Scientific Computing - Homework 3

蔡子若 018071910011

1 Page 148, 3.30 (Scientific Computing: An Introductory Survey)

When $a_1 > 0$, choose $\alpha = -\|a\|_2$.

When $a_1 < 0$, choose $\alpha = \|a\|_2$.

2 Page 148, 3.33

(a) For LU factorization of $A \in \mathbb{R}^{n\times n}$ by Gaussian elimination with partial pivoting, we need a vector of length at most $2 \times n$ to record the indexs of the rows that are permutated.
(b) For QR factorization of $A \in \mathbb{R}^{m\times n}$ by Householder transformations, we need a vector of length $\min(m, n)$ to save the most upper nonzero elements of matrix Q, since they are overlapped with the most lower nonzero elements of matrix R.

In addition, when doing the Householder transformations $Ha = a - 2\frac{2v^T a}{v^T v}v$, we need a vector of length m to record the original vector a in order to compute Ha.

3 Page 149, 3.44

QR factorization: Suppose $A = QR$. Then the number of the nonzero rows of R is the rank of A.

Singular Value Decomposition: Suppose $A = U\Sigma V^T$. Then the number of nonzero diagonal elements of Σ is the rank of A.

4 Page 149, 3.45

(a) QR factorization by Householder transformation requires about $mn^2 - n^3/3 = \frac{5}{3}n^3$ multiplications.
(b) Forming normal equations and solving resulting linear system requires about $n^2m/2 + n^3/6 = \frac{7}{6}n^3$ multiplications.
(c) Singular value decomposition requires more than $4(mn^2 + n^3) = 12n^3$.

In summary, the normal equation method is the fastest and SVD is the slowest.

5 Page 149, 3.1

(a)

$$\begin{bmatrix} 1 & 10 \\ 1 & 15 \\ 1 & 20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 11.6 \\ 11.85 \\ 12.25 \end{bmatrix}$$

(b) The system is not consistent.

$$\begin{bmatrix} 1 & 10 \\ 1 & 15 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 11.6 \\ 11.85 \end{bmatrix} \Rightarrow x_1 = 11.1, \quad x_2 = 0.05$$

$$\begin{bmatrix} 1 & 15 \\ 1 & 20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 11.85 \\ 12.25 \end{bmatrix} \Rightarrow x_1 = 10.65, \quad x_2 = 0.08$$

$$\begin{bmatrix} 1 & 10 \\ 1 & 20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 11.6 \\ 12.25 \end{bmatrix} \Rightarrow x_1 = 10.95, \quad x_2 = 0.065$$

I prefer the third one since the condition number of the matrix $\begin{bmatrix} 1 & 10 \\ 1 & 20 \end{bmatrix}$ is the smallest of these three matrixs.

(c) $A^T Ax = A^T b \Rightarrow$

$$\begin{bmatrix} 3 & 45 \\ 45 & 725 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 35.7 \\ 538.75 \end{bmatrix} \Rightarrow x_1 = 10.925, \quad x_2 = 0.065$$

Comparing with b, it is indeed reasonable to choose the matrix which has the smallest condition number.

6 Page 149, 3.3

$$\begin{bmatrix} 1 & e \\ 2 & e^2 \\ 3 & e^3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$$

7 Page 150, 3.14

$H^T = I - 2\frac{(vv^T)^T}{v^T v} = I - 2\frac{vv^T}{v^T v} = H$, hence the matrix is symmetric.

$H^2 = (I - 2\frac{vv^T}{v^T v})^2 = I - 4\frac{vv^T}{v^T v} + 4\frac{vv^T vv^T}{v^T vv^T v} = I - 4\frac{vv^T}{v^T v} + 4\frac{vv^T}{v^T v} = I \Rightarrow H^T H = I$, hence the matrix is orthogonal.

8 Page 152, 3.31

Suppose the singular value decomposition of matrix $A \in \mathbb{R}^{m\times n}$ is $A = U\Sigma V^T$, then the solution of the least squares problem $Ax \cong b$ is the same as $A^T Ax = A^T b \Rightarrow (V\Sigma^T \Sigma V^T)x = V\Sigma^T U^T b$.

