Topic 4 - Positive Definite Matrices

AMA3724 Further Mathematical Methods(2024/25 Semester 1)
Lecturer: Jianbo Cui

- Symmetric matrices
- The Spectral Theorem
- Positive definite matrices
- Singular Value Decomposition

THE HONG KONG POLYTECHNIC UNIVERSITY 香港理工大學

Symmetric matrices

Symmetric matrices

A square matrix A is said to be **symmetric** if

$$A^T = A.$$

Example The following are symmetric matrices:

[1]:
```python
import sympy as sp; from sympy import I;
A = sp.Matrix([[1,-3],[-3,1]]);
B = sp.Matrix([[6,-2,-1],[-2,6,-1],[-1,-1,5]]);
C = sp.Matrix([[3,-2,1+I],[-2,6,2],[1+I,2,3]]);
D = sp.Matrix([[3,2,2,2],[2,3,2,2],[2,2,3,2],[2,2,2,3]]);
```

[2]:
```python
A
```

[2]: $\begin{bmatrix} 1 & -3 \\ -3 & 1 \end{bmatrix}$

[3]:
```python
B
```

[3]: $\begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$

[4]:
```python
C
```

[4]: $\begin{bmatrix} 3 & -2 & 1+i \\ -2 & 6 & 2 \\ 1+i & 2 & 3 \end{bmatrix}$

[5]:
```python
D
```

[5]: $\begin{bmatrix} 3 & 2 & 2 & 2 \\ 2 & 3 & 2 & 2 \\ 2 & 2 & 3 & 2 \\ 2 & 2 & 2 & 3 \end{bmatrix}$

[6]:
```python
A.eigenvals()
```

[6]: `{4: 1, -2: 1}`

[7]:
```python
B.eigenvals()
```

[7]: `{8: 1, 6: 1, 3: 1}`

[8]:
```python
C.eigenvals()
```

[8]:
```
{4 + I: 1, 4 - I/2 - sqrt(47 +
8*I)/2: 1, 4 - I/2 + sqrt(47
+ 8*I)/2: 1}
```

[9]:
```python
D.eigenvals()
```

[9]: `{9: 1, 1: 3}`

Symmetric matrices

Symmetric matrix with real entries

Let A be an $n \times n$ symmetric matrix with real entries only. Then all eigenvalues of A are real numbers, and the corresponding eigenvectors can be chosen to be real vectors.

Proof. Suppose $\lambda = \alpha + i\beta$ is a complex eigenvalue of A and $\mathbf{x} = \mathbf{y} + i\mathbf{z}$ is a corresponding eigenvector, where $\alpha, \beta \in \mathbb{R}$ and $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$. Then

$$A\mathbf{x} = \lambda\mathbf{x} \quad \Longrightarrow \quad A(\mathbf{y} + i\mathbf{z}) = (\alpha + i\beta)(\mathbf{y} + i\mathbf{z}).$$

Notice that $\mathbf{y}^T A\mathbf{z} = (\mathbf{y}^T A\mathbf{z})^T = \mathbf{z}^T A^T \mathbf{y} = \mathbf{z}^T A\mathbf{y}$. Therefore,

$$\begin{aligned}
(\overline{\mathbf{y} + i\mathbf{z}})^T A(\mathbf{y} + i\mathbf{z}) &= (\alpha + i\beta)(\overline{\mathbf{y} + i\mathbf{z}})^T(\mathbf{y} + i\mathbf{z}) \\
(\mathbf{y} - i\mathbf{z})^T A(\mathbf{y} + i\mathbf{z}) &= (\alpha + i\beta)(\mathbf{y} - i\mathbf{z})^T(\mathbf{y} + i\mathbf{z}) \\
\mathbf{y}^T A\mathbf{y} + \mathbf{z}^T A\mathbf{z} + i\mathbf{y}^T A\mathbf{z} - i\mathbf{z}^T A\mathbf{y} &= (\alpha + i\beta)(\mathbf{y}^T\mathbf{y} + \mathbf{z}^T\mathbf{z} + i\mathbf{y}^T\mathbf{z} - i\mathbf{z}^T\mathbf{y}) \\
(\mathbf{y}^T A\mathbf{y} + \mathbf{z}^T A\mathbf{z}) + i(\mathbf{y}^T A\mathbf{z} - i\mathbf{z}^T A\mathbf{y}) &= (\alpha + i\beta)(\mathbf{y}^T\mathbf{y} + \mathbf{z}^T\mathbf{z}) \\
\mathbf{y}^T A\mathbf{y} + \mathbf{z}^T A\mathbf{z} &= \alpha(\mathbf{y}^T\mathbf{y} + \mathbf{z}^T\mathbf{z}) + i\beta(\mathbf{y}^T\mathbf{y} + \mathbf{z}^T\mathbf{z}).
\end{aligned}$$

Then $\beta = 0$ and so $\lambda = \alpha$ is real. Furthermore, $A(\mathbf{y} + i\mathbf{z}) = \alpha(\mathbf{y} + i\mathbf{z})$ implies $A\mathbf{y} = \alpha\mathbf{y}$ and $A\mathbf{z} = \alpha\mathbf{z}$, so one can choose real vectors to be the corresponding eigenvectors.

Symmetric matrices

Orthogonal set of eigenvectors

Let A be an $n \times n$ symmetric matrix with real entries only. If $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are real eigenvectors that correspond to real and distinct eigenvalues $\lambda_1, \ldots, \lambda_k$, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ forms an orthogonal set.

Proof. Notice that for any $1 \le i < j \le k$,

$$\mathbf{v}_i^T A\mathbf{v}_j = \mathbf{v}_i^T(\lambda_j\mathbf{v}_j) = \lambda_j\mathbf{v}_i^T\mathbf{v}_j$$

and

$$\mathbf{v}_j^T A\mathbf{v}_i = \mathbf{v}_j^T(\lambda_i\mathbf{v}_i) = \lambda_i\mathbf{v}_j^T\mathbf{v}_i.$$

Since A is symmetric,

$$\mathbf{v}_i^T A\mathbf{v}_j = \left(\mathbf{v}_i^T A\mathbf{v}_j\right)^T = \mathbf{v}_j^T A^T\mathbf{v}_i = \mathbf{v}_j^T A\mathbf{v}_i.$$

Then we have

$$\lambda_j\mathbf{v}_i^T\mathbf{v}_j = \lambda_i\mathbf{v}_j^T\mathbf{v}_i \quad \Longrightarrow \quad \mathbf{v}_j^T\mathbf{v}_i = 0 \quad \Longrightarrow \quad \langle\mathbf{v}_j, \mathbf{v}_i\rangle = 0.$$

Thus, $\mathbf{v}_i$ and $\mathbf{v}_j$ are orthogonal. Since this works for all $1 \le i < j \le k$, the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ forms an orthogonal set.

