奇异值分析

主 讲 人: 马 丽 艳
办 公 室: 计 1013
Email: liyanma@t.shu.edu.cn
计算机工程与科学学院

主要内容

- 数值稳定性与条件数
- 奇异值分解
- 奇异值的工程应用案例

数值稳定性与条件数

方程Ax=b的解x，当A和b做微小扰动时，其相对误差如何？

若b有微小扰动，有

$$A(x+\delta x)=b+\delta b,\quad \delta x=A^{-1}\delta b$$

考虑扰动对精度的影响！

应用范数的性质有: $\|\delta x\|_2 \le \|A^{-1}\|_2\|\delta b\|_2$

对于$Ax=b$,应用范数的性质有,$\|b\|_2 \le \|A\|_2\|x\|_2$

此时，解的相对扰动为

$$\frac{\|\delta x\|_2}{\|x\|_2} \le (\|A\|_2\cdot\|A^{-1}\|_2)\frac{\|\delta b\|_2}{\|b\|_2}$$

数值稳定性与条件数

若A有微小扰动，有

$$(A+\delta A)(x+\delta x)=b,$$

$$\delta x=(A+\delta A)^{-1}b-x=(A+\delta A)^{-1}b-A^{-1}b$$

$$=A^{-1}A(A+\delta A)^{-1}b-A^{-1}(A+\delta A)(A+\delta A)^{-1}b$$

$$=A^{-1}(A-(A+\delta A))(A+\delta A)^{-1}b=-A^{-1}\delta A(x+\delta x)$$

应用范数的性质有: $\|\delta x\|_2 \le \left\|A^{-1}\right\|_2 \|\delta A\|_2 \|x+\delta x\|_2$

若A有微小扰动，则解的相对扰动为

$$\frac{\|\delta x\|_2}{\|x+\delta x\|_2} \le (\|A\|_2 \cdot \|A^{-1}\|_2)\frac{\|\delta \mathrm{A}\|_2}{\|\mathrm{A}\|_2}$$

数值稳定性与条件数

若b有微小扰动，则解的相对扰动为

$$\frac{\|\delta x\|_2}{\|x\|_2} \le (\|A\|_2 \cdot \|A^{-1}\|_2)\frac{\|\delta b\|_2}{\|b\|_2}$$

若A有微小扰动，则解的相对扰动为

$$\frac{\|\delta x\|_2}{\|x+\delta x\|_2} \le (\|A\|_2 \cdot \|A^{-1}\|_2)\frac{\|\delta \mathrm{A}\|_2}{\|\mathrm{A}\|_2}$$

数值稳定性

定义：矩阵A的条件数为

$$cond(A) = \|A\|_2 \cdot \|A^{-1}\|_2$$

数值稳定性与条件数

- 矩阵A是良态矩阵：

  若稀疏矩阵A很小的一个扰动，只引起解向量x很小的扰动。

- 矩阵A是病态矩阵：

  若稀疏矩阵A很小的一个扰动，引起解向量x很大的扰动。

- 条件数刻画了求解线性方程组时，误差经过矩阵A的传播。

数值稳定性与条件数

条件数的性质

（1）cond(A)≥1. 特别地，若A为正交矩阵或酉矩阵，则cond(A)=1.

（2）$cond(aA)=cond(A),(a\neq 0)$

（3）$cond(A^H A)=[cond(A)]^2$.

（4）$cond(QA)=cond(A)$, Q为酉矩阵.

从属范数(诱导范数)：$\|A\|_P=\max_{\|x\|_P=1}\|Ax\|_P$

- ◆ 列和范数：$\|A\|_1=\max_j\sum_{i=1}^m|a_{i,j}|$
- ◆ 谱范数：$\|A\|_2=\sqrt{\lambda_{\max}(A^H A)}=\|A\|_{spec}$
- ◆ 行和范数：$\|A\|_\infty=\max_i\sum_{j=1}^n|a_{i,j}|$

Matlab：
norm(X,1)： 1-norm of X.
norm(X,2)： 2-norm of X.
norm(X,Inf)： infinity norm of X.
norm(X,‘fro’)： Frobenius norm of X.

数值稳定性与条件数

```
clear all
 A=rand(5,5);
A=A+A';
cond(A)
[u,d]=eig(A)
d(4,4) = 0.0001;
B=u*d*u'
[u,d]=eig(B)
cond(B)
```

B =

0.7238	1.0638	0.9920	1.0196	1.2275
1.0638	0.5530	1.4702	1.3816	1.0046
0.9920	1.4702	1.3617	1.4289	1.6967
1.0196	1.3816	1.4289	1.5830	1.8911
1.2275	1.0046	1.6967	1.8911	1.3534

u =

-0.3463	-0.2963	0.4054	-0.7855	0.1044
-0.3807	0.6218	0.3772	0.0524	-0.5687
-0.4790	-0.4089	0.4261	0.6136	0.2129
-0.5071	-0.3534	-0.6307	-0.0309	-0.4682
-0.4986	0.4832	-0.3376	-0.0526	0.6334

d =

6.5262	0	0	0	0
0	-0.9253	0	0	0
0	0	0.1480	0	0
0	0	0	0.0001	0
0	0	0	0	-0.1740

cond(B) = 6.5262e+04 cond(A) = 44.1033

数值稳定性与条件数

```
clear all
 A=rand(5,5);
A=A+A';
cond(A)
[u,d]=eig(A)
d(4,4) = 0.0001;
B=u*d*u'
[u,d]=eig(B)
cond(B)
B=B+0.01*eye(5)
[u,d]=eig(B)
cond(B)
```

```
B =
    0.7338    1.0638    0.9920    1.0196    1.2275
    1.0638    0.5630    1.4702    1.3816    1.0046
    0.9920    1.4702    1.3717    1.4289    1.6967
    1.0196    1.3816    1.4289    1.5930    1.8911
    1.2275    1.0046    1.6967    1.8911    1.3634
u =
   -0.3463   -0.2963    0.4054   -0.7855    0.1044
   -0.3807    0.6218    0.3772    0.0524   -0.5687
   -0.4790   -0.4089    0.4261    0.6136    0.2129
   -0.5071   -0.3534   -0.6307   -0.0309   -0.4682
   -0.4986    0.4832   -0.3376   -0.0526    0.6334
d =
    6.5362       0         0         0         0
         0   -0.9153       0         0         0
         0       0    0.1580         0         0
         0       0         0    0.0101         0
         0       0         0         0   -0.1640
>> cond(B) =  647.1503
```

例：计算条件数与矩阵奇异性之间的关系

```
%基于dd生成一个矩阵序列
A=rand(5,5);
A=A+A';
[u,d]=eig(A);

N=100;
Y = zeros(N,1);
dd=1;
for i =1:N
    d = diag([5,4,3,2,dd/i]);
    A=u*d*u';
    %计算矩阵条件数
    Y(i)=cond(A);
end
%画Y的函数图象
plot(Y);
xlabel('d的取值. 越小其性能越接近奇异矩阵');
ylabel('条件数');
```

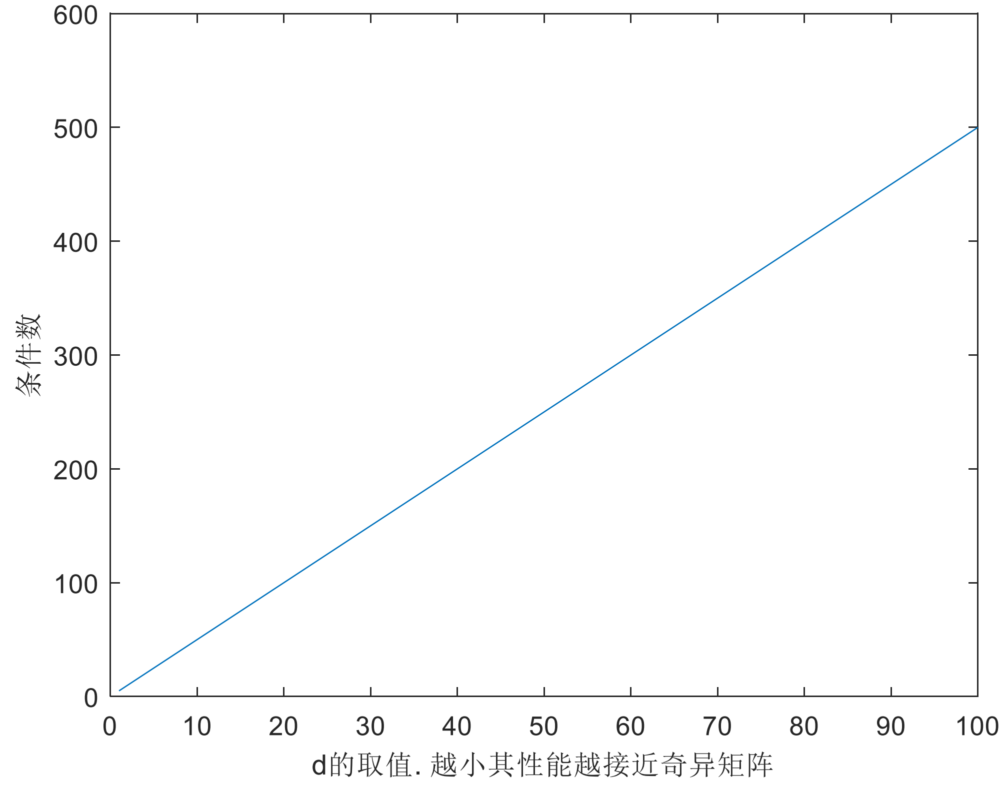


图2. 矩阵条件数与奇异性之间的关系

结论：矩阵越“靠近”奇异矩阵，条件数越大.

例：条件数与方程解的精度之间的关系

```
%基于dd生成一个矩阵序列
A=rand(5,5);  A=A+A';
[u,d]=eig(A);
N=100;
Y = zeros(N,1);
YY = zeros(N,1);
dd=1; b=rand(5,1);
dA=0.01*rand(5,5);
for i=1:N
    d = diag([5,4,3,2,dd/i]);
   A=u*d*u';  Y(i)=cond(A);
    %%%%
    X=A\b;    XX=(A+dA)\b;
    a1=norm(XX);
    a2=norm(XX-X);
    YY(i)=a2/a1;
end
plot(YY,Y);
```

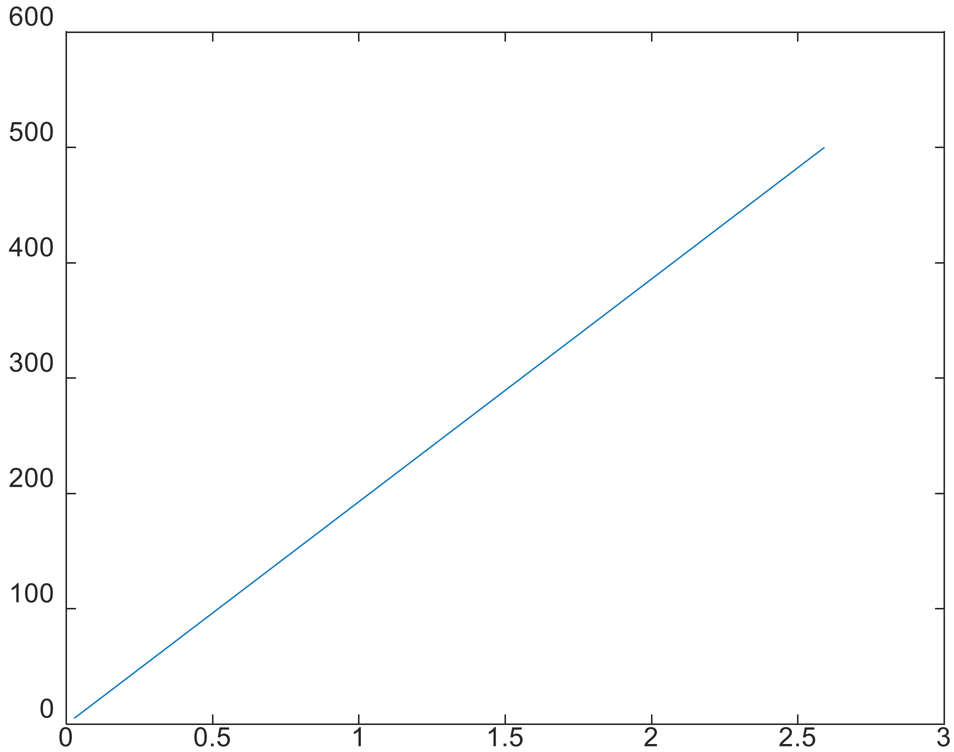


图3. 矩阵条件数与精度之间的关系

结论：相对精度与条件数成正比！

奇异值分解

矩阵的奇异值分解是指，将一个非零的 $m \times n$ 矩阵A， $A \in R^{m \times n}$ ，表示为以下三个实矩阵乘积形式的运算，即进行矩阵的因子分解：

$$A = U\Sigma V^T$$ （如果是复空间为 $U\Sigma V^H$ ）

其中U是m阶正交矩阵(orthogonal matrix)，V是一个n阶正交矩阵，Σ 是由降序排列的非负的对角线元素组成的 $m \times n$ 矩阵对角矩阵(rectangular diagonal matrix)，满足： $UU^T = I$ $VV^T = I$

$$\Sigma = diag(\sigma_1, \sigma_2, \cdots, \sigma_n)$$

$$\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_p \ge 0$$

$$p = \min(m, n)$$

奇异值分解

- $U\Sigma V^{\mathrm{T}}$称为矩阵A的奇异值分解（singular value decomposition，SVD）
- σ_i 称为矩阵 A的奇异值（singular value）
- U的列向量称为左奇异向量（left singular vector）
- V 的列向量称为右奇异向量（right singular vector）
- **注意：** 奇异值分解不要求矩阵A是方阵，事实上矩阵的奇异值分解可以看作是方阵的对角化的推广。

奇异值分解

例：

- 给定一个5×4 矩阵A

$$A=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 0 & 4\\ 0 & 3 & 0 & 0\\ 0 & 0 & 0 & 0\\ 2 & 0 & 0 & 0\end{pmatrix}$$

- 它的奇异值分解由三个矩阵的乘积 $U\Sigma V^T$ 给出，矩阵U、Σ 、V分别为

奇异值分解

$$U=\begin{pmatrix} 0 & 0 & \sqrt{0.2} & 0 & \sqrt{0.8} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \sqrt{0.8} & 0 & -\sqrt{0.2} \end{pmatrix} \quad \Sigma=\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad V^T=\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

- 矩阵Σ是对角矩阵，对角线外的元素都是0，对角线上的元素**非负**，按降序排列。
- 矩阵U和V是正交矩阵，它们与各自的转置矩阵相乘是单位矩阵，即 $U^TU=I, V^TV=I.$

奇异值分解

矩阵的奇异值分解不是唯一的。在此例中如果选择U为

$$U=\begin{pmatrix} 0 & 0 & \sqrt{0.2} & \sqrt{0.4} & -\sqrt{0.4} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{0.5} & \sqrt{0.5} \\ 0 & 0 & \sqrt{0.8} & -\sqrt{0.1} & -\sqrt{0.1} \end{pmatrix}$$

