2.17

(a)因为

$$f(x)=\sum_{i=1}^{n}a_ix_i+b$$

所以

$$\frac{\partial f}{\partial x_k}=a_k$$

因此

$$\begin{aligned}\nabla f(x)&=\begin{bmatrix}\frac{\partial f}{\partial x_1}\\ \vdots\\ \frac{\partial f}{\partial x_n}\end{bmatrix}\\&=\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}\\&=a\end{aligned}$$

(b)因为

$$f(x)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_ia_{ij}x_j$$

所以

$$\frac{\partial f}{\partial x_k}=\sum_{j=1}^{n}a_{kj}x_j+\sum_{i=1}^{n}a_{ik}x_i$$

因此

$$
\begin{aligned}
\nabla f(x) &= \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{j=1}^n a_{1j}x_j + \sum_{i=1}^n a_{i1}x_i \\ \vdots \\ \sum_{j=1}^n a_{nj}x_j + \sum_{i=1}^n a_{in}x_i \end{bmatrix} \\
&= \begin{bmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{nj}x_j \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^n a_{i1}x_i \\ \vdots \\ \sum_{i=1}^n a_{in}x_i \end{bmatrix} \\
&= Ax + A^T x \\
&= (A + A^T)x
\end{aligned}
$$

(c)由$A^T = A$和上一题可得

$$
\nabla f(x) = 2Ax
$$

3.13

对于行满秩矩阵$A \in \mathbb{R}^{m\times n}$，如果$A^T$的$QR$分解为

$$
\begin{aligned}
A^T &= QR \in \mathbb{R}^{n\times m} \\
Q^T Q = I_n&, Q \in \mathbb{R}^{n\times m} \\
R &\in \mathbb{R}^{m\times m}
\end{aligned}
$$

其中R可逆，取

$$
\begin{aligned}
B^T &= R^{-1}Q^T \\
B &= Q(R^T)^{-1}
\end{aligned}
$$

那么

$$
\begin{aligned}
AB &= (B^T A^T)^T \\
&= \left(R^{-1}Q^T QR\right)^T \\
&= (R^{-1}R)^T \\
&= I_m
\end{aligned}
$$

(a)此时只要计算

$$
A_1 = \begin{bmatrix} -1 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}
$$

的右逆B_1，然后对B_1的第二行插入0即可，注意A_1依然行满秩，所以仍然存在右逆，编写程序后得到：

```
import numpy as np

A = np.array([
        [-1, 0, 0, -1, 1],
        [0, 1, 1, 0, 0],
        [1, 0, 0, 1, 0]
        ])
n, m = A.shape

#### (a)
A1 = np.c_[A[:, 0], A[:, 2:]]
#QR分解
Q, R = np.linalg.qr(A1.T)
#计算右逆
B1 = Q.dot(np.linalg.inv(R.T))
#计算最终结果
B = np.r_[B1[0, :], np.zeros(n)].reshape(2, n)
B = np.r_[B, B1[1:, :]]
print(B)
#验证结果
print(A.dot(B))
```

```
[[9.72785220e-18 0.00000000e+00 5.00000000e-01]
 [0.00000000e+00 0.00000000e+00 0.00000000e+00]
 [0.00000000e+00 1.00000000e+00 0.00000000e+00]
 [1.91026513e-16 0.00000000e+00 5.00000000e-01]
 [1.00000000e+00 0.00000000e+00 1.00000000e+00]]
[[ 1.00000000e+00  0.00000000e+00 -2.22044605e-16]
 [ 0.00000000e+00  1.00000000e+00  0.00000000e+00]
 [ 2.00754365e-16  0.00000000e+00  1.00000000e+00]]
```

不考虑浮点数产生的误差，我们有

$$B=\left[\begin{array}{ccc}0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \\ 1 & 0 & 1\end{array}\right]$$

(b)不可能，原因如下：

由条件可得

$$\operatorname{rank}(B)=3-1=2$$

但是因为

$$AB=I_3$$

所以

$$\operatorname{rank}(AB)=3\le\operatorname{rank}(B)=2$$

这就产生了矛盾。

(c)不可能，如果B的第三列为0，那么AB的第三列为0，与条件矛盾。

(d)记

$$A=\begin{bmatrix}\tilde a_1^T\\ \tilde a_2^T\\ \tilde a_3^T\end{bmatrix},B=\begin{bmatrix}b_1 & b_2 & b_3\end{bmatrix}$$

