Scientific Computing - Homework 5

蔡子若 018071910011

1 Page 301-302, 6.2, 6.3, 6.7, 6.9, 6.12

1.1 Page 301, 6.2

(a) $f(x) = x^2$ is strictly convex on $\mathbb{R}$.

(b) $f(x) = x^3$ is nonconvex on $\mathbb{R}$.

(c) $f(x) = e^{-x}$ is strictly convex on $\mathbb{R}$.

(d) $f(x) = |x|$ is convex on $\mathbb{R}$.

1.2 Page 301, 6.3

(a) The first-order and second-order optimality conditions say that 0 is a minimum on $\mathbb{R}$.

(b) The second-order optimality condition says 0 is not a minimum of $f(x) = x^2$ on $\mathbb{R}$.

(c) The first-order and second-order optimality conditions say that 0 is a minimum on $\mathbb{R}$.

(d) The first-order and second-order optimality conditions say that 0 is a maximum on $\mathbb{R}$.

1.3 Page 302, 6.7

$L(x, \lambda) = f(x) + \lambda g(x) = x_1^2 - 2x_1 + x_2^2 - x_3^2 + 4x_3 + \lambda(x_1 - x_2 + 2x_3 - 2)$.

$$\nabla L(x, \lambda) = \begin{bmatrix} 2x_1 - 2 + \lambda \\ 2x_2 - \lambda \\ -2x_3 + 4 + 2\lambda \\ x_1 - x_2 + 2x_3 - 2 \end{bmatrix} \Rightarrow x^* = [2.5, -1.5, -1]^T,\ \lambda^* = -3$$

$$H_L(x, \lambda) = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & -2 & 2 \\ 1 & -1 & 2 & 0 \end{bmatrix}$$

The null space of $J_g(x^*)$ is $Z = \{x \in \mathbb{R}^3 : [1, -1, 2]x = 0\}$.

Choose a basis matrix $z = \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & -1 \end{bmatrix}$ s.t. $J_g(x^*)z = 0$.

Then $z^T \nabla_{xx} L(x^*, \lambda^*) z = \begin{bmatrix} 4 & 4 \\ 4 & 6 \end{bmatrix}$ is a positive definite matrix.

Hence x^* is a constrained local minimum for this function.

1.4 Page 302, 6.9

(a) $\nabla f(x) = Ax - b$, $H_f(x) = A^T$.

Then after one iteration of Newton's method, $x_1 = x_0 - (A^T)^{-1}(Ax_0 - b)$, x_1 satisfies

$\nabla f(x_1) = Ax_1 - b = Ax_0 - A(A^T)^{-1}(Ax_0 - b) - b = 0$, which means x_1 is exactly the solution for this method.

(b) If x^* is the solution and $A(x_0 - x^*) = \lambda(x_0 - x^*)$, then $x^* = x^* - \alpha_k(Ax^* - b)$.

For steepest descent method $x_{k+1} = x_k - \alpha_k(Ax_k - b)$,

we get $x_{k+1} - x^* = x_k - x^* - \alpha_k A(x_k - x^*) = (I - \alpha_k A)(x_k - x^*)$.

If $A(x_k - x^*) = \lambda(x_k - x^*)$, then from the method of induction we find $A(x_{k+1} - x^*) = \lambda(x_{k+1} - x^*)$, which means $x_{k+1} - x^*$ is still an eigenvector of A corresponding to the same eigenvalue λ.

Hence $x_{k+1} - x^* = (1 - \alpha_k\lambda)(x_k - x^*)$, whether x_k will converges to x^* depends on the value of $\alpha_k\lambda$.

1.5 Page 302, 6.12

(a) If x is the local minimum of the convex function f on the convex set $S \subset \mathbb{R}^n$ and x is not a global minimum, then there exists $y \in S$ s.t. $f(y) < f(x)$.

Then $\forall \alpha \in [0,1]$, $\alpha x + (1-\alpha)y \in S$ and $f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha)f(y)$.