- 而 Σ 和V不变，那么 $U\Sigma V^T$ 也是A的一个奇异值分解 。

奇异值分解

```
>> A=randi(10,4,3)
     7     7     7
     1     8     2
     9     8     8
    10     4     1
>> B=A‘ * A
   231   169   133
   169   193   133
   133   133   118

>> [u,d]=eig(B)
u =
    0.0854    0.7515    0.6541
    0.5471   -0.5841    0.5996
   -0.8327   -0.3067    0.4611
d =
   16.9676         0          0
         0   45.3819          0
         0         0   479.6505
```

```
>> [u,s,v]=svd(B)
u =
   -0.6541    0.7515   -0.0854
   -0.5996   -0.5841   -0.5471
   -0.4611   -0.3067    0.8327
s =
  479.6505         0          0
         0   45.3819          0
         0         0    16.9676
v =
   -0.6541    0.7515   -0.0854
   -0.5996   -0.5841   -0.5471
   -0.4611   -0.3067    0.8327
```

奇异值分解

```
>> A=randi(10,4,4);
>> B=A'+A
     4    13    12    17
    13    16     5    11
    12     5     4    11
    17    11    11     6
>> [u,d]=eig(B)
u =
    0.7990   -0.2169    0.1804    0.5311
   -0.1168    0.2443   -0.7936    0.5449
   -0.1693    0.7459    0.5178    0.3833
   -0.5651   -0.5805    0.2639    0.5235
d =
  -12.4667         0         0         0
         0   -6.4123         0         0
         0         0    6.1244         0
         0         0         0   42.7546
```

```
>> [u,s,v]=svd(B)
u =
   -0.5311    0.7990   -0.2169   -0.1804
   -0.5449   -0.1168    0.2443    0.7936
   -0.3833   -0.1693    0.7459   -0.5178
   -0.5235   -0.5651   -0.5805   -0.2639
s =
   42.7546         0         0         0
         0   12.4667         0         0
         0         0    6.4123         0
         0         0         0    6.1244
v =
   -0.5311   -0.7990    0.2169   -0.1804
   -0.5449    0.1168   -0.2443    0.7936
   -0.3833    0.1693   -0.7459   -0.5178
   -0.5235    0.5651    0.5805   -0.2639
```

奇异值分解

```
>> A=randn(4,4);

>> [u,s,v]=svd(A)
u =
   -0.2807   -0.9030    0.1851   -0.2673
    0.0492   -0.1888   -0.9766   -0.0901
   -0.9423    0.3036   -0.0967   -0.1030
   -0.1758   -0.2381   -0.0509    0.9538
s =
    3.7970         0         0         0
         0    1.5679         0         0
         0         0    1.4289         0
         0         0         0    0.0828
v =
    0.2876    0.7680    0.5542    0.1424
    0.7967   -0.3341    0.1712   -0.4736
   -0.3536   -0.4596    0.8144    0.0236
    0.3968   -0.2955   -0.0196    0.8688
```

```
>> [u,d]=eig(A)
d =
  2.2591 + 0.0000i   0.0000 + 0.0000i   0.0000 + 0.0000i   0.0000 + 0.0000i
  0.0000 + 0.0000i  -1.3741 + 1.1072i   0.0000 + 0.0000i   0.0000 + 0.0000i
  0.0000 + 0.0000i   0.0000 + 0.0000i  -1.3741 - 1.1072i   0.0000 + 0.0000i
  0.0000 + 0.0000i   0.0000 + 0.0000i   0.0000 + 0.0000i   0.1001 + 0.0000i
>> abs(d)
ans =
    2.2591         0         0         0
         0    1.7647         0         0
         0         0    1.7647         0
         0         0         0    0.1001
```

奇异值分解

```
>> A=randn(4,3)+i*randn(4,3);

>> [u,s,v]=svd(A)
u =
  -0.1205 - 0.7855i   0.4087 - 0.1237i   0.0064 + 0.0144i  -0.0465 + 0.4287i
  -0.1034 - 0.4901i  -0.2965 + 0.3757i  -0.1966 + 0.3659i   0.2888 - 0.5139i
  -0.1159 + 0.0096i  -0.0522 - 0.3540i  -0.7360 + 0.1554i  -0.5277 - 0.1188i
  -0.1992 + 0.2539i   0.2744 + 0.6207i  -0.5070 - 0.0659i   0.2116 + 0.3594i
s =
    4.4518         0         0
         0    2.7072         0
         0         0    1.0172
         0         0         0
v =
   0.6886 + 0.0000i  -0.2025 + 0.0000i   0.6963 + 0.0000i
  -0.1507 - 0.7063i  -0.1052 - 0.2447i   0.1185 + 0.6273i
   0.0169 + 0.0635i  -0.7440 - 0.5783i  -0.2331 - 0.2309i
```

奇异值分解

奇异值分解基本定理

若A为一个 $m\times n$ 复矩阵，$A\in C^{m\times n}$，A 的奇异值分解存在

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

其中U是m阶酉矩阵，V是n阶酉矩阵，Σ 是 m x n 矩形对角矩阵，其对角线元素非负，且按降序排列。

奇异值分解

证明：证明是构造性的，对给定的矩阵A，构造出其奇异值分解的各个矩阵。

为了方便，不妨假设m≥n，如果m<n证明仍然成立。

（1）确定V和 Σ

首先构造n阶酉矩阵V和 m x n 矩形对角实矩阵 Σ 。

矩阵A是$m\times n$复矩阵，则矩阵 A^HA 是 n 阶Hermitian矩阵。

因而 A^HA 的特征值都是实数，并且存在一个 n 阶酉矩阵 V 实现 A^HA 的对角化，使得 $V^H\left(A^HA\right)V=\Lambda$ 成立。其中是 Λ 阶对角矩阵，其对角线元素由 A^HA 的特征值组成。

奇异值分解

而且，　$A^H A$ 的特征值都是非负的。事实上，λ 是 $A^H A$ 的一个特征值，x 是对应的特征向量，则

$$\|Ax\|^2 = (Ax)^H Ax = x^H (A^H A) x = \lambda x^H x = \lambda \|x\|^2$$

于是

$$\lambda = \frac{\|Ax\|^2}{\|x\|^2} \geq 0$$

可以假设酉矩阵V的列的排列使得对应的特征值形成降序排列

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

计算特征值的平方根（就是A的奇异值）$\sigma_j = \sqrt{\lambda_j}, j = 1,2,\cdots,n$

$$A^H A = (U\Sigma V^H)^H U\Sigma V^H = V\Sigma^H U^H U\Sigma V^H = V\Sigma^H \Sigma V^H$$

奇异值分解

现在设矩阵 A 的秩是 r，则矩阵 A^HA 的秩也是 r。由于 A^HA 是对称矩阵，它的秩等于正的特征值的个数，所以

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0, \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$$

对应地有 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$

令 $V_1 = [v_1, v_2, \cdots, v_r], V_2 = [v_{r+1}, v_{r+2}, \cdots, v_n]$

其中 $v_1, \cdots, v_r$ 为 A^HA 的**正特征值**对应的特征向量，$v_{r+1}, \cdots, v_n$ 为 0 特征值对应的特征向量，则 $V = [V_1 \quad V_2]$

这就是矩阵A的奇异值分解中的n阶酉矩阵V。

奇异值分解

令

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}$$

则 Σ_1 是一个 r 阶对角矩阵，其对角线元素为按降序排列的正的 $\sigma_1, \cdots, \sigma_r$，于是 m x n 矩形对角矩阵 Σ 可以表为

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix}$$

这就是矩阵A的奇异值分解中的 m x n 矩形对角矩阵 Σ 。

奇异值分解

V_2的列向量是 A^HA 对应于特征值为0的特征向量。因此

$$\left(Av_j\right)^H Av_j = v_j{}^H A^H Av_j = v_j{}^H \lambda v_j = v_j{}^H 0 v_j = 0$$

有 $Av_j = 0$ 。所以V_2的列向量构成A的零空间的一组标准正交基。因此，$AV_2 = 0$

由于V是酉矩阵，可得

$$I = VV^H = V_1V_1^H + V_2V_2^H$$

$$A = AI = AV_1V_1^H + AV_2V_2^H = AV_1V_1^H$$

奇异值分解

(2)确定U

接着构造 m 阶酉矩阵U。 令

$$u_j = \frac{1}{\sigma_j} A v_j, j = 1,2,\cdots,r \quad U_1 = [u_1, u_2, \cdots, u_r]$$

则有 $$AV_1 = U_1 \Sigma_1$$

U_1的列向量构成了一组标准正交集，并且

$$u_i^H u_j = \left(\frac{1}{\sigma_i} v_i^H A^H\right)\left(\frac{1}{\sigma_j} A v_j\right) = \frac{1}{\sigma_i \sigma_j} v_i^H \left(A^H A v_j\right)$$

$$= \frac{\sigma_j}{\sigma_i} v_i^H v_j = \delta_{ij}, i = 1,2,\cdots,r; j = 1,2,\cdots,r$$

故$\{u_1, u_2, \cdots, u_r\}$是$AA^H$的单位正交特征向量。

奇异值分解

令 $\{u_{r+1}, u_{r+2}, \ldots, u_m\}$ 为一组标准正交基，且满足

$$A^H u_j = 0, j = r+1, \cdots, m$$

并令 $U_2 = [u_{r+1} \quad u_{r+2} \quad \cdots \quad u_m]$，$U = [U_1 \quad U_2]$

有

$$u_j^H u_i = u_j^H \left(\frac{1}{\sigma_i} A v_i \right) = \frac{1}{\sigma_j} \left(A^H u_j \right)^H v_i = 0,$$

$$i = 1, 2, \cdots, r; j = r+1, \cdots, m$$

- 则 $u_1, u_2, \cdots, u_m$ 构成了 C^m 的一组标准正交基。因此，U是m阶酉矩阵。这就是矩阵A的奇异值分解中的m阶酉矩阵。

奇异值分解

（3）证明 $U\Sigma V^H = A$

综上，有

$$U\Sigma V^H = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix} = U_1 \Sigma_1 V_1^H = A V_1 V_1^H = A$$

至此证明了矩阵A存在奇异值分解。

奇异值分解的计算

步骤：

（1）求 A^HA 的特征值和特征向量

计算对称矩阵 $W = A^H A$

求解特征方程 $(W - \lambda I)x = 0$

得到特征值 λ_j ，并将特征值由大到小排列 $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge 0$

将特征值 $\lambda_i(i = 1,2,\cdots,n)$ 代入特征方程求得对应的特征向量

（2）求 n 阶酉矩阵V

将特征向量单位化，得到单位特征向量 $v_1, v_2, \cdots, v_n$ ，构成 n 阶酉矩阵V：

$$V = [v_1, v_2, \cdots, v_n]$$

奇异值分解的计算

(3) 求 $m\times n$ 对角矩阵

- 计算A的奇异值 $\sigma_i=\sqrt{\lambda_i}, i=1,2,\cdots,n$
- 构造 $m\times n$ 矩形对角矩阵 $\sum$，主对角线元素是奇异值，其余元素是零：$\boldsymbol{\Sigma} = diag(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \cdots, \boldsymbol{\sigma}_n)$

(4) 求 m 阶酉矩阵U

- 对A的前r个正奇异值，令 $u_j=\dfrac{1}{\sigma_j}Av_j, j=1,2,\cdots,r$ 得到 $U_1=\begin{bmatrix}u_1 & u_2 & \cdots & u_r\end{bmatrix}$
- 求A^H的零空间的一组标准正交基 $\{u_{r+1},u_{r+2},\cdots,u_m\}$，令

$U_2=\begin{bmatrix}u_{r+1} & u_{r+2} & \cdots & u_m\end{bmatrix}$ 并令 $U=\begin{bmatrix}U_1 & U_2\end{bmatrix}$。

(5) 得到奇异值分解 $A=U\Sigma V^H$

奇异值分解的计算

例：试求矩阵

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

的奇异值分解。

奇异值分解的计算

- （1）求 $A^T A$ 的特征值和特征向量

$$A^T A=\begin{bmatrix}1 & 2 & 0\\ 1 & 2 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 2 & 2\\ 0 & 0\end{bmatrix}=\begin{bmatrix}5 & 5\\ 5 & 5\end{bmatrix}$$

$$\left(A^T A-\lambda I\right)x=0$$

- 得到齐次线性方程组

$$\begin{cases}(5-\lambda)x_1+5x_2=0\\ 5x_1+(5-\lambda)x_2=0\end{cases}$$

奇异值分解的计算

- 该方程有非零解的充要条件是

$$\begin{vmatrix} 5-\lambda & 5 \\ 5 & 5-\lambda \end{vmatrix} = 0$$

$$\lambda^2 - 10\lambda = 0$$

- 解此方程，得矩阵A^{T}A的特征值 $\lambda_1 = 10$ 和 $\lambda_2 = 0$ 。
- 将特征值代入线性方程组，得到对应的单位特征向量

$$v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, v_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

奇异值分解的计算

（2）求正交矩阵V

- 构造正交矩阵V

$$V=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}$$

（3）求对角矩阵 Σ

- 奇异值为 $\sigma_1=\sqrt{\lambda_1}=\sqrt{10}$ 和 $\sigma_2=0$
- 构造对角矩阵

$$\Sigma=\begin{bmatrix}\sqrt{10} & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}$$

奇异值分解的计算

- （4）求正交矩阵U

基于A的正奇异值计算得到列向量u_1

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ 0 \end{bmatrix}$$

- 列向量u_2，u_3是A^T的零空间$N(A^T)$的一组标准正交基

奇异值分解的计算

- 求解

$$A^T x = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_1 + 2x_2 + 0x_3 = 0$$
$$\Rightarrow x_1 = -2x_2 + 0x_3$$

- 分别取(x_2, x_3)为(1, 0)和(0, 1)，得到$N(A^T)$的基

$$(-2,1,0)^T, (0,0,1)^T$$

- $N(A^T)$的一组标准正交基是

$$u_2 = \left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0\right)^T, u_3 = (0,0,1)^T$$

- 构造正交矩阵U

$$U = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} & 0 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

奇异值分解的计算

- （5）矩阵A的奇异值分解

$$A = U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} & 0 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

奇异值分解

```
>> A=randn(4,3)+i*randn(4,3);

>> [u,s,v]=svd(A)
u =
  -0.1205 - 0.7855i   0.4087 - 0.1237i   0.0064 + 0.0144i  -0.0465 + 0.4287i
  -0.1034 - 0.4901i  -0.2965 + 0.3757i  -0.1966 + 0.3659i   0.2888 - 0.5139i
  -0.1159 + 0.0096i  -0.0522 - 0.3540i  -0.7360 + 0.1554i  -0.5277 - 0.1188i
  -0.1992 + 0.2539i   0.2744 + 0.6207i  -0.5070 - 0.0659i   0.2116 + 0.3594i
s =
    4.4518         0         0
         0    2.7072         0
         0         0    1.0172
         0         0         0
v =
   0.6886 + 0.0000i  -0.2025 + 0.0000i   0.6963 + 0.0000i
  -0.1507 - 0.7063i  -0.1052 - 0.2447i   0.1185 + 0.6273i
   0.0169 + 0.0635i  -0.7440 - 0.5783i  -0.2331 - 0.2309i
```

