Приближение табличных функций

Градиентный метод

Шерухин Кирилл

5030102/30001
СПБПУ
2024

Содержание

1 Формулировка задачи и ее формализация

Дано:

- Симметричная невырожденная произвольная система линейных алгебраических уравнений из n уравнений и n неизвестных. ($AX = B; A = A^T$)
- Число $\epsilon > 0$, означающее допустимую ошибку решения СЛАУ.

.
Цель: Численно получить решение СЛАУ с заданной точностью ϵ

План выполнения:

1. Представить данную СЛАУ в матричном виде.
2. Получить решение градиентым методом методом с необходимой точностью.

2 Алгоритм и условия его применимости

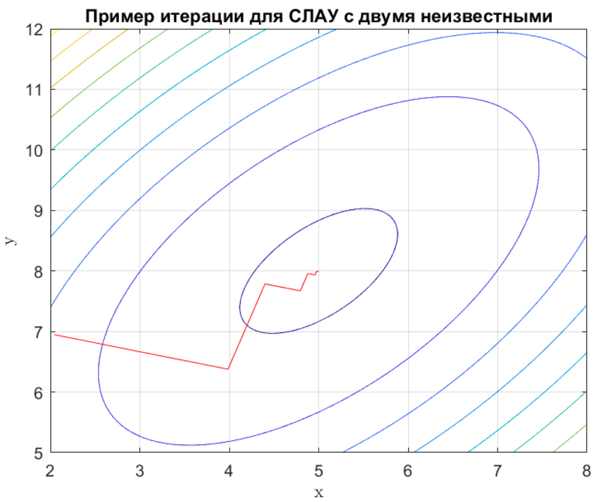
2.1 Градиентный метод

$\sqsupset A \in \mathbb{R}^{n\times n}\ \ \&\ \ B \in \mathbb{R}^n:\ \ AX = B$

$$F(X) := (AX; X) - 2(B; X)$$
$$g^{(k)} := \nabla F(X^{(k)})$$
$$X \text{ — решение } \iff X \text{ — минимум } F(X)$$
$$X^{(k+1)} = X^{(k)} - \alpha_k \nabla F(X^{(k)})$$
$$\alpha_k = \frac{(g^{(k)}; g^{(k)})}{2(Ag^{(k)}; g^{(k)})}$$

Условия остановки:

$$M := \frac{cond(A) - 1}{cond(A) + 1}$$
$$\left|\left|X^{(k+1)} - X\right|\right| \leq \frac{M}{1 - M}\left|\left|X^{(k+1)} - X^{(k)}\right|\right|$$



2.2 Условия применимости

Для градиентного метода необходимо, чтобы коэффициентная матрица A была симметричная и положительно определённая ($A = A^T\ \ \&\ \ \forall X \in \mathbb{R}^n : X^T AX > 0$).

3 Предварительный анализ задачи

Перед исследованием градиентного метода стоит убедиться, что матрицы, на которых мы будем его использовать удволетворяют условиям применимости (т.е. симметричные и положительно определённые), так же стоит помнить, что при работе с большими системами число обусловленности может значительно замедлить сходимость алгоритма. Для удовлетворения успловий применимости будем создавать матрицы по формуле:

$$A \coloneqq Q * \operatorname{diag}(\lambda_0, \cdots, \lambda_n) * Q^T$$

Где Q — ортогональная матрица полученная из произвольной матрицы $n \times n$ QR разложением. Все нормы векторного пространства будем считать энергетическими нормами по сейчас исследуемой матрице A. Нормы эвклидовы, ведь A — положительно определённая.

4 Тестовый пример

$$\sqsupset A \coloneqq \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix};\ X^{(1)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix};\ B = \begin{bmatrix} 1 \\ 2 \end{bmatrix};\ \epsilon = 0.5;\ cond(A) = 3 \implies M = 0.5$$

$$1 : g^{(1)} = 2(AX^{(1)} - B) = 2(\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix}) = \begin{bmatrix} -2 \\ -4 \end{bmatrix}$$

$$\alpha_1 = \frac{(g^{(1)}; g^{(1)})}{2(Ag^{(1)}; g^{(1)})} = \frac{\begin{bmatrix} -2 & -4 \end{bmatrix} \begin{bmatrix} -2 \\ -4 \end{bmatrix}}{2 \begin{bmatrix} -2 & -4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ -4 \end{bmatrix}} = \frac{20}{2 * 56} \approx 0.1786$$

$$X^{(2)} \coloneqq X^{(1)} - \alpha_1 g^{(1)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - 0.1786 * \begin{bmatrix} -2 \\ -4 \end{bmatrix} = \begin{bmatrix} 0.3571 \\ 0.7143 \end{bmatrix}$$

$$\frac{M}{1 - M} \left\| X^{(2)} - X^{(1)} \right\| = \left\| \begin{bmatrix} 0.3571 \\ 0.7143 \end{bmatrix} \right\| = \sqrt{\begin{bmatrix} 0.3571 & 0.7143 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0.3571 \\ 0.7143 \end{bmatrix}} =$$

$$= \sqrt{1.7856} = 1.3363 > \epsilon \implies iterate$$

$$2 : g^{(2)} = 2(AX^{(2)} - B) = 2(\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0.3571 \\ 0.7143 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix}) = 2 \begin{bmatrix} 1.4285 - 1 \\ 1.7857 - 2 \end{bmatrix} = \begin{bmatrix} 0.8570 \\ -0.4286 \end{bmatrix}$$

$$\alpha_2 = \frac{(g^{(2)}; g^{(2)})}{2(Ag^{(2)}; g^{(2)})} = \frac{\begin{bmatrix} 0.8570 & -0.4286 \end{bmatrix} \begin{bmatrix} 0.8570 \\ -0.4286 \end{bmatrix}}{2 \begin{bmatrix} 0.4285 & -0.2143 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0.8570 \\ -0.4286 \end{bmatrix}} = \frac{0.2295}{2 * 0.2754} \approx 0.4167$$

$$X^{(3)} \coloneqq X^{(2)} - \alpha_2 g^{(2)} = \begin{bmatrix} 0.3571 \\ 0.7143 \end{bmatrix} - 0.4167 * \begin{bmatrix} 0.8570 \\ -0.4286 \end{bmatrix} = \begin{bmatrix} 0.0000 \\ 0.8929 \end{bmatrix}$$

$$\frac{M}{1 - M} \left\| X^{(3)} - X^{(2)} \right\| = \left\| \begin{bmatrix} 0.3571 \\ -0.1786 \end{bmatrix} \right\| = \sqrt{\begin{bmatrix} 0.3571 & -0.1786 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0.3571 \\ -0.1786 \end{bmatrix}} =$$

$$= \sqrt{0.1913} = 0.4374 < \epsilon \implies result$$

$$X \approx \begin{bmatrix} 0.0000 & 0.8929 \end{bmatrix}^T$$

5 Подготовка контрольных тестов

Исследовать алгоритм будем по следующим характеристикам:

1. Зависимость нормы ошибки решения от количества итераций для матриц с разными числами обусловленности.
   Ожидаются убывающие зависимости, скорости которых так же зависят от числа обусловленности матриц.

2. Зависимость фактической ошибки Err от заданной допустимой ϵ для матриц с разными числами обусловленности.
   Ожитается, что для всех матриц выполняется $Err(\epsilon) < \epsilon$.

Для расчёта фактической ошибки необходимо знать точное решение. Прямо зададим его и будем генерировать матрицы A и B, исходя из решения:

$$X = \begin{bmatrix} 1, & \cdots, & 30 \end{bmatrix}^T$$

6 Модульная структура программы

Программный код оформим на языке C, он будет состоять из:

Структуры для более удобного хранения СЛАУ:

```
typedef struct slae {
    int size;
    double** A;
    double* B;
} SLAE_t;
```

Самого градиентного метода, принимающего СЛАУ, число обусловленности матрицы, допустимую ошибку, первое предположение решения и дополнительные параметры для логирования решения в каждой итерации:

```
int gradSol(SLAE_t eq, double condA, double eps, double* X_0,
            FILE* logFile, char to_log);
```

Теста, принимающего информацию о СЛАУ в бинарном формате из файла по пути *in* и возвращающего множество ренений для каждой итерации по пути *out* в csv формате:

```
void Sol(char* filename, char* resName, double eps);
```

Теста, принимающего информацию о СЛАУ в бинарном формате из файла по пути *in* и возвращающего множество решений для различных ϵ в файл по пути *out* в csv формате:

```
void testConvergence(char* in, char* out);
```

7 Исследование метода

Из рис. 1 можно заметить, что, как и ожидалось, все зависимости убывающие. Число обусловленности влияет на скорость сходимости алгоритма, это не удивительно, можно вспомнить условие остановки, где есть $\frac{M}{1-M}$. Интересна визуальная "непрерывность" функции по значению $||dX||_{H_A}$, а значит подобное стоит ожидать и для нормы $||\cdot||_2$ и прочих равносильных им топологически в $\mathbb{R}^n$. Скорее всего, это связано с движением против градиента на малое расстояние, для практических целей это наблюдение неважно.

Из рис. 2 можно заметить, что, для всех матриц фактическая ошибка попадает в диапазон, позволенный заданным ϵ. Это минимальное условие для использования любого итерационного алгоритма. Интересно заметить, что большее число обусловленности даёт меньшую фактическую ошибку.

8 Вывод

Данная реализация градиентного метода обладает довольно строгими условиями: симметричность и положительная опреденённость матрицы A. Хоть он и показывает хорошие линейные ассимптотики, при таких условиях стоит рассмотреть использование метода сопряженных направлений, характеристики которого ещё лучше при тех же условиях применимости на матрицу A.

9 Визуальные приложения

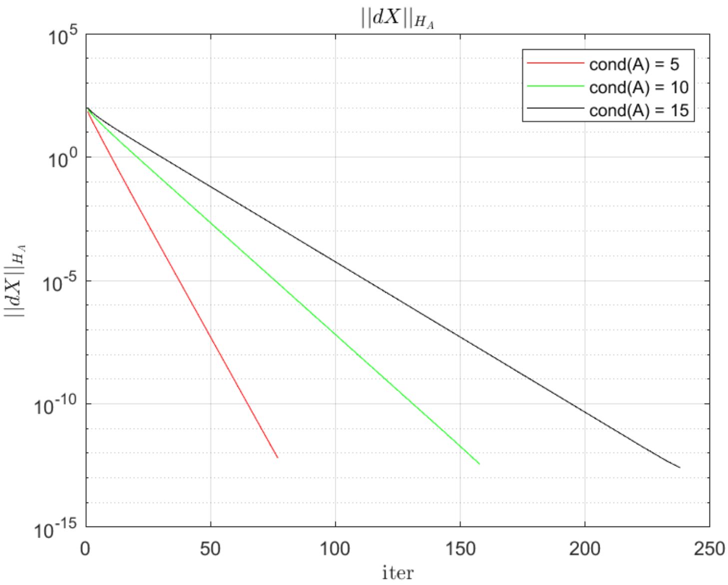


Рис. 1: Зависимость нормы ошибки решения от числа итераций ($\epsilon = 10^{-12}$)

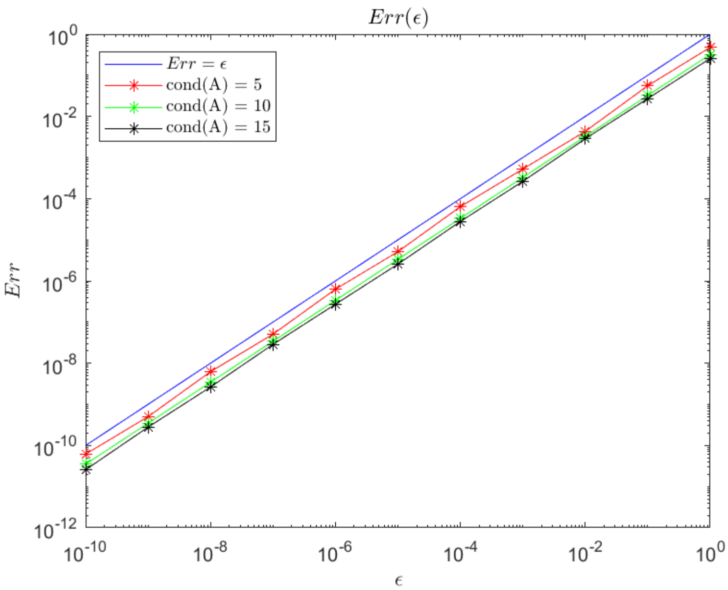


Рис. 2: Зависимость нормы фактический ошибки решения от ϵ