Then $\forall \alpha \in [0,1)$, $f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha)f(y) < f(x)$.

Since x is a local minimum, then $\exists \delta > 0$ s.t. $\forall \alpha \in [1-\delta, 1]$, $f(\alpha x + (1-\alpha)y) \geqslant f(x)$, which is a contradiction.

Hence x is a global minimum of f on S.

(b) Suppose x is the local minimum of the strictly convex function f on the convex set $S \subset \mathbb{R}^n$.

According to (a), x is also a global minimum of f on S.

If x is not the unique global minimum, then suppose $y \neq x$ is also the global minimum and $f(x) = f(y)$.

Then $\forall \alpha \in [0,1]$, $\alpha x + (1-\alpha)y \in S$ and $f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha)f(y) = f(x)$,

hence $\forall \alpha \in [0,1]$, $f(\alpha x + (1-\alpha)y) = \alpha f(x) + (1-\alpha)f(y)$, which means f is not strictly convex, contradiction.

Hence x is the unique global minimum.

2 For a sparse positive definite martix A, if you have a way to find a sparse lower triangle matrix L so that L^TAL has much a smaller condition number than A, how would you solve the linear equation $Ax = b$ with Conjugate Gradient method?

Solving linear equation $Ax = b$ can be regarded as solving the minimization problem of $f(x) = \frac{1}{2}x^TAx - b^Tx + c$, $x \in \mathbb{R}^n$, since $\nabla f(x) = Ax - b$.

If L is a sparse lower triangle matrix s.t. L^TAL has much smaller condition number than A,

then let $x = Ly$ and $Ax = b$ is equivalent to $L^TALy = L^Tb$, which can be transformed into solving the unconstrained minimization problem $g(y) = \frac{1}{2}y^TL^TALy - b^TLy + c$, $y \in \mathbb{R}^n$.

In this case, the error will be significantly smaller than the original problem.

3 Computer problem (in C or C++): Write the functions to achieve (1) CG method with a linear search method; (2) CG method for linear systems with positive definite matrix. Test these algorithms for a few matrix and check the orthogonal properties for the residuals for a large scale linear system.

(a) Define $A = \begin{bmatrix} 0.649894 & 0.851725 & 0.717294 & 0.75354 & 0.613836 \\ 0.851725 & 1.21141 & 0.972604 & 1.00657 & 0.767806 \\ 0.717294 & 0.972604 & 2.06925 & 1.38739 & 1.5804 \\ 0.75354 & 1.00657 & 1.38739 & 1.68572 & 1.69407 \\ 0.613836 & 0.767806 & 1.5804 & 1.69407 & 1.88082 \end{bmatrix}$, $b = \begin{bmatrix} 0.526929 \\ 0.0919649 \\ 0.653919 \\ 0.415999 \\ 0.701191 \end{bmatrix}$.

We start from the point $x_0 = [0, 0, 0, 0, 0]^T$ and use CG method, obtaining the solution $\begin{bmatrix} 10.0271 \\ -8.18685 \\ 1.09698 \\ 2.44558 \\ -2.6821 \end{bmatrix}$

(b) Define $A = \begin{bmatrix} 0.5 & 0 \\ 0 & 2.5 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, we start from $x_0 = [5, 1]^T$,
and we get the solution $[4.44089 \times 10^{-16}, 1.11022 \times 10^{-16}]^T$.

4 Computer problem (in C or C++): Achieve the interpolants of Runge's function at equally spaced points.

We use Newton interpolation to achieve Runge's function $f(x) = \dfrac{1}{1 + 25x^2}$ at equally spaced points (5 points, 10 points and 15 points) and obtain the following polynomials:

