Лекция 14. Итерационные методы для решения систем линейных уравнений

Максим Рахуба

ФКН ВШЭ
Основы матричных вычислений 20/21

Май 11, 2021

Введение: Прямые vs. итерационные методы

Дано: невырожденная $\mathbf{A} \in \mathbb{R}^{n \times n}$, вектор $\mathbf{b} \in \mathbb{R}^{n}$.
Цель: найти $\mathbf{x} \in \mathbb{R}^{n}$: $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Введение: Прямые vs. итерационные методы

Дано: невырожденная $\mathbf{A} \in \mathbb{R}^{n\times n}$, вектор $\mathbf{b} \in \mathbb{R}^n$.
Цель: найти $\mathbf{x} \in \mathbb{R}^n$: $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Прямые методы (LU разложение)

- Подходят и для плотных, и для разреженных $\mathbf{A}$.
- Для плотных: $\mathcal{O}(n^3)$ FLOPs и $\mathcal{O}(n^2)$ память

n	10^3	10^4	10^5
Time	0.4 s	6 s	1.7 h
Mem($\mathbf{A}$)	8 Mb	0.8 Gb	80 Gb

- Для разреженных матриц обычно невозможно добиться оптимальной сложности $\mathcal{O}(n)$.

Введение: Прямые vs. итерационные методы

Дано: невырожденная $\mathbf{A} \in \mathbb{R}^{n \times n}$, вектор $\mathbf{b} \in \mathbb{R}^n$.
Цель: найти $\mathbf{x} \in \mathbb{R}^n$: $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Прямые методы (LU разложение)

- Подходят и для плотных, и для разреженных $\mathbf{A}$.
- Для плотных: $\mathcal{O}(n^3)$ FLOPs и $\mathcal{O}(n^2)$ память

n	10^3	10^4	10^5
Time	0.4 s	6 s	1.7 h
Mem($\mathbf{A}$)	8 Mb	0.8 Gb	80 Gb

Сначала закончится память

- Для разреженных матриц обычно невозможно добиться оптимальной сложности $\mathcal{O}(n)$.

Итерационные методы

Пример: метод простой итерации (Richardson iteration):

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau(\overbrace{\mathbf{b} - \mathbf{A}\mathbf{x}_k}^{r_k}), \quad \tau \in \mathbb{R}.$$

- Сложность определяется вычислением $\mathbf{A}\mathbf{x}_k$.
- Количество итераций $K = K(n, \varepsilon)$

Введение: Прямые vs. итерационные методы

Дано: невырожденная $\mathbf{A} \in \mathbb{R}^{n\times n}$, вектор $\mathbf{b} \in \mathbb{R}^n$.
Цель: найти $\mathbf{x} \in \mathbb{R}^n$: $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Прямые методы (LU разложение)

- Подходят и для плотных, и для разреженных $\mathbf{A}$.
- Для плотных: $\mathcal{O}(n^3)$ FLOPs и $\mathcal{O}(n^2)$ память

n	10^3	10^4	10^5
Time	0.4 s	6 s	1.7 h
Mem$(\mathbf{A})$	8 Mb	0.8 Gb	80 Gb

Сначала закончится память

- Для разреженных матриц обычно невозможно добиться оптимальной сложности $\mathcal{O}(n)$.

Итерационные методы

Пример: метод простой итерации (Richardson iteration):

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau(\mathbf{b} - \mathbf{A}\mathbf{x}_k), \quad \tau \in \mathbb{R}.$$

- Сложность определяется вычислением $\mathbf{A}\mathbf{x}_k$.
- Количество итераций $K = K(n, \varepsilon)$

Введение: матвеки

Матрично-векторное произведение (матвек):

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

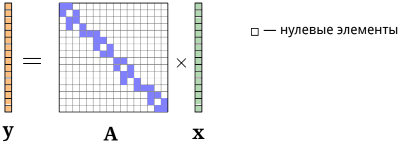
- Произвольная плотная $\mathbf{A}$: $\mathcal{O}(n^2)$ операций и памяти.

Введение: матвеки

Матрично-векторное произведение (матвек):

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

- Произвольная плотная $\mathbf{A}$: $\mathcal{O}(n^2)$ операций и памяти.
- Разреженная $\mathbf{A}$: $\mathcal{O}(\mathtt{nnz}(\mathbf{A}))$ операций и памяти, где $\mathtt{nnz}(\mathbf{A})$ — число ненулевых элементов (#nonzeros).

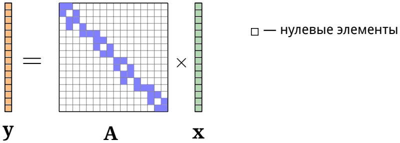


Введение: матвеки

Матрично-векторное произведение (матвек):

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

- Произвольная плотная $\mathbf{A}$: $\mathcal{O}(n^2)$ операций и памяти.
- Разреженная $\mathbf{A}$: $\mathcal{O}(\mathtt{nnz}(\mathbf{A}))$ операций и памяти, где $\mathtt{nnz}(\mathbf{A})$ — число ненулевых элементов (#nonzeros).



- Сложность меньше $\mathcal{O}(n^2)$ для некоторых структурированных матриц (Фурье, теплицевы, циркулянты, ганкелевы, ...), а также для матриц с блоками малого ранга ($\mathcal{H}$, $\mathcal{H}^2$ матрицы).

Одношаговые итерационные методы

Сегодня будем рассматривать одношаговые методы:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \mathbf{P}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k),$$

где $\tau_k \in \mathbb{R}$ и $\mathbf{P} \in \mathbb{R}^{n \times n}$ подбираются для ускорения сходимости.

Одношаговые итерационные методы

Сегодня будем рассматривать одношаговые методы:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \mathbf{P}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k),$$

где $\tau_k \in \mathbb{R}$ и $\mathbf{P} \in \mathbb{R}^{n \times n}$ подбираются для ускорения сходимости.

- Если $\tau_k \equiv 1$, $\mathbf{P} = \mathbf{A}$, то сойдется за 1 итерацию, но умножить на $\mathbf{A}^{-1}$ так же сложно, как и решить $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Одношаговые итерационные методы

Сегодня будем рассматривать одношаговые методы:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \mathbf{P}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k),$$

где $\tau_k \in \mathbb{R}$ и $\mathbf{P} \in \mathbb{R}^{n \times n}$ подбираются для ускорения сходимости.