Since $U = (u_1, u_2, \cdots u_m)$, $V = (v_1, v_2, \cdots u_n)$, $m > n$, $\sigma_i, i = 1, \cdots n$ are the diagonal elements of matrix Σ, the above equation can be written as $(\sum_{i=1}^{n} \sigma_1^2 v_i v_i^T)x = (\sum_{i=1}^{n} \sigma_i v_i u_i^T)b$.

Multiply $v_i^T, i = 1, \cdots n$ both on the left of the equation, we get $\sigma_i^2 v_i^T x = \sigma_i u_i^T b$. When $\sigma_i = 0$, we can permutate the matrix A to ensure the first $k, k < n$ diagonal elements of the matrix Σ are nonzero and the other elements are zero. In this case the problem turns into a $m \times k$ dimension least squares problem.

When $\sigma_i \neq 0, i = 1, \cdots n$, since $\{v_i\}_{i=1}^n$ is a orthonormal basis of $\mathbb{R}^n$, $x = \sum_{i=1}^{i=n} \langle v_i, x\rangle v_i = \sum_{i=1}^{i=n} v_i^T x v_i = \sum_{i=1}^{i=n} \frac{u_i^T b}{\sigma_i} v_i$.

In conclusion, the solution x is given by $x = \sum_{\sigma_i \neq 0} \frac{u_i^T b}{\sigma_i} v_i$.

9 Page 152, 3.32

Suppose the singular value decomposition of A is $A = U\Sigma V^T$. Hence $A^+ = V\Sigma^+ U^T$.

(a) $AA^+A = U\Sigma V^T V\Sigma^+ U^T U\Sigma V^T = U\Sigma\Sigma^+\Sigma V^T$. Since $\Sigma\Sigma^+\Sigma = \Sigma$, then $AA^+A = U\Sigma V^T = A$.

(b) $A^+AA^+ = V\Sigma^+ U^T U\Sigma V^T V\Sigma^+ U^T = V\Sigma^+\Sigma\Sigma^+ U^T$. Since $\Sigma^+\Sigma\Sigma^+ = \Sigma^+$, then $A^+AA^+ = V\Sigma^+ U^T = A^T$.

(c) $(AA^+)^T = (U\Sigma\Sigma^+ U^T)^T = U(\Sigma\Sigma^+)^T U^T = U\Sigma\Sigma^+ U^T = AA^+$.

(d) $(A^+A)^T = (V\Sigma^+\Sigma V^T)^T = V(\Sigma^+\Sigma)^T V^T = V\Sigma^+\Sigma V^T = A^+A$.

10 Computer problem (in C or C++): Using Householder transform to achieve the QR decomposition with and without a column pivoting. Then, using QR decomposition to finish 3.8 on page 154

The problem is

$$\begin{bmatrix} 1 & t_1 & t_1^2 & \dots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \dots & t_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

For $m = 21, n = 12, \epsilon = 10^{-10}$, after pertubating y_i, the solution of $A^TAx = A^Tb$ using Cholesky method and the solution using QR factorization are respectively $\begin{bmatrix} 1 \\ 1.00001 \\ 0.999809 \\ 1.00274 \\ 0.979571 \\ 1.09006 \\ 0.749554 \\ 1.45135 \\ 0.473759 \\ 1.38311 \\ 0.841686 \\ 1.02835 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0.999997 \\ 1.00002 \\ 0.999942 \\ 1.00014 \\ 0.999784 \\ 1.00023 \\ 0.999844 \\ 1.00006 \\ 0.999989 \end{bmatrix}$. The first solution shows that the normal equation method is more sensitive to the perturbation we introduced into the data. Solving the least squares system using QR factorization comes closer to recovering the x that we used to generate the data.

The solutions differ indeed affect our ability to fit the data points (t_i, y_i) closely by the polynomial. Since the noises of the data points always exist and we need a method which is more insensitive to the perturbation of y_i.