18.10.19 ИО II

Михаил Рубинович

Метод Ньютона

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

Разложение в окр-ти x:

$$f(x+h) \approx f(x) + \nabla f^T(x) h + \frac{1}{2} h^T \nabla^2 f(x) h$$

ф-ия f вып-я и $\nabla^2 f > 0$, тогда модель имеет минимум

$$\nabla f^T(x) h + \frac{1}{2} h^T \nabla^2 f(x) h \to \min_{h \in \mathbb{R}^n}$$

$$\nabla f(x) + \nabla^2 f(x) h = 0, \quad h = -\nabla^2 f(x)^{-1} \nabla f(x) = \Delta x_N$$

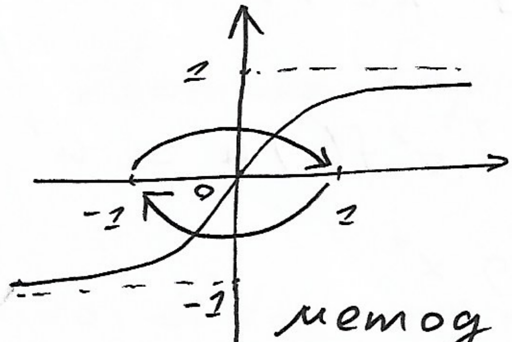
ньютоновское направление

x: $x^+ = x + \alpha \Delta x_N$

Минимизация ф-ий $f(x) \to \min_{x \in \mathbb{R}^n}$ экв-на решению уравнений $\nabla f(x) = 0$

Пример: $f(x) = \dfrac{x}{\sqrt{1+x^2}}$ невыпуклая!

корни?



метод не сх-ся!

$x_0 = 1$

$$x^+ = x - f'(x)^{-1} f(x)$$

$$f'(x) = \frac{1}{\sqrt{1+x^2}} - \frac{\cancel{2}}{2} \frac{x^2}{(1+x^2)^{\frac{3}{2}}} = \frac{1}{(1+x^2)^{3/2}}$$

$f(1) = \frac{1}{\sqrt{2}}$, $f'(1) = \frac{1}{2^{3/2}}$, $x^+ = 1 - 2^{3/2} \frac{1}{\sqrt{2}} = 1 - 2 = -1$

Пример 2:
$$f(x) = \frac{1}{4}x^4 - x, \quad f'(x) = x^3 - 1 = 0, \quad x = 1$$

$$x^0 = 0, \quad f'(0) = -1, \quad f''(0) = 3x^2\Big|_{x=0} = 0$$

ф-ия выпуклая, но $f'' \equiv 0$, вырожденный гессиан, направления Ньютона нет!

Newton, $\underbrace{\text{Raphson}}$

ученик Ньютона, шаговый мн-ль α

$\underbrace{\text{Инв-ть}}$ по отношению к лин. преобр. пр-ва

$$x \to x^+, \quad x^+ = F(x), \quad \exists\ \underbrace{x = Tz}_{\text{лин. преобр.}}$$

$$f(x) = f(Tz) = \tilde{f}(z), \quad z^+ = \hat{F}(z)$$

Тогда $x^+ = Tz^+$, (образы связаны лин. преобр.)

$$\underbrace{\nabla f^T(x)\, h}_{\text{пр-я по напр-ю } h} = \nabla f^T(\cancel{x}Tz)\, Tu = (T^T \nabla f(Tz))^T u =$$
$$= \nabla_z \hat{f}(z)^T u$$

пр-я по напр-ю h
$\exists\ h = Tu$

$$\boxed{\nabla_z \tilde{f}(z) = T^T \nabla_x f(x)}$$

$$h^T \nabla^2 f(x) h = u^T T^T \nabla^2 f(Tz)\, Tu = u^T \nabla^2 \tilde{f}(z)\, u$$

$$\boxed{\nabla_z^2 f(z) = T^T \nabla_x^2 f(x)\, T}$$

$?_{\,''}\ x^+$

$$Tz^+ = \underbrace{Tz}_{x} - T\nabla^2 \tilde{f}(z)^{-1} \nabla \tilde{f}(z) =$$
$$= x - T(T^T \nabla^2 f(x) T)^{-1} T^T \nabla f(x) =$$
$$= x - T(T^{-1} \nabla^2 f(x)^{-1} (T^T)^{-1} T^T \nabla f(x) = x - \nabla^2 f(x)^{-1} \nabla f(x)$$

$1 - \frac{1}{2} \approx 0$, для $\nabla^2 f(x)^{-1} \approx I$

Исследование сх-ти сх-ти метода

$A_x = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & & \\ & & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}$, $x = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ {Задача из прошлой лекции}

$\|x\| = \sqrt{n}$, $Ax = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$, $\|Ax\| = \sqrt{2}$

$Ax = \lambda x$, $\frac{\|Ax\|}{\|x\|} = \lambda$, $\lambda = \min\limits_x \frac{\|Ax\|}{\|x\|} \le \sqrt{\frac{2}{n}} \xrightarrow[n\to\infty]{} 0$

x с.в.