- Если $\tau_k \equiv 1$, $\mathbf{P} = \mathbf{A}$, то сойдется за 1 итерацию, но умножить на $\mathbf{A}^{-1}$ так же сложно, как и решить $\mathbf{A}\mathbf{x} = \mathbf{b}$.
- $\mathbf{P}$ надо выбирать “близким” к $\mathbf{A}$, но, при этом, чтобы на $\mathbf{P}^{-1}$ было быстро умножать. Например,

Одношаговые итерационные методы

Сегодня будем рассматривать одношаговые методы:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \mathbf{P}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k),$$

где $\tau_k \in \mathbb{R}$ и $\mathbf{P} \in \mathbb{R}^{n \times n}$ подбираются для ускорения сходимости.

- Если $\tau_k \equiv 1$, $\mathbf{P} = \mathbf{A}$, то сойдется за 1 итерацию, но умножить на $\mathbf{A}^{-1}$ так же сложно, как и решить $\mathbf{A}\mathbf{x} = \mathbf{b}$.
- $\mathbf{P}$ надо выбирать “близким” к $\mathbf{A}$, но, при этом, чтобы на $\mathbf{P}^{-1}$ было быстро умножать. Например,
  - $\mathbf{P}$ – диагональная часть $\mathbf{A}$ (метод Якоби при $\tau_k = 1$)

Одношаговые итерационные методы

Сегодня будем рассматривать одношаговые методы:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \mathbf{P}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k),$$

где $\tau_k \in \mathbb{R}$ и $\mathbf{P} \in \mathbb{R}^{n\times n}$ подбираются для ускорения сходимости.

- Если $\tau_k \equiv 1$, $\mathbf{P} = \mathbf{A}$, то сойдется за 1 итерацию, но умножить на $\mathbf{A}^{-1}$ так же сложно, как и решить $\mathbf{A}\mathbf{x} = \mathbf{b}$.
- $\mathbf{P}$ надо выбирать "близким" к $\mathbf{A}$, но, при этом, чтобы на $\mathbf{P}^{-1}$ было быстро умножать. Например,
  - $\mathbf{P}$ – диагональная часть $\mathbf{A}$ (метод Якоби при $\tau_k = 1$)
  - $\mathbf{P}$ – нижнетреугольная часть $\mathbf{A}$ (Гаусс-Зейдель при $\tau_k = 1$)

Одношаговые итерационные методы

Сегодня будем рассматривать одношаговые методы:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \mathbf{P}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k),$$

где $\tau_k \in \mathbb{R}$ и $\mathbf{P} \in \mathbb{R}^{n\times n}$ подбираются для ускорения сходимости.

- Если $\tau_k \equiv 1$, $\mathbf{P} = \mathbf{A}$, то сойдется за 1 итерацию, но умножить на $\mathbf{A}^{-1}$ так же сложно, как и решить $\mathbf{A}\mathbf{x} = \mathbf{b}$.
- $\mathbf{P}$ надо выбирать “близким” к $\mathbf{A}$, но, при этом, чтобы на $\mathbf{P}^{-1}$ было быстро умножать. Например,
  - $\mathbf{P}$ – диагональная часть $\mathbf{A}$ (метод Якоби при $\tau_k = 1$)
  - $\mathbf{P}$ – нижнетреугольная часть $\mathbf{A}$ (Гаусс-Зейдель при $\tau_k = 1$)
- Будем рассматривать $\tau_k = \tau$ (метод простой итерации) и переменное τ_k (градиентный спуск, метод Чебышёва).

План

Метод простой итерации

Градиентный спуск

Чебышёвский итерационный метод

Метод простой итерации

Положим $\tau_k \equiv \tau$ и перепишем в следующем виде:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau\,\mathbf{P}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k) \equiv \mathbf{G}\mathbf{x}_k + \mathbf{g},$$

где $\mathbf{G} = (\mathbf{I} - \tau\,\mathbf{P}^{-1}\mathbf{A})$ и $\mathbf{g} = \tau\mathbf{P}^{-1}\mathbf{b}$.

Теорема

Итерационный процесс $\mathbf{x}_{k+1} = \mathbf{G}\mathbf{x}_k + \mathbf{g}$, *сходится к единственному решению для любого начального приближения* $\mathbf{x}_0$ *и любого* $\mathbf{g}$ *тогда и только тогда, когда* $\rho(\mathbf{G}) < 1$.

$$x_1 = G x_0 + g$$

$$x_2 = G x_1 + g = G(G x_0 + g) + g = G^2 x_0 + G g + g$$

$$\vdots$$

$$x_k = G^k x_0 + \underbrace{G^{k-1} g + \ldots + G g + G}_{\sum_{i=0}^{k-1} G^i g}$$

Следствие

Если $\|G\| < 1$ *для некоторой матричной* $\|\cdot\|$, *то* $\mathbf{x}_{k+1} = \mathbf{G}\mathbf{x}_k + \mathbf{g}$ *сходится к единственному решению для любого* $\mathbf{x}_0$.

Метод простой итерации

Пусть $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$, где $\mathbf{D} = \mathrm{diag}(\mathbf{A})$, а $\mathbf{L}, \mathbf{U}$ – нижне- и верхнетреугольная части $\mathbf{A}$ без диагонали.

Метод простой итерации

Пусть $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$, где $\mathbf{D} = \mathrm{diag}(\mathbf{A})$, а $\mathbf{L}, \mathbf{U}$ – нижне- и верхнетреугольная части $\mathbf{A}$ без диагонали.

Утверждение (сходимость метода Якоби)

Пусть матрица $\mathbf{A}$ обладает свойством строгого строчного диагонального преобладания: $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$, тогда процесс:

$$\mathbf{x}_{k+1} = \underbrace{\cancel{\mathbf{x}_k} + \mathbf{D}^{-1}(\mathbf{b} - \overset{= L + \cancel{D} + U}{\mathbf{A}}\mathbf{x}_k)}_{= D^{-1}(b - (L+U)x_k)}, \quad \textit{(метод Якоби)}$$

— альтернативная запись

сходится для любого $\mathbf{x}_0$.

$$\square \quad \|I - D^{-1}A\|_\infty = \|D^{-1}(L+U)\|_\infty = \frac{\sum_{j\neq i} |a_{ij}|}{|a_{ii}|} < 1$$

Метод простой итерации

Пусть $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$, где $\mathbf{D} = \mathrm{diag}(\mathbf{A})$, а $\mathbf{L}, \mathbf{U}$ – нижне- и верхнетреугольная части $\mathbf{A}$ без диагонали.

Утверждение (сходимость метода Якоби)

Пусть матрица $\mathbf{A}$ обладает свойством строгого строчного диагонального преобладания: $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$, тогда процесс:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{D}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k), \quad \textit{(метод Якоби)}$$

сходится для любого $\mathbf{x}_0$.

Утверждение (сходимость метода Гаусса-Зейделя)

Пусть $\mathbf{A} = \mathbf{A}^\top > 0$. Тогда процесс:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + (\mathbf{L} + \mathbf{D})^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_k), \quad \textit{(метод Гаусса-Зейделя)}$$

сходится для любого $\mathbf{x}_0$.

$\square$ Идея $\|I - (L+D)^{-1}A\|_A$, $\|x\|_A = \sqrt{x^T A x}$

(см. семинар)

Метод простой итерации

Рассмотрим итерационный процесс

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau(\mathbf{b} - \mathbf{A}\mathbf{x}_k).$$

Утверждение (сходимость метода простой итерации)

Пусть $\mathbf{A} = \mathbf{A}^\intercal > 0$ с собственными значениями $\lambda_1 \geq \dots \geq \lambda_n > 0$.

1. *Если $0 < \tau < 2/\lambda_1$, то $\{\mathbf{x}_k\}$ сходится для любого $\mathbf{x}_0$.*
2. *$\tau = \tau_{\text{opt}} \equiv \frac{2}{\lambda_1 + \lambda_n}$ минимизирует $\|\mathbf{I} - \tau\mathbf{A}\|_2$, и для ошибки $\mathbf{e}_k \equiv \mathbf{x} - \mathbf{x}_k$ справедливо*

// 100

$$\|\mathbf{e}_{k+1}\|_2 \leq \frac{\operatorname{cond}_2(\mathbf{A}) - 1}{\underbrace{\operatorname{cond}_2(\mathbf{A}) + 1}}\|\mathbf{e}_k\|_2$$

$\frac{99}{100} = 0.99$

Метод простой итерации

Рассмотрим итерационный процесс

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau(\mathbf{b} - \mathbf{A}\mathbf{x}_k).$$

Утверждение (сходимость метода простой итерации)

Пусть $\mathbf{A} = \mathbf{A}^\mathsf{T} > 0$ с собственными значениями $\lambda_1 \geq \cdots \geq \lambda_n > 0$.

1. *Если $0 < \tau < 2/\lambda_1$, то $\{\mathbf{x}_k\}$ сходится для любого $\mathbf{x}_0$.*
2. $\tau = \tau_{\text{opt}} \equiv \frac{2}{\lambda_1 + \lambda_n}$ *минимизирует $\|\mathbf{I} - \tau\mathbf{A}\|_2$, и для ошибки $\mathbf{e}_k \equiv \mathbf{x} - \mathbf{x}_k$ справедливо*

$$\|\mathbf{e}_{k+1}\|_2 \leq \frac{\operatorname{cond}_2(\mathbf{A}) - 1}{\operatorname{cond}_2(\mathbf{A}) + 1}\|\mathbf{e}_k\|_2$$

Доказательство.

- $\|\mathbf{x} - \mathbf{x}_{k+1}\|_2 = \|\mathbf{x} - \mathbf{x}_k - \tau(\overset{=Ax}{\mathbf{b}} - \mathbf{A}\mathbf{x}_k)\|_2 = \|\mathbf{x} - \mathbf{x}_k - \tau(\mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{x}_k)\|_2 = \|(\mathbf{I} - \tau\mathbf{A})(\mathbf{x} - \mathbf{x}_k)\|_2 \leq \|\mathbf{I} - \tau\mathbf{A}\|_2\|\mathbf{x} - \mathbf{x}_k\|_2$.

$\square$

Метод простой итерации

Рассмотрим итерационный процесс

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau(\mathbf{b} - \mathbf{A}\mathbf{x}_k).$$

Утверждение (сходимость метода простой итерации)

Пусть $\mathbf{A} = \mathbf{A}^\intercal > 0$ *с собственными значениями* $\lambda_1 \geq \cdots \geq \lambda_n > 0$.

1. *Если* $0 < \tau < 2/\lambda_1$, *то* $\{\mathbf{x}_k\}$ *сходится для любого* $\mathbf{x}_0$.
2. $\tau = \tau_{\text{opt}} \equiv \frac{2}{\lambda_1 + \lambda_n}$ *минимизирует* $\|\mathbf{I} - \tau\mathbf{A}\|_2$, *и для ошибки* $\mathbf{e}_k \equiv \mathbf{x} - \mathbf{x}_k$ *справедливо*

$$\|\mathbf{e}_{k+1}\|_2 \leq \frac{\mathrm{cond}_2(\mathbf{A}) - 1}{\mathrm{cond}_2(\mathbf{A}) + 1} \|\mathbf{e}_k\|_2$$

Доказательство.

- $\|\mathbf{x} - \mathbf{x}_{k+1}\|_2 = \|\mathbf{x} - \mathbf{x}_k - \tau(\mathbf{b} - \mathbf{A}\mathbf{x}_k)\|_2 = \|\mathbf{x} - \mathbf{x}_k - \tau(\mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{x}_k)\|_2 = \|(\mathbf{I} - \tau\mathbf{A})(\mathbf{x} - \mathbf{x}_k)\|_2 \leq \|\mathbf{I} - \tau\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{x}_k\|_2$.
- $\|\mathbf{I} - \tau\mathbf{A}\|_2 = \max_i |1 - \tau\lambda_i| < 1 \;\Rightarrow\; -1 < 1 - \tau\lambda_i < 1 \;\Rightarrow\; 0 < \tau < 2/\lambda_i \, \forall i$.

□

Метод простой итерации

Рассмотрим итерационный процесс

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau(\mathbf{b} - \mathbf{A}\mathbf{x}_k).$$

Утверждение (сходимость метода простой итерации)

Пусть $\mathbf{A} = \mathbf{A}^\intercal > 0$ *с собственными значениями* $\lambda_1 \geq \cdots \geq \lambda_n > 0$.

1. *Если* $0 < \tau < 2/\lambda_1$, *то* $\{\mathbf{x}_k\}$ *сходится для любого* $\mathbf{x}_0$.
2. $\tau = \tau_{\text{opt}} \equiv \frac{2}{\lambda_1 + \lambda_n}$ *минимизирует* $\|\mathbf{I} - \tau\mathbf{A}\|_2$, *и для ошибки* $\mathbf{e}_k \equiv \mathbf{x} - \mathbf{x}_k$ *справедливо*

$$\|\mathbf{e}_{k+1}\|_2 \leq \frac{\operatorname{cond}_2(\mathbf{A}) - 1}{\operatorname{cond}_2(\mathbf{A}) + 1}\|\mathbf{e}_k\|_2$$

Доказательство.