奇异值分解

```
>> [u,d]=eig(A'*A)
u =
   0.4947 - 0.4901i   0.1599 - 0.1242i   0.1774 - 0.6654i
   0.5256 + 0.3622i   0.2333 + 0.1286i  -0.7213 - 0.0363i
  -0.3281 + 0.0000i   0.9424 + 0.0000i   0.0657 + 0.0000i
d =
    1.0346         0         0
         0    7.3291         0
         0         0   19.8189
>> sqrt(d)
ans =
    1.0172         0         0
         0    2.7072         0
         0         0    4.4518
```

```
s =
    4.4518         0         0
         0    2.7072         0
         0         0    1.0172
         0         0         0
```

奇异值分解

```
>> A=randn(3,4)

>> [u,s,v]=svd(A)
u =
   -0.6274   -0.7778    0.0367
    0.5587   -0.4169    0.7170
   -0.5424    0.4703    0.6961
s =
    3.8939         0         0         0
         0    1.5319         0         0
         0         0    0.4594         0
v =
   -0.2501   -0.8285    0.3443   -0.3641
    0.0815    0.4525    0.2454   -0.8535
   -0.0893   -0.1876   -0.9061   -0.3685
   -0.9606    0.2715    0.0154    0.0566
```

```
>> [u,d]=eig(A‘ *A)
u =
   -0.3641    0.3443    0.8285    0.2501
   -0.8535    0.2454   -0.4525   -0.0815
   -0.3685   -0.9061    0.1876    0.0893
    0.0566    0.0154   -0.2715    0.9606
d =
    0.0000         0         0         0
         0    0.2110         0         0
         0         0    2.3468         0
         0         0         0   15.1627
>> sqrt(d)
ans =
    0.0000         0         0         0
         0    0.4594         0         0
         0         0    1.5319         0
         0         0         0    3.8939
```

奇异值分解

```
>> A=randn(3,4)

>> [u,s,v]=svd(A)
u =
  -0.6274   -0.7778    0.0367
   0.5587   -0.4169    0.7170
  -0.5424    0.4703    0.6961
s =
   3.8939        0         0         0
        0   1.5319         0         0
        0        0    0.4594         0
v =
  -0.2501   -0.8285    0.3443   -0.3641
   0.0815    0.4525    0.2454   -0.8535
  -0.0893   -0.1876   -0.9061   -0.3685
  -0.9606    0.2715    0.0154    0.0566
```

```
>> [u,d]=eig(A*A‘ )
u =
  -0.0367   -0.7778    0.6274
  -0.7170   -0.4169   -0.5587
  -0.6961    0.4703    0.5424
d =
   0.2110        0         0
        0   2.3468         0
        0        0   15.1627
>> sqrt(d)
ans =
   0.4594        0         0
        0   1.5319         0
        0        0    3.8939
```

奇异值分解

```
>> A=randn(4,3)+i*randn(4,3);

>> [u,s,v]=svd(A)
u =
  -0.3284 - 0.2340i  -0.5609 + 0.2423i  -0.3160 - 0.5051i   0.2753 + 0.1825i
  -0.4886 - 0.2244i   0.1197 + 0.4996i   0.4934 + 0.0993i  -0.3742 + 0.2316i
   0.4885 + 0.5343i  -0.3616 + 0.4520i   0.2474 - 0.2048i  -0.1941 + 0.0072i
   0.1201 - 0.0991i  -0.0504 + 0.1586i   0.4675 + 0.2647i   0.8110 + 0.0423i

>> [u,d]=eig(A*A')
u =
  -0.2845 - 0.1679i  -0.5239 - 0.2838i  -0.4006 - 0.4613i   0.1045 + 0.3895i
   0.3617 - 0.2508i   0.4783 - 0.1567i  -0.4400 + 0.2653i   0.2342 + 0.4839i
   0.1934 - 0.0173i   0.1143 - 0.3001i  -0.5402 - 0.2079i  -0.0369 - 0.7230i
  -0.8121 + 0.0000i   0.5372 + 0.0000i  -0.1664 + 0.0000i  -0.1557 + 0.0000i
```

奇异值分解

```
>> A=randn(4,3)+i*randn(4,3);
>> [u,s,v]=svd(A)
>> abs(u)
ans =
   0.4033    0.6110    0.5958    0.3303
   0.5376    0.5138    0.5033    0.4401
   0.7239    0.5788    0.3211    0.1942
   0.1557    0.1664    0.5372    0.8121

>> [u,d]=eig(A*A')
>> abs(u)
ans =
   0.3303    0.5958    0.6110    0.4033
   0.4401    0.5033    0.5138    0.5376
   0.1942    0.3211    0.5788    0.7239
   0.8121    0.5372    0.1664    0.1557
```

奇异值分解

```
>> A=randn(4,3)+i*randn(4,3)；

>> [u,s,v]=svd(A)
v =

  -0.6673 + 0.0000i   0.6139 + 0.0000i  -0.4216 + 0.0000i
   0.0392 - 0.6791i   0.1805 - 0.3278i   0.2007 + 0.5975i
   0.2330 + 0.1942i   0.6874 + 0.1022i   0.6323 - 0.1586i

>> [u,d]=eig(A'*A)
u =
  -0.4089 - 0.1026i   0.6073 - 0.0903i  -0.5126 + 0.4273i
   0.0494 + 0.6284i   0.1303 - 0.3508i  -0.4047 - 0.5467i
   0.6519 + 0.0000i   0.6950 + 0.0000i   0.3033 + 0.0000i
```

奇异值分解

```
>> A=randn(4,3)+i*randn(4,3)；

>> [u,s,v]=svd(A)
>> abs(v)
ans =
   0.6673   0.6139   0.4216
   0.6802   0.3742   0.6303
   0.3033   0.6950   0.6519

>> [u,d]=eig(A'*A)
>> abs(u)
ans =
   0.4216   0.6139   0.6673
   0.6303   0.3742   0.6802
   0.6519   0.6950   0.3033
```

奇异值分解

紧奇异值分解与截断奇异值分解

- $A = U\Sigma V^H$ 又称为矩阵的完全奇异值分解（full singular value decomposition）。
- 实际常用的是奇异值分解的紧凑形式和截断形式。
- 紧奇异值分解是与原始矩阵等秩的奇异值分解。
- 截断奇异值分解是比原始矩阵低秩的奇异值分解。

紧奇异值分解

定义：设A为 $m \times n$ 实矩阵，其秩为rank(A)=r，r≤min(m,n)，则 $U_r \Sigma_r V_r^H$ 称为A的紧奇异值分解（compact singular value decomposition)，即

$$A = U_r \Sigma_r V_r^H$$

其中：U_r是 $m \times r$ 矩阵，V_r是 $n \times r$ 矩阵，Σ_r 是r阶对角矩阵；矩阵U_r、V_r、Σ_r 分别由完全奇异值分解中的U的前r列、矩阵V的前r列、矩阵 $\boldsymbol{\Sigma}$ 的前r个对角线元素得到。紧奇异值分解的对角矩阵的秩与原始矩阵A的秩相等。

紧奇异值分解

例： 矩阵A的秩 r = 3

$$A=\begin{pmatrix}1&0&0&0\\0&0&0&4\\0&3&0&0\\0&0&0&0\\2&0&0&0\end{pmatrix}$$

- A的紧奇异值分解是 $A=U_r\Sigma_rV_r^T$

其中，

$$U_r=\begin{bmatrix}0&0&\sqrt{0.2}\\1&0&0\\0&1&0\\0&0&0\\0&0&\sqrt{0.8}\end{bmatrix},\Sigma_r=\begin{bmatrix}4&0&0\\0&\sqrt{3}&0\\0&0&\sqrt{5}\end{bmatrix},V_r^T=\begin{bmatrix}0&0&0&1\\0&1&0&0\\1&0&0&0\end{bmatrix}$$

紧奇异值分解

B =

6.9100	2.8900	6.3700	3.0100	3.3400
6.5000	2.8000	5.6000	2.9000	3.2000
3.4100	1.5000	2.7800	1.5500	1.7000
5.0000	2.0600	4.4600	2.2400	2.3600
2.3600	0.8700	1.5500	1.2800	0.9200
2.5300	1.0800	1.7200	1.2700	1.1800
4.9200	1.4600	4.0200	2.6400	1.5600

```
>> [u,s,v]=svd(B)
```

u =

-0.5552	0.3252	-0.3703	-0.6092	0.0945	-0.1548	0.2120
-0.5139	0.2447	0.2010	0.6257	0.3840	0.0649	0.3042
-0.2662	0.1013	0.3246	0.0319	-0.0266	-0.6311	-0.6430
-0.3970	0.1257	-0.1295	0.1900	-0.7101	0.4404	-0.2747
-0.1704	-0.3296	0.4692	-0.1099	-0.4783	-0.2959	0.5602
-0.1877	-0.1229	0.6025	-0.4254	0.2720	0.5402	-0.2000
-0.3673	-0.8274	-0.3424	0.0853	0.1897	-0.0144	-0.1406

紧奇异值分解

```
s =
   19.4601        0         0         0         0
         0   1.0516         0         0         0
         0        0    0.6767         0         0
         0        0         0    0.0000         0
         0        0         0         0    0.0000
         0        0         0         0         0
         0        0         0         0         0
v =
   -0.6554   -0.3313    0.2275    0.4048   -0.4951
   -0.2645    0.3882    0.4015   -0.7206   -0.3143
   -0.5647    0.2238   -0.7704   -0.1531    0.1185
   -0.3026   -0.6042    0.2114   -0.3612    0.6067
   -0.3001    0.5695    0.3857    0.4036    0.5234
>> norm(B-u(:,1:3)*s(1:3,1:3)*v(:,1:3)','fro')
ans =   3.6839e-15
```

截断奇异值分解

在矩阵的奇异值分解中，只取最大的k个奇异值（k<r，r为矩阵的秩）对应的部分，就得到矩阵的截断奇异值分解。实际应用中提到矩阵的奇异值分解时，通常指截断奇异值分解。

- 截断奇异值分解

定义：设A为 $m\times n$ 复矩阵，其秩rank(A)=r，且 $0<k<r$，则称 $U_k\Sigma_k V_k^H$为矩阵A的截断奇异值分解（truncated singular value decomposition），有 $A\approx U_k\Sigma_k V_k^H$

其中 U_k 是 $m\times k$ 矩阵，V_k 是 $n\times k$矩阵，Σ_k 是k阶对角矩阵，分别由完全奇异值分解U的前k列、V的前k列、Σ 的前k个对角线元素得到的。对角矩阵 Σ_k 的秩比A低。

截断奇异值分解

矩阵A

$$A=\begin{pmatrix}1&0&0&0\\0&0&0&4\\0&3&0&0\\0&0&0&0\\2&0&0&0\end{pmatrix}$$

秩3，若取k=2则其截断奇异值是 $A \approx A_2 = U_2\Sigma_2 V_2^T$

截断奇异值分解

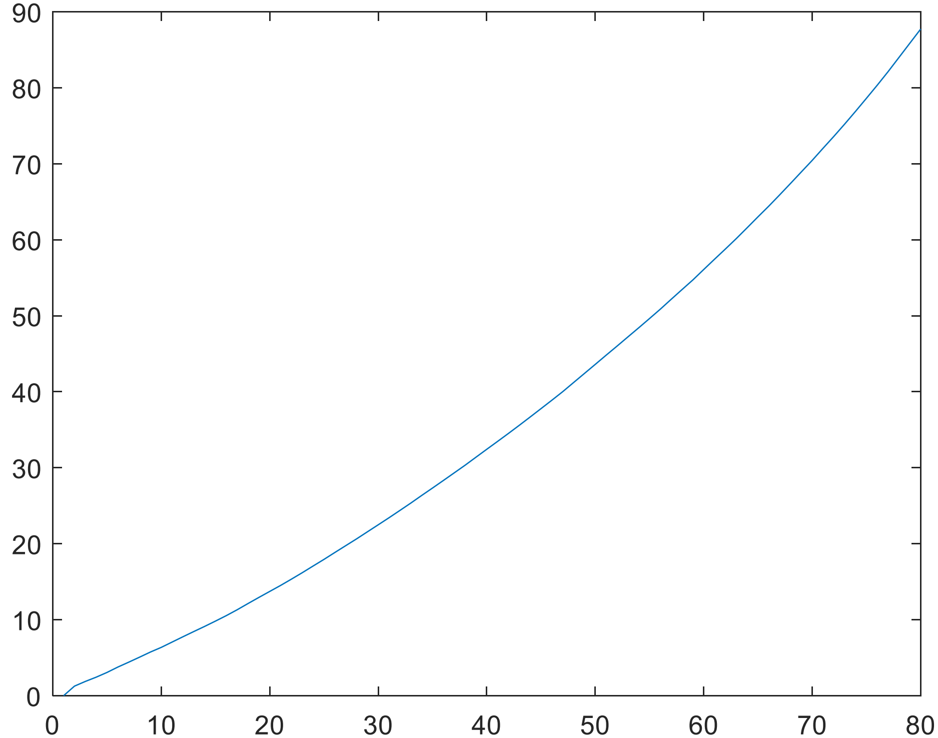
其中 $$U_2=\begin{bmatrix}0&0\\1&0\\0&1\\0&0\\0&0\end{bmatrix},\Sigma_2=\begin{bmatrix}4&0\\0&3\end{bmatrix},V_2^T=\begin{bmatrix}0&0&0&1\\0&1&0&0\end{bmatrix},$$

$$A_2=U_2\Sigma_2V_2^T=\begin{bmatrix}0&0&0&0\\0&0&0&4\\0&3&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

$$A=\begin{bmatrix}1&0&0&0\\0&0&0&4\\0&3&0&0\\0&0&0&0\\2&0&0&0\end{bmatrix}$$

截断奇异值分解

```
A=randn(100,80); [u,s,v] = svd(A);    k=79;   err=zeros(k+1,1);
for i=1:k+1
   err(i)=norm(A-u(:,1:80-i+1)*s(1:80-i+1,1:80-i+1)*v(:,1:80-i+1) ','fro');
end
plot(err)
```



奇异值分解的几何解释

从线性变换的角度理解奇异值分解，$m \times n$ 矩阵A表示从n维空间R^n 到m维空间 R^m 的一个线性变换,$T : x \to Ax$,

$x \in R^n, Ax \in R^m$ 分别是各自空间的向量。

即：将矩阵和空间中的线性变换视为同样的事物。

线性变换可以分解为三个简单的变换：

- 一个坐标系的旋转或反射变换
- 一个坐标轴的缩放变换
- 另一个坐标系的旋转或反射变换

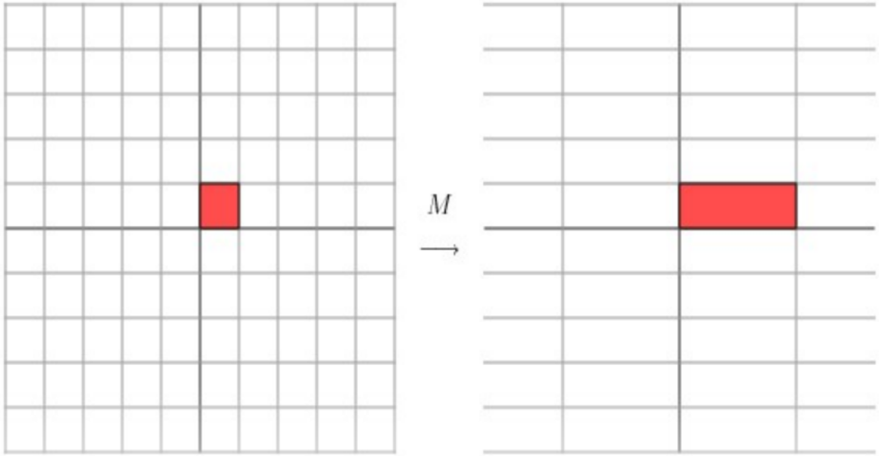
奇异值定理保证这种分解一定存在。这就是奇异值分解的几何解释。

奇异值分解的几何解释

对角矩阵

$$M=\begin{bmatrix}3 & 0\\0 & 1\end{bmatrix}, \quad \begin{bmatrix}3 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}3x\\y\end{bmatrix}.$$

