$Ax = b$ $\quad A \in S_+^d \quad b \in \mathbb{R}^d$

решить $\quad x^* \in \mathbb{R}^d$

**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

$$p_i^T A p_j = 0.$$

**Линейная независимость сопряженных направлений**

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

Док-во: от противного

$\exists \quad p_j = \sum_{i \neq j} \lambda_i p_i \qquad \lambda_i \in \mathbb{R}$

$\forall \quad A p_m \cdot \qquad m \neq j$

$$\underbrace{p_m^T A p_j}_{=0} = \sum_{i \neq j} \lambda_i \underbrace{p_m^T A p_i}_{i \neq m \;\; 0}$$

в силу сопряж. (в силу сопр.)

$$= p_m^T A p_m \cdot \lambda_m$$

$\lambda_m = 0 \qquad m \neq j$

$\forall i \neq j \quad \lambda_i = 0 \Rightarrow p_j = 0$ противоречие

Тривиальное решение:

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

$p_j^T A \cdot$

$$p_j^T \underbrace{A x^*}_{b} = \sum_{i=0}^{n-1} \lambda_i \underbrace{p_j^T A p_i}_{j \neq i \quad 0} = \lambda_j p_j^T A p_j$$

$$\boxed{\lambda_j = \frac{p_j^T b}{p_j^T A p_j}}$$

"Дешевая" итерация

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

Суть: добавляем новое направление
находим $\alpha_k$ для $x^*(?)$

$$\alpha_k = \lambda_k ?$$

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

$x_0 = 0 \qquad \lambda_k = \alpha_k$

$$\lambda_j^0 = \frac{p_j^T (A x^0)}{p_j^T A p_j}$$

$$x^* = \sum_{i=0}^{n-1} \lambda_i p_i = \underset{\uparrow \atop \sum \lambda_i^0 p_i}{x^0} + \sum_{i=0}^{n-1} \alpha_i p_i$$

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

$$= \frac{p_i^T (b - Ax^0)}{p_i^T A p_i}$$

Заметим $\quad p_k^T A(\underset{\uparrow \atop \sum_{i=0}^{k-1} p_i \alpha_i}{x^k - x^0}) = 0$

$$p_k^T A p_i = 0$$

$$r_k = Ax^k - b$$

$$p_k^T A x^0 = p_k^T A x^k$$

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

Другой способ $\alpha_k$:

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

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

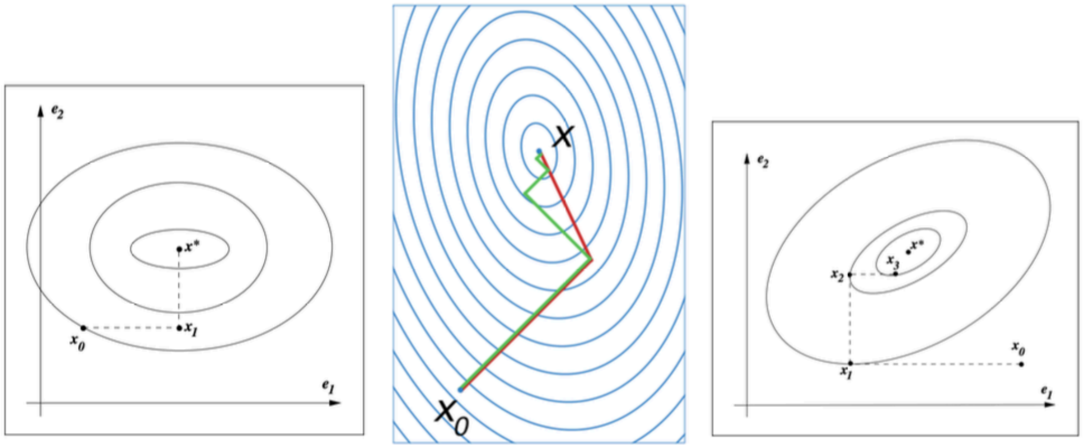
$$\nabla f(x^*) = Ax^* - b = 0$$

$$g(\alpha) = f(x^k + \alpha p_k) \to \min_{\alpha}$$

$$\alpha^* = -\frac{p_k^T r_k}{p_k^T A p_k}$$

$\uparrow$

мы подбираем наилучшим образом




**Физический смысл $p$**

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

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

Док-во:

• База: $r_1 = Ax^1 - b = A(x^0 + \alpha_0 p_0) - b$

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

$$= r_0 + \alpha_0 A p_0$$

$$\alpha_0 = -\frac{p_0^T r_0}{p_0^T A p_0}$$

$$= r_0 - \frac{p_0^T r_0 A p_0}{p_0^T A p_0}$$

Д: $r_1^T p_0 \stackrel{?}{=} 0$

$\in \mathbb{R}$

$$p_0^T r_1 = p_0^T \left( r_0 - \frac{p_0^T r_0 A p_0}{p_0^T A p_0} \right)$$

$$= p_0^T r_0 - \frac{p_0^T r_0 \, \cancel{p_0^T A p_0}}{\cancel{p_0^T A p_0}} = 0$$

- Предположение: верно для $i \le k$
- Переход: докажем для $k+1$

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

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

$$r_{k+1}^T p_i \overset{?}{=} 0 \qquad i = 0 \dots k$$

1) $i = k$

$$p_k^T r_{k+1} = p_k^T r_k + \alpha_k p_k^T A p_k = 0$$

$$\alpha_k = -\frac{p_k^T r_k}{p_k^T A p_k}$$

2) $i < k$

$$p_i^T r_{k+1} = \underbrace{p_i^T r_k}_{0 \text{ по индукции}} + \alpha_k \underbrace{p_i^T A p_k}_{0} = 0$$