- $\|\mathbf{x}-\mathbf{x}_{k+1}\|_2 = \|\mathbf{x}-\mathbf{x}_k-\tau(\mathbf{b}-\mathbf{A}\mathbf{x}_k)\|_2 = \|\mathbf{x}-\mathbf{x}_k-\tau(\mathbf{A}\mathbf{x}-\mathbf{A}\mathbf{x}_k)\|_2 = \|(\mathbf{I} - \tau\mathbf{A})(\mathbf{x} - \mathbf{x}_k)\|_2 \leq \|\mathbf{I} - \tau\mathbf{A}\|_2\|\mathbf{x} - \mathbf{x}_k\|_2$.
- $\|\mathbf{I} - \tau\mathbf{A}\|_2 = \max_i |1 - \tau\lambda_i| < 1 \;\Rightarrow\; -1 < 1 - \tau\lambda_i < 1 \;\Rightarrow\; 0 < \tau < 2/\lambda_i \forall i$.
- $\|I - \tau_{\text{opt}}\mathbf{A}\|_2 = \max_i \frac{|\lambda_1+\lambda_n-2\lambda_i|}{\lambda_1+\lambda_n} = \frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n} = \frac{\lambda_1/\lambda_n-1}{\lambda_1/\lambda_n+1} = \frac{\operatorname{cond}_2(\mathbf{A})-1}{\operatorname{cond}_2(\mathbf{A})+1}$.

$\square$

План

Метод простой итерации

Градиентный спуск τ

Чебышёвский итерационный метод

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

Предположение: $\mathbf{A} = \mathbf{A}^\intercal > 0$

Хотим, чтобы $\mathbf{x}_* = \mathbf{A}^{-1}\mathbf{b}$ было также решением некоторой задачи оптимизации:

$$\mathbf{x}_* = \arg\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}).$$

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

Предположение: $\mathbf{A} = \mathbf{A}^{\intercal} > 0$

Хотим, чтобы $\mathbf{x}_* = \mathbf{A}^{-1}\mathbf{b}$ было также решением некоторой задачи оптимизации:

$$\mathbf{x}_* = \arg\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}).$$

Выбор $f(\mathbf{x})$:

- $\|\mathbf{x} - \mathbf{x}_*\|_2^2$ — необходимо знать $\mathbf{x}_*$.

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

Предположение: $\mathbf{A} = \mathbf{A}^\intercal > 0$

Хотим, чтобы $\mathbf{x}_* = \mathbf{A}^{-1}\mathbf{b}$ было также решением некоторой задачи оптимизации:

$$\mathbf{x}_* = \arg\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}).$$

Выбор $f(\mathbf{x})$:

- $\|\mathbf{x} - \mathbf{x}_*\|_2^2$ — необходимо знать $\mathbf{x}_*$.
- $\|\mathbf{x} - \mathbf{x}_*\|_\mathbf{A}^2$, где $\|\mathbf{x}\|_\mathbf{A} \overset{\text{def}}{=} \sqrt{\mathbf{x}^\intercal\mathbf{A}\mathbf{x}}$:

$(x, y)_A = x^T A y$

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

Предположение: $\mathbf{A} = \mathbf{A}^\intercal > 0$

Хотим, чтобы $\mathbf{x}_* = \mathbf{A}^{-1}\mathbf{b}$ было также решением некоторой задачи оптимизации:

$$\mathbf{x}_* = \arg\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}).$$

Выбор $f(\mathbf{x})$:

- $\|\mathbf{x} - \mathbf{x}_*\|_2^2$ — необходимо знать $\mathbf{x}_*$.
- $\|\mathbf{x} - \mathbf{x}_*\|_\mathbf{A}^2$, где $\|\mathbf{x}\|_\mathbf{A} \overset{\text{def}}{=} \sqrt{\mathbf{x}^\intercal\mathbf{A}\mathbf{x}}$:

$$\begin{aligned}\|\mathbf{x} - \mathbf{x}_*\|_\mathbf{A}^2 &= (\mathbf{x} - \mathbf{x}_*)^\intercal\mathbf{A}(\mathbf{x} - \mathbf{x}_*) = \\ &= \mathbf{x}^\intercal\mathbf{A}\mathbf{x} - 2\mathbf{x}^\intercal\mathbf{A}\mathbf{x}_* + \mathbf{x}_*^\intercal\mathbf{A}\mathbf{x}_* = \\ &= \mathbf{x}^\intercal\mathbf{A}\mathbf{x} - 2\mathbf{x}^\intercal\mathbf{b} + \text{const}.\end{aligned}$$

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

Предположение: $\mathbf{A} = \mathbf{A}^\intercal > 0$

Хотим, чтобы $\mathbf{x}_* = \mathbf{A}^{-1}\mathbf{b}$ было также решением некоторой задачи оптимизации:

$$\mathbf{x}_* = \arg\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}).$$

Выбор $f(\mathbf{x})$:

- $\|\mathbf{x} - \mathbf{x}_*\|_2^2$ — необходимо знать $\mathbf{x}_*$.
- $\|\mathbf{x} - \mathbf{x}_*\|_{\mathbf{A}}^2$, где $\|\mathbf{x}\|_{\mathbf{A}} \overset{\text{def}}{=} \sqrt{\mathbf{x}^\intercal\mathbf{A}\mathbf{x}}$:

$$\begin{aligned}
\|\mathbf{x} - \mathbf{x}_*\|_{\mathbf{A}}^2 &= (\mathbf{x} - \mathbf{x}_*)^\intercal\mathbf{A}(\mathbf{x} - \mathbf{x}_*) = \\
&= \mathbf{x}^\intercal\mathbf{A}\mathbf{x} - 2\mathbf{x}^\intercal\mathbf{A}\mathbf{x}_* + \mathbf{x}_*^\intercal\mathbf{A}\mathbf{x}_* = \\
&= \mathbf{x}^\intercal\mathbf{A}\mathbf{x} - 2\mathbf{x}^\intercal\mathbf{b} + \text{const}.
\end{aligned}$$

Таким образом, будем использовать $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\intercal\mathbf{A}\mathbf{x} - \mathbf{x}^\intercal\mathbf{b}$.

