14.16

(a)注意到

$$(A^T A)_{ii} = \sum_{k=1}^{n} A_{ki}^2$$

所以

$$\begin{aligned}\operatorname{Tr} A^T A &= \sum_{i=1}^{n}\sum_{k=1}^{n} A_{ki}^2 \\ &= \sum_{i,j} |A_{ij}|^2\end{aligned}$$

因此

$$\|A\|_{\mathrm{F}} = \left(\sum_{i,j} |A_{ij}|^2\right)^{1/2}$$

(b)

$$\begin{aligned}\operatorname{Tr}\big((UA)^T(UA)\big) &= \operatorname{Tr}\big(A^T U^T U A\big) \\ &= \operatorname{Tr}\big(A^T A\big) \\ \operatorname{Tr}\big((AV)^T(AV)\big) &= \operatorname{Tr}\big(V^T A^T A V\big) \\ &= \operatorname{Tr}\big(V V^T A^T A\big) \\ &= \operatorname{Tr}\big(A^T A\big)\end{aligned}$$

所以

$$\|UA\|_{\mathrm{F}} = \|AV\|_{\mathrm{F}} = \|A\|_{\mathrm{F}}$$

(c)假设A的满奇异值分解为

$$A = U\Sigma V^T$$

由(b)可得

$$\|A\|_{\mathrm{F}} = \|\Sigma\|_{\mathrm{F}} = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$$

所以

$$\sigma_{\max}(A) \le \|A\|_{\mathrm{F}} \le \sqrt{r}\sigma_{\max}(A)$$

14.26

(a)将卷积写成矩阵形式，记

$$h=\begin{bmatrix}h_2\\ \vdots\\ h_{2n}\end{bmatrix},c=\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix},w=\begin{bmatrix}w_1\\ \vdots\\ w_n\end{bmatrix},C=\begin{bmatrix}A\\ B\end{bmatrix},\hat{h}=\begin{bmatrix}h_{n+1-k}\\ \vdots\\ h_{n+1+k}\end{bmatrix}=Dh$$

其中

$$A=\begin{bmatrix}c_1 & 0 & 0 & 0\\ c_2 & c_1 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots\\ c_n & c_{n-1} & \ldots & c_1\end{bmatrix}$$
$$B=\begin{bmatrix}0 & c_n & c_{n-1} & \ldots & c_2\\ 0 & 0 & c_n & \ldots & c_3\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & \ldots & 0 & 0 & c_n\end{bmatrix}$$
$$D=\begin{bmatrix}0_{(2k+1)\times(n-k-1)} & I_{2k+1} & 0_{(2k+1)\times(n-k-1)}\end{bmatrix}$$

所以

$$h=Cw$$

另一方面

$$\hat{h}=DCw$$

所以

$$\begin{aligned}E_{\text{tot}}&=h^Th\\ &=w^TC^TCw\\ E_{\text{des}}&=\hat{h}^T\hat{h}\\ &=w^TC^TD^TDCw\end{aligned}$$

我们的目标是最大化

$$\frac{E_{\text{des}}}{E_{\text{tot}}}=\frac{w^TC^TD^TDCw}{w^TC^TCw}$$

将其化为条件约束问题

$$\begin{aligned}\max\quad & w^TC^TD^TDCw\\ \text{s.t}\quad & w^TC^TCw=1\end{aligned}$$

构造拉格朗日乘子

$$L(w,\lambda)=w^TC^TD^TDCw-\lambda\left(w^TC^TCw-1\right)$$

求梯度可得

$$
\begin{aligned}
\nabla L_w(w,\lambda) &= 2C^TD^TDCw - 2\lambda C^TCw = 0 \\
\nabla L_\lambda(w,\lambda) &= -w^TC^TCw + 1 = 0
\end{aligned}
$$

第一个式子说明

$$
C^TD^TDCw = \lambda C^TCw
$$

带入原式得到

$$
w^TC^TD^TDCw = \lambda wC^TCw = \lambda
$$

如果C^TC可逆，那么第一个式子可以化为

$$
(C^TC)^{-1}C^TD^TDCw \triangleq Ew = \lambda w
$$

所以λ是E的特征值，因此

$$
\max \frac{E_{\text{des}}}{E_{\text{tot}}} = \max \lambda\left((C^TC)^{-1}C^TD^TDC\right)
$$

(b)

```
c = [  0.0455; -0.2273; -0.0455;  0.2727;  0.4545;  0.4545;  0.2727; -0.0455; -0.2273;
 0.0455;];
k = 1;
n = length(c);
A = zeros(n, n);
B = zeros(n - 1, n);
for i = 1: n
    for j = 1:i
        A(i, j) = c(i + 1 - j)
    end
end

for i = 1: (n - 1)
    for j = 1 : (n - i)
        B(i, j + i) = c(n + 1 - j)
    end
end

C = [A; B];
D = [zeros(2 * k + 1, n - k - 1), eye(2 * k + 1), zeros(2 * k + 1, n - k - 1)];
%E = inv(C' * C) * (C' * D' * D * C);
E = (C' * C) \ (C' * D' * D * C);
Eig = eig(E);
res = max(Eig)
```

```
0.9375
```

15.2

(a)回顾定义

$$\kappa(A) = \|A\|\|A^{-1}\| = \sigma_{\max}(A)/\sigma_{\min}(A)$$

显然

$$\kappa(A) \geq 1$$

所以

$$\kappa(A) = 1$$

等价于

$$\sigma_{\max}(A) = \sigma_{\min}(A) = \sigma$$

等价于A的SVD为

$$A = U\sigma IV^T = \sigma UV^T$$

显然UV^T为正交矩阵，所以结论成立。

15.3

假设A的SVD为

$$A = U\Sigma V^T$$

其中

$$\sigma_1 \geq \ldots \geq \sigma_n$$

那么A^{-1}的SVD为

$$A^{-1} = (U\Sigma V^T)^{-1} = V\Sigma^{-1}U^T$$

取

$$\begin{aligned} x &= u_1 \\ y &= \sigma_1 v_1 \\ \delta x &= \frac{1}{\sigma_n} v_n \\ \delta y &= u_n \end{aligned}$$

不难看出

$$\begin{aligned} Ax &= \sigma_1 v_1 \\ &= y \\ A\delta x &= v_n \\ &= \delta y \end{aligned}$$

所以

$$
\begin{aligned}
\frac{\|\delta x\|}{\|x\|} &= \frac{1}{\sigma_n} \\
\frac{\|\delta y\|}{\|y\|} &= \frac{1}{\sigma_1}
\end{aligned}
$$

注意到

$$
\kappa(A) = \|A\|\|A^{-1}\| = \frac{\sigma_1}{\sigma_n}
$$

所以等号可以成立。

15.6

记

$$
Y = \begin{bmatrix} y_1 & \dots & y_N \end{bmatrix}
$$

要使得ρ最小化，等价于最小化

$$
\begin{aligned}
\sum_{i=1}^N \left(q^T y_i\right)^2 &= \sum_{i=1}^N q^T y_i y_i^T q \\
&= q^T \left(\sum_{i=1}^N y_i y_i^T\right) q \\
&= q^T Y Y^T q
\end{aligned}
$$

假设Y的奇异值分解为

$$
Y = U\Sigma V^T
$$

那么

$$
\begin{aligned}
\sum_{i=1}^N \left(q^T y_i\right)^2 &= q^T Y Y^T q \\
&= q^T U\Sigma V^T V \Sigma^T U^T q \\
&= q^T U \Sigma^2 U^T q
\end{aligned}
$$

要使得上式最大，只要取$q = u_n$即可，此时

$$
\sum_{i=1}^N \left(q^T y_i\right)^2 = \sigma_n^2
$$

15.8

(a)

$$x(t)=e^{tA}x(0)$$

对于固定的t，对e^{tA}做奇异值分解

$$e^{tA}=U\Sigma V^T$$

注意到约束条件为$\|x(0)\|=1$，所以要使得$x(t)$模最大上式最大，必然有

$$x(0)=v_1$$

要使得$x(t)$模最大上式最小，必然有

$$x(0)=v_r$$

(b)

```
expA = expm(3 * A);
[U, S, V] = svd(expA);
% (a)
x0_1 = V(:, 1);

% (b)
x0_2 = V(:, 5);
```

15.10

(a)注意到，如果

$$\|As_1-As_2\|\ge 2V_{\max}$$

那么可以利用距离判别。

如果

$$\|y-As_1\|<\|y-As_2\|$$

则输出结果为s_1，否则输出结果为s_2。

注意到上式等价于

$$\begin{gathered}\|y-As_1\|^2<\|y-As_2\|^2\\ y^Ty-2y^TAs_1+\|As_1\|^2<y^Ty-2y^TAs_2+\|As_2\|^2\\ 2y^T(As_2-As_1)<\|As_2\|^2-\|As_1\|^2\end{gathered}$$

我们希望在下式最小的情形下达到最开始的条件

$$P_{\max}=\max\{\|s_1\|,\|s_2\|\}$$

假设A的SVD为

$$A=U\Sigma V^T$$

那么利用SVD的性质可得，只要选择

$$s_1 = ku_1, s_2 = -ku_1, k > 0$$

即可，带入原式可得

$$\|As_1 - As_2\| = \|2k\sigma_1 v_1\| = 2k\sigma_1 \geq 2V_{\max} \Rightarrow k \geq \frac{V_{\max}}{\sigma_1}$$

所以

$$s_1 = \frac{V_{\max}}{\sigma_1}u_1, s_2 = -\frac{V_{\max}}{\sigma_1}u_1$$

此时

$$P_{\max} = \max\{\|s_1\|, \|s_2\|\} = \frac{V_{\max}}{\sigma_1}$$

(b)

```
A = [2 4 5 4 5;    0 5 7 7 1;    7 8 0 6 7;    7 0 4 9 4;    9 1 1 8 7];
Vmax = 3;
[U, S, V] = svd(A);
k = Vmax / S(1, 1);
s1 = k * V(:, 1)
s2 = - k * V(:, 1)
```

```
s1 =
   -0.0606
   -0.0373
   -0.0312
   -0.0746
   -0.0549
s2 =
    0.0606
    0.0373
    0.0312
    0.0746
    0.0549
```

15.11

(a)回顾结论，我们有

$$x(t) = A^t x(0) + \mathcal{C}_t \begin{bmatrix} u(t-1) \\ \vdots \\ u(0) \end{bmatrix}$$

其中

$$\mathcal{C}_t = [\,B \quad AB \quad \cdots \quad A^{t-1}B\,]$$

记

$$\mathcal{R}_t = \mathrm{range}(\mathcal{C}_t)$$

要使得$x(T)=0$，只要

$$A^T x(0) + \mathcal{C}_T \begin{bmatrix} u(T-1) \\ \vdots \\ u(0) \end{bmatrix} = 0$$

即

$$A^T x(0) \in \mathcal{R}_T$$

可用如下方式判定

$$\mathrm{rank}(\mathcal{R}_T) == \mathrm{rank}([\mathcal{R}_T, A^T x(0)])$$

找到T之后，现在要求下式的最小范数解

$$\mathcal{C}_T \begin{bmatrix} u(T-1) \\ \vdots \\ u(0) \end{bmatrix} = -A^T x(0)$$

利用SVD的性质，我们可得

$$\begin{bmatrix} u(T-1) \\ \vdots \\ u(0) \end{bmatrix} = -\mathcal{C}_T^{\dagger} A^T x(0)$$

```
A = [1, 0, 0, 0; 1, 1, 0, 0; 0, 1, 1, 0; 1, 0, 0, 0];
B = [0, 1; 0, 1; 1, 0; 0, 0];
x0 = [1; 0; -1; 1];
T = 1;
C = B;
tmp = B;
x = A * x0;

while true
    if rank(C) == rank([C, x])
        break
    end
    x = A * x;
    tmp = A * tmp;
    C = [C, tmp];
    T = T + 1;
end
```

```
% (a)
u = - pinv(C) * x;
J1 = norm(u) ^ 2
```

(b)利用第7,8讲的内容求解该问题。

记

$$U = \begin{bmatrix} u(9) \\ \vdots \\ u(0) \end{bmatrix}$$

构造损失函数

$$J = \|x(10)\|^2 + \rho \|U\|^2$$

对每个ρ，我们最小化该损失函数。注意我们有

$$x(10) = A^{10} x(0) + \mathcal{C}_{10} U$$

所以

$$J = \left\| \begin{bmatrix} \mathcal{C}_{10} \\ \sqrt{\rho} I \end{bmatrix} U + \begin{bmatrix} A^{10} \\ 0 \end{bmatrix} x(0) \right\|^2$$

最优解为

$$U = -\left(\mathcal{C}_{10}^T \mathcal{C}_{10} + \rho I\right)^{-1} \mathcal{C}_{10}^T A^{10} x(0)$$

我们找到ρ，使得

$$\|x(10)\| = 0.1$$

然后计算相应的U即可。

```
% (b)
C = [];
tmp = B;
x10 = x0;
for i = 1:10
    C = [C, tmp];
    tmp = A * tmp;
    x10 = A * x10;
end

P = C;
v = - x10;
[m, n] = size(P);
N = 100;
Lambda = logspace(1, -1, N);
res = zeros(1, N);
for i = 1: N
```

```
    lambda = Lambda(i);
    u = inv(eye(n) + lambda * P' * P) * lambda * P' * v;
    res(i) = norm(P * u - v) - 0.1;
    if i > 1 && res(i) * res(i - 1) < 0
        u_res = u;
        %break;
    end
end

J9 = norm(u_res) ^ 2;
plot(res);
```