注意到

$$\begin{aligned}\tilde a_2^Tb_1&=b_{21}+b_{31}=2b_{21}=0\\ \tilde a_2^Tb_2&=b_{22}+b_{32}=2b_{22}=1\\ \tilde a_2^Tb_3&=b_{23}+b_{33}=2b_{23}=0\end{aligned}$$

所以

$$\begin{aligned}b_{21}&=b_{31}=0\\ b_{22}&=b_{32}=\frac{1}{2}\\ b_{23}&=b_{33}=0\end{aligned}$$

注意到此时有

$$\tilde a_2^TB=\begin{bmatrix}0&1&1&0&0\end{bmatrix}\begin{bmatrix}b_{11}&b_{12}&b_{13}\\ 0&\frac12&0\\ 0&\frac12&0\\ b_{41}&b_{42}&b_{43}\\ b_{51}&b_{52}&b_{53}\end{bmatrix}=\begin{bmatrix}0&1&0\end{bmatrix}$$

所以只要考虑

$$A_1=\begin{bmatrix}\tilde a_1^T\\ \tilde a_3^T\end{bmatrix}$$

求出满足条件的B，使得

$$A_1B=\begin{bmatrix}1&0&0\\0&0&1\end{bmatrix}$$

即可。又因为A_1的$2,3$列为0，所以只要考虑

$$A_2 = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{41} & b_{42} & b_{43} \\ b_{51} & b_{52} & b_{53} \end{bmatrix}$$

求B_2，使得

$$A_2 B_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

求解该方程组得到

$$B_2 = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \\ 1 & 0 & 1 \end{bmatrix}$$

最终的结果为

$$B = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \\ 1 & 0 & 1 \end{bmatrix}$$

最后验证结果：

```
#### (d)
B = np.array([
        [0, 0, 0.5],
        [0, 0.5, 0],
        [0, 0.5, 0],
        [0, 0, 0.5],
        [1, 0, 1]
        ])
print(A.dot(B))
```

```
[[1. 0. 0.]
 [0. 1. 0.]
 [0. 0. 1.]]
```

(e)此时

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{33} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

那么

$$
\begin{aligned}
AB &= \begin{bmatrix} -1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{33} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} -b_{11} & -b_{12} & -b_{13} \\ 0 & b_{22} & b_{23}+b_{33} \\ b_{11} & b_{12} & b_{13} \end{bmatrix} \\
&= I_3
\end{aligned}
$$

所以

$$
-b_{11} = 1, b_{11} = 0
$$

这就产生了矛盾。

(f)此时

$$
B = \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \\ b_{51} & b_{52} & b_{53} \end{bmatrix}
$$

那么

$$
\begin{aligned}
AB &= \begin{bmatrix} -1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \\ b_{51} & b_{52} & b_{53} \end{bmatrix} \\
&= \begin{bmatrix} -b_{11}-b_{41}+b_{51} & -b_{42}+b_{52} & -b_{43}+b_{53} \\ b_{21}+b_{31} & b_{22}+b_{32} & b_{33} \\ b_{11}+b_{41} & b_{42} & b_{43} \end{bmatrix} \\
&= I_3
\end{aligned}
$$

所以可以取

$$
B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}
$$

验证结果：

```
#### (f)
B = np.array([
        [0, 0, 0],
        [0, 0, 0],
        [0, 1, 0],
        [0, 0, 1],
        [1, 0, 1]
        ])
print(A.dot(B))
```

```
[[1 0 0]
 [0 1 0]
 [0 0 1]]
```

4.1

将U扩张为正交矩阵：

$$\tilde U = [\,U \quad U_1\,]$$

那么

$$\|x\| = \|\tilde U^T x\|$$

注意到

$$\begin{aligned}\tilde U^T x &= \begin{bmatrix} U^T \\ U_1^T \end{bmatrix} x \\ &= \begin{bmatrix} U^T x \\ U_1^T x \end{bmatrix}\end{aligned}$$

因此

$$\|\tilde U^T x\|^2 = \|U^T x\|^2 + \|U_1^T x\|^2 \ge \|U^T x\|^2$$

所以

$$\|x\| = \|\tilde U^T x\| \ge \|U^T x\|$$

当且仅当

$$U_1^T x = 0$$

时等号成立，由x的任意性可得必然有$k = n$。

4.2

(a)

$$
\begin{aligned}
(UV)^T(UV) &= V^TU^TUV \\
&= V^TV \\
&= I
\end{aligned}
$$

(b)

$$
\begin{aligned}
(U^{-1})^TU^{-1} &= (UU^T)^{-1} \\
&= I
\end{aligned}
$$

(c)由正交矩阵的特点可得

$$
\begin{aligned}
u_1^Tu_1 &= 1 \\
u_2^Tu_2 &= 1
\end{aligned}
$$

所以不妨设

$$
U = \begin{bmatrix} \cos\alpha & \cos\beta \\ \sin\alpha & \sin\beta \end{bmatrix}
$$

又因为

$$
u_1^Tu_2 = 0
$$

所以

$$
\begin{aligned}
u_1^Tu_2 &= \cos\alpha\cos\beta + \sin\alpha\sin\beta \\
&= \cos(\beta-\alpha) \\
&= 0
\end{aligned}
$$

所以

$$
\beta - \alpha = \frac{\pi}{2} + k\pi, k \in \mathbb{Z}
$$

如果

$$
\beta - \alpha = \frac{\pi}{2} + 2k\pi, k \in \mathbb{Z}
$$

那么

$$
\begin{aligned}
U &= \begin{bmatrix} \cos\alpha & \cos\beta \\ \sin\alpha & \sin\beta \end{bmatrix} \\
&= \begin{bmatrix} \cos\alpha & \cos(\alpha + \frac{\pi}{2} + 2k\pi) \\ \sin\alpha & \sin(\alpha + \frac{\pi}{2} + 2k\pi) \end{bmatrix} \\
&= \begin{bmatrix} \cos\alpha & \sin(-\alpha) \\ \sin\alpha & \cos(-\alpha) \end{bmatrix} \\
&= \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}
\end{aligned}
$$

此时为旋转矩阵。

如果

$$
\beta - \alpha = \frac{\pi}{2} + (2k+1)\pi, k \in \mathbb{Z}
$$

那么

$$
\begin{aligned}
U &= \begin{bmatrix} \cos\alpha & \cos\beta \\ \sin\alpha & \sin\beta \end{bmatrix} \\
&= \begin{bmatrix} \cos\alpha & \cos(\alpha + \frac{\pi}{2} + (2k+1)\pi) \\ \sin\alpha & \sin(\alpha + \frac{\pi}{2} + (2k+1)\pi) \end{bmatrix} \\
&= \begin{bmatrix} \cos\alpha & \sin(-\alpha - \pi) \\ \sin\alpha & \cos(-\alpha - \pi) \end{bmatrix} \\
&= \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}
\end{aligned}
$$

此时为反射矩阵。

4.3

(a)

$$
\begin{aligned}
(I-P)^T &= I - P^T \\
&= I - P \\
(I-P)^2 &= I - 2P + P \\
&= I - P
\end{aligned}
$$

(b)

$$
\begin{aligned}
(UU^T)^T &= UU^T \\
(UU^T)^2 &= UU^TUU^T \\
&= UU^T
\end{aligned}
$$

(c)

$$
\begin{aligned}
\left(A\left(A^TA\right)^{-1}A^T\right)^T &= A\left(\left(A^TA\right)^{-1}\right)^TA^T \\
&= A\left(A^TA\right)^{-1}A^T \\
\left(A\left(A^TA\right)^{-1}A^T\right)^2 &= A\left(A^TA\right)^{-1}A^TA\left(A^TA\right)^{-1}A^T \\
&= A\left(A^TA\right)^{-1}A^T
\end{aligned}
$$

(d)$\forall z_0 = Pz \in \mathcal{R}(P)$，那么

$$
\begin{aligned}
\|x-z_0\|^2 &= \|x-Pz\|^2 \\
&= \|x-Px+Px-Pz\|^2 \\
&= \|(I-P)x+P(x-z)\|^2 \\
&= ((I-P)x+P(x-z))^T\left((I-P)x+P(x-z)\right) \\
&= ((I-P)x)^T\left((I-P)x\right)+2(P(x-z))^T(I-P)x+(P(x-z))^T\left(P(x-z)\right) \\
&= \|x-Px\|^2+2(x-z)^TP^T(I-P)x+\|Px-Pz\|^2 \\
&= \|x-Px\|^2+2(x-z)^T(P-P^2)x+\|Px-Pz\|^2 \\
&= \|x-Px\|^2+\|Px-Pz\|^2 \\
&\geq \|x-Px\|^2
\end{aligned}
$$