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

Предположение: $\mathbf{A} = \mathbf{A}^\intercal > 0$

Хотим, чтобы $\mathbf{x}_* = \mathbf{A}^{-1}\mathbf{b}$ было также решением некоторой задачи оптимизации:

$$\mathbf{x}_* = \arg\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}).$$

Выбор $f(\mathbf{x})$:

- $\|\mathbf{x} - \mathbf{x}_*\|_2^2$ — необходимо знать $\mathbf{x}_*$.
- $\|\mathbf{x} - \mathbf{x}_*\|_{\mathbf{A}}^2$, где $\|\mathbf{x}\|_{\mathbf{A}} \overset{\text{def}}{=} \sqrt{\mathbf{x}^\intercal\mathbf{A}\mathbf{x}}$:

$$\begin{aligned}\|\mathbf{x} - \mathbf{x}_*\|_{\mathbf{A}}^2 &= (\mathbf{x} - \mathbf{x}_*)^\intercal\mathbf{A}(\mathbf{x} - \mathbf{x}_*) = \\ &= \mathbf{x}^\intercal\mathbf{A}\mathbf{x} - 2\mathbf{x}^\intercal\mathbf{A}\mathbf{x}_* + \mathbf{x}_*^\intercal\mathbf{A}\mathbf{x}_* = \\ &= \mathbf{x}^\intercal\mathbf{A}\mathbf{x} - 2\mathbf{x}^\intercal\mathbf{b} + \text{const}.\end{aligned}$$

Таким образом, будем использовать $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\intercal\mathbf{A}\mathbf{x} - \mathbf{x}^\intercal\mathbf{b}$.

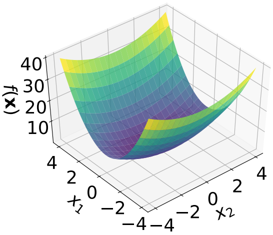
$\mathbf{x}_*$ – точка глобального мин. $\|\mathbf{x} - \mathbf{x}_*\|_{\mathbf{A}}^2$ и $f(\mathbf{x})$ по построению.

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

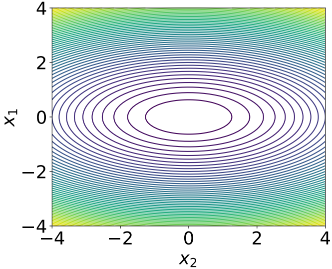
$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\intercal \mathbf{A}\mathbf{x} - \mathbf{x}^\intercal \mathbf{b}.$$

Пример

$$\mathbf{A} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad f(\mathbf{x}) = \frac{1}{2}\left(4x_1^2 + x_2^2\right) - 0.$$



Функция $f(\mathbf{x})$



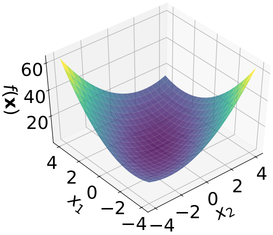
Изолинии
$f(\mathbf{x}) = \text{const}$

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

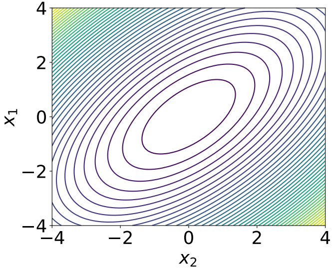
$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\intercal \mathbf{A}\mathbf{x} - \mathbf{x}^\intercal \mathbf{b}.$$

Пример

$$\mathbf{A} = \mathbf{U}\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}\mathbf{U}^\intercal, \quad \mathbf{U} = \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$



Функция $f(\mathbf{x})$



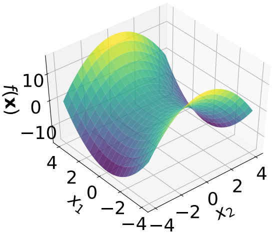
Изолинии
$f(\mathbf{x}) = \text{const}$

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

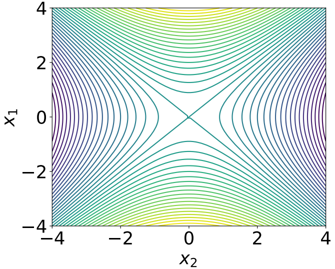
$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\intercal \mathbf{A}\mathbf{x} - \mathbf{x}^\intercal \mathbf{b}.$$

Пример

$$\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad f(\mathbf{x}) = \frac{1}{2}\left(2x_1^2 - 2x_2^2\right) - 0$$



Функция $f(\mathbf{x})$



Изолинии
$f(\mathbf{x}) = \text{const}$

$\mathbf{Ax} = \mathbf{b}$ как задача минимизации

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

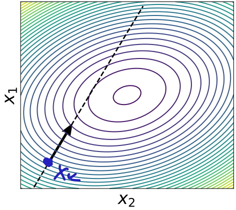
непрерывно дифференцируема и $\nabla f = \mathbf{Ax} - \mathbf{b}$. Действительно,

$$\begin{aligned} f(\mathbf{x}+\mathbf{h}) &= \frac{1}{2}(\mathbf{x}+\mathbf{h})^\intercal \mathbf{A}(\mathbf{x}+\mathbf{h}) - (\mathbf{x}+\mathbf{h})^\intercal \mathbf{b} = \\ &= \frac{1}{2}\mathbf{x}^\intercal \mathbf{A}\mathbf{x} + \mathbf{x}^\intercal \mathbf{A}\mathbf{h} + \frac{1}{2}\mathbf{h}^\intercal \mathbf{A}\mathbf{h} - \mathbf{x}^\intercal \mathbf{b} - \mathbf{h}^\intercal \mathbf{b} = \qquad (\clubsuit) \\ &= f(\mathbf{x}) + (\mathbf{Ax} - \mathbf{b})^\intercal \mathbf{h} + \frac{1}{2}\mathbf{h}^\intercal \mathbf{A}\mathbf{h}. \end{aligned}$$

Градиентный спуск

Хотим улучшить $\mathbf{x}_k$, используя направление $\mathbf{d}_k$ и длину шага τ_k:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \mathbf{d}_k.$$



Вспомним формулу (♣):

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^\intercal \mathbf{h} + \frac{1}{2}\mathbf{h}^\intercal \mathbf{A}\mathbf{h}, \quad \nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}.$$

Выбор направления $\mathbf{d}_k$

Хотим найти $\mathbf{d} : \|\mathbf{d}\|_2 = 1$, который минимизирует производную по направлению:

$$\nabla_{\mathbf{d}} f(\mathbf{x}) = \lim_{\tau \to 0} \frac{f(\mathbf{x}+\tau\mathbf{d}) - f(\mathbf{x})}{\tau} \overset{(♣)}{=} \nabla f(\mathbf{x})^\intercal \mathbf{d} = \|\nabla f(\mathbf{x})\|_2 \cos\theta \to \min_\theta$$

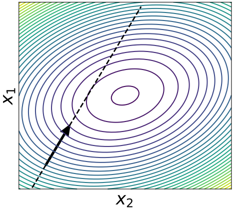
Значит, $\mathbf{d}_k = -\nabla f(\mathbf{x}_k) = \mathbf{b} - \mathbf{A}\mathbf{x}_k = \mathbf{r}_k$ — вектор невязки.