Исследование шагового мн-ля

$\lambda^2 = \Delta x_N^T \nabla^2 f(x) \Delta x_N = (-\nabla f(x)^T (\nabla^2 f(x))^{-1}) \nabla^2 f(x) (-(\nabla^2 f(x))^{-1} \nabla f)$
$= \nabla f(x) \nabla^2 f(x)^{-1} \nabla f(x)$

{здесь λ не min с.зн.}

$\nabla f(x)^T \Delta x_N = -\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x) = -\lambda^2$

Демпфированный или метод Ньютона-Рафсона

$f(x + \alpha \Delta x_N) - f(x) \le \frac{\alpha}{4} \lambda^2$

анализ метода Армихо выбора шага α

1) $f(x)$ сильно вып. с μ! $h^T \nabla^2 f(x) h \ge \mu \|h\|^2$

и $h^T \nabla^2 f(x) h \le L \|h\|^2$

3) $\exists M : \forall x, y:$

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \le M \|x - y\|$$

$$\|A\| = \sqrt{\max_{\|x\|=1} \langle Ax, Ax \rangle}$$ спектральная норма

I Демпфированная фаза

$$f(x + \alpha \Delta x_N) \le f(x) + \alpha \underbrace{\nabla f(x)^T \Delta x_N}_{-\lambda^2} + \frac{\alpha^2 L}{2} \|\Delta x_N\|^2 \le$$

оценка сверху!

$$\le f(x) - \alpha \lambda^2 + \frac{\alpha^2 L}{2} \frac{1}{\mu} \underbrace{\Delta x_N^T \nabla^2 f(x) \Delta x_N}_{\lambda^2} =$$

$$= f(x) - \alpha \lambda^2 + \frac{\alpha^2 L}{2\mu} \lambda^2 = f(x) - \lambda^2 \left(\alpha - \frac{\alpha^2 L}{2\mu}\right) =$$

$$\| \quad \alpha^* = \frac{\mu}{L}$$

$$= f(x) - \frac{\mu}{2L} \lambda^2$$, $\lambda^2 \to 0$ на итерациях метода

Отсюда линейная ск-ть сх-ти и оценка α в нер-ве Армихо

II Вторая фаза (основная)

$$f(x + \alpha \Delta x_N) = f(x) - \alpha \lambda^2 + \frac{\alpha^2}{2} \Delta x_N^T \nabla^2 f(\xi) \Delta x_N =$$

$$\{\pm \Delta x_N^T \nabla^2 f(x) \Delta x_N\} \qquad \xi \in (x, x + \alpha \Delta x_N)$$

$$= f(x) - \alpha \lambda^2 + \frac{\alpha^2}{2} \lambda^2 + \frac{\alpha^2}{2} \Delta x_N^T (\nabla^2 f(\xi) - \nabla^2 f(x)) \Delta x_N \le$$

$$\le f(x) - \alpha \lambda^2 + \frac{\alpha^2}{2} + \frac{\alpha^3}{2} M \|\Delta x_N\|^3 = \{\alpha = 1\} =$$

$$\{\|\xi - x\| \le \alpha \|\Delta x_N\|\}$$

$$\|\Delta x_N\|^3 \le \frac{\lambda^3}{\mu^{3/2}} \quad \left\{ \text{из} \quad \begin{array}{l} h^T \nabla^2 f(x) h \ge \mu \|h\|^2 \\ \text{и } \lambda^2 = \Delta x_N^T \nabla^2 f(x) \Delta x_N \end{array} \right\}$$

$$\stackrel{\{\alpha = 1\}}{=} \quad f(x) - \lambda^2 + \frac{\lambda^2}{2} + \frac{M}{2} \frac{\lambda^3}{\mu^{3/2}} =$$

$$= f(x) - \lambda^2 \underbrace{\left(1 - \frac{1}{2} - \frac{M\lambda}{2\mu^{3/2}}\right)}_{\frac{1}{2} - \frac{M\lambda}{2\mu^{3/2}}} \le$$

$$\le f(x) - \frac{\lambda^2}{4} \qquad \frac{M\lambda}{2\mu^{3/2}} \to 0$$

$$\exists \quad \frac{M\lambda}{2\mu^{3/2}} < \frac{1}{4}$$

Оценка метода Армихо для $\alpha = 1$

1) Boyd, Stearen: Convex Optimization

2) Michelle Menu