Как находить $p$? Будем искать в упрощ. виде

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

$\beta_k$? $p_k$ и $p_{k-1}$ были сопряж.

$$0 = p_k^T A p_{k-1} = -r_k^T A p_{k-1} + \beta_k p_{k-1}^T A p_{k-1}$$

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

**Алгоритм 3** Метод сопряженных градиентов

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

**for** $k = 0, 1, \dots, K-1$ **do**
  $\alpha_k = -\frac{r_k^T p_k}{p_k^T A p_k}$
  $x^{k+1} = x^k + \alpha_k p_k$
  $r_{k+1} = Ax^{k+1} - b$
  $\beta_{k+1} = \frac{r_{k+1}^T A p_k}{p_k^T A p_k}$
  $p_{k+1} = -r_{k+1} + \beta_{k+1} p_k$
**end for**

**Выход:** $x^K$

стартуем в т. $x^0$

Мы знаем, что $p_k^T A p_{k-1} = 0$
но это не гарантирует, что $\{p_i\}$ – система сопряж. направлений

Хотим
$\forall k \geq 1$ и $\forall i < k$ показать $p_i^T A p_k = 0$

Начнём док-во:

- База: $p_0$ и $p_1$ – сопряж (по построению алгоритма)
- Предположение $\{p_i\}_{i=0}^k$ – сопряж. векторы
- Переход: докажем для $k+1$

1) $p_{k+1}$ и $p_k$ – сопряжены (по построению)

2) $p_{k+1}$ и $p_i$ $(i<k)$

$$p_{k+1}^T A p_i \overset{p_{k+1} \text{ из алгоритма}}{=} (-r_{k+1} + \beta_{k+1} p_k)^T A p_i$$

$$= -r_{k+1}^T A p_i + \beta_{k+1} \underbrace{p_k^T A p_i}_{0 \text{ по индукции}}$$

Нужно доказать $-r_{k+1}^T A p_i = 0 \quad i<k$

Вспомогательный факт:

$$\operatorname{span}\{r_0 \dots r_k\} = \operatorname{span}\{r_0 \dots A^k r_0\}$$
$$\operatorname{span}\{p_0 \dots p_k\} = \operatorname{span}\{r_0 \dots A^k r_0\}$$

Док-во вс. факта:

- База: $r_0 = r_0$, $r_0 = -p_0$ — из иниц. алгоритма
- Предположение: верно $i \le k$
- Переход: докажем для $k+1$

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

$r_k \in \operatorname{span}\{r_0 \dots A^k r_0\}$ по предположению

$p_k \in \operatorname{span}\{r_0 \dots A^k r_0\}$ по предпол.

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

$$r_{k+1} \in \operatorname{span}\{r_0 \dots A^k r_0, A^{k+1} r_0\}$$

$$\operatorname{span}\{r_0 \dots r_{k+1}\} \subseteq \operatorname{span}\{r_0 \dots A^{k+1} r_0\}$$

вложение в одну сторону

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

$\in \operatorname{span}\{p_0 \dots p_k\}$ по предп.

$$A p_i = \frac{r_{i+1} - r_i}{\alpha_i}$$

$$A^{k+1} r_0 \in \operatorname{span}\{r_0 \dots r_{k+1}\}$$

$$\operatorname{span}\{r_0 \dots A^{k+1} r_0\} \subseteq \operatorname{span}\{r_0 \dots r_{k+1}\}$$

вложение в обр. сторону

∎ первый факт доказан.

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

по правилу подсчета $p_{k+1}$

по предположению 2:

$$\operatorname{span}\{p_0 \dots p_{k+1}\} = \operatorname{span}\{r_0 \dots A^k r_0, r_{k+1}\}$$

по предположению 1:

$$\operatorname{span}\{p_0 \dots p_{k+1}\} = \operatorname{span}\{r_0 \dots r_k, r_{k+1}\}$$

по доказанному

$$\operatorname{span}\{p_0 \dots p_{k+1}\} = \operatorname{span}\{r_0 \dots A^{k+1} r_0\}$$

∎ второй факт доказан

---

Продолжение док-ва:

Нужно доказать $-r_{k+1}^T A p_i = 0 \quad i < k$

$$p_i \in \operatorname{span}\{r_0 \dots A^i r_0\}$$

$$A p_i \in \underbrace{\operatorname{span}\{A r_0 \dots A^{i+1} r_0\}}_{\subseteq \operatorname{span}\{p_0 \dots p_{i+1}\}}$$

$$r_{k+1}^T p_i = 0 \quad i \le k$$

(по инд. предп.-си. выше)

∎

**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.

Изначально метод придумывался как точный

**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:

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

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

сходится как тяжелый шарик и Нестеров для кв. задачи

Обобщим на произв. задачу

**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**

$\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k}$

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

$r_{k+1} = r_k + \alpha_k A p_k$

$\beta_{k+1} = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$

$p_{k+1} = -r_{k+1} + \beta_{k+1} p_k$

**end for**

**Выход:** $x^K$

$\nabla f(x^k) = r_k$

меняем

**Алгоритм 8** 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:** $x^K$