Найти $x \in \mathbb{R}^d$: $Ax = b$ — СЛУ

$\in \mathbb{R}^{d\times d}$ – положит. опред. (симм.) $\quad b \in \mathbb{R}^d$

Сопряженные направления

Определение сопряженных направлений

Множество векторов $\{p_i\}_{i=0}^{n-1}$ будем называть сопряженным относительно положительно определенной матрицы A, если для любых $i \neq j \in \{0, \dots n-1\}$ следует

$$p_i^T A p_j = 0. \quad \leftarrow \text{ортог. отн. } A$$

$\langle p_i ; p_j \rangle = 0 \quad p_i^T p_j = 0$ (ортог.)

Линейная независимость сопряженных направлений

Сопряженных векторы $\{p_i\}_{i=0}^{n-1}$ является линейно независимыми.

Док-во: от противного

$\exists\ \exists p_i : \quad p_i = \sum_{j \neq i} \lambda_j p_j$ где $\exists \lambda_j \in \mathbb{R}$

по опред. сопр. направл. $(A p_m$ или $p_m^T A^T)$, $m \neq i$, $A^T = A$

$$\underbrace{p_m^T A p_i}_{=0,\ m\neq i} = \sum_{j \neq i} \lambda_j\, p_m^T A p_j \underset{\substack{j\neq m \\ p_m A p_j = 0}}{=} \lambda_m \underbrace{p_m^T A p_m}_{>0}$$

$\lambda_m = 0$

по всем m $\quad \lambda_m = 0 \Rightarrow$ противоречие

$\{p_i\}$ — базис

$$x^* = \sum_{i=0}^{d-1} \lambda_i p_i$$

x^* — решение СЛУ

Из сопряж. ($p_j^T A \cdot$):

$$\underbrace{p_j^T A x^*}_{\in\mathbb{R}} = \sum_{i=0}^{d-1} \lambda_i \, p_j^T A p_i \overset{i\neq j,\ p_j^T A p_i = 0}{=} \underbrace{\lambda_j}_{\in\mathbb{R}} \underbrace{p_j^T A p_j}_{\in\mathbb{R}}$$

$$\lambda_j = \frac{p_j^T A x^*}{p_j^T A p_j} \overset{Ax^*=b}{=} \frac{p_j b}{p_j^T A p_j}$$

Итеративный метод

$$\boxed{x^{k+1} = x^k + \alpha_k p_k}$$

итеративно добавляю в x новое направление

$\alpha = \lambda$?

$$x^d = x^0 + \sum_{i=0}^{d-1} \alpha_i p_i$$

было

$$x^d = x^* = \sum_{i=0}^{d-1} \lambda_i p_i$$

- если $x^0 = 0$ $\quad \alpha_i = \lambda_i$
- если $x^0 \neq 0$, то нужно правило для α_i

x^0 по p_i

$$x^0 = \sum_{i=0}^{d-1} \tilde{\lambda}_i p_i \quad \Big| \ p_j^T A \cdot$$

$$\tilde{\lambda}_i = \frac{p_i^T A x^0}{p_i^T A p_i} \quad \Big| \ p_j^T A x^0 = p_j^T A p_j \cdot \tilde{\lambda}_j$$

$$x^* = \underbrace{x^0 + \sum_{i=0}^{d-1} \alpha_i p_i} = \underbrace{\sum_{i=0}^{d-1} \lambda_i p_i}$$

$$\sum_{i=0}^{d-1} \frac{p_i^T A x^0}{p_i^T A p} p_i + \sum_{i=0}^{d-1} \alpha_i p_i = \sum_{i=0}^{d-1} \frac{p_i^T b}{p_i^T A p_i} p_i$$

$$\alpha_i = \frac{p_i^T (b - A x^0)}{p_i^T A p_i}$$ убедим

$$p_k^T A (x^k - x^0) = 0 \quad \Big| \quad \underbrace{x^k - x^0}_{p_k^T A} = \underbrace{\sum_{i=0}^{k-1} \alpha_i p_i}_{p_k^T A}$$

$p_k^T A p_i = 0$ — сопряж.

$$\Downarrow$$

$$-p_i^T A x^0 = -p_i^T A x^i$$

$$\Downarrow$$

$$\alpha_i = \frac{p_i^T (b - A x^i)}{p_i^T A p_i} = \boxed{-\frac{p_i^T r_i}{p_i^T A p_i}}, \quad r_i = A x^i - b$$