Symmetric matrices

[The 2×2 Case] Let A be a 2×2 real matrix with real eigenvalues. Suppose λ is a real eigenvalue with a real eigenvector $\mathbf{x}$ with $\|\mathbf{x}\| = 1$. Then $A\mathbf{x} = \lambda\mathbf{x}$. Let $\mathbf{y}$ be a unit vector in $\mathrm{Span}\{\mathbf{x}\}^{\perp}$, that is $\|\mathbf{y}\| = 1$ and $\langle \mathbf{y}, \mathbf{x} \rangle = 0$. Then $\{\mathbf{x}, \mathbf{y}\}$ forms an orthornormal basis for $\mathbb{R}^2$. Therefore, one may write

$$A\mathbf{y} = c\mathbf{x} + d\mathbf{y} \quad \text{for some } c, d \in \mathbb{R}.$$

Define $U = \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix}$ and $T = \begin{bmatrix} \lambda & c \\ 0 & d \end{bmatrix}$. Then U is an orthogonal matrix and T is a upper triangular matrix.

$$AU = A\begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix} = \begin{bmatrix} A\mathbf{x} & A\mathbf{y} \end{bmatrix} = \begin{bmatrix} \lambda\mathbf{x} & c\mathbf{x} + d\mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix} \begin{bmatrix} \lambda & c \\ 0 & d \end{bmatrix} = UT.$$

Thus, $AU = UT$ implies

$$A = UTU^T = U \begin{bmatrix} \lambda & c \\ 0 & d \end{bmatrix} U^T.$$

The Schur decomposition

Let A be an $n \times n$ real matrix with real eigenvalues. Then there exist an orthogonal matrix U and an upper triangular matrix T such that

$$A = UTU^T.$$

Proof. By mathematical induction.

The Spectral Theorem

The Spectral Theorem

Let A be an $n \times n$ real symmetric matrix with real eigenvalues $\lambda_1, \ldots, \lambda_n$.

1. There are an orthogonal matrix U and a diagonal matrix D such that
$$A = UDU^T.$$
2. A has n eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$, which form an orthonormal basis.
3. $A = \sum_{j=1}^{n} \lambda_j \mathbf{u}_j \mathbf{u}_j^T = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T.$

Proof. (1) Since all eigenvalues of A are real, by the Schur decomposition, there exist an orthogonal matrix U and an upper triangular matrix T such that

$$A = UTU^T.$$

Since A is symmetric, i.e., $A = A^T$, then

$$UTU^T = A = A^T = (UTU^T)^T = UT^TU^T \quad \Longrightarrow \quad T = T^T.$$

Therefore, T is both upper triangular and lower triangular, hence, T is a diagonal matrix.

(2) By the Diagonalization Theorem, the columns of U are n linearly independent eigenvectors of A. Since U is an orthogonal matrix, these n linearly independent eigenvectors indeed form an orthonormal basis.

(3) Let $D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$ and $U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix}$. Then

$$\begin{aligned} A = UDU^T &= \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\mathbf{u}_1 & \cdots & \lambda_n\mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \\ &= \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T = \sum_{j=1}^{n} \lambda_j \mathbf{u}_j\mathbf{u}_j^T. \end{aligned}$$

Example 4.1 Let $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$.

Standard calculations found that A has

- eigenvalue $\lambda_1 = 3$ with eigenspace
  $\text{Span}\{\mathbf{v}_1\} = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$.
- eigenvalue $\lambda_2 = 6$ with eigenspace
  $\text{Span}\{\mathbf{v}_2\} = \text{Span}\left\{\begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix}\right\}$,
- eigenvalue $\lambda_3 = 8$ with eigenspace
  $\text{Span}\{\mathbf{v}_3\} = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\}$,

Notice that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ forms an orthogonal basis.

[1]:
```
import sympy as sp;
import numpy as np;
```

[2]:
```
A = sp.Matrix([[6,-2,-1],
[-2,6,-1], [-1,-1,5]]);A
```

[2]: $\begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$

[3]:
```
A.eigenvects()
```

[3]:
```
[(3,
  1,
  [Matrix([
   [1],
   [1],
   [1]])]),
 (6,
  1,
  [Matrix([
   [-1/2],
   [-1/2],
   [   1]])]),
 (8,
  1,
  [Matrix([
   [-1],
   [ 1],
   [ 0]])])]
```

Define

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \ \mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \text{ and } \mathbf{u}_3 = \frac{1}{\|\mathbf{v}_3\|}\mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

Then $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ forms an orthonormal basis in $\mathbb{R}^3$. Let

$$U = \begin{bmatrix} 1/\sqrt{3} & -1\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 2/\sqrt{6} & 0 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 8 \end{bmatrix}.$$

Then U is orthogonal, i.e., $U^T U = I_n$ and

$$A = UDU^{-1} = UDU^T.$$

[4]:
```
A = np.array([[6,-2,-1], [-2,6,-1], [-1,-1,5]]);A
```

[4]:
```
array([[ 6, -2, -1],
       [-2,  6, -1],
       [-1, -1,  5]])
```

[5]:
```
np.linalg.eig(A)
```

[5]:
```
(array([3., 8., 6.]),
 array([[ 5.77350269e-01,  7.07106781e-01, -4.08248290e-01],
        [ 5.77350269e-01, -7.07106781e-01, -4.08248290e-01],
        [ 5.77350269e-01,  9.06835124e-17,  8.16496581e-01]]))
```


Example 4.2 Let $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

Standard calculations found that A has

- eigenvalue $\lambda_1 = -2$ with eigenspace
$$\text{Span}\{\mathbf{v}_1\} = \text{Span}\left\{\begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}\right\},$$
- eigenvalue $\lambda_2 = 7$ with eigenspace
$$\text{Span}\{\mathbf{v}_2, \mathbf{v}_3\} = \text{Span}\left\{\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\}.$$

Notice that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ does NOT form an orthogonal basis.

