0 Повторение результатов про град. методы

Сходимость

GD: $O\left(\frac{1}{k}\right)$ вып. ; $O\left(\left(1-\frac{\mu}{L}\right)^k\right)$ сильно вып. $= \frac{\varkappa - 1}{\varkappa}$

SD $O\left(\frac{1}{\sqrt{k}}\right)$ $O\left(\frac{1}{k}\right)$ ← улучшить нельзя

AGD $O\left(\frac{1}{k^2}\right)$; $O\left(\left(1-\sqrt{\frac{\mu}{L}}\right)^k\right)$ $\frac{\sqrt{\varkappa}-1}{\sqrt{\varkappa}}$

квадратичн. задача

$A \in \mathbb{S}^n_{++}$

$$f(x) = \frac{1}{2}x^T A x - b^T x + c$$

$$\nabla f(x) = Ax - b$$

$$\nabla^2 f = A$$

$\lambda_{min}(A) = \mu > 0$

$\lambda_{max}(A) = L > 0$

$\varkappa = \frac{L}{\mu}$

лин. поиск шага для направления d в квадр. задаче

$$x^{k+1} = x^k + \alpha \cdot d \qquad -\nabla f(x) = r$$

$$\alpha = \underset{\alpha > 0}{\operatorname{argmin}} f(x^{k+1})$$

$$-r^{k+1\,T} d = 0 \qquad (\alpha A d - r^k) d = 0$$

$$r^{k+1} = b - Ax^k - \alpha A d = r^k - \alpha A d$$

$$\boxed{\alpha = \frac{d^T r^k}{d^T A d^T}}$$

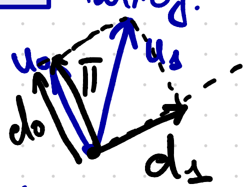
Ортогонализация Грамма - Шмидта

Вход: $u_0, \dots u_{n-1}$ - n ЛНЗ векторов

Выход: $d_0, \dots, d_{n-1}$ - n ЛНЗ $\perp$ векторов

$$d_0 = u_0$$

$$d_1 = u_1 - \Pi_{d_0}(u_1)$$



$$\Pi_{d_0}(u_1) = \frac{\langle u_1, d_0\rangle}{\langle d_0, d_0\rangle} d_0$$

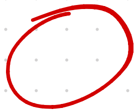
$$d_2 = u_2 - \Pi_{d_0}(u_2) - \Pi_{d_1}(u_2)$$

$$d_k = u_k + \sum_{i=0}^{k-1} \beta_{ik} \cdot d_i \qquad \beta_{ik} = \frac{-\langle d_i, u_k\rangle}{\langle d_i, d_i\rangle}$$

① Постановка задачи + интуиция

$$r^k = b - Ax^k$$
$$e^k = x^k - x^*$$
$$r^k = -Ae^k$$

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}x^TAx - b^Tx + c$$



A-ортогональность

② Метод сопряжённых направлений:

и A-ортогональных направлений (по граммшмидту)

\+

лин. поиск по каждому

Метод сопр. напр.

$$X_i = X_0 + \sum_{j=0}^{i-1} \alpha_j d_j$$

$d_0,\dots,d_{n-1}$ - все попарно A-сопряжены

$$\alpha_j = \frac{d_j^T r_j}{d_j^T A d_j}$$

из наиск. спуска вдоль d_j

Напр ↑ получались из $U_0,\dots,U_{n-1}$ ЛНЗ путём Gram-Schmidt.

$$\boxed{d_i = U_i + \sum_{j=0}^{i-1} \beta_{ij} d_j} \quad (GS)$$

из u_i вычитаем проекции d_j на u_i

$$\beta_{ij} = -\frac{u_i^T A d_j}{d_j^T A d_j} \quad (B) \quad \forall i \in [0, n-1]$$

③ Докажем, что при таком выборе α_k и $d_0 \dots d_n$, мы сойдемся за n шагов.

$$\overset{x_0 - x^*}{\overset{''}{e_0}} = \sum_{i=0}^{n-1} \delta_i \cdot \overset{d}{d_i}$$

коэф. разложения e_0 в базис

фиксируем k-ую коорд.

$d_k^T A$.

$$x_0 - x^* = \sum_{i=0}^{n-1} \delta_i d_i \qquad -\alpha_i$$

$$x^* = x_0 - \sum_{i=0}^{n-1} \delta_i d_i$$

$$d_k^T A e_0 = \sum_{i=0}^{n-1} \delta_i \cdot d_k^T A d_i$$

$$d_k^T A e_0 = \delta_k d_k^T A d_k$$

$$\delta_k = \frac{d_k^T A e_0}{d_k^T A d_k} = \frac{d_k^T A \overbrace{\left(e_0 + \sum_{i=0}^{k-1} \alpha_i d_i\right)}^{e_k}}{d_k^T A d_k} = - \frac{d_k^T r_k}{d_k^T A d_k} = -\alpha_k .$$

④ Идея CG: невязки в качестве $u_0, \dots u_{n-1}$ + + показать, что для рассчёта d_k необходимо знать лишь d_{k-1}