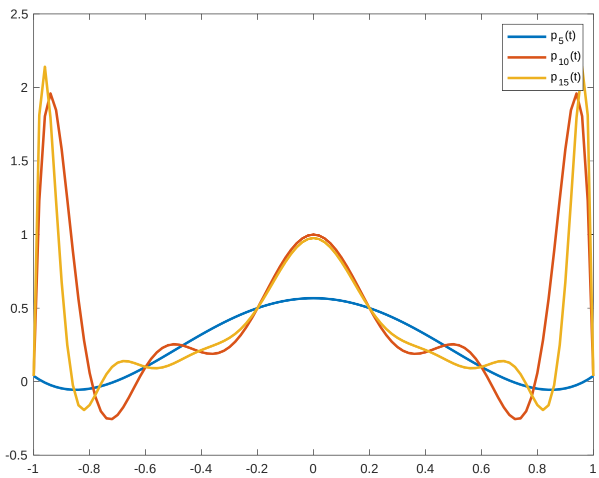

Figure 1: Polynomial interpolants of Runge's function at equally spaced points

5 Computer problem (in C or C++): Write the functions to find the natural cubic spline function using the shooting method and the B-spline method.

Shooting method

We use shooting method to solve the natural cubic spline interpolation by 11 points $(t_i, y_i), i = 1, 2 \cdots, 11$ and plot this figure:

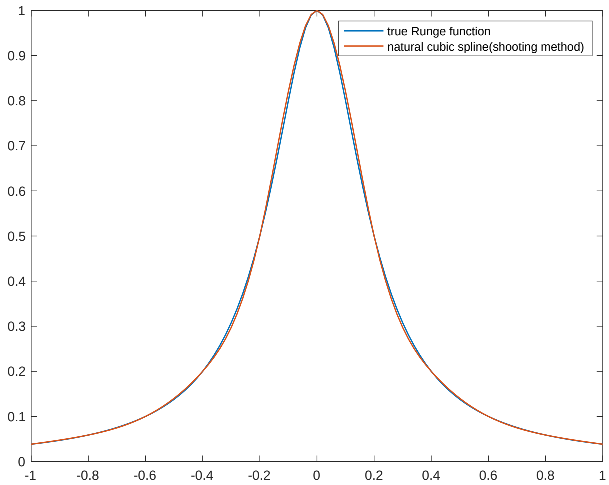

Figure 2: Interpolate Runge's function by shooting method

This figure shows the result of cubic spline interpolation is satisfyingly good for interpolating Runge's function.

B-spline method

Suppose $t_{-2} < t_{-1} < \cdots < t_{n+3}$ and the basis of spline functions of degree 3 on the interval $[t_1, t_n]$ are $\{B^3_{-2}(t), B^3_{-1}(t), \cdots, B^3_{n-1}(t)\}$, where $B^3_i(t)$ is only nonzero in $[t_i, t_{i+4}]$, $i = -2, \cdots, n-1$.
Our target is to determine the parameters $a_{-2}, a_{-1}, \cdots, a_{n-1}$ such that $f(x) = \sum_{i=-2}^{n-1} a_i B^3_i(t)$ interpolates the n given data points $(t_i, y_i), i = 1, 2, \cdots, n$ and satisfies $f''(t_1) = f''(t_n) = 0$. We utilize this method to interpolate Runge's function and get this figure:

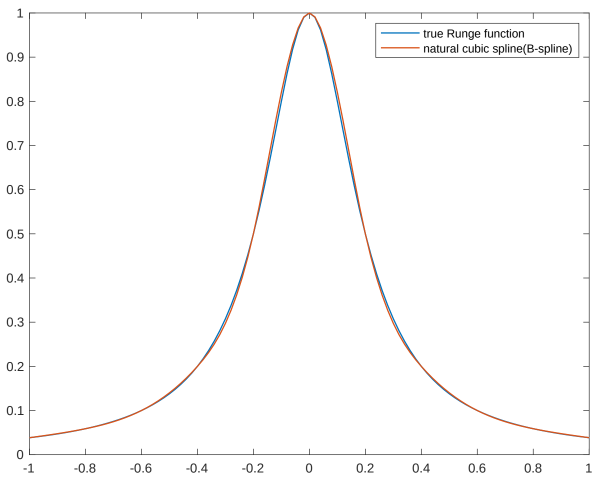

Figure 3: Interpolate Runge's function by B-spline