Физический смысл α

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

$$\nabla f(x^*) = A x^* - b = 0$$

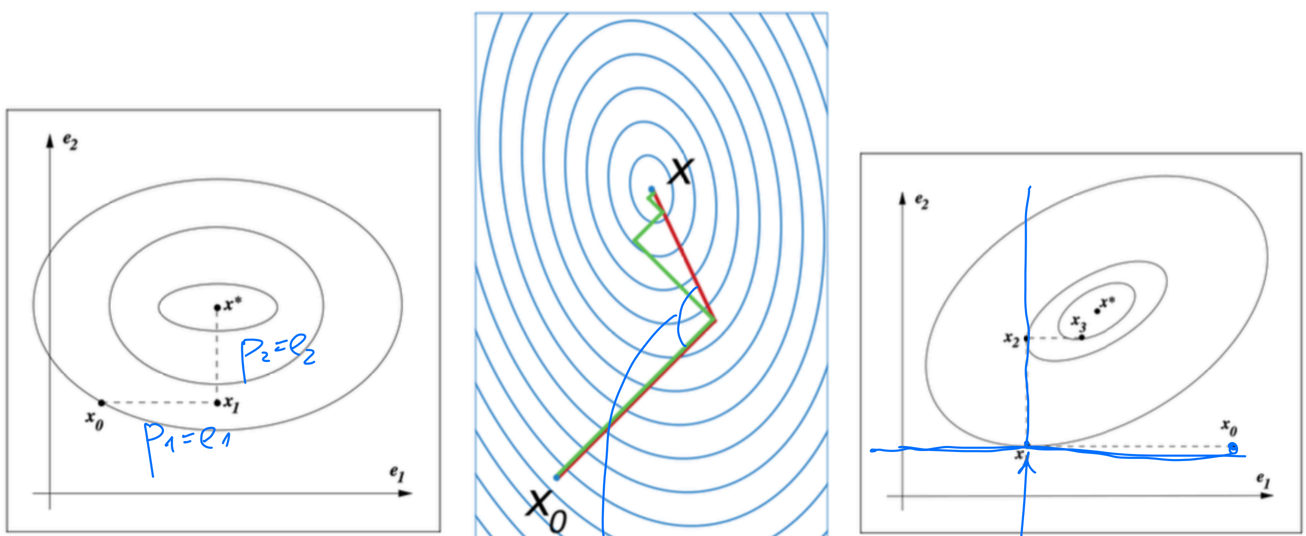
поиск решения СЛУ

$$g(\alpha) = f(x^k + \alpha p_k) \quad \Big| \quad f(x^k - \alpha \nabla f(x^k))$$

$\uparrow$ $\underset{\alpha}{\operatorname{argmin}}$

$$\alpha^* = \operatorname{argmin} g(\alpha) = \frac{p_k^T (b - A x^k)}{p_k^T A p_k}$$

$\leftarrow$ как в методе выше!



сопр. напр. не $\perp$

α - min

здесь $\perp$, но не сопр. отн. A

за 2 шага не сход.

Физический смысл p

Физический смысл p

Если $\{p_i\}_{i=0}^k$ сопряженные направления, то для любого $k \geq 0$ и $i \leq k$ справедливо:

$$r_{k+1}^T p_i = 0 \text{ то же самое, что } \langle \nabla f(x^{k+1}), p_i \rangle = 0.$$

градиент в новой точке $\perp$ всем p_i
значит вдоль p_i все уж. опт.

$$r_k = Ax^k - b \qquad \alpha_k = -\frac{p_k^T r_k}{p_k^T A p_k}$$

$$x^{k+1} = x^k + \alpha_k p_k$$

Док-во: по индукции

База: $p_0^T r_1 = 0$?

$$r_1 = Ax^1 - b = \underbrace{Ax^0 - b}_{r_0} + \alpha_0 A p_0 = r_0 + \alpha_0 A p_0$$

$$p_0^T r_1 = p_0^T (r_0 + \alpha_0 A p_0) = \cancel{p_0^T r_0} - \frac{\cancel{p_0^T r_0}}{\cancel{p_0^T A p_0}} \cdot \cancel{p_0^T A p_0} = 0$$

ПИ: верно $i \leq k$

Переход: $p_k^T r_{k+1}$ — ? $\quad p_i^T r_{k+1}$ — ?

$$r_{k+1} \underset{\substack{\uparrow \\ r}}{=} Ax^{k+1} - b \underset{\substack{\uparrow \\ x}}{=} \underbrace{Ax^k - b}_{r_k} + \alpha_k A p_k = r_k + \alpha_k A p_k$$

$$p_k^T r_{k+1} = p_k^T r_k + \alpha_k p_k^T A p_k \underset{\substack{\uparrow \\ \alpha_k}}{=} 0 \quad (+)$$