буквально β_{ij} почти везде 0

④.1

Fun fact:

$$e_i = e_0 + \sum_{j=0}^{i-1} \alpha_j d_j = \qquad x_i = x_0 + \sum_{j=0}^{i-1} \alpha_j d_j$$

$$= \sum_{j=0}^{n-1} -\alpha_j d_j + \sum_{j=0}^{i-1} \alpha_j d_j = \qquad e_i = x_i - x^*$$

$$e_0 = x_0 - x^* =$$

$$= \sum_{j=i}^{n-1} -\alpha_j d_j . \qquad (ER) \qquad = x_0 - \left(x_0 + \sum_{j=0}^{n-1} \alpha_j d_j\right)$$

4.2

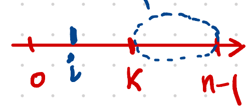
Теперь (ER записано для фикс. k) $\cdot \left(-d_i^T A\right)$

I

$$e_k = \sum_{j=k}^{n-1} -\alpha_j d_j$$

A-ортогона

$$-d_i^T A \cdot e_k = \sum_{j=k}^{n-1} \alpha_j d_i^T A d_j$$



$$d_i^T r_k = 0$$

если $i<k$

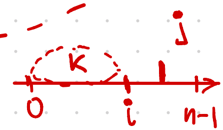
Таким образом: r_k перпендикулярна всем предыдущим направлениям d_i

$$\boxed{d_i^T r_k = 0}\ , i<k.$$

4.3

$(GS)^T \cdot r_j$

$$d_i^T r_j = u_i^T r_j + \sum_{k=0}^{i-1} \beta_{ik} d_k^T r_j$$



$$\underset{0}{\overset{\shortparallel}{d_i^T r_j}} = u_i^T r_j + 0$$

$i<j$

$$\boxed{u_i^T r_j = 0} \longleftarrow \underline{i<j}$$

Пусть $i=j$

$$\boxed{d_i^T r_i = u_i^T r_i}$$

4.4 $r_{i+1} = -A \cdot e_{i+1} = -A(e_i + \alpha_i d_i) = -Ae_i - \alpha_i A d_i =$

$= r_i - \alpha_i A d_i$.

II $\to$ $r_i^T r_j = 0$, $i<j$

$r_i^T r_j = 0$, $i \neq j$.

Если теперь $u_i = r_i$, то

CG:

Рассмотрим:

(B) $$\beta_{ij} = -\frac{u_i^T A d_j}{d_j^T A d_j}$$

$u_i = r_i$

$$-\frac{r_i^T A d_j}{d_j^T A d_j}$$

Рассмотрим $\forall i,j$

$$r_i^T r_{j+1} \overset{\text{III}}{=} r_i^T (r_j - \alpha_j A d_j) = r_i^T r_j - \alpha_j r_i^T A d_j$$

$$\alpha_j r_i^T A d_j = r_i^T r_j - r_i^T r_{j+1}$$

$$r_i^T A d_j = \begin{cases} \frac{1}{\alpha_i} \cdot r_i^T r_i, & i=j \\ -\frac{1}{\alpha_{i-1}} r_i^T \cdot r_i, & i=j+1,\ j=i-1 \\ 0, & \text{иначе} \end{cases}$$

$j < i$

ИТОГО

$$\beta_{ij} = \frac{-r_i^T A d_j}{d_j^T A d_j} = +\frac{1}{\alpha_{i-1}} \cdot \frac{r_i^T r_i}{d_{i-1}^T A d_{i-1}} =$$

$$= \frac{\cancel{d_{i-1}^T A d_{i-1}}}{d_{i-1}^T r_{i-1}} \frac{r_i^T r_i}{\cancel{d_{i-1}^T A d_{i-1}}}$$

возрадуемся!

$$\beta_{ij} = \frac{r_i^T r_i}{r_{i-1}^T r_{i-1}} \quad \text{для } j = i-1, \text{ иначе } 0$$

$$f(x) = \frac{1}{2} x^T A x - b^T x + c \to \min_{x \in \mathbb{R}^n}$$

$$A \in \mathbb{S}^n_{++}$$

Алгоритм

CG.

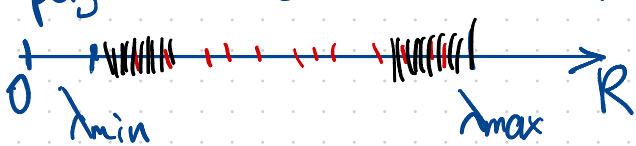
есть x_0

1. $d_0 = r_0 = b - Ax_0$.
2. For $i = 0, \dots$ — Ad_i = хрень$_i$
3. $\alpha_i = \dfrac{r_i^T r_i}{d_i^T A d_i}$ — Вместо формулы – линейный поиск (backtrack)
4. $x_{i+1} = x_i + \alpha_i d_i$
5. $r_{i+1} = r_i - \alpha_i A d_i$ — можно: $r_{i+1} = b - Ax_{i+1}$
6. $\beta_{i+1} = \dfrac{r_{i+1}^T r_{i+1}}{r_i^T r_i}$
7. $d_{i+1} = r_{i+1} + \beta_{i+1} d_i$

Сходимость: в идеальной арифметике

сходится точно за m итераций,

где m – число различных собственных чисел



В реальности

метод итерационный

$$\|e_k\|_A^2 \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k} \|e_0\|_A^2$$

4

< 1

$$\kappa = \|A\| \cdot \|A^{-1}\| = \frac{\sigma_{max}}{\sigma_{min}} = \frac{\lambda_{max}}{\lambda_{min}} = \frac{L}{\mu}$$