[6]:
```
A = sp.Matrix([[3,-2,4], [-2,6,2], [4,2,3]]);A
```

[6]: $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

[7]:
```
A.eigenvects()
```

[7]:
```
[(-2,
  1,
  [Matrix([
   [  -1],
   [-1/2],
   [   1]])]),
 (7,
  2,
  [Matrix([
   [-1/2],
   [   1],
   [   0]]),
   Matrix([
   [1],
   [0],
   [1]])])]
```

[8]:
```
P,D = A.diagonalize();P,D
```

[8]:
```
(Matrix([
 [-2, -1, 1],
 [-1,  2, 0],
 [ 2,  0, 1]]),
 Matrix([
 [-2, 0, 0],
 [ 0, 7, 0],
 [ 0, 0, 7]]))
```

The Spectral Theorem

Applying the Gram-Schmidt algorithm on $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ yields

$$\mathbf{u}_1 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -\sqrt{5}/5 \\ 2\sqrt{5}/5 \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 4\sqrt{5}/15 \\ 2\sqrt{5}/15 \\ \sqrt{5}/3 \end{bmatrix}.$$

Here,

- $\mathrm{Span}\{\mathbf{u}_1\} = \mathrm{Span}\{\mathbf{v}_1\} = E_{-2}(A)$,
- $\mathrm{Span}\{\mathbf{u}_2, \mathbf{u}_3\} = \mathrm{Span}\{\mathbf{v}_2, \mathbf{v}_3\} = E_7(A)$,
- $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ forms an orthonormal basis for $\mathbb{R}^3$.

Let

$$U = \begin{bmatrix} -2/3 & -\sqrt{5}/5 & 4\sqrt{5}/15 \\ -1/3 & 2\sqrt{5}/5 & 2\sqrt{5}/15 \\ 2/3 & 0 & \sqrt{5}/3 \end{bmatrix}$$

and

$$D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

Then $A = UDU^T$.

[9]:
```
P
```

[9]:

$$\begin{bmatrix} -2 & -1 & 1 \\ -1 & 2 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

[10]:
```
P.QRdecomposition()
```

[10]:
```
(Matrix([
 [-2/3,  -sqrt(5)/5, 4*sqrt(5)/15],
 [-1/3, 2*sqrt(5)/5, 2*sqrt(5)/15],
 [ 2/3,           0,    sqrt(5)/3]]),
 Matrix([
 [3,       0,            0],
 [0, sqrt(5),   -sqrt(5)/5],
 [0,       0, 3*sqrt(5)/5]]))
```


The Spectral Theorem

Example 4.2 [Re-visited]

Let $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$. Standard calculations show that A is diagonalizable and $A = PDP^{-1}$ with

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

and

$$D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

Next, P has a QR decomposition with

$$Q = \begin{bmatrix} -2/3 & -\sqrt{5}/5 & 4\sqrt{5}/15 \\ -1/3 & 2\sqrt{5}/5 & 2\sqrt{5}/15 \\ 2/3 & 0 & \sqrt{5}/3 \end{bmatrix}.$$

Then $A = QDQ^T$.

[6]:
```
A = sp.Matrix([[3,-2,4], [-2,6,2],
↪[4,2,3]]);A
```

[6]:

$$\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

[8]:
```
P,D = A.diagonalize();P,D
```

[8]:
```
(Matrix([
 [-2, -1, 1],
 [-1,  2, 0],
 [ 2,  0, 1]]),
 Matrix([
 [-2, 0, 0],
 [ 0, 7, 0],
 [ 0, 0, 7]]))
```

[10]:
```
P.QRdecomposition()
```

[10]:
```
(Matrix([
 [-2/3,  -sqrt(5)/5, 4*sqrt(5)/15],
 [-1/3, 2*sqrt(5)/5, 2*sqrt(5)/15],
 [ 2/3,           0,    sqrt(5)/3]]),
 Matrix([
 [3,       0,            0],
 [0, sqrt(5),   -sqrt(5)/5],
 [0,       0, 3*sqrt(5)/5]]))
```

The Spectral Theorem

Example 4.3

Let $A = \begin{bmatrix} 3 & 2 & 2 & 2 \\ 2 & 3 & 2 & 2 \\ 2 & 2 & 3 & 2 \\ 2 & 2 & 2 & 3 \end{bmatrix}$. Standard calculations show that A is diagonalizable and $A = PDP^{-1}$ with

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

and

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}.$$

Next, P has a QR decomposition with

$$Q = \begin{bmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & \frac{1}{2} \\ 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{6} & \frac{1}{2} \\ 0 & 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

Then $A = QDQ^T$.

[11]:
```
A = sp.Matrix([[3,2,2,2], [2,3,2,2],␣
↪[2,2,3,2], [2,2,2,3]]);A
```

[11]: $\begin{bmatrix} 3 & 2 & 2 & 2 \\ 2 & 3 & 2 & 2 \\ 2 & 2 & 3 & 2 \\ 2 & 2 & 2 & 3 \end{bmatrix}$

[12]:
```
P,D = A.diagonalize(); P,D
```

[12]:
```
(Matrix([
 [-1, -1, -1, 1],
 [ 1,  0,  0, 1],
 [ 0,  1,  0, 1],
 [ 0,  0,  1, 1]]),
 Matrix([
 [1, 0, 0, 0],
 [0, 1, 0, 0],
 [0, 0, 1, 0],
 [0, 0, 0, 9]]))
```

[13]:
```
Q,R = P.QRdecomposition();Q
```

[13]: $\begin{bmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & \frac{1}{2} \\ 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{6} & \frac{1}{2} \\ 0 & 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$

The Spectral Theorem

Given an $n \times n$ symmetric matrix $A = [a_{ij}]$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$. The term

$$\mathbf{x}^T A \mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

is called the quadratic form.