从几何上讲，M是将二维平面上的点(x,y)经过线性变换到另外一个点的变换矩阵，变换的效果如下图所示，变换后的平面仅仅是沿 X 水平方面进行了拉伸3倍，垂直方向是并没有发生变化。



奇异值分解的几何解释

对矩阵A进行奇异值分解，得到 $A = U\Sigma V^T$

其中V和U都是正交矩阵。

V的列向量v_1，v_2，…，v_n构成 R^n 空间的一组标准正交基，表示R中的正交坐标系的旋转或反射变换。

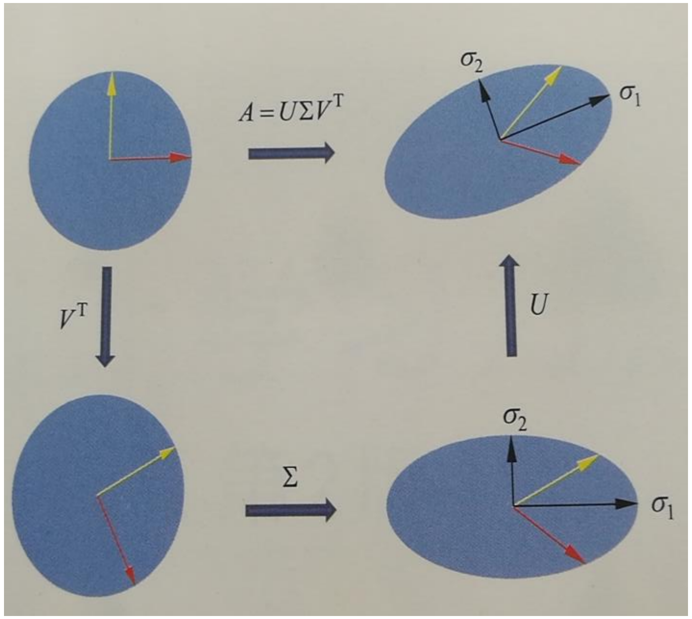
U的列向量u_1，u_2，…，u_m构成 R^m 空间的一组标准正交基，表示 R^m 中的正交坐标系的旋转或反射变换。

Σ 的对角元素 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 是一组非负实数，表示R^n中的原始正交坐标系坐标轴的 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 倍的缩放变换。

奇异值分解的几何解释

- 任意一个向量 $x \in R^n$ ，经过基于 $A = U\Sigma V^T$ 的线性变换，等价于经过坐标系的旋转或反射变换V^T，坐标轴的缩放变换 Σ，以及坐标系的旋转或反射变换U，得到向量 $Ax \in R^m$ 。

- 原始空间的标准正交基，经过坐标系的旋转变换V^T、坐标轴的缩放变换 Σ 、坐标系的旋转变换U，得到和经过线性变换A等价的结果。



奇异值分解的几何解释

- 给定一个2阶矩阵

$$A=\begin{bmatrix}3 & 1\\ 2 & 1\end{bmatrix}$$

- 其奇异值分解为

$$U=\begin{bmatrix}0.8174 & -0.5760\\ 0.5760 & 0.8174\end{bmatrix},\ \Sigma=\begin{bmatrix}3.8643 & 0\\ 0 & 0.2588\end{bmatrix},$$

$$V^T=\begin{bmatrix}0.9327 & 0.3606\\ -0.8174 & 0.9327\end{bmatrix}$$

奇异值分解的几何解释

- 观察基于矩阵A的奇异值分解将R^2的标准正交基

$$e_1=\begin{bmatrix}1\\0\end{bmatrix}, e_2=\begin{bmatrix}0\\1\end{bmatrix}$$

进行线性转换的情况。

- 首先，V^T表示一个旋转变换，将标准正交基e_1，e_2旋转，得到向量$V^T e_1$，$V^T e_2$：

$$V^Te_1=\begin{bmatrix}0.9327\\-0.3606\end{bmatrix}, V^Te_2=\begin{bmatrix}0.3606\\0.9327\end{bmatrix}$$

奇异值分解的几何解释

- 其次，Σ表示一个缩放变换，将向量$V^T e_1$，$V^T e_2$ 在坐标轴方向缩放 σ_1 倍和 σ_2 倍，得到向量 $\Sigma V^T e_1, \Sigma V^T e_2$：

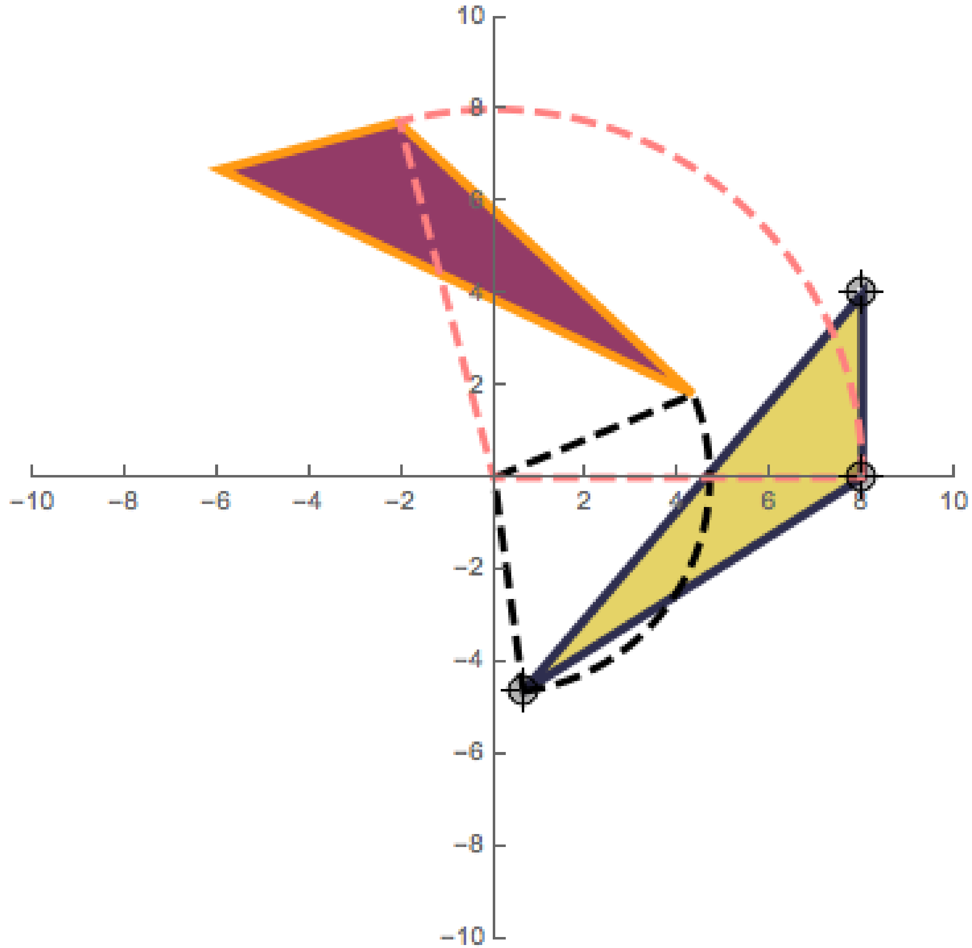
$$\Sigma V^T e_1 = \begin{bmatrix} 3.6042 \\ -0.0933 \end{bmatrix}, \Sigma V^T e_2 = \begin{bmatrix} 1.3935 \\ 0.2414 \end{bmatrix}$$

- 最后，U表示一个旋转变换，再将向量 $\Sigma V^T e_1, \Sigma V^T e_2$ 旋转，得到向量 $U\Sigma V^T e_1, U\Sigma V^T e_2$，也就是向量 Ae_1, Ae_2 ：

$$Ae_1 = U\Sigma V^T e_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, Ae_2 = U\Sigma V^T e_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

奇异值分解的几何解释

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} -0.259 & -0.966 \\ 0.966 & -0.259 \end{pmatrix}$$



奇异值分解的几何解释

```
>> A=[1 0; 1 0]
A =
     1     0
     1     0

>> [u,s,v] = svd(A)
u =
   -0.7071   -0.7071
   -0.7071    0.7071

s =
    1.4142         0
         0         0

v =
    -1     0
     0     1
```

```
>> u(:,1)=u(:,1)*-1
u =

    0.7071   -0.7071
    0.7071    0.7071

>> v(:,1)=v(:,1)*-1
v =

     1     0
     0     1

>> u*s*v'
ans =

    1.0000         0
    1.0000         0
```

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

奇异值分解的主要性质

（1）设矩阵A的奇异值分解为$A=U\Sigma V^T$，则一下关系成立：

$$A^T A=\left(U\Sigma V^T\right)^T\left(U\Sigma V^T\right)=V\left(\Sigma^T\Sigma\right)V^T$$

$$AA^T=\left(U\Sigma V^T\right)\left(U\Sigma V^T\right)^T=U\left(\Sigma^T\Sigma\right)U^T$$

就是说，矩阵A^TA和AA^T的特征分解存在，且可以由矩阵A的奇异值分解的矩阵表示。V的列向量是A^TA的特征向量，U的列向量是AA^T的特征向量。Σ 的奇异值是A^TA和AA^T的特征值的平方根。

$$A^H A=(U\Lambda V^H)^H U\Lambda V^H=V\Lambda^H\Lambda V^H$$

$$AA^H=U\Lambda V^H(U\Lambda V^H)^H=U\Lambda\Lambda^H U^H$$

奇异值分解的主要性质

- （2）在矩阵A的奇异值分解中，奇异值、左奇异向量和右奇异向量之间存在对应关系。

假设m≥n由 $A = U\Sigma V^H$ 易知 $AV = U\Sigma$ 。

比较这一等式两端的第j列，得到 $Av_j = \sigma_j u_j, j = 1,2,\cdots,n$ 。

这是矩阵A的右奇异向量和奇异值、 左奇异向量的关系。

类似地，由 $A^H U = \Sigma^H V$ 得到

$$A^H u_j = \sigma_j v_j, j = 1,2,\cdots,n; A^H u_j = 0, j = n+1, n+2,\cdots,m$$

这是矩阵A的左奇异向量和奇异值、右奇异向量的关系。

奇异值分解的主要性质

（3）矩阵A的奇异值分解中，奇异值 $\sigma_1,\sigma_2,\cdots,\sigma_n$ 是唯一的，而矩阵U和V不是唯一的。

（4）矩阵A和 Σ 的秩相等，等于正奇异值 σ_i 的个数r（包含重复的奇异值）。

（5）矩阵A的r个右奇异向量v_1，v_2，…，v_r构成A^H的值域$R(A^H)$的一组标准正交基。

（因为矩阵A^H是从C^m映射到C^n的线性变换，则A^H的值域$R(A^H)$和A^H列空间是相同的，v_1，v_2，…，v_r是A^H的一组标准正交基，因而也是$R(A^H)$的一组标准正交基。）

奇异值分解

- 奇异值分解也是一种矩阵近似的方法，这个近似是在弗罗贝尼乌斯范数（Frobenius norm）意义下的近似。

- 矩阵的弗罗贝尼乌斯范数是向量的L_2范数的直接推广，对应着机器学习中的平方损失函数。

- **定义**（弗罗贝尼乌斯范数）设矩阵 $A \in C^{m\times n}, A=\left[a_{ij}\right]_{m\times n}$，定义矩阵A的弗罗贝尼乌斯范数为

$$\|A\|_F=\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}\right|^2\right)^{\frac{1}{2}}$$

奇异值分解

- 引理：

设矩阵 $A \in C^{m \times n}$，A 的奇异值分解为 $U\Sigma V^H$，其中 $\Sigma = diag(\sigma_1, \sigma_2, \cdots, \sigma_n)$，则

$$\|A\|_F = \left(\sigma_1{}^2 + \sigma_2{}^2 \cdots + \sigma_n{}^2\right)^{\frac{1}{2}}$$

奇异值分解

证明：

一般地，若Q是 m 阶酉矩阵，则有 $\|QA\|_F = \|A\|_F$

因为

$$\|QA\|_F^2 = \|(Qa_1, Qa_2, \cdots, Qa_n,)\|_F^2 = \sum_{i=1}^{n} \|Qa_i\|_2^2 = \sum_{i=1}^{n} \|a_i\|_2^2 = \|A\|_F^2$$

同样，若P是 n 阶酉矩阵，则有 $\|AP^H\|_F = \|A\|_F$.