$i < k$

$$p_i^T r_{k+1} = \underbrace{p_i^T r_k}_{0 \ (\text{ПИ})} + \underbrace{\alpha_k p_i^T A p_k}_{0 \ (\text{сопр})} = 0 \quad (+)$$

Цель: считать p_k итеративно

„дешево" — не хочется хранить p_k
не хочется пересч. по всем p_i

$$p_k = -\underbrace{r_k}_{\nabla f(x^k)} + \underbrace{\beta_k}_{\in \mathbb{R}} p_{k-1}$$

Найти: p_k и p_{k-1} сопряженные

$$\underset{\text{сопр.}}{0} = p_{k-1}^T A p_k = -\underbrace{p_{k-1}^T A r_k}_{\in \mathbb{R}} + \underbrace{\beta_k}_{\in \mathbb{R}} \underbrace{p_{k-1}^T A p_{k-1}}_{\in \mathbb{R}}$$

$$\beta_k = \frac{p_{k-1}^T A r_k}{p_{k-1}^T A p_{k-1}}$$

Алгоритм 1 Метод сопряженных градиентов

Вход: стартовая точка $x^0 \in \mathbb{R}^d$, $r_0 = Ax_0 - b$, $p_0 = -r_0$ количество итераций K

1: **for** $k = 0, 1, \ldots, K-1$ **do**
2: $\quad \alpha_k = -\frac{r_k^T p_k}{p_k^T A p_k}$
3: $\quad x^{k+1} = x^k + \alpha_k p_k$
4: $\quad r_{k+1} = Ax^{k+1} - b$ $\quad \nabla f$
5: $\quad \beta_{k+1} = \frac{r_{k+1}^T A p_k}{p_k^T A p_k}$
6: $\quad p_{k+1} = -r_{k+1} + \beta_{k+1} p_k$
7: **end for**

Выход: x^K

Сходимость

$$x^{k+1} = x^k + \alpha_k p_k \qquad r_{k+1} = Ax^{k+1} - b$$

Нужно, что $\{p_i\}$ — сопряженные направл. ?

вот это докажем

База: p_0 и p_1 — сопряжены (подбор β_1)

ПИ: $\{p_i\}_{i=0}^{k}$ — сопр. направл.

<u>Переход</u>:

- $i = k$ $\quad p_{k+1}$ и p_i – спряжени (подбор β_{k+1})
- $i < k$

$$p_{k+1}^T A p_i \underset{\substack{\uparrow \\ p_{k+1}}}{=} -r_{k+1}^T A p_i + \beta_{k+1} \underbrace{p_k^T A p_i}_{\substack{0 \\ (\text{ПИ})}} = -r_{k+1}^T A p_i \overset{?}{=} 0$$

<u>Важен факт</u>

$$\operatorname{span}\{r_0 \ldots r_k\} = \operatorname{span}\{r_0 \ldots A^k r_0\}$$

$$\operatorname{span}\{p_0, \ldots p_k\} = \operatorname{span}\{r_0 \ldots A^k r_0\}$$

<u>База</u>: $\quad r_0 = -p_0$ (инициализация алг. 1)

<u>ПИ</u>: $\quad$ вярно $i \leq k$

<u>Переход</u>:

$$p_k \in \operatorname{span}\{r_0 \ldots A^k r_0\}$$

$$\Downarrow$$

$$A p_k \in \operatorname{span}\{A r_0 \ldots A^{k+1} r_0\}$$

$$r_k \in \operatorname{span}\{r_0 \ldots A^k r_0\}$$

$$r_{k+1} = A x^{k+1} - b = \underbrace{A x^k - b}_{r_k} + \alpha_k A p_k = r_k + \alpha_k A p_k$$

$$r_{k+1} \in \operatorname{span}\{r_k; A p_k\} \subseteq \{r_0 \ldots A^{k+1} r_0\}$$

- $\operatorname{span}\{r_0 \ldots r_{k+1}\} \underset{\substack{\uparrow \\ \text{нужно обратно}}}{\subseteq} \operatorname{span}\{r_0 \ldots A^{k+1} r_0\}$ $\quad \oplus$

$$A^{k+1} r_0 = A(\underbrace{A^k r_0}_{\in \operatorname{span}\{p_0 \ldots p_k\}}) \in \operatorname{span}\{A p_0 \ldots A p_k\}$$

$$\frac{r_{i+1} - r_i}{\alpha_i} = A p_i \longrightarrow \in \operatorname{span}\{r_0 \ldots r_{k+1}\}$$

• $\operatorname{span}\{r_0 \ldots r_{k+1}\} \supseteq \operatorname{span}\{r_0 \ldots A^{k+1} r_0\}$ ⊕