Eigenvalues of real symmetric matrix

Let A be an $n \times n$ real symmetric matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ (in ascending order). Then

$$\lambda_1 \leq \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \lambda_n \quad \text{for any nonzero } \mathbf{x} \in \mathbb{R}^n.$$

Furthermore,

$$\lambda_1 = \min_{\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \quad \text{and} \quad \lambda_n = \max_{\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

[Courant-Fischer] Let S denote a subspace of $\mathbb{R}^n$. Then

$$\lambda_k = \min_{S:\dim S = k} \max_{\mathbf{x} \in S, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \max_{S:\dim S = n-k+1} \min_{\mathbf{x} \in S, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

The Spectral Theorem

Proof. Since A is real symmetric,

$$A = \sum_{j=1}^{n} \lambda_j \mathbf{u}_j \mathbf{u}_j^T,$$

where $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are eigenvectors of A and form an orthonormal basis for $\mathbb{R}^n$. Then for any $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = \sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle \mathbf{u}_j$. Therefore,

$$\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2.$$

Furthermore,

$$\begin{aligned} \mathbf{x}^T A \mathbf{x} &= \left(\sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle \mathbf{u}_j^T \right) \left(\sum_{k=1}^{n} \lambda_k \mathbf{u}_k \mathbf{u}_k^T \right) \left(\sum_{l=1}^{n} \langle \mathbf{x}, \mathbf{u}_l \rangle \mathbf{u}_l^T \right) \\ &= \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} \lambda_k \langle \mathbf{x}, \mathbf{u}_j \rangle \langle \mathbf{x}, \mathbf{u}_l \rangle (\mathbf{u}_j^T \mathbf{u}_k)(\mathbf{u}_k^T \mathbf{u}_l) = \sum_{j=1}^{n} \lambda_j \langle \mathbf{x}, \mathbf{u}_j \rangle^2. \end{aligned}$$

The Spectral Theorem

Now,

$$\frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\sum_{j=1}^{n} \lambda_j \langle \mathbf{x}, \mathbf{u}_j \rangle^2}{\sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2} = \frac{\lambda_1 \langle \mathbf{x}, \mathbf{u}_1 \rangle^2 + \cdots + \lambda_n \langle \mathbf{x}, \mathbf{u}_n \rangle^2}{\langle \mathbf{x}, \mathbf{u}_1 \rangle^2 + \cdots + \langle \mathbf{x}, \mathbf{u}_n \rangle^2}$$

Then

$$\frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\sum_{j=1}^{n} \lambda_j \langle \mathbf{x}, \mathbf{u}_j \rangle^2}{\sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2} \geq \frac{\sum_{j=1}^{n} \lambda_1 \langle \mathbf{x}, \mathbf{u}_j \rangle^2}{\sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2} = \frac{\lambda_1 \sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2}{\sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2} = \lambda_1.$$

Furthermore, when $\mathbf{x} = \mathbf{u}_1$,

$$\frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\lambda_1 \langle \mathbf{x}, \mathbf{u}_1 \rangle^2}{\langle \mathbf{x}, \mathbf{u}_1 \rangle^2} = \lambda_1.$$

Similarly,

$$\frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\sum_{j=1}^{n} \lambda_j \langle \mathbf{x}, \mathbf{u}_j \rangle^2}{\sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2} \leq \frac{\sum_{j=1}^{n} \lambda_n \langle \mathbf{x}, \mathbf{u}_j \rangle^2}{\sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2} = \frac{\lambda_n \sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2}{\sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_j \rangle^2} = \lambda_n.$$

Furthermore, when $\mathbf{x} = \mathbf{u}_n$,

$$\frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\lambda_n \langle \mathbf{x}, \mathbf{u}_1 \rangle^2}{\langle \mathbf{x}, \mathbf{u}_1 \rangle^2} = \lambda_n.$$

Thus, $\lambda_1 = \min_{\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$. and $\lambda_n = \max_{\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$.

The Spectral Theorem

[Courant-Fischer] Let $T = \text{Span}\{\mathbf{u}_k, \ldots, \mathbf{u}_n\}$. Then $\dim T = n - k + 1$ and for any $\mathbf{x} \in T$,

$$\langle \mathbf{x}, \mathbf{u}_j \rangle = 0 \quad \text{for } j = 1, \ldots, k-1.$$

Then

$$\min_{\mathbf{x}\in T, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \min_{\mathbf{x}\in T, \mathbf{x}\neq \mathbf{0}} \frac{\sum_{j=1}^n \lambda_j \langle \mathbf{x}, \mathbf{u}_j\rangle^2}{\sum_{j=1}^n \langle \mathbf{x}, \mathbf{u}_j\rangle^2} = \min_{\mathbf{x}\in T, \mathbf{x}\neq \mathbf{0}} \frac{\sum_{j=k}^n \lambda_j \langle \mathbf{x}, \mathbf{u}_j\rangle^2}{\sum_{j=k}^n \langle \mathbf{x}, \mathbf{u}_j\rangle^2} = \lambda_k.$$

For any k-dimensional subspace S of $\mathbb{R}^n$, since $\dim S = k$, $\dim T = n - k + 1$, and $\dim(S+T) \leq n$,

$$\dim S \cap T = \dim S + \dim T - \dim(S+T) \geq k + (n-k+1) - n = 1.$$

The subspace $S \cap T$ has at least dimension one, and so it is nonempty. Then

$$\max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \geq \max_{\mathbf{x}\in S\cap T, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \geq \min_{\mathbf{x}\in S\cap T, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \geq \min_{\mathbf{x}\in T, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_k.$$

Furthermore, when $S = \text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$,

$$\max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\sum_{j=1}^n \lambda_j \langle \mathbf{x}, \mathbf{u}_j\rangle^2}{\sum_{j=1}^n \langle \mathbf{x}, \mathbf{u}_j\rangle^2} = \max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\sum_{j=1}^k \lambda_j \langle \mathbf{x}, \mathbf{u}_j\rangle^2}{\sum_{j=1}^k \langle \mathbf{x}, \mathbf{u}_j\rangle^2} = \lambda_k.$$

Thus,

$$\min_{S:\dim S = k} \max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \geq \lambda_k.$$

The Spectral Theorem

[1]:
```
import numpy as np;
A = np.array([[6,3,2,5], [3,8,5,6], [2,5,4,5], [5,6,5,2]]);A
```

[1]:
```
array([[6, 3, 2, 5],
       [3, 8, 5, 6],
       [2, 5, 4, 5],
       [5, 6, 5, 2]])
```

