Assignment 2

Name: 陈子蔚 SID: 12332440

1

a

$$
\begin{aligned}
11.60 &= x_1 + 10x_2 && (1)\\
11.85 &= x_1 + 15x_2 && (2)\\
12.25 &= x_1 + 20x_2 && (3)\\
\Leftrightarrow \begin{bmatrix} 11.60 \\ 11.85 \\ 12.25 \end{bmatrix} &= \begin{bmatrix} 1 & 10 \\ 1 & 15 \\ 1 & 20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} && (4)
\end{aligned}
$$

Denote it as

$$
\mathrm{y} = \mathrm{A}\mathrm{x} \tag{5}
$$

b

- Choose (1)&(2), $\mathrm{x} = [11.1, 0.05]^\top$
- Choose (1)&(3), $\mathrm{x} = [10.95, 0.065]^\top$
- Choose (2)&(3), $\mathrm{x} = [10.65, 0.08]^\top$
  Not consistent obviously. There's no particular reason to prefer one of these results over the others without additional context or constraints.

c

$$
\mathrm{x} = (\mathrm{A}^\top \mathrm{A})^{-1}\mathrm{A}^\top y = [10.925, 0.065]^\top \tag{6}
$$

Comparing this result with those obtained in part (b), we see that the least squares solution gives a different result from any pair of solutions obtained by considering only two of the three points. This least squares solution is preferred when dealing with an overdetermined system because it takes into account all available data to minimize the MSE.

2

a

$$
A \in \mathbb{R}^{m\times n} \wedge m > n \tag{7}
$$

Thus,

$$
A = Q \begin{bmatrix} R \\ 0 \end{bmatrix} \Leftrightarrow \#H = \#\text{column} = 3 \tag{8}
$$

b

To avoid cancellation,

$$-\|c_1\|e_1 = [-2, 0, 0, 0]^\top \tag{9}$$

c

The first column would not change.

$$[-2, 0, 0, 0]^\top \tag{10}$$

3

$$A \in \mathbb{R}^{m\times n} \wedge m > n \tag{11}$$

Thus,

$$\begin{aligned} & A = QR \wedge Q \in \mathbb{R}^{m\times n} \wedge R \in \mathbb{R}^{n\times n} \\ \Leftrightarrow\ & Q^\top Q = I \in \mathbb{R}^{n\times n} \end{aligned} \tag{12}$$

As for the equation,

$$Ax = QRx = b \tag{13}$$

$$\Leftrightarrow Rx = Q^\top b \tag{14}$$

Output

```
# Q
[[ 5.77350269e-01  2.04124145e-01  3.53553391e-01]
 [ 0.00000000e+00  6.12372436e-01  3.53553391e-01]
 [ 0.00000000e+00  0.00000000e+00  7.07106781e-01]
 [-5.77350269e-01  4.08248290e-01 -1.17756934e-16]
 [-5.77350269e-01 -2.04124145e-01  3.53553391e-01]
 [ 0.00000000e+00 -6.12372436e-01  3.53553391e-01]]
# R
[[ 1.73205081 -0.57735027 -0.57735027]
 [ 0.          1.63299316 -0.81649658]
 [ 0.          0.          1.41421356]]
# x
[1236. 1943. 2416.]
```

4

Firstly prove the $\|A\|_2 = \max_i \sqrt{\lambda_i} = \max_i \sigma_i$,

$$\|A\|_2 = \sup_{\|x\|_2 = 1} \|Ax\|_2 \tag{15}$$

Since $A^\top A$ is semi-positive,

$$A^\top A = H^\top \mathrm{diag}(\lambda_i) H \tag{16}$$

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \end{bmatrix}$$

$$\operatorname{diag}(\lambda_i) = \begin{bmatrix} 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \tag{17}$$

$\forall x \in \{x \| \|x\|_2 = 1\}$, denote

$$Hx = y = (y_1, y_2, \dots, y_n)^\top \wedge \|y\|_2 = \|x\|_2 = 1 \tag{18}$$

$$\|Ax\|_2^2 = x^\top A^\top Ax = (Hx)^\top \operatorname{diag}(\lambda_i)(Hx) = \sum_{i=1}^{n} \lambda_i y_i^2 \le \|y\|_2^2 \max_{1\le i\le n} \lambda_i = \max_{1\le i\le n} \lambda_i \tag{19}$$

That is,

$$\|A\|_2 = \max \sigma(A) \tag{20}$$

$$\Leftrightarrow \|A^{-1}\|_2 = \max \sigma(A^{-1}) = \max \frac{1}{\sigma(A)} = \frac{1}{\min \sigma(A)} \tag{21}$$

5

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \tag{22}$$

Please refer to `as2_5.py` for more information.

The iteration history for the function $f(x) = x^2 - 1$ using Newton's method with $x_0 = 2$ for 4 iterations is:

0. $x_0 = 2, h = -\frac{3}{4}$
1. $x_1 = \frac{5}{4}, h = -\frac{9}{40}$
2. $x_2 = \frac{41}{40}, h = -\frac{81}{3280}$
3. $x_3 = \frac{3281}{3280}, h = -\frac{6561}{21523360}$
4. $x_3 = \frac{21523361}{21523360}$

When we start with $x_0 = -2$, the iteration history is:

0. $x_0 = -2, h = \frac{3}{4}$
1. $x_1 = -\frac{5}{4}, h = \frac{9}{40}$
2. $x_2 = -\frac{41}{40}, h = \frac{81}{3280}$
3. $x_3 = -\frac{3281}{3280}, h = \frac{6561}{21523360}$
4. $x_3 = -\frac{21523361}{21523360}$

The sequences converge to the roots of the function, which are $x = 1$ and $x = -1$, respectively. The sequences are symmetrical because the function is symmetrical and the initial guesses are symmetrical about the origin.

6

$$x_{k+1} = x_k - f(x_k)\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \tag{23}$$

Please refer to `as2_6.py` for more information.

After applying the secant method to the function $f(x) = x^2 - 1$ with starting points $x_0 = 3$ and $x_1 = 2$, and performing 6 iterations, the sequence of approximations to the root is:

0. $x_1 = 2, h = -0.6$
1. $x_2 = 1.4, h = -0.2823529411764706$
2. $x_3 = 1.1176470588235294, h = -0.0989554700384827$
3. $x_4 = 1.0186915887850467, h = -0.017662252706817205$
4. $x_5 = 1.0010293360782296, h = -0.0010198100459649776$
5. $x_6 = 1.0000095260322646, h = -9.52113206547989 \times 10^{-6}$
6. $x_7 = 1.0000000049001991$

As we can see, the approximations are converging towards 1.

7

Core code

```python
# Define the Newton's method function
def newton(f, df, a, b, iterations=5):
    x_n = (a + b) / 2
    for _ in range(iterations):
        fx = f(x_n)
        dfx = df(x_n)
        x_n = x_n - fx / dfx

    assert a <= x_n <= b # Ensure the root is within the interval
    return x_n

# Define the Bisection method function
def bisect(f, a, b, iterations=15):
    for _ in range(iterations):
        c = (a + b) / 2
        if f(a) * f(c) < 0:
            b = c
        else:
            a = c
    return c
```

Output

```
# roots obtained by Bisection method function
[-0.90618896484375, -0.53846435546875, -1.220703124988898e-05, 0.53846435546875, 0.90618896484375]
# roots obtained by Newton's method function
[-0.9062605812803672, -0.538469310105683, 0.0, 0.538469310105683, 0.9062605812803672]
```