当且仅当$z=x$时等号成立。

5.1

因为

$$
x_{\mathrm{ls}} = \left(A^TA\right)^{-1}A^Ty
$$

所以

$$
y_{\mathrm{ls}} = Ax_{\mathrm{ls}} = A\left(A^TA\right)^{-1}A^Ty \triangleq Py
$$

由上一题可知P为对称矩阵，因此

$$
\begin{aligned}
\|r\|^2 &= \|y - y_{\text{ls}}\|^2 \\
&= \|y - Py\|^2 \\
&= \|(I - P)y\|^2 \\
&= y^T(I - P)^T(I - P)y \\
&= y^T(I - P)^2 y \\
&= y^T(I - P)y \\
&= y^T y - y^T P y \\
&= y^T y - y^T P^2 y \\
&= y^T y - y^T P^T P y \\
&= \|y\|^2 - \|Py\|^2 \\
&= \|y\|^2 - \|y_{\text{ls}}\|^2
\end{aligned}
$$

6.9

```
A = zeros(N, n_pixels^2);
for i = 1 : N
    data = line_pixel_length(lines_d(i),lines_theta(i),n_pixels);
    data = data(:);
    A(i, :) = data;
end

% v = inv(A' * A) * A' * y;
v = A \ y;
X = reshape(v, n_pixels, n_pixels);
figure(1)      % display the original image
colormap gray
imagesc(X)
axis image
```

补充题

1

```
N = 40;
x =[0.0197;    0.0305;    0.0370;    0.1158;    0.2778;    0.3525;    0.3974;    0.3976;
 0.4053;    0.4055;    0.4623;    0.5444;    0.7057;    0.8114;    0.8205;    0.8373;
 0.8894;    0.8902;    0.9129;    0.9320;    0.9720;    1.0503;    1.2076;    1.2137;
 1.2309;    1.3443;    1.4764;    1.4936;    1.5242;    1.5839;    1.6263;    1.6428;
 1.6924;    1.7826;    1.7873;    1.8338;    1.8436;    1.8636;    1.8709;    1.9003];
y =[-0.0339;   -0.1022;   -0.0165;   -0.0532;   -0.2022;   -0.1149;   -0.1310;   -0.1924;
-0.1768;   -0.1845;   -0.2210;   -0.1994;   -0.3058;   -0.1916;   -0.3097;   -0.3011;
-0.2657;   -0.3162;   -0.3295;   -0.3710;   -0.3247;   -0.4274;   -0.3756;   -0.3323;
-0.4545;   -0.4242;   -0.4710;   -0.6230;   -0.6332;   -0.5694;   -0.6458;   -0.6025;
-0.6313;   -0.7051;   -0.6799;   -0.7489;   -0.7310;   -0.8675;   -0.8146;   -0.8469];

% (a)
x1 = [x, ones(N, 1)];
% A1 = inv(x1' * x1) * x1' * y;
A1 = x1 \ y

% (b)
x2 = [x.^3, x.^2, x, ones(N, 1)];
% A3 = inv(x2' * x2) * x2' * y;
A2 = x2 \ y
```

输出为

```
A1 =
   -0.3881
    0.0014
A2 =
   -0.1476
    0.3045
   -0.4632
   -0.0320
```

2

```
N = 1000;
R1 = zeros(N, 1);
R2 = zeros(N, 1);

for i = 1:N
    % (a)生成数据
    A = randn(50, 20);
    v = 0.1 * randn(50, 1);
    x = randn(20, 1);
    y = A * x + v;

    % (b)最小二乘
    xls = A \ y;
    r1 = norm(xls - x) / norm(x);
```

```
    % (c)
    y_trunc = y(1:20, :);
    A_trunc = A(1:20, :);
    xjem = A_trunc \ y_trunc;
    r2 = norm(xjem - x) / norm(x);

    R1(i) = r1;
    R2(i) = r2;
end

mean(R1)
mean(R2)
```

```
ans =
    0.0189
ans =
    0.8128
```