}{\lambda}$$

$$f^* - f(x) \geq -\frac{\|\nabla f(x)\|^2}{\lambda} + \lambda\frac{\|\nabla f(x)\|^2}{2\lambda} = -\frac{\|\nabla f(x)\|^2}{2\lambda}$$

$$2\lambda\,(f^* - f(x)) \geq -\|\nabla f(x)\|^2$$

Тогда: $f(x+\alpha h) - f(x) \leq -\frac{1}{2L}\|\nabla f(x)\|^2 \leq$

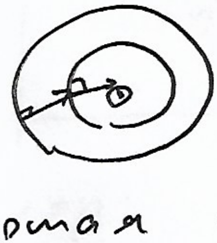
$$\leq \frac{\lambda}{L}(f^* - f(x))$$

$$f(x+\alpha h) - f^* \leq f(x) - f^* - \frac{\lambda}{L}(f(x) - f^*) =$$

$$= \left(1 - \frac{\lambda}{L}\right)(f(x) - f^*) \leq \left(1 - \frac{\lambda}{L}\right)^k (f(x^0) - f^*)$$

геометрическая прогрессия, линейная ск-ть сх-ти, число обусловленности $\frac{\lambda}{L}$

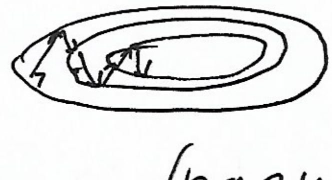
$\lambda << L$



$1 - \frac{\lambda}{L} << 1$

хорошая сх-ть

$\lambda \sim L$



$1 - \frac{\lambda}{L} \sim 1$

пилообразный эффект

плохая сх-ть

многошаговые градиентные методы

Шаговый множитель $\alpha = \frac{1}{L}$

Правило Армихо:

$$f(x - \alpha \nabla f(x)) - f(x) \leq - \frac{\alpha}{2} \|\nabla f(x)\|^2$$

проверяем условие

иначе дробим $\alpha \leftarrow \frac{\alpha}{2}$

Геометрия ф-ии $\frac{\lambda}{L}$, <u>оценка</u>

$$L \leq \max_i \sum_j |a_{ij}| \quad , \{\nabla^2 f(x)\}_{ij} = \{a\}_{ij}$$

Условие диагонального преобладания, матрицы специального вида

$$\begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & -1 & 2 & & \\ & & & \ddots & -1 \\ 0 & & & -1 & 2 \end{pmatrix} \quad \begin{matrix} L \leq 4 \\ \min \mu \ ? \end{matrix}$$