первый факт ⊕

$$\operatorname{span}\{p_0 \ldots p_{k+1}\} = \operatorname{span}\{p_0 \ldots p_k, r_{k+1}\}$$

ПИ

$$= \operatorname{span}\{r_0 \ldots A^k r_0, r_{k+1}\}$$

ПИ

$$= \operatorname{span}\{r_0 \ldots r_k, r_{k+1}\}$$

$$= \operatorname{span}\{r_0 \ldots A^k r_0\}$$

только

$\measuredangle$: $-r_{k+1}^T A p_i = 0$? ⊕

вспомним сразу смысл p_i:

$\langle r_k; p_i \rangle = 0$

$i < k$

1) $p_i \in \operatorname{span}\{r_0 \ldots A^i r_0\}$

$Ap_i \in \operatorname{span}\{Ar_0 \ldots A^{i+1} r_0\}$

2) $\operatorname{span}\{r_0 \ldots A^{i+1} r_0\} = \operatorname{span}\{p_0 \ldots p_{i+1}\}$

$Ap_i \in \operatorname{span}\{p_0 \ldots p_{i+1}\}$ ∎

Теорема о сходимости метода сопряженных градиентов

Метод сопряженных градиентов для решения системы линейных уравнений с квадратной положительно определенной матрицей размера d находит точное решение за не более чем d итераций.

$O(d^3)$ = точное решение с помощью сопряж. град.

Еще сходимость

Теорема о сходимости метода сопряженных градиентов

Метод сопряженных градиентов для решения системы линейных уравнений с квадратной положительно определенной матрицей размера d находит точное решение за не более чем r итераций, где r – число уникальных собственных значений матрицы.

Теорема о сходимости метода сопряженных градиентов

Метод сопряженных градиентов для решения системы линейных уравнений с квадратной положительно определенной матрицей размера d имеет следующую оценку сходимости:

$$\|x^k - x^*\|_A \le 2\left(\frac{\sqrt{\kappa(A)}-1}{\sqrt{\kappa(A)}+1}\right)^k \|x^0 - x^*\|_A.$$

Здесь $\|x\|_A^2 = x^T A x$ и $\kappa(A) = \lambda_{\max}(A)/\lambda_{\min}(A)$.

работает не хуже Нестерова

Обобщение для произв. f

Алгоритм 5 Метод сопряженных градиентов (классическая версия)

Вход: стартовая точка $x^0 \in \mathbb{R}^d$, $r_0 = Ax_0 - b$, $p_0 = -r_0$ количество итераций K

for $k = 0, 1, \ldots, K-1$ **do**

- $\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k}$
- $x^{k+1} = x^k + \alpha_k p_k$
- $\cancel{r_{k+1} = r_k + \alpha_k A p_k}$
- $\beta_{k+1} = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$
- $p_{k+1} = -r_{k+1} + \beta_{k+1} p_k$

end for

Выход: x^K

$r = \nabla f$

Алгоритм 6 Метод сопряженных градиентов (Флетчер - Ривс)

Вход: стартовая точка $x^0 \in \mathbb{R}^d$, $p_0 = -\nabla f(x_0)$ количество итераций K

for $k = 0, 1, \ldots, K-1$ **do**

- $\alpha_k = ?$ $\arg\min_{\alpha} f(x^k + \alpha p_k)$
- $x^{k+1} = x^k + \alpha_k p_k$
- $\beta_{k+1} = \frac{\langle \nabla f(x^{k+1}), \nabla f(x^{k+1})\rangle}{\langle \nabla f(x^k), \nabla f(x^k)\rangle}$
- $p_{k+1} = -\nabla f(x^{k+1}) + \beta_{k+1} p_k$

end for

Выход: x^K

Алгоритм 9 Метод сопряженных градиентов (Полак - Рибьер)

Вход: стартовая точка $x^0 \in \mathbb{R}^d$, $p_0 = -\nabla f(x_0)$ количество итераций K

for $k = 0, 1, \ldots, K-1$ **do**

  $\alpha_k =$ Линейый поиск

  $x^{k+1} = x^k + \alpha_k p_k$

  $\beta_{k+1} = \frac{\langle \nabla f(x^{k+1}), \nabla f(x^{k+1}) - \nabla f(x^k) \rangle}{\langle \nabla f(x^k), \nabla f(x^k) \rangle}$

  $p_{k+1} = -\nabla f(x^{k+1}) + \beta_{k+1} p_k$

end for

Выход: x^K

⊕ дешевый решатель на ранних итерациях

⊕ физика

⊖ гарантии в общем случае плохие

⊖ нужны рестарты ($\beta_{k+1} = 0$)