有 $\|A\|_F = \|U\Sigma V^H\|_F = \|\Sigma\|_F$

即

$$\|A\|_F = \left(\sigma_1{}^2 + \sigma_2{}^2 \cdots + \sigma_n{}^2\right)^{\frac{1}{2}}$$

矩阵的最优近似

- 定理：设 $A \in C^{m\times n}$ 复矩阵且秩rank(A)=r，有奇异值分解 $A=U\Sigma V^H$，并设M为$C^{m\times n}$ 中所有秩不超过 k 的矩阵集合，$0<k<r$，存在一个秩为k的矩阵 $X\in M$，满足 $\|A-X\|_F=\min\limits_{S\in M}\|A-S\|_F$

则 $\|A-X\|_F=\left(\sigma_{k+1}^2+\sigma_{k+2}^2+\sigma_n^2\right)^{\frac{1}{2}}$

特别地，若 $A'=U\Sigma' V^H$，其中

$$\Sigma'=\begin{bmatrix}\sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0\end{bmatrix}=\begin{bmatrix}\Sigma_k & 0\\ 0 & 0\end{bmatrix}$$

则

$$\|A-A'\|_F=\left(\sigma_{k+1}^2+\sigma_{k+2}^2+\cdots+\sigma_n^2\right)^{\frac{1}{2}}=\min_{S\in M}\|A-S\|_F$$

矩阵的最优近似

- 证明：
- 令 $X \in M$ 满足式 $\|A-X\|_F = \min_{S\in M}\|A-S\|_F$ 的一个矩阵。由于

$$\|A-X\|_F \le \|A-A'\|_F = (\sigma_{k+1}{}^2 + \sigma_{k+2}{}^2 + \cdots + \sigma_n{}^2)^{\frac{1}{2}}$$

- 下面证明 $\|A-X\|_F \ge (\sigma_{k+1}{}^2 + \sigma_{k+2}{}^2 + \cdots + \sigma_n{}^2)^{\frac{1}{2}}$ ，于是式 $\|A-X\|_F = \left(\sigma_{k+1}^2 + \sigma_{k+2}^2 + \sigma_n^2\right)^{\frac{1}{2}}$ 成立。

上海大学 SHANGHAI UNIVERSITY

矩阵的最优近似

- 设X的奇异值分解为 $X = Q\Omega P^{\mathrm{H}}$

其中

$$\Omega = \begin{pmatrix} \omega_1 & & & & & \\ & \ddots & & & & \\ & & \omega_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = \begin{bmatrix} \Omega_k & 0 \\ 0 & 0 \end{bmatrix}$$

- 若令矩阵B=Q$^{\mathrm{H}}$AP，则A=QBP$^{\mathrm{H}}$。由此得到

$$\|A - X\|_F = \|Q(B - \Omega)P^H\|_F = \|B - \Omega\|_F$$

矩阵的最优近似

- 用 Ω 分块方法对B分块

$$B=\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

- 其中B_{11}是 $k\times k$ 矩阵，B_{12}是 $k\times(n-k)$ 的子矩阵，B_{21}是 $(m-k)\times k$ 矩阵，B_{22} 是 $(m-k)\times(n-k)$ 子矩阵。可得

$$\|A-X\|_F^2=\|B-\Omega\|_F^2=\|B_{11}-\Omega_k\|_F^2+\|B_{12}\|_F^2+\|B_{21}\|_F^2+\|B_{22}\|_F^2$$

矩阵的最优近似

- 现证$B_{12}=0$，　$B_{21}=0$。　用反证法。若$B_{12}\neq 0$，令

$$Y=Q\begin{pmatrix}\Omega_k & B_{12}\\ 0 & 0\end{pmatrix}P^H$$

- 则 $Y\in M$ ，且

$$\|A-Y\|_F^2=\|B_{11}-\Omega_k\|_F^2+\|B_{21}\|_F^2+\|B_{22}\|_F^2<\|A-X\|_F^2$$

- 这与X的定义式 $\|A-X\|_F=\min\limits_{S\in M}\|A-S\|_F$ 矛盾
- 因此$B_{12}=0$，　同样可证$B_{21}=0$。于是

$$\|A-X\|_F^2=\|B_{11}-\Omega_k\|_F^2+\|B_{22}\|_F^2$$

矩阵的最优近似

- 再证 $B_{11}=\Omega_k$ ，为此令

$$Z=Q\begin{pmatrix} B_{11} & 0 \\ 0 & 0 \end{pmatrix}P^H$$

- 则 $Z\in M$ ，且 $\|A-Z\|_F^2=\|B_{22}\|_F^2\le\|B_{11}-\Omega_k\|_F^2+\|B_{22}\|_F^2=\|A-X\|_F^2$
- 由 $\|A-X\|_F=\min\limits_{S\in M}\|A-S\|_F$ 知，$\|B_{11}-\Omega_k\|_F^2=0$ 即 $B_{11}=\Omega_k$
- 最后看B_{22}。若 $(m-k)\times(n-k)$ 子矩阵B_{22}有奇异值分解 $U_1\Lambda V_1^H$ 则 $\|A-X\|_F=\|B_{22}\|_F=\|\Lambda\|_F$

矩阵的最优近似

证明 Λ 的对角线元素为A的奇异值。为此，令

$$U_2=\begin{pmatrix} I_k & 0 \\ 0 & U_1 \end{pmatrix}, V_2=\begin{pmatrix} I_k & 0 \\ 0 & V_1 \end{pmatrix}$$

其中I_k是k阶单位矩阵，U_2，V_2的分块与B的分块一致

注意到B及B_{22}的奇异值分解，即得

$$U_2^H Q^H A P V_2=\begin{pmatrix} \Omega_k & 0 \\ 0 & \Lambda \end{pmatrix} \quad A=(QU_2)\begin{pmatrix} \Omega_k & 0 \\ 0 & \Lambda \end{pmatrix}(PV_2)^H$$

由此可知 Λ 的对角线元素为A的奇异值，故有

$$\|A-X\|_F=\|\Lambda\|_F \geq \left(\sigma_{k+1}^2+\sigma_{k+2}^2+\cdots+\sigma_n^2\right)^{\frac{1}{2}}$$

$$\|A-X\|_F=\left(\sigma_{k+1}^2+\sigma_{k+2}^2+\cdots+\sigma_n^2\right)^{\frac{1}{2}}=\|A-A'\|_F$$

矩阵的最优近似

- 在秩不超过k的 $m \times n$ 矩阵的集合中，存在矩阵A的弗罗贝尼乌斯范数意义下的最优近似矩阵X。
- $A' = U\Sigma' V^H$ 是达到最优值的一个矩阵。
- 紧奇异值分解是在弗罗贝尼乌斯范数意义下的无损压缩。
- 截断奇异值分解是有损压缩。
- 截断奇异值分解得到的矩阵的秩为k，通常远小于原始矩阵的秩r，所以是由低秩矩阵实现了对原始矩阵的压缩。

矩阵的外积展开式

- 矩阵A的奇异值分解$U\Sigma V^H$也可以由外积形式表示

$$A=\sigma_1 u_1 v_1{}^H+\sigma_2 u_2 v_2{}^H+\cdots+\sigma_n u_n v_n{}^H \qquad A=\sum_{i=1}^{n}\sigma_i u_i v_i^H = U\Lambda V^H$$

- 若将A的奇异值分解看成矩阵$U\Sigma$ 和V^H的乘积，将$U\Sigma$按列向量分块，将 V^H 按行向量分块，即得

$$U\Sigma=\begin{bmatrix}\sigma_1 u_1 & \sigma_2 u_2 & \cdots & \sigma_n u_n\end{bmatrix}$$

$$V^H=\begin{bmatrix} v_1{}^H \\ v_2{}^H \\ \vdots \\ v_n{}^H \end{bmatrix}$$

矩阵的外积展开式

由矩阵A的外积展开式知，若A的秩为n，则

$$A=\sigma_1 u_1 v_1^H+\sigma_2 u_2 v_2^H+\cdots+\sigma_n u_n v_n^H$$

设矩阵 $A_k=\sigma_1 u_1 v_1^H+\sigma_2 u_2 v_2^H+\cdots+\sigma_k u_k v_k^H$

则A_k的秩为k，并且A_k是秩为k矩阵在弗罗贝尼乌斯范数意义A的最优近似矩阵。

矩阵A_k就是A的截断奇异值分解。

由于通常奇异值 σ_i 递减很快，所以k取很小值时，A_k也可以对A有很好的近似。

矩阵的外积展开式

例

- 给出矩阵A

$$A=\begin{pmatrix}1&0&0&0\\0&0&0&4\\0&3&0&0\\0&0&0&0\\2&0&0&0\end{pmatrix}$$

的秩为3，求A的秩为2的最优近似。

SHANGHAI UNIVERSITY

矩阵的外积展开式

- 从前例已知

$$u_1=\begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix},u_2=\begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix},v_1=\begin{bmatrix}0\\0\\0\\1\end{bmatrix},v_2=\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\sigma_1=4,\sigma_2=3$$

- 于是得到A的最优近似

$$A_2=\sigma_1u_1v_1^T+\sigma_2u_2v_2^T=\begin{pmatrix}0&0&0&0\\0&0&0&4\\0&3&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

奇异值分解的主要性质

定理（Eckart-Young）：设 $A\in C^{m\times n}$ 的奇异值分解为 $A=U\Sigma V^H$ 且 $\sigma_1\ge\sigma_2\ge\cdots\ge\sigma_r>0, r=rank(A)$ ．令 $A_k=\sum_{i=1}^{k}\sigma_i u_i v_i^H, k\le r$

则

$$\min_{rank(X)\le k}\|A-X\|_{spec}=\|A-A_k\|_{spec}=\sigma_{k+1}$$（谱范数逼近）

$$\min_{rank(X)\le k}\|A-X\|_F=\|A-A_k\|_F=\sqrt{\sigma_{k+1}^2+\cdots+\sigma_r^2}$$（F范数逼近）

推论：设 $A\in C^{m\times n}$ 的奇异值为 $\sigma_1\ge\cdots\ge\sigma_r>0, r=rank(A)$，则

$$\sigma_k=\min_{E\in C^{m\times n}}\left\{\|E\|_{spec}:rank(A+E)\le k-1\right\}, k=1,\cdots,r$$

且最佳扰动矩阵为 $E_k=-\sum_{i=k}^{r}\sigma_i u_i v_i^H$　　矩阵的低秩逼近

此时，$\|E\|_{spec}=\sigma_k$ 且 $rank(A+E_k)=k-1$.

奇异值分解的主要性质

矩阵的奇异值与条件数的关系： $cond(A)=\dfrac{\sigma_1}{\sigma_p}, p=\min\{m,n\}$

因此，（1）条件数是一个大于或等于1的整数.

（2）奇异矩阵的条件数为无穷大.

（3）条件数虽不是无穷大，但却很大时，A是接近奇异的，即此时其行或列线性相关性很强.

（4）$cond\left(A^H A\right)=\sigma_1^2/\sigma_p^2=[cond(A)]^2$

奇异值与矩阵范数的关系

（1）谱范数：$\|A\|_{spec}=\sigma_1$

（2）Frobenius范数：$\|A\|_F=\left[\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right]^{1/2}=\sqrt{\sigma_1^2+\sigma_2^2+\cdots+\sigma_r^2}$

奇异值分解和特征值分解

奇异值分解与特征值之间的联系与区别

（1）奇异值分解适用于任何长方形矩阵，特征值分解只适用于正方矩阵.

（2）即使是同一个n×n非Hermitian矩阵A，奇异值和特征值的定义也是完全不同的：奇异值定义为

最佳扰动矩阵

$$\sigma_k = \min_{E \in C^{m \times n}} \left\{ \|E\|_{spec} : rank(A+E) \le k-1 \right\}$$

$$\left(k \le \min\{m,n\}\right)$$

而特征值定义为特征多项式 $\det(A-\lambda I)=0$ 的根.

同一个正方矩阵的奇异值和特征值之间没有内在的关系，但是m×n矩阵A的非零奇异值是n×n Hermitian矩阵 $A^H A$ 或m×m Hermitian矩阵 AA^H 的非零特征值的正平方根.

奇异值分解和特征值分解

设n×n正方矩阵A的特征值为

$$\lambda_1, \lambda_2, \cdots, \lambda_n \qquad \left(|\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_n|\right)$$

奇异值为

$$\sigma_1, \sigma_2, \cdots, \sigma_n \qquad \left(\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_n \ge 0\right)$$

则有

$$\sigma_1 \ge |\lambda_i| \ge \sigma_n \qquad \left(i = 1, 2, \cdots, n\right)$$

$$cond(A) \ge \frac{|\lambda_1|}{|\lambda_n|}$$

特别地，当A（共轭）对称时，L2-范数下的条件数

$$cond(A) = \frac{|\lambda_1|}{|\lambda_n|}$$

奇异值分解和特征值分解

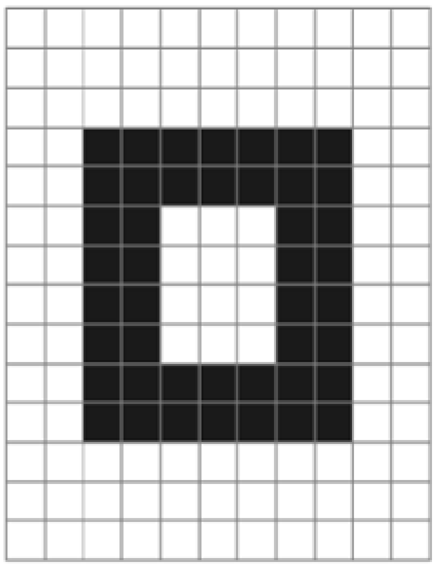
（3）m×n矩阵A的左、右奇异向量 $u_i^H A v_i = \sigma_i$

而n×n矩阵A的左、右特征向量 $u^H A = \lambda_i u^H, Av_i = \lambda_i v_i$

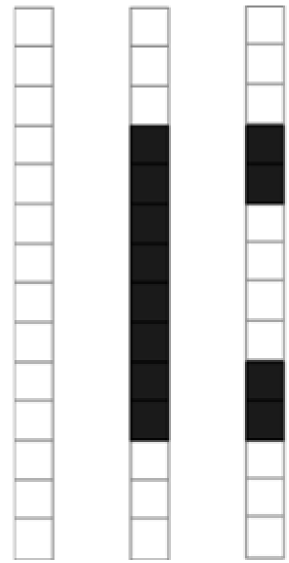
因此，对于同一个n×n非Hermitian矩阵A，其（左和右）奇异向量与（左和右）特征向量之间没有内在的关系.

矩阵 $A \in C^{m\times n}$ 的左奇异向量 u_i 是m×m Hermitian矩阵AA^H 的特征向量. 类似地，矩阵$A \in C^{m\times n}$ 的右奇异向量v_i 是n×nHermitian矩阵 $A^H A$ 的特征向量.

基于奇异值分解的图像压缩



14×11图像



3个列基矢量

基于奇异值分解的图像压缩

$$A=\begin{bmatrix}1&1&1&1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1&1&1&1\\1&1&0&0&0&0&0&0&0&1&1\\1&1&0&0&0&0&0&0&0&1&1\\1&1&0&0&1&1&1&0&0&1&1\\1&1&0&0&1&1&1&0&0&1&1\\1&1&0&0&1&1&1&0&0&1&1\\1&1&0&0&1&1&1&0&0&1&1\\1&1&0&0&0&0&0&0&0&1&1\\1&1&0&0&0&0&0&0&0&1&1\\1&1&1&1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1&1&1&1\end{bmatrix}_{14\times11}$$

基于奇异值分解的图像压缩

上述原始图像需要用14×11=154个数值来表示与存储.

对A进行奇异值分解

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

$$\sigma_1 = 9.7065 \quad \sigma_2 = 3.3983 \quad \sigma_3 = 2.0579 \implies A = \sum_{i=1}^{3} \sigma_i u v^T$$

这样，该图像可以用 $\{(\sigma_i, u_i, v_i), i = 1,2,3\}$ 来完全刻画．此时，所需要的元素个数仅为(14+11+1)×3=78.

推而广之，一幅m×n 图像的像素矩阵A具有奇异值分解$A = U\Sigma V^T$，其中奇异值按从大到小顺序排列．从中选择k个大奇异值及与之对应的左、右奇异向量，便可以共所使用k(m+n+1)个数值近似代替原来的m×n个图像数据.

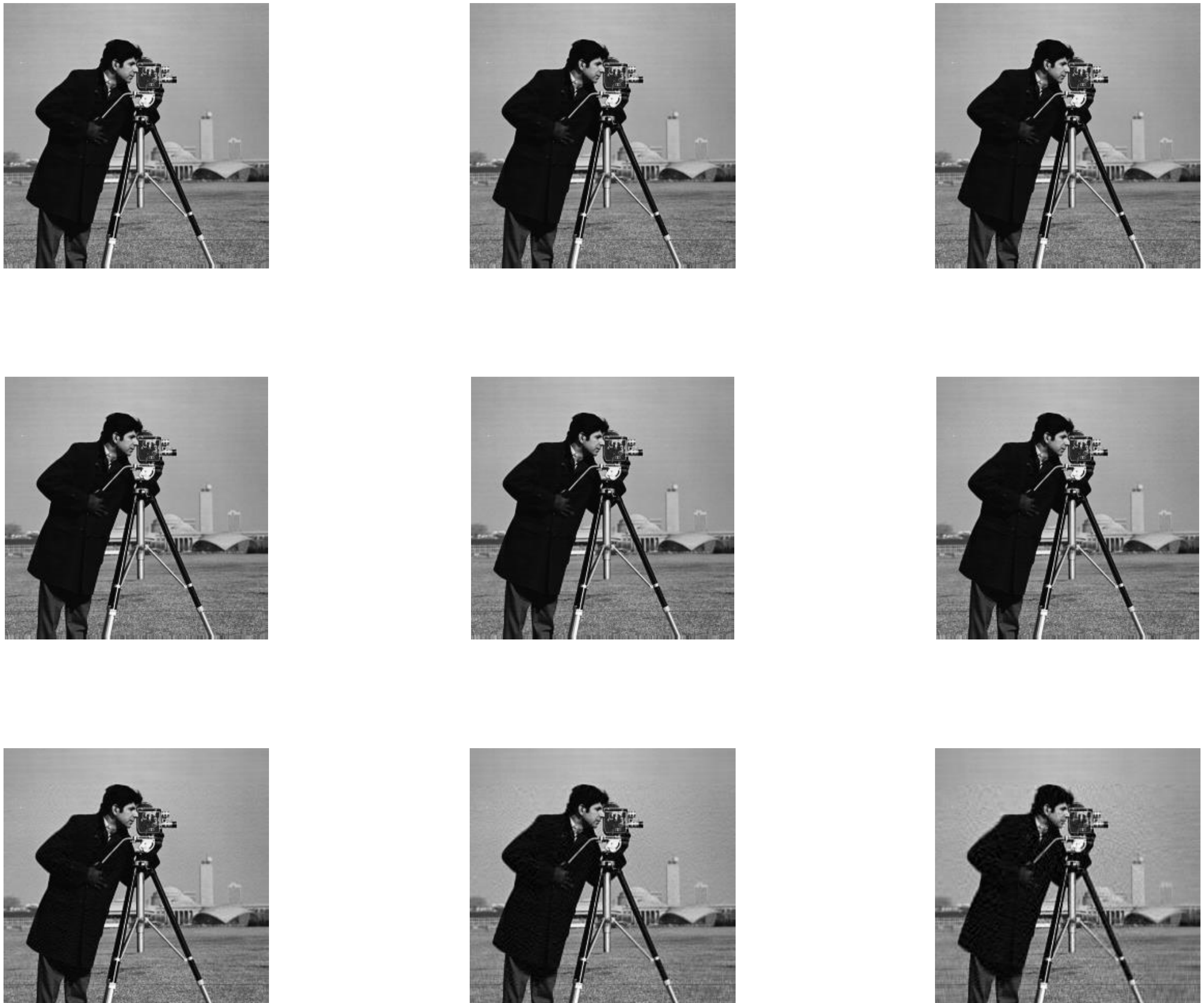
压缩比： $\rho = \dfrac{m \times n}{k(m+n+1)}$

基于奇异值分解的图像压缩

```
clear all
close all
I = imread('cameraman.tif');

figure(1)
subplot(3,3,1)
imshow(I,[0,255]);

[U,D,V] = svd(double(I));
cc = min(m,n);
for j = 9:-1:2
    dnum = floor(j*10*cc/100);%利用最大的i*10%的奇异值重建图像
    I_svd = U(:,1:dnum)*D(1:dnum,1:dnum)*V(:,1:dnum)';
    subplot(3,3,9-j+2)
    imshow(uint8(I_svd),[0,255])
end
```



基于奇异值分解的数字水印技术

为了版权保护和防止篡改，嵌入在数字载体(多媒体、文档、软件等)当中的标识信息；该标志信息的嵌入不影响原载体的使用，也不容易被探知和再次修改，但可以被生产方识别和辨认.

按特性划分:

鲁棒水印(Robust Watermarking):用于在数字作品中标识著作权信息

脆弱水印(Fragile Watermarking):用于数字作品的完整性保护和认证

按用途划分:

票证防伪水印、版权保护水印、篡改提示水印和隐蔽标识水印

此三幅图像来源网络，仅做教学观摩使用，版权属于原作者

数字水印技术模型主要有三部分组成：

（1）水印信息

（2）水印嵌入算法

（3）水印提取、检测和验证算法

嵌入过程：

$$\begin{cases} A \Rightarrow USV^T \\ L \Leftarrow S + aW \\ L \Rightarrow U_1 S_1 V_1^T \\ A_w \Leftarrow U S_1 V^T \end{cases}$$

A：原始图像
W：水印
A_W: 含有数字水印的图像
a: 水印强度系数
L：过渡矩阵

提取过程：

$$\begin{cases} P \Rightarrow U_p S_p V_p^T \\ F \Leftarrow U_1 S_p V_1^T \\ W_E \Leftarrow (F - S)/a \end{cases}$$

a, U1,V1,S: 保留参数

P：待检测图像
F：过渡矩阵
W_E：水印提取数据

检测和验证过程：

基于保留参数库（a, U1,V1,S,W)判断W_E是否水印、是什么水印

```
AA=imread('A.jpg');
WW=imread('test.jpg');
A(:,:)=AA(:,:,1);
W=rgb2gray(WW);SW=size(W);
figure(1),imagesc(AA);
[U,S,V]=svd(double(A));
SS=size(S);
WWW=zeros(SS(1),SS(2));
WWW(1:SW(1),1:SW(2))=W(:,:);
figure(2),imagesc(WWW);
a=0.1;
%嵌入数字水印
L=a*double(WWW)+S;
[U1,S1,V1]=svd(double(L));
AW=U*S1*V';
AA(:,:,1)=AW(:,:);
figure(3),imagesc(AA);
%提取数字水印
[Up,Sp,Vp]=svd(double(AW));
F=U1*Sp*V1';
WE=(F-S)/a;
figure(4),imagesc(WE);
```

test6.m

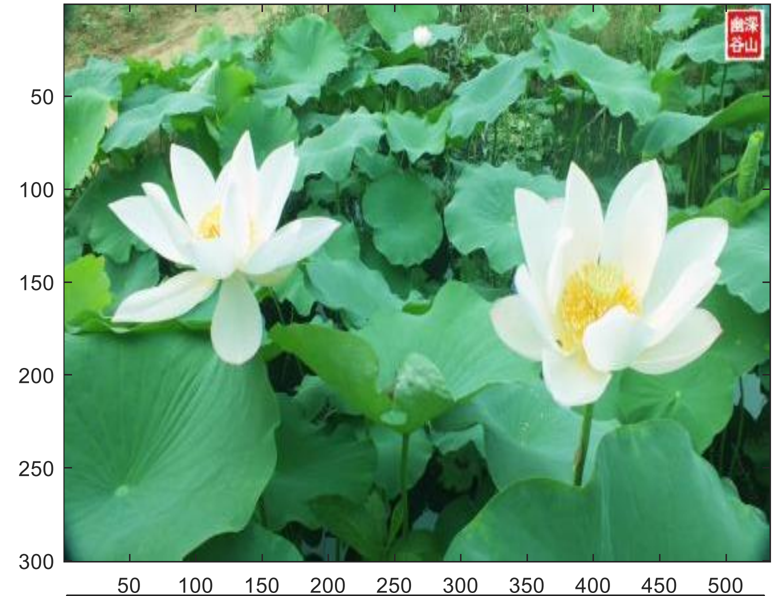


图5.原图

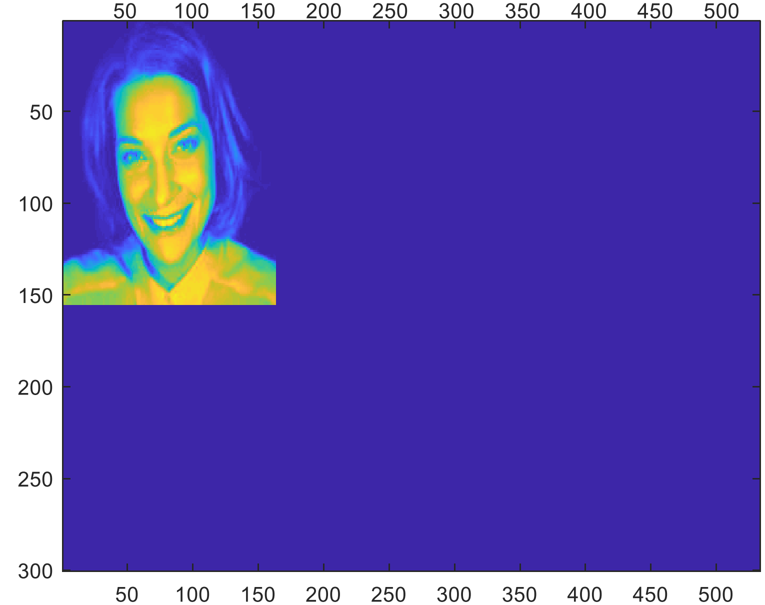


图6.水印图像及其摆放位置

图7. 含有水印的图像，a=0.1

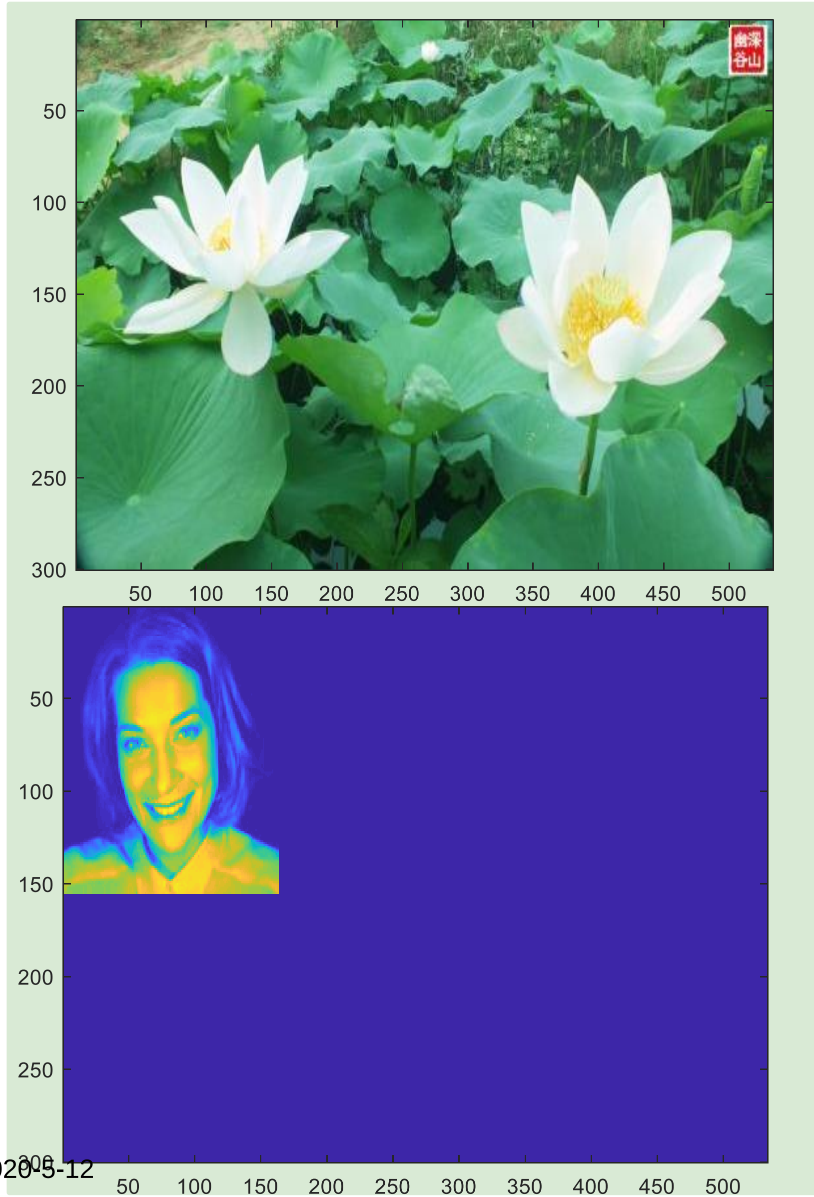
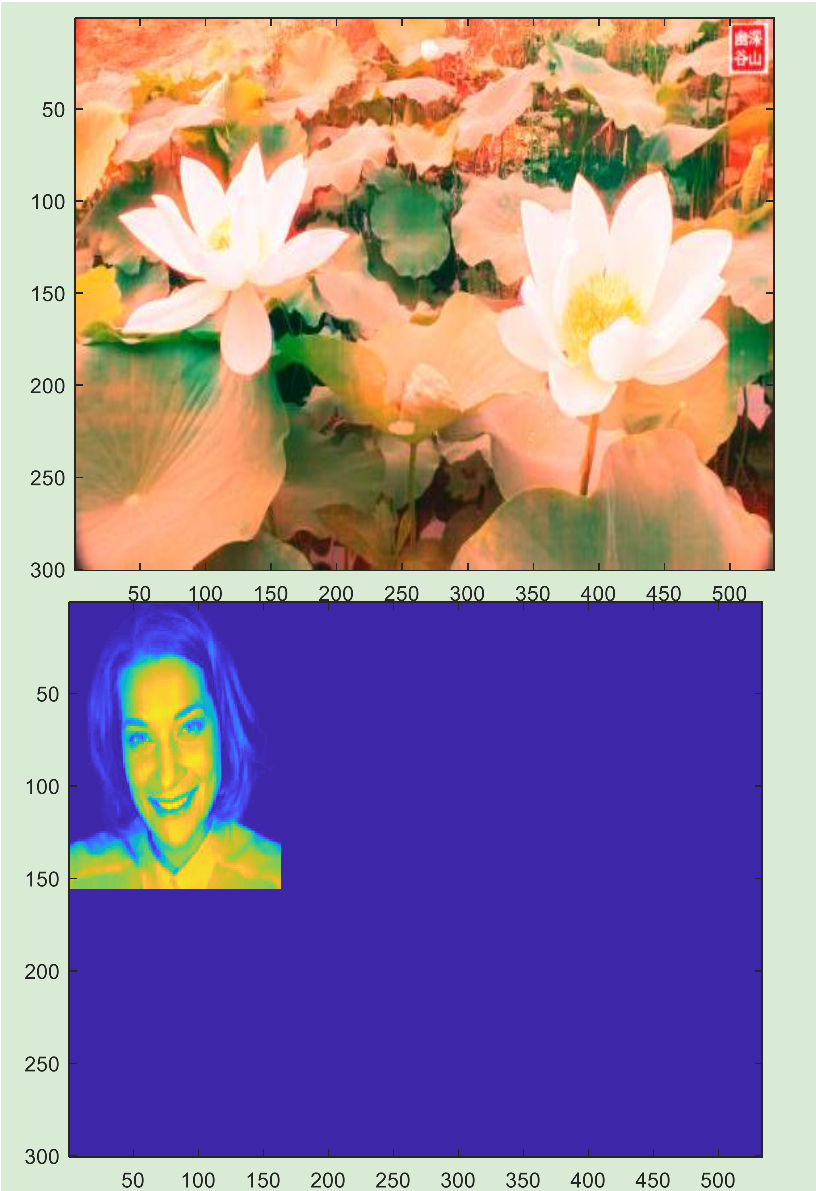


图8 含有水印的图像，a=10.1



现象：水印强度非常小的时候，也能检测出来.

数据的主成分分析

pca

princomp(Matlab2006a Before)

Principal component analysis of raw data

Syntax

```
coeff = pca(X)
coeff = pca(X,Name,Value)
[coeff,score,latent] = pca(___)
[coeff,score,latent,tsquared] = pca(___)
[coeff,score,latent,tsquared,explained,mu] = pca(___)
```

Matlab2012

```
C = cov(X);
[EV,~,~] = svd(C);
eigenvecs = EV(:,1:d);
```

K-SVD

K-SVD是一种经典的字典训练算法，依据误差最小原则，对误差项进行SVD分解，选择使误差最小的分解项作为更新的字典原子和对应的原子系数，经过不断的迭代从而得到优化的解。

K-SVD—稀疏表示

- 用较少的基本信号的线性组合来表达大部分或者全部的原始信号。每个矩阵的列向量可看成一个信号，一个矩阵则是信号的集合。其中，基本信号可成为原子信号。

设矩阵Y为样本集，由N个样本组成，每个样本由n个特征表示，即Y的尺寸为(n ＊ N)。稀疏表示，就是找到一组向量基(一组原子信号)，将此组向量基进行线性组合类表示矩阵(样本集)： $Y = DX$

这里的D是 $n*K$ 的矩阵，称为字典，D是由K个n维原子组成。且n<<P, 此字典称为超完备字典（字典的原子数量大于特征维度）。X 由N个k维系数向量组成，N对应样本个数，k对应字典个数。

K-SVD

主要目标

寻找最佳的字典D，同时使X系数矩阵达到稀疏最大。系数矩阵中，0元素越多，越稀疏，即目标是用更少的原子线性组合来逼近原始矩阵。这里系数矩阵指的是待优化的稀疏矩阵。

即 $$\min\|X\|_0$$

$$s.t. Y = DX$$

K-SVD

- **K-SVD算法算法原理、步骤**

K-SVD是由一系列原子来线性组合逼近，因此相比K-means更适用于压缩，编码等应用。

K-SVD算法也分两步：字典的优化及系数矩阵的优化。优化系统矩阵时，字典固定；优化字典时，系数矩阵同时跟着优化。

K-SVD

初始化字典是随机抽取样本作为字典，且本身字典的行数小于列数，因此存在一些重复，所以整体误差 $E = Y - DX$ 仍需要进一步优化。

K-SVD采取的字典更新算法是，对原子向量进行逐个更新。

算法原理：

- （1）剥离字典中第k个原子向量对D_K的贡献，如果不计入第k个原子向量的贡献，则误差矩阵为：$E_k = Y - \sum_{i \neq k} d_i x_i$

然后优化 d_k, x_k，使得E_k 最小。即使 $d_k x_k \approx E_k$.

- 注意此处 d_k, x_k 是一个列向量乘以一个行向量，x_k 是系数矩阵的行，而非列向量。

K-SVD

- （2）去除E_k中对应 x_k 为0项对应的部分。

优点：很大部分的减少E_k的大小，简化计算；使优化后系数X中的为零项仍旧为0，保证稀疏性。

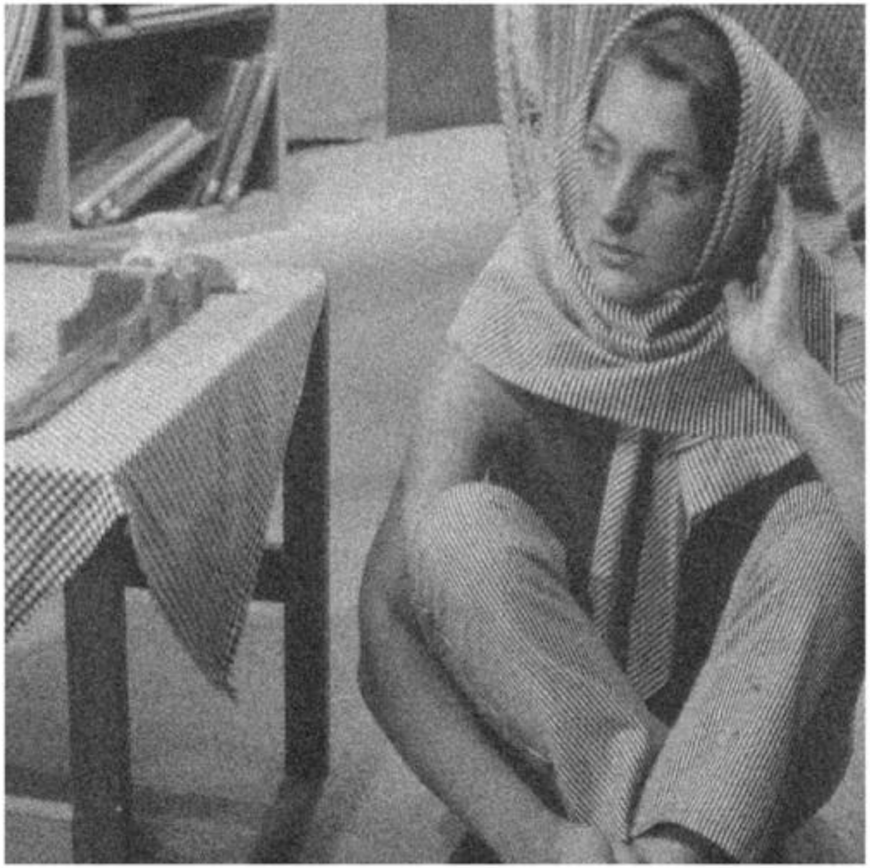
（3）对E_k进行SVD分解，即 $E_k = U\Delta V^T$ ，令U 的第一列为d_k的优化值，V 的第一列(V^T的第一行)乘以第一个奇异值即Δ[1, 1] 作为x_k的优化值。这里Δ[1, 1] 为一个常数，而非向量。

- （4）在所有原子向量都优化过后，判断是否达到停止条件，满足则退出优化，否则继续迭代系数矩阵及字典优化。

K-SVD

Original Image

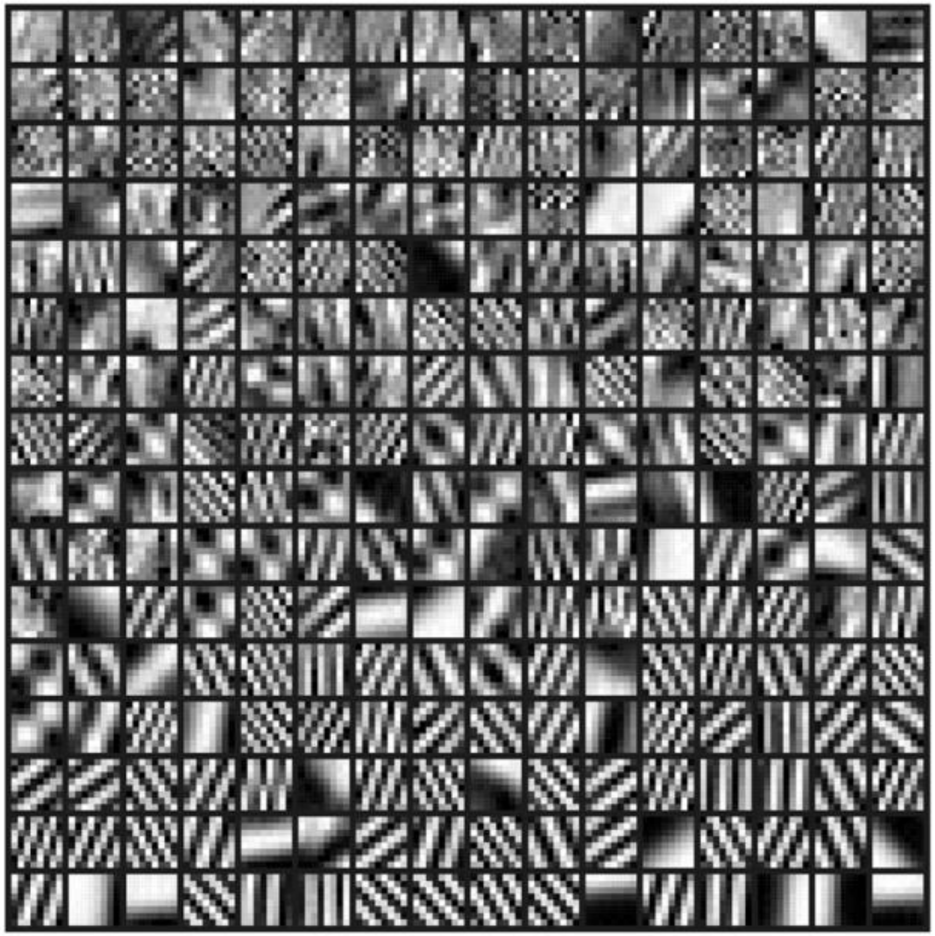
Noisy Image (22.1307 dB, σ=20)



K-SVD

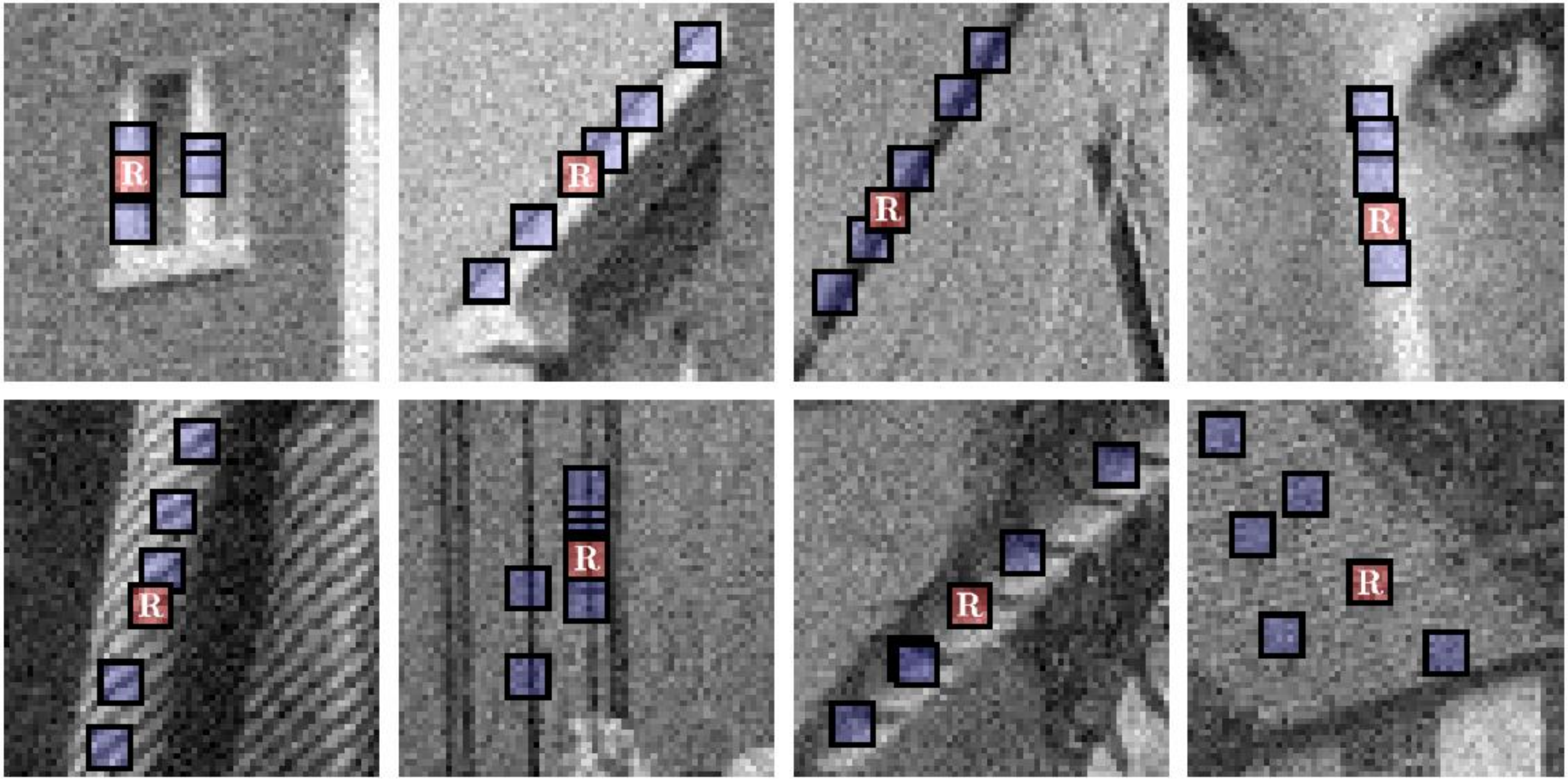
Denoised Image Using
Adaptive Dictionary (30.8295 dB)

Created Adaptive Dictionary

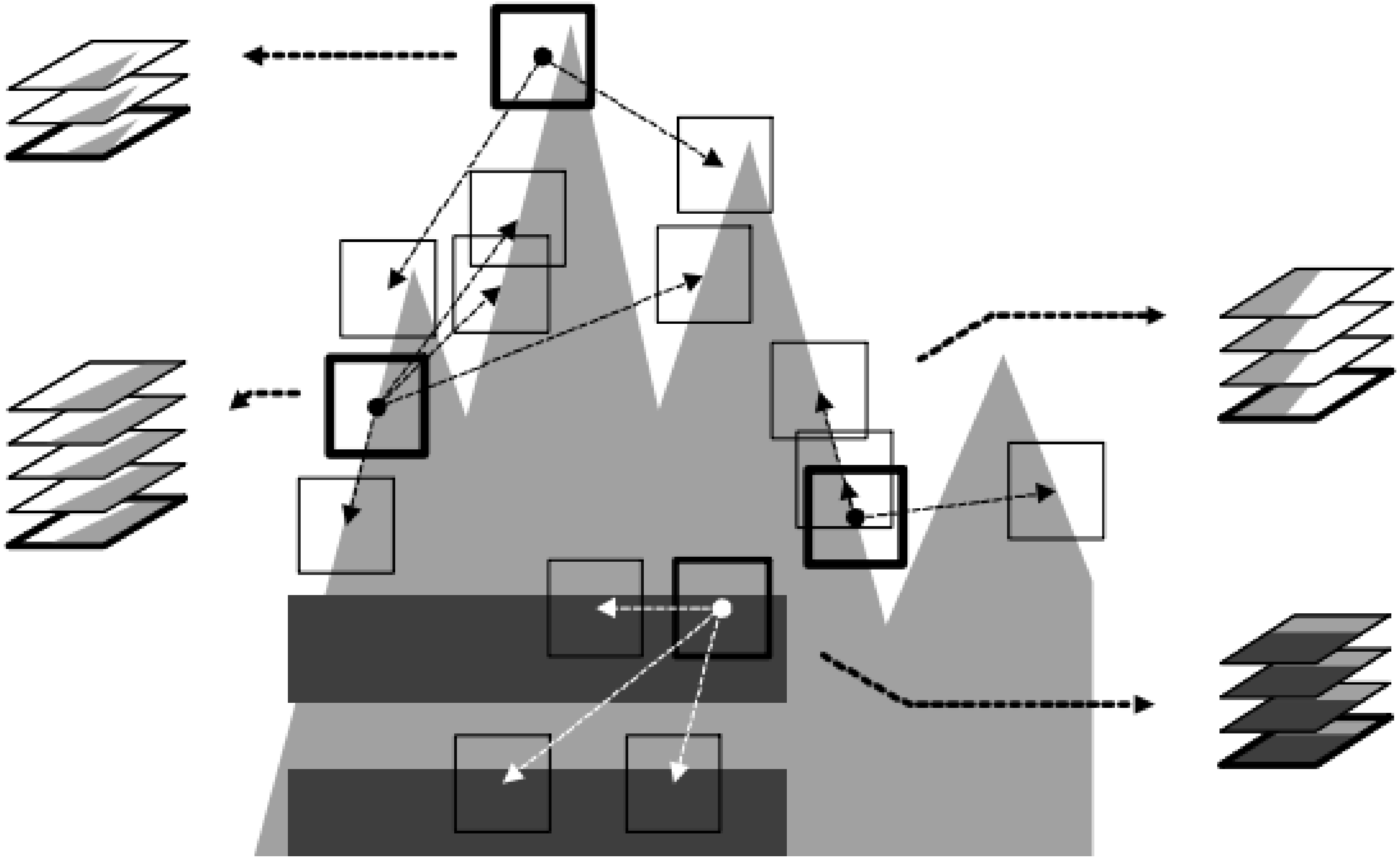


上海大学 SHANGHAI UNIVERSITY

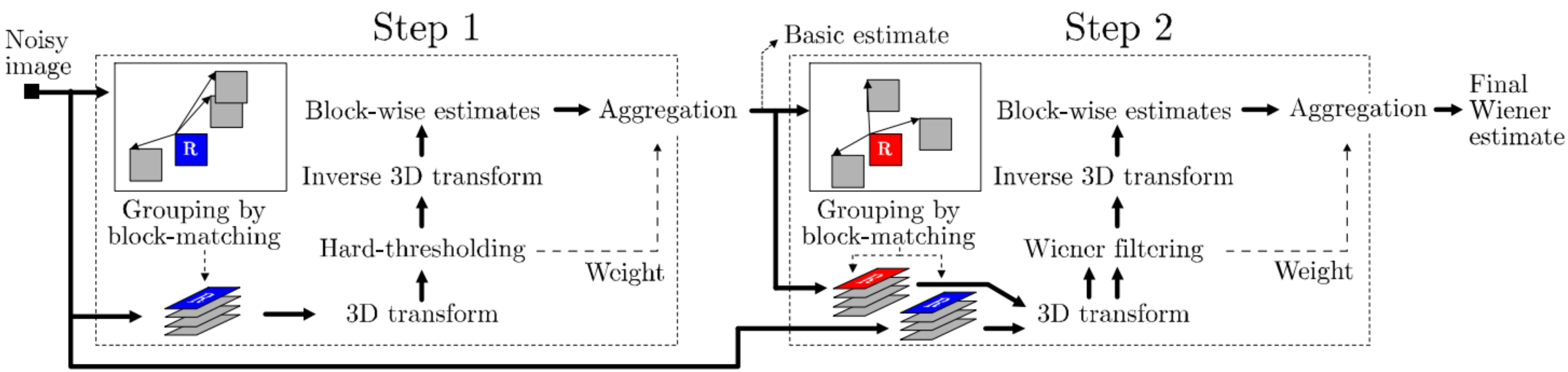
BM3D



BM3D



BM3D

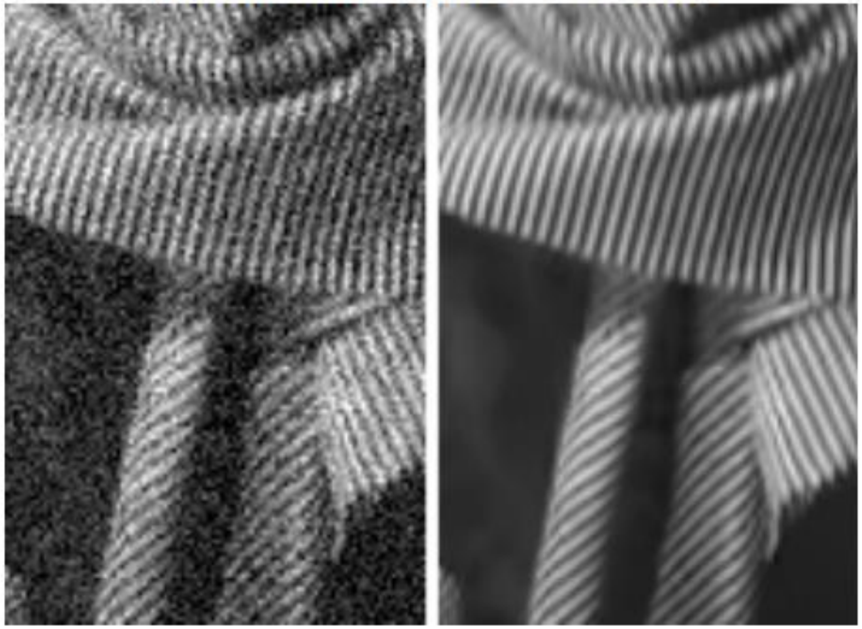


Noisy image
Step 1
Block-wise estimates
Aggregation
Inverse 3D transform
Hard-thresholding
3D transform
Grouping by block-matching
Weight
R
Basic estimate
Step 2
Block-wise estimates
Aggregation
Final Wiener estimate
Inverse 3D transform
Wiener filtering
3D transform
Grouping by block-matching
Weight

BM3D

(a) *Lena* (PSNR 32.08 dB)

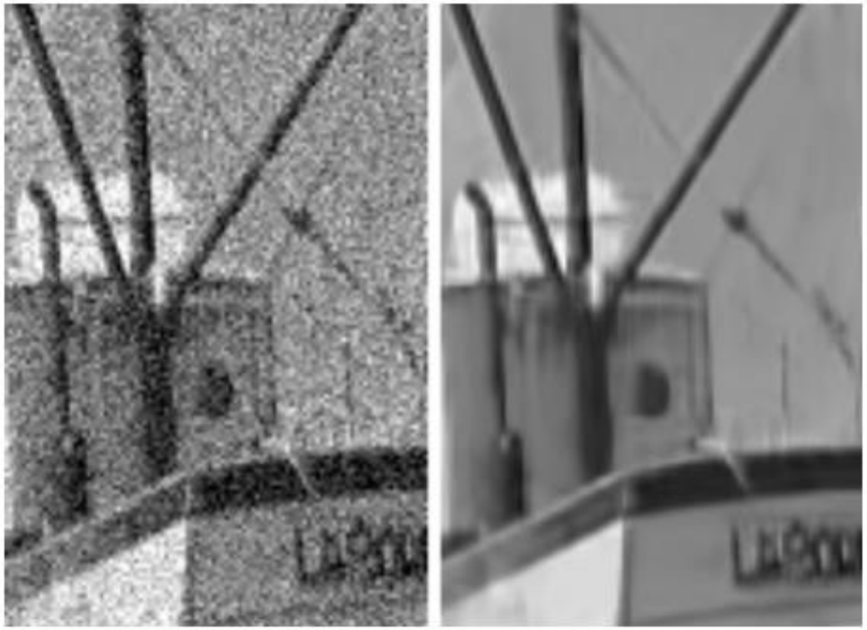
(b) *Barbara* (PSNR 30.73 dB)



BM3D

(d) *Man* (PSNR 29.62 dB)

(e) *Boats* (PSNR 29.91 dB)



Low rank

That is, the solution of

$$\hat{X} = \arg\min_{X} \|\boldsymbol{Y} - \boldsymbol{X}\|_F^2 + \lambda \|\boldsymbol{X}\|_*, \tag{1}$$

where λ is a positive constant, can be obtained by

$$\hat{X} = \boldsymbol{U}\mathcal{S}_\lambda(\boldsymbol{\Sigma})\boldsymbol{V}^T, \tag{2}$$

where $\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ is the SVD of $\boldsymbol{Y}$ and $\mathcal{S}_\lambda(\boldsymbol{\Sigma})$ is the soft-thresholding function on diagonal matrix $\boldsymbol{\Sigma}$ with parameter λ. For each diagonal element $\boldsymbol{\Sigma}_{ii}$ in $\boldsymbol{\Sigma}$, there is

$$\mathcal{S}_\lambda(\boldsymbol{\Sigma})_{ii} = \max(\boldsymbol{\Sigma}_{ii} - \lambda, 0). \tag{3}$$

Low rank

$$\|\boldsymbol{X}\|_{\boldsymbol{w},*} = \sum_i |w_i \sigma_i(\boldsymbol{X})|_1, \qquad (4)$$

where $\boldsymbol{w} = [w_1, \ldots, w_n]$ and $w_i \geq 0$ is a non-negative weight assigned to $\sigma_i(\boldsymbol{X})$.

$$\min_{\boldsymbol{X}} \|\boldsymbol{Y} - \boldsymbol{X}\|_F^2 + \|\boldsymbol{X}\|_{\boldsymbol{w},*}.$$

The global optimal solution under the order constraints $0 \leq w_1 \leq w_2 \leq \cdots \leq w_n$ is given by

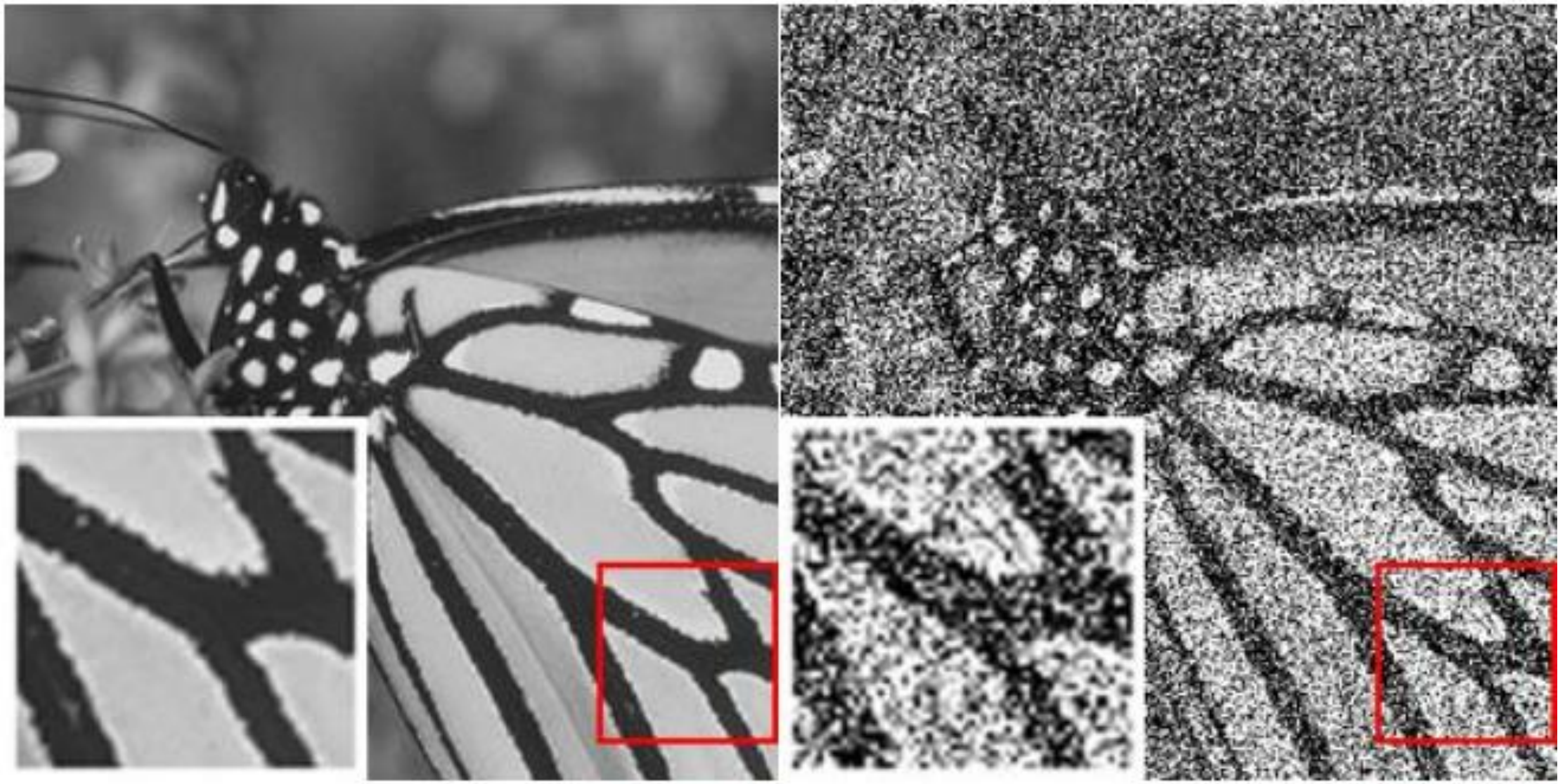
$$X^* = UDV^{\mathrm{T}},$$

where $Y = U\Sigma V^{\mathrm{T}}$

$$\Sigma = \begin{pmatrix} \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n) \\ 0 \end{pmatrix}, D = \begin{pmatrix} \mathrm{diag}(d_1, d_2, \ldots, d_n) \\ 0 \end{pmatrix},$$

and $d_i = \max(\sigma_i - w_i, 0)$, $i = 1, \ldots, n$. Further, if all the nonzero singular values of Y are distinct, then X^* is the unique optimal solution.

Low rank



(a) Ground truth (b) Noisy image (PSNR: 8.10dB)

Low rank


(c) BM3D (PSNR: 22.52dB)


(h) WNNM (PSNR: 22.91dB)

作 业

矩阵论及其工程应用，p.129

① 4.2

② 4.3

③ 4.7

④ 4.12

⑤ 4.13

作 业

4.2 分别计算矩阵

$$A=\begin{bmatrix} 1 & -1 \\ 3 & -3 \\ -3 & 3 \end{bmatrix} \quad 和 \quad A=\begin{bmatrix} 3 & 4 & 5 \\ 2 & 1 & 7 \end{bmatrix}$$

的奇异值分解。

4.3 已知矩阵

$$A=\begin{bmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & 9 & -25 \end{bmatrix}$$

求 $\boldsymbol{A}$ 的奇异值以及与最小奇异值 σ_1 相对应的左、右奇异向量。

4.7 证明：若 $\boldsymbol{A}$ 为 $n\times n$ 正定矩阵，则 $\boldsymbol{A}$ 的奇异值与 $\boldsymbol{A}$ 的特征值相同。

作 业

4.12 用矩阵 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ $(m \geqslant n)$ 的奇异向量表示 $\begin{bmatrix} \boldsymbol{O} & \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{A} & \boldsymbol{O} \end{bmatrix}$ 的特征向量。

4.13 利用 MATLAB 函数 $[\mathrm{U},\mathrm{S},\mathrm{V}] = \mathrm{svd}(X)$ 求解方程 $\boldsymbol{Ax} = \boldsymbol{b}$，其中

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & 3 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 1 \\ 4 \\ 3 \\ 2 \end{bmatrix}$$

上海大学
SHANGHAI UNIVERSITY

谢　谢！