[2]:
```
np.linalg.eig(A)
```

[2]:
```
(array([18.37878448,  4.2761543 , -3.37813216,  0.72319338]),
 array([[-0.42046042, -0.84845279,  0.31313495, -0.07271451],
        [-0.6197679 ,  0.43924433,  0.21547076, -0.61361593],
        [-0.44479769,  0.28301455,  0.34945395,  0.77455775],
        [-0.49117796, -0.08423208, -0.85638822,  0.13508654]]))
```

[3]:
```
B = A[0:3,0:3]; B
```

[3]:
```
array([[6, 3, 2],
       [3, 8, 5],
       [2, 5, 4]])
```

[4]:
```
np.linalg.eig(B)
```

[4]:
```
(array([13.19251931,  4.19280421,  0.61467648]),
 array([[ 0.44813055,  0.89394688,  0.00616301],
        [ 0.74077971, -0.36747195, -0.56232535],
        [ 0.50042426, -0.2565606 ,  0.82689311]]))
```

```
[5]: C = A[1:4,1:4]; C
```

```
[5]: array([[8, 5, 6],
            [5, 4, 5],
            [6, 5, 2]])
```

```
[6]: np.linalg.eig(C)
```

```
[6]: (array([15.78059973,  0.75095036, -2.53155008]),
      array([[-0.70641567, -0.65152327, -0.27657607],
             [-0.50869908,  0.73903826, -0.44164204],
             [-0.49214037,  0.17128887,  0.8534975 ]]))
```

Cauchy Interlacing Inequality

Let A be an $n \times n$ real symmetric matrix and B be the $(n-1) \times (n-1)$ leading principal submatrix of A, i.e.,

$$A = \begin{bmatrix} B & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Denote $\lambda_1(A) \leq \cdots \leq \lambda_n(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_{n-1}(B)$ the eigenvalues of A and B respectively. Then

$$\begin{aligned} &\lambda_1(A) \leq \lambda_1(B) \leq \lambda_2(A) \leq \lambda_2(B) \leq \cdots \leq \lambda_k(A) \leq \lambda_k(B) \\ &\qquad \leq \lambda_{k+1}(A) \leq \cdots \leq \lambda_{n-1}(A) \leq \lambda_{n-1}(B) \leq \lambda_n(A). \end{aligned}$$

Proof. Let $E = \text{Span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n-1}\}$. Then $\mathbf{x} \in E$ if and only if $\mathbf{x} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}$ with $\mathbf{y} \in \mathbb{R}^{n-1}$. Then

$$\mathbf{x}^T A \mathbf{x} = \begin{bmatrix} \mathbf{y}^T & 0 \end{bmatrix} \begin{bmatrix} B & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{y}^T B & \mathbf{y}^T A_{12} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} = \mathbf{y}^T B \mathbf{y}.$$

$$\begin{aligned} \lambda_k(A) &= \min_{S \subseteq \mathbb{R}^n : \dim S = k} \ \max_{\mathbf{x} \in S, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \\ &\leq \min_{S \subseteq E : \dim S = k} \ \max_{\mathbf{x} \in S, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \min_{T \subseteq \mathbb{R}^{n-1} : \dim T = k} \ \max_{\mathbf{y} \in T, \mathbf{y} \neq \mathbf{0}} \frac{\mathbf{y}^T B \mathbf{y}}{\mathbf{y}^T \mathbf{y}} = \lambda_k(B). \end{aligned}$$

On the other hand,

$$\begin{aligned} \lambda_{k+1}(A) &= \max_{S \subseteq \mathbb{R}^n : \dim S = n-(k+1)+1} \ \min_{\mathbf{x} \in S, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \\ &\geq \max_{S \subseteq E : \dim S = n-(k+1)+1} \ \min_{\mathbf{x} \in S, \mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \\ &= \max_{T \subseteq \mathbb{R}^{n-1} : \dim T = (n-1)-k+1} \ \min_{\mathbf{y} \in T, \mathbf{y} \neq \mathbf{0}} \frac{\mathbf{y}^T B \mathbf{y}}{\mathbf{y}^T \mathbf{y}} = \lambda_k(B). \end{aligned}$$

Therefore,

$$\lambda_k(A) \leq \lambda_k(B) \leq \lambda_{k+1}(A) \quad \text{for all } k = 1, \ldots, n-1.$$

Positive definite matrices

Positive definite matrix

Given an $n \times n$ real symmetric matrix A.

- A is said to be positive definite (positive semi-definite) if

$$\mathbf{x}^T A\mathbf{x} > 0 \ (\geq 0) \quad \text{for all nonzero } \mathbf{x} \in \mathbb{R}^n.$$

- A is said to be negative definite (negative semi-definite) if

$$\mathbf{x}^T A\mathbf{x} < 0 \ (\leq 0) \quad \text{for all nonzero } \mathbf{x} \in \mathbb{R}^n.$$

- A is said to be indefinite if it is neither positive semi-definite nor negative semi-definite.
- A is negative definite (negative semi-definite) if and only if $-A$ is positive definite (positive semi-definite).

Positive definite matrices

Example Let $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -3 & 0 \\ -3 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$.

For any nonzero $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$,

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= 2x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2 \\ &= x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2 > 0. \end{aligned}$$

So A is a positive definite matrix.

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -3 & 0 \\ -3 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 2 \quad \text{and} \quad \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & -3 & 0 \\ -3 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = -2.$$

So B is an indefinite matrix.

Positive definite $\Longrightarrow$ Nonsingular

Suppose $A = [a_{ij}]$ is a positive definite matrix. Then

1. A is nonsingular, and
2. all diagonal entries a_{jj} are positive.

Proof.

1. Suppose A is singular, then the system $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution $\mathbf{x}$. Thus,
$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T(\mathbf{0}) = 0$$
and so A is not positive definite.
2. Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis for $\mathbb{R}^n$. Take $\mathbf{x} = \mathbf{e}_j$. Then
$$a_{jj} = \mathbf{e}_j^T A\mathbf{e}_j > 0.$$

Remark In general, $\mathbf{e}_i^T A\mathbf{e}_j = a_{ij}$ for $1 \le i, j \le n$.

Eigenvalues of positive definite matrices

A symmetric matrix A is positive definite (semi-definite) if and only if all eigenvalues of A are positive (non-negative).

Proof. Suppose A has eigenvalues $\lambda_1 \le \cdots \le \lambda_n$. Then

$$\min_{\mathbf{x}\in\mathbb{R}^n,\, \mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_1 \le \cdots \le \lambda_n.$$

It is clear that A is positive semi-definite if and only if all eignvalues of A are non-negative.