Метод наискорейшего спуска

(градиентный спуск с оптимальным выбором τ_k)

Мы выбрали $\mathbf{d}_k = -\nabla f(\mathbf{x}_k) = \mathbf{r}_k$.
Найдем τ_k (line search):

$$\tau_k = \arg\min_{\tau \in \mathbb{R}} f(\mathbf{x}_k + \tau \mathbf{r}_k).$$



Вспомним формулу (♣):

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^\intercal \mathbf{h} + \frac{1}{2}\mathbf{h}^\intercal \mathbf{A}\mathbf{h}, \quad \nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}.$$

Поиск τ_k

$$f(\mathbf{x}_k + \tau \mathbf{r}_k) \overset{(♣)}{=} f(\mathbf{x}_k) - \tau \mathbf{r}_k^\intercal \mathbf{r}_k + \frac{\tau^2}{2} \mathbf{r}_k^\intercal \mathbf{A} \mathbf{r}_k.$$

$$\frac{\partial f(\mathbf{x}_k + \tau \mathbf{r}_k)}{\partial \tau} = -\mathbf{r}_k^\intercal \mathbf{r}_k + \tau \mathbf{r}_k^\intercal \mathbf{A} \mathbf{r}_k \overset{!}{=} 0 \quad \Rightarrow \quad \tau_k = \frac{\mathbf{r}_k^\intercal \mathbf{r}_k}{\mathbf{r}_k^\intercal \mathbf{A} \mathbf{r}_k}.$$

$$\frac{\partial^2 f(\mathbf{x}_k + \tau \mathbf{r}_k)}{\partial \tau^2} = \mathbf{r}_k^\intercal \mathbf{A} \mathbf{r}_k > 0.$$

Метод наискорейшего спуска

Метод можно записать как:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \big(\underbrace{\mathbf{b} - \mathbf{A}\mathbf{x}_k}_{\mathbf{r}_k} \big), \quad \tau_k = \frac{\mathbf{r}_k^\intercal \mathbf{r}_k}{\mathbf{r}_k^\intercal \mathbf{A}\mathbf{r}_k}.$$

Метод наискорейшего спуска

Метод можно записать как:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \big(\underbrace{\mathbf{b} - \mathbf{A}\mathbf{x}_k}_{\mathbf{r}_k} \big), \quad \tau_k = \frac{\mathbf{r}_k^\mathsf{T} \mathbf{r}_k}{\mathbf{r}_k^\mathsf{T} \mathbf{A} \mathbf{r}_k}.$$

Эффективный алгоритм

- Проблема: 2 матвека ($\mathbf{A}\mathbf{x}_k$ и $\mathbf{A}\mathbf{r}_k$) на каждой итерации.
- Наблюдение: $\mathbf{r}_{k+1} = \mathbf{r}_k - \tau_k \mathbf{A}\mathbf{r}_k$.

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$
for $k = 0, 1, \ldots$ until convergence **do**

$$\tau_k = \frac{\mathbf{r}_k^\mathsf{T} \mathbf{r}_k}{\mathbf{r}_k^\mathsf{T} \mathbf{A} \mathbf{r}_k}$$
$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k \mathbf{r}_k$$
$$\mathbf{r}_{k+1} = \mathbf{r}_k - \tau_k \mathbf{A}\mathbf{r}_k$$

Критерий остановки: $k > k_{\max}$ или $\|\mathbf{A}\mathbf{x}_k - \mathbf{b}\|_2 \leq \mathtt{tol}$ $\| x_{k+1} - x_k \|$

Сходимость метода

Теорема (Сходимость наискорейшего спуска)

Пусть $\mathbf{A} = \mathbf{A}^\intercal > 0$ *и* $\mathbf{x}_k$ *сгенерировано с помощью метода наискорейшего спуска, тогда для ошибки* $\mathbf{e}_k := \mathbf{x}_* - \mathbf{x}_k$ *справедливо:*

$$\|\mathbf{e}_{k+1}\|_{\mathbf{A}} \leq \frac{\mathrm{cond}_2(\mathbf{A}) - 1}{\mathrm{cond}_2(\mathbf{A}) + 1} \|\mathbf{e}_k\|_{\mathbf{A}},$$

Доказательство.

1. $2f(\mathbf{x}) + \mathrm{const} = \|\mathbf{x} - \mathbf{x}_*\|_{\mathbf{A}}^2$, так что

$$\tau_k = \arg\min_{\tau} f(\mathbf{x}_k + \tau \mathbf{r}_k) = \arg\min_{\tau} \|\mathbf{x}_k + \tau \mathbf{r}_k - \mathbf{x}_*\|_{\mathbf{A}}^2.$$

□

Сходимость метода

Теорема (Сходимость наискорейшего спуска)

Пусть $\mathbf{A} = \mathbf{A}^\intercal > 0$ *и* $\mathbf{x}_k$ *сгенерировано с помощью метода наискорейшего спуска, тогда для ошибки* $\mathbf{e}_k := \mathbf{x}_* - \mathbf{x}_k$ *справедливо:*

$$\|\mathbf{e}_{k+1}\|_{\mathbf{A}} \le \frac{\operatorname{cond}_2(\mathbf{A}) - 1}{\operatorname{cond}_2(\mathbf{A}) + 1}\|\mathbf{e}_k\|_{\mathbf{A}},$$

Доказательство.

1. $2f(\mathbf{x}) + \text{const} = \|\mathbf{x} - \mathbf{x}_*\|_{\mathbf{A}}^2$, так что

$$\tau_k = \arg\min_\tau f(\mathbf{x}_k + \tau\mathbf{r}_k) = \arg\min_\tau \|\mathbf{x}_k + \tau\mathbf{r}_k - \mathbf{x}_*\|_{\mathbf{A}}^2.$$

$x_k - x_* = e_k$, $r_k = -Ae_k$

2. $\|\mathbf{e}_{k+1}\|_{\mathbf{A}}^2 \overset{1.}{=} \min_\tau\|\mathbf{x}_k+\tau\mathbf{r}_k-\mathbf{x}_*\|_{\mathbf{A}}^2 \overset{\forall t}{\le} \|\mathbf{x}_k+t\mathbf{r}_k-\mathbf{x}_*\|_{\mathbf{A}}^2 = \|(\mathbf{I}-t\mathbf{A})\mathbf{e}_k\|_{\mathbf{A}}^2$
$= \|\mathbf{A}^{1/2}(\mathbf{I} - t\mathbf{A})\mathbf{e}_k\|_2^2 = \|(\mathbf{I} - t\mathbf{A})\mathbf{A}^{1/2}\mathbf{e}_k\|_2^2 \le \|\mathbf{I} - t\mathbf{A}\|_2^2\,\|\mathbf{e}_k\|_{\mathbf{A}}^2.$

$\|y\|_A^2 = y^T A y = y^T A^{1/2} A^{1/2} y = (A^{1/2} y)^T A^{1/2} y = \|A^{1/2} y\|_2^2$ □

Сходимость метода

Теорема (Сходимость наискорейшего спуска)

Пусть $\mathbf{A} = \mathbf{A}^\intercal > 0$ *и* $\mathbf{x}_k$ *сгенерировано с помощью метода наискорейшего спуска, тогда для ошибки* $\mathbf{e}_k := \mathbf{x}_* - \mathbf{x}_k$ *справедливо:*

$$\|\mathbf{e}_{k+1}\|_\mathbf{A} \le \frac{\mathrm{cond}_2(\mathbf{A}) - 1}{\mathrm{cond}_2(\mathbf{A}) + 1} \|\mathbf{e}_k\|_\mathbf{A},$$

Доказательство.

1. $2f(\mathbf{x}) + \mathrm{const} = \|\mathbf{x} - \mathbf{x}_*\|^2_\mathbf{A}$, так что

$$\tau_k = \arg\min_\tau f(\mathbf{x}_k + \tau \mathbf{r}_k) = \arg\min_\tau \|\mathbf{x}_k + \tau\mathbf{r}_k - \mathbf{x}_*\|^2_\mathbf{A}.$$

2. $\|\mathbf{e}_{k+1}\|^2_\mathbf{A} \overset{1.}{=} \min\limits_\tau \|\mathbf{x}_k + \tau\mathbf{r}_k - \mathbf{x}_*\|^2_\mathbf{A} \overset{\forall t}{\le} \|\mathbf{x}_k + t\mathbf{r}_k - \mathbf{x}_*\|^2_\mathbf{A} = \|(\mathbf{I} - t\mathbf{A})\mathbf{e}_k\|^2_\mathbf{A}$
$= \|\mathbf{A}^{1/2}(\mathbf{I} - t\mathbf{A})\mathbf{e}_k\|^2_2 = \|(\mathbf{I} - t\mathbf{A})\mathbf{A}^{1/2}\mathbf{e}_k\|^2_2 \le \|\mathbf{I} - t\mathbf{A}\|^2_2 \, \|\mathbf{e}_k\|^2_\mathbf{A}.$

3. Выберем $t = \frac{2}{\lambda_1 + \lambda_n}$.

□

План

Метод простой итерации

Градиентный спуск

Чебышёвский итерационный метод

Чебышёвский итерационный метод

Хотим ускорить сходимость

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k(\mathbf{b} - \mathbf{A}\mathbf{x}_k), \qquad A = A^T > 0$$

по сравнению с простой итерацией и наискорейшим спуском.

Идея

Будем оптимизировать по всем итерациям:

$$\cancel{\le \|I - \tau_{k-1} A\| \|e_{k-1}\|}$$

$$\|\mathbf{e}_k\|_2 = \|(\mathbf{I} - \tau_{k-1}\mathbf{A})\mathbf{e}_{k-1}\|_2 = \|\overbrace{(\mathbf{I} - \tau_{k-1}\mathbf{A}) \dots (\mathbf{I} - \tau_0\mathbf{A})}^{p_k(A)}\mathbf{e}_0\|_2 =$$

$$= \|p_k(\mathbf{A})\mathbf{e}_0\|_2 \le \|p_k(\mathbf{A})\|_2 \|\mathbf{e}_0\|_2, \quad p_k \in \mathbb{P}_k, \quad p_k(0) = 1.$$

Чебышёвский итерационный метод

Хотим ускорить сходимость

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k(\mathbf{b} - \mathbf{A}\mathbf{x}_k),$$

по сравнению с простой итерацией и наискорейшим спуском.

Идея

Будем оптимизировать по всем итерациям:

$$\|\mathbf{e}_k\|_2 = \|(\mathbf{I} - \tau_{k-1}\mathbf{A})\mathbf{e}_{k-1}\|_2 = \|(\mathbf{I} - \tau_{k-1}\mathbf{A})\dots(\mathbf{I} - \tau_0\mathbf{A})\mathbf{e}_0\|_2 =$$
$$= \|p_k(\mathbf{A})\mathbf{e}_0\|_2 \le \|p_k(\mathbf{A})\|_2\|\mathbf{e}_0\|_2, \quad p_k \in \mathbb{P}_k, \quad p_k(0) = 1.$$

Из собственного разложения $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$ получим

$$\|p_k(\mathbf{A})\|_2 = \|\mathbf{U}p_k(\mathbf{\Lambda})\mathbf{U}^*\|_2 = \|p_k(\mathbf{\Lambda})\|_2 = \max_i |p_k(\lambda_i)| \le \max_{\lambda\in[\lambda_n,\lambda_1]} |p_k(\lambda)|.$$

Чебышёвский итерационный метод

Хотим ускорить сходимость

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k(\mathbf{b} - \mathbf{A}\mathbf{x}_k),$$

по сравнению с простой итерацией и наискорейшим спуском.

Идея

Будем оптимизировать по всем итерациям:

$$\|\mathbf{e}_k\|_2 = \|(\mathbf{I} - \tau_{k-1}\mathbf{A})\mathbf{e}_{k-1}\|_2 = \|(\mathbf{I} - \tau_{k-1}\mathbf{A})\dots(\mathbf{I} - \tau_0\mathbf{A})\mathbf{e}_0\|_2 =$$
$$= \|p_k(\mathbf{A})\mathbf{e}_0\|_2 \le \|p_k(\mathbf{A})\|_2\|\mathbf{e}_0\|_2, \quad p_k \in \mathbb{P}_k, \quad p_k(0) = 1.$$

Из собственного разложения $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$ получим

$$\|p_k(\mathbf{A})\|_2 = \|\mathbf{U}p_k(\mathbf{\Lambda})\mathbf{U}^*\|_2 = \|p_k(\mathbf{\Lambda})\|_2 = \max_i |p_k(\lambda_i)| \le \max_{\lambda\in[\lambda_n,\lambda_1]} |p_k(\lambda)|.$$

Итог – хотим решить минимакс задачу:

$$\min_{\substack{p_k\in\mathbb{P}_k \\ p_k(0)=1}} \max_{\lambda\in[\lambda_n,\lambda_1]} |p_k(\lambda)|,$$

то есть найти полином, наименее отклоняющийся от 0 на $[\lambda_n, \lambda_1]$.

Чебышёвский итерационный метод

Решение задачи

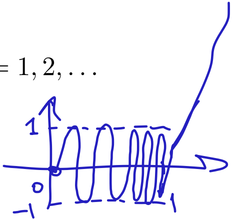
$$\min_{\substack{p_k \in \mathbb{P}_k \\ p_k(0)=1}} \max_{\lambda \in [\lambda_n, \lambda_1]} |p_k(\lambda)|,$$

можно выразить с помощью полиномов Чебышёва.

Полиномы Чебышёва

$$T_0(x) = 1, \quad T_1(x) = x,$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, \dots$$

Альтернативные формулы:



1. $T_n(x) = \begin{cases} \cos(n \arccos x), & |x| \le 1 \\ \mathrm{ch}(n\, \mathrm{arcch} x), & |x| > 1 \end{cases}$
2. $T_n(x) = \frac{1}{2}\left((x + \sqrt{x^2 - 1})^n + (x + \sqrt{x^2 - 1})^{-n}\right) \ \forall x \in \mathbb{C}$

Чебышёвский итерационный метод

Теорема

Пусть $\xi \notin [a, b]$. *Решение задачи*

$$\min_{\substack{p_k \in \mathbb{P}_k \\ p_k(\xi) = M}} \max_{\lambda \in [a,b]} |p_k(\lambda)|,$$

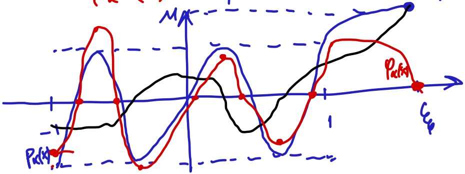
имеет вид

$$p_k(\lambda) = \frac{M}{T_k\left(\frac{2\xi - a - b}{b - a}\right)} T_k\left(\frac{2\lambda - a - b}{b - a}\right)$$

$\square$ $[-1,1] = [a,b]$, $\xi \notin [-1,1]$, $t = \dfrac{2x - a - b}{b - a}$

$\exists\, \tilde{p}_k(x): \quad \tilde{E} = \max\limits_{x \in [-1,1]} |\tilde{p}_k(x)| < E = \max\limits_{x \in [-1,1]} |p_k(x)|$

$\varphi_k(x) = p_k(x) - \tilde{p}_k(x)$



$n+1$ корней у $\varphi_k(x)$

$\Downarrow$

$\varphi_k(x) \equiv 0$ $\blacksquare$

Чебышёвский итерационный метод

Чебышёвский итерационный метод

- Обозначим $\sigma(x) = x + \sqrt{x^2 - 1}$, тогда $T_k(x) = \frac{1}{2}(\sigma(x)^k + \sigma(x)^{-k})$

Чебышёвский итерационный метод

- Обозначим $\sigma(x) = x + \sqrt{x^2 - 1}$, тогда $T_k(x) = \frac{1}{2}(\sigma(x)^k + \sigma(x)^{-k})$
- Применяя теорему при $M = 1$, $[a, b] = [\lambda_n, \lambda_1]$ и учитывая, что $|T_k(x)| \le 1$ при $x \in [-1, 1]$, получим:

$$p_k(\lambda) \le \frac{1}{\left|T_k\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)\right|} = \frac{2\sigma\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)^k}{1+\sigma\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)^{2k}} \le 2\sigma\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)^k,$$

$$= \frac{1}{\frac{1}{2}(\sigma^k + \sigma^{-k})}$$

где

$$\sigma \equiv \sigma\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right) = \frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n} + \sqrt{\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)^2 \cancel{+1} - 1} = \frac{\sqrt{\frac{\lambda_1}{\lambda_n}} - 1}{\sqrt{\frac{\lambda_1}{\lambda_n}} + 1}$$

Чебышёвский итерационный метод

- Обозначим $\sigma(x) = x + \sqrt{x^2 - 1}$, тогда $T_k(x) = \frac{1}{2}(\sigma(x)^k + \sigma(x)^{-k})$
- Применяя теорему при $M = 1$, $[a, b] = [\lambda_n, \lambda_1]$ и учитывая, что $|T_k(x)| \leq 1$ при $x \in [-1, 1]$, получим:

$$p_k(\lambda) \leq \frac{1}{T_k\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)} = \frac{2\sigma\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)^k}{1+\sigma\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)^{2k}} \leq 2\sigma\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)^k,$$

где

$$\sigma \equiv \sigma\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right) = \frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n} + \sqrt{\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)^2 \cancel{+1} \; {-1}} = \frac{\sqrt{\frac{\lambda_1}{\lambda_n}}-1}{\sqrt{\frac{\lambda_1}{\lambda_n}}+1}$$

$\|e_k\| \leq \frac{\mathrm{cond}-1}{\mathrm{cond}+1}\|e_{k-1}\| \leq \ldots \leq \left(\frac{\mathrm{cond}-1}{\mathrm{cond}+1}\right)^k \|e_0\|$

было

- Итог:

$$\|\mathbf{e}_k\|_2 \leq 2\left(\frac{\sqrt{\mathrm{cond}_2(\mathbf{A})}-1}{\sqrt{\mathrm{cond}_2(\mathbf{A})}+1}\right)^k \|\mathbf{e}_0\|_2,$$

где $1/\tau_k$ – корни $p_k(\lambda)$.

Литература

- Saad Y., Iterative methods for sparse linear systems. SIAM, 2003.
- Лебедев В.И. Функциональный анализ и вычислительная математика, Физматлит, 2005.
- Тыртышников Е. Е. Методы численного анализа. – Академия, 2007. – 320 с.