If A is positive definite, then A is nonsingular and so all eigenvalues of A are nonzero, hence, must be positive. On the other hand, if $\lambda_1 > 0$, then $\mathbf{x}^T A\mathbf{x} > 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^n$ and so A is positive definite.

Nonsingular positive semi-definite $=$ Positive definite

A positive semi-definite matrix is positive definite if and only if it is nonsingular.

Positive definite matrices

Leading principal submatrix

For any $n \times n$ matrix $A = \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kk} & \cdots & a_{kn} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{nn} \end{bmatrix}$ and $1 \leq k \leq n$, the $k \times k$ matrix A_k is formed by deleting the last $n-k$ rows and columns of A, i.e.,

$$A_k = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}.$$

The matrices $A_1, A_2, \ldots, A_n$ are called leading principal submatrices of A.

Necessary and sufficient condition for positive definite matrix

An $n \times n$ symmetric matrix A is positive definite if and only if all leading principal submatrices have positive determinant, i.e.,

$$\det(A_k) > 0 \quad \text{for all} \quad k = 1, \ldots, n.$$

Proof. Omitted.

Positive definite matrices

Example Let $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$. Then

$$A_1 = \begin{bmatrix} 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \quad \text{and} \quad A_3 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Notice that

$$\det A_1 = 2, \quad \det A_2 = 3, \quad \text{and} \quad \det A_3 = 4.$$

Therefore, A is positive definite.

Necessary and sufficient condition for negative definite matrix

An $n \times n$ symmetric matrix A is negative definite if and only if

$$(-1)^k \det(A_k) > 0 \quad \text{for all} \quad k = 1, \ldots, n.$$

Proof. A is negative definite if and only if $-A$ is positive definite, which is equivalent to

$$(-1)^k \det(A_k) = \det(-A_k) > 0 \quad \text{for all} \quad k = 1, \ldots, n.$$

Positive definite matrices

Eigenvalues of $A + B$

Let A be an $n \times n$ real symmetric matrix and let B be an $n \times n$ positive semi-definite matrix. Then

$$\lambda_k(A+B) \geq \lambda_k(A) \quad \text{for all } k = 1, \ldots, n.$$

The inequality is strict if B is positive definite.

Proof.

$$\begin{aligned}
\max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T(A+B)\mathbf{x}}{\mathbf{x}^T\mathbf{x}} &= \max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \left(\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} + \frac{\mathbf{x}^T B\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \right) \\
&\geq \max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} + \min_{\mathbf{x}\in \mathbb{R}^n, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T B\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \\
&= \max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} + \lambda_1(B) \geq \max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}.
\end{aligned}$$

Then for any $k = 1, \ldots, n$,

$$\lambda_k(A+B) = \min_{S:\dim S = k} \max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T(A+B)\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \geq \min_{S:\dim S = k} \max_{\mathbf{x}\in S, \mathbf{x}\neq \mathbf{0}} \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_k(A).$$

When B is positive definite, $\lambda_1(B) > 0$, and so the inequality is strict.

Singular Value Decomposition

Singular Value Decomposition

Let A be an $m \times n$ rectangular matrix. Then there exist

- some nonnagative numbers $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min\{m,n\}} \geq 0$,
- an $m \times m$ orthogonal matrix $U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_m \end{bmatrix}$,
- an $n \times n$ orthogonal matrix $V = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}$, and
- an $m \times n$ matrix Σ with the (j,j)th entry σ_j and zero elsewhere

such that

(1) When $m = n$,

$$\begin{aligned}
A &= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \\
&= \begin{bmatrix} u_{11} & \cdots & u_{1m} \\ \vdots & \ddots & \vdots \\ u_{m1} & \cdots & u_{mm} \end{bmatrix} \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_m \end{bmatrix} \begin{bmatrix} v_{11} & \cdots & v_{n1} \\ \vdots & \ddots & \vdots \\ v_{1n} & \cdots & v_{nn} \end{bmatrix} \\
&= U\Sigma V^T
\end{aligned}$$

Singular Value Decomposition

(2) When $m > n$,

$$\begin{aligned} A &= \left[\begin{array}{ccc} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \\ \hline \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{array}\right] \\ &= \begin{bmatrix} u_{11} & \cdots & \cdots & u_{1m} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ u_{m1} & \cdots & \cdots & u_{mm} \end{bmatrix} \left[\begin{array}{ccc} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \\ \hline \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{array}\right] \begin{bmatrix} v_{11} & \cdots & v_{n1} \\ \vdots & \ddots & \vdots \\ v_{1n} & \cdots & u_{nn} \end{bmatrix} \\ &= U\Sigma V^T \end{aligned}$$

Singular Value Decomposition

(3) When $m < n$,

$$\begin{aligned} A &= \left[\begin{array}{ccc|cc} a_{11} & \cdots & a_{1m} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mm} & \cdots & a_{mn} \end{array}\right] \\ &= \begin{bmatrix} u_{11} & \cdots & u_{1m} \\ \vdots & \ddots & \vdots \\ u_{m1} & \cdots & u_{mm} \end{bmatrix} \left[\begin{array}{ccc|cc} \sigma_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & \sigma_m & \cdots & 0 \end{array}\right] \begin{bmatrix} v_{11} & \cdots & \cdots & v_{n1} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ v_{1n} & \cdots & \cdots & v_{nn} \end{bmatrix} \\ &= U\Sigma V^T \end{aligned}$$

In all cases,

$$A = \sum_{j=1}^{d} \sigma_j \mathbf{u}_j \mathbf{v}_j^T = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_d \mathbf{u}_d \mathbf{v}_d^T \quad \text{where } d = \min\{m, n\}.$$

The negative numbers $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d$ are called the singular values of A.
Notice that $\sigma_1^2 \geq \cdots \geq \sigma_d^2$ are the eigenvalues of $A^T A$.

Singular Value Decomposition

Examples

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \underbrace{\frac{1}{\sqrt{10}}\begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix}}_{U} \cdot \underbrace{\begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix}}_{\Sigma} \cdot \underbrace{\frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{bmatrix}}_{V^T}$$

$$A = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} = \underbrace{\frac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ 2 & -2 & -1 \end{bmatrix}}_{U} \cdot \underbrace{\begin{bmatrix} \sqrt{360} & 0 \\ 0 & \sqrt{90} \\ 0 & 0 \end{bmatrix}}_{\Sigma} \cdot \underbrace{\frac{1}{\sqrt{10}}\begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}}_{V^T}$$

$$A = \begin{bmatrix} 8 & 0 & 8 & 4 & -12 \\ 16 & 9 & 7 & -1 & -15 \\ 16 & 18 & -2 & -10 & -6 \end{bmatrix}$$

$$= \underbrace{\frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}}_{U} \cdot \underbrace{\begin{bmatrix} 36 & 0 & 0 & 0 & 0 \\ 0 & 18 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{\Sigma} \cdot \underbrace{\frac{1}{30}\begin{bmatrix} 20 & 15 & 5 & -5 & -15 \\ 0 & -15 & 15 & 15 & -15 \\ 20 & -15 & 5 & -5 & 15 \\ 9 & 8 & 9 & 25 & -7 \\ 12 & -6 & 12 & 0 & 24 \end{bmatrix}}_{V^T}$$

Singular Value Decomposition

Proof. Suppose the $n \times n$ matrix A^TA has r positive eigenvalues and $n-r$ zero eigenvalues, i.e.,

$$\lambda_1 \geq \cdots \geq \lambda_r > 0 \quad \text{and} \quad \lambda_{r+1} = \cdots = \lambda_n = 0,$$

and $\mathbf{v}_1, \ldots, \mathbf{v}_r, \ldots, \mathbf{v}_n$ are the n corresponding orthonormal eigenvectors, which form a basis for $\mathbb{R}^n$. Notice that

$$\langle A\mathbf{v}_i, A\mathbf{v}_j \rangle = \mathbf{v}_i^T A^T A \mathbf{v}_j = \lambda_j \mathbf{v}_i^T \mathbf{v}_j = 0 \quad \text{if } i \neq j$$

and

$$\|A\mathbf{v}_j\|^2 = \langle A\mathbf{v}_j, A\mathbf{v}_j \rangle = \mathbf{v}_j^T A^T A \mathbf{v}_j = \lambda_j \mathbf{v}_j^T \mathbf{v}_j = \lambda_j \begin{cases} > 0 & \text{if } j = 1, \ldots, r, \\ = 0 & \text{if } j = r+1, \ldots, n. \end{cases}$$

The first r vectors $A\mathbf{v}_1, \ldots, A\mathbf{v}_r$ form an orthogonal set in $\mathbb{R}^m$ and $A\mathbf{v}_{r+1} = \cdots = A\mathbf{v}_n = \mathbf{0}$. Define

$$\sigma_j = \begin{cases} \sqrt{\lambda_j} = \|A\mathbf{v}_j\| & j = 1, \ldots, r, \\ 0 & j = r+1, \ldots, \min\{m, n\}, \end{cases}$$

and

$$\mathbf{u}_j = \frac{1}{\|A\mathbf{v}_j\|} A\mathbf{v}_j = \frac{1}{\sqrt{\lambda_j}} A\mathbf{v}_j = \frac{1}{\sigma_j} A\mathbf{v}_j \quad j = 1, \ldots, r.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ forms an orthonormal set in $\mathbb{R}^m$. Now expand this orthonormalset to an **orthonormal basis** $\{\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\}$ (how?).

Singular Value Decomposition

Set

$U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_r & \mathbf{u}_{r+1} & \cdots & \mathbf{u}_m \end{bmatrix}$ and $V = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_r & \mathbf{v}_{r+1} & \cdots & \mathbf{v}_n \end{bmatrix}$.

Then both U and V are orthogonal. Now

$$\begin{aligned}
AV &= \left[\begin{array}{ccc|ccc} A\mathbf{v}_1 & \cdots & A\mathbf{v}_r & A\mathbf{v}_{r+1} & \cdots & A\mathbf{v}_n \end{array}\right] \\
&= \left[\begin{array}{ccc|ccc} \sigma_1\mathbf{u}_1 & \cdots & \sigma_r\mathbf{u}_r & \mathbf{0} & \cdots & \mathbf{0} \end{array}\right] \\
&= \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_r \end{bmatrix} \left[\begin{array}{ccc|ccc} \sigma_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r & 0 & \cdots & 0 \end{array}\right] \\
&= \left[\begin{array}{ccc|ccc} \mathbf{u}_1 & \cdots & \mathbf{u}_r & \mathbf{u}_{r+1} & \cdots & \mathbf{u}_m \end{array}\right] \left[\begin{array}{ccc|ccc} \sigma_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ \hline \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right] \\
&= U\Sigma.
\end{aligned}$$

Therefore,

$$AV = U\Sigma \quad \Longrightarrow \quad A = U\Sigma V^T.$$

Singular Value Decomposition

Example 4.4 Let $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$ has eigenvalues

- $\lambda_1 = 360$ with $E_{360}(A^TA) = \text{Span}\{\mathbf{v}_1\} = \text{Span}\left\{\frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\right\}$,
- $\lambda_2 = 90$ with $E_{90}(A^TA) = \text{Span}\{\mathbf{v}_2\} = \text{Span}\left\{\frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}\right\}$,
- $\lambda_3 = 0$ with $E_0(A^TA) = \text{Span}\{\mathbf{v}_3\} = \text{Span}\left\{\frac{1}{3}\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}\right\}$.

Then $\sigma_1 = \sqrt{360}$ and $\sigma_2 = \sqrt{90}$. Also

$$A\mathbf{v}_1 = \begin{bmatrix} 18 \\ 6 \end{bmatrix} \quad \text{and} \quad A\mathbf{v}_2 = \begin{bmatrix} 3 \\ -9 \end{bmatrix}.$$

Define

$$\mathbf{u}_1 = \frac{1}{\|A\mathbf{v}_1\|}A\mathbf{v}_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \frac{1}{\|A\mathbf{v}_2\|}A\mathbf{v}_2 = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ -3 \end{bmatrix}.$$

Set

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix},$$

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}.$$

Then

$$A = U\Sigma V^T = \underbrace{\frac{1}{\sqrt{10}} \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix}}_{U} \cdot \underbrace{\begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix}}_{\Sigma} \cdot \underbrace{\frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{bmatrix}}_{V^T}.$$

Example 4.5 Let $A = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 333 & 81 \\ 81 & 117 \end{bmatrix}$ has eigenvalues

- $\lambda_1 = 360$ with $E_{360}(A^TA) = \mathrm{Span}\{\mathbf{v}_1\} = \mathrm{Span}\left\{ \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}$,
- $\lambda_2 = 90$ with $E_{90}(A^TA) = \mathrm{Span}\{\mathbf{v}_2\} = \mathrm{Span}\left\{ \frac{1}{\sqrt{10}} \begin{bmatrix} -1 \\ 3 \end{bmatrix} \right\}$.

Then $\sigma_1 = \sqrt{360}$ and $\sigma_2 = \sqrt{90}$. Also

$$A\mathbf{v}_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 20 \\ 40 \\ 40 \end{bmatrix} \quad \text{and} \quad A\mathbf{v}_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} 20 \\ 10 \\ -20 \end{bmatrix}.$$

Define

$$\mathbf{u}_1 = \frac{1}{\|A\mathbf{v}_1\|} A\mathbf{v}_1 = \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \frac{1}{\|A\mathbf{v}_2\|} A\mathbf{v}_2 = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}.$$

Now we pick an unit vector in $(\mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2\})^\perp = (\mathrm{Col}\, W)^\perp = \mathrm{Nul}\, W^T$,

where $W = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{bmatrix}$.

Singular Value Decomposition

Direct computations show that $\operatorname{Nul} W^T = \operatorname{Span}\left\{\begin{bmatrix}2\\-2\\1\end{bmatrix}\right\}$. So take

$\mathbf{u}_3 = \frac{1}{3}\begin{bmatrix}2\\-2\\1\end{bmatrix}$. Set

$$
\begin{aligned}
U &= \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3\end{bmatrix} = \frac{1}{3}\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & -2\\ 2 & -2 & 1\end{bmatrix},\\
\Sigma &= \begin{bmatrix}\sqrt{\lambda_1} & 0\\ 0 & \sqrt{\lambda_2}\\ 0 & 0\end{bmatrix} = \begin{bmatrix}\sqrt{360} & 0\\ 0 & \sqrt{90}\\ 0 & 0\end{bmatrix},\\
V &= \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} = \frac{1}{\sqrt{10}}\begin{bmatrix}3 & -1\\ 1 & 3\end{bmatrix}.
\end{aligned}
$$

Then

$$
A = U\Sigma V^T = \underbrace{\frac{1}{3}\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & -2\\ 2 & -2 & 1\end{bmatrix}}_{U} \cdot \underbrace{\begin{bmatrix}\sqrt{360} & 0\\ 0 & \sqrt{90}\\ 0 & 0\end{bmatrix}}_{\Sigma} \cdot \underbrace{\frac{1}{\sqrt{10}}\begin{bmatrix}3 & 1\\ -1 & 3\end{bmatrix}}_{V^T}.
$$

The Spectral Theorem

[1]:
```
import sympy as sp
import numpy as np
```

[2]:
```
A = sp.Matrix([[4,11,14], [8,7,-2]]);A
```

[2]: $\begin{bmatrix}4 & 11 & 14\\ 8 & 7 & -2\end{bmatrix}$

[3]:
```
# Singular Values
A.singular_values()
```

[3]: [6*sqrt(10), 3*sqrt(10), 0]

[4]:
```
# Singular Value Decomposition @ SymPy
U,S,V = A.singular_value_decomposition()
```

[5]:
```
U
```

[5]: $\begin{bmatrix}-\frac{\sqrt{10}}{10} & \frac{3\sqrt{10}}{10}\\ \frac{3\sqrt{10}}{10} & \frac{\sqrt{10}}{10}\end{bmatrix}$

[6]:
```
S
```

[6]: $\begin{bmatrix}3\sqrt{10} & 0\\ 0 & 6\sqrt{10}\end{bmatrix}$

[7]:
```
V
```

[7]: $\begin{bmatrix}\frac{2}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{2}{3}\\ -\frac{2}{3} & \frac{2}{3}\end{bmatrix}$

[8]:
```
B = A.T;B
```

[8]: $\begin{bmatrix}4 & 8\\ 11 & 7\\ 14 & -2\end{bmatrix}$

[9]:
```
U,S,V = B.singular_value_decomposition();U
```

[9]: $\begin{bmatrix}\frac{2}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{2}{3}\\ -\frac{2}{3} & \frac{2}{3}\end{bmatrix}$

[10]:
```
S
```

[10]: $\begin{bmatrix}3\sqrt{10} & 0\\ 0 & 6\sqrt{10}\end{bmatrix}$

[11]:
```
V
```

[11]: $\begin{bmatrix}-\frac{\sqrt{10}}{10} & \frac{3\sqrt{10}}{10}\\ \frac{3\sqrt{10}}{10} & \frac{\sqrt{10}}{10}\end{bmatrix}$

The Spectral Theorem

[12]:
```
A = np.array([[4,11,14], [8,7,-2]]);A
```

[12]:
```
array([[ 4, 11, 14],
       [ 8,  7, -2]])
```

[13]:
```
# Singular Value Decomposition @ NumPy
np.linalg.svd(A)
```

[13]:
```
(array([[-0.9486833 , -0.31622777],
        [-0.31622777,  0.9486833 ]]),
 array([18.97366596,  9.48683298]),
 array([[-0.33333333, -0.66666667, -0.66666667],
        [ 0.66666667,  0.33333333, -0.66666667],
        [-0.66666667,  0.66666667, -0.33333333]]))
```

[14]:
```
B = A.T;B
```

[14]:
```
array([[ 4,  8],
       [11,  7],
       [14, -2]])
```

[15]:
```
np.linalg.svd(B)
```

[15]:
```
(array([[-0.33333333, -0.66666667, -0.66666667],
        [-0.66666667, -0.33333333,  0.66666667],
        [-0.66666667,  0.66666667, -0.33333333]]),
 array([18.97366596,  9.48683298]),
 array([[-0.9486833 , -0.31622777],
        [ 0.31622777, -0.9486833 ]]))
```


Singular Value Decomposition

Some Facts about Singular Value Decomposition

Let A be an $m \times n$ rectangular matrix with rank r and $A = U\Sigma V^T$.

- $\sigma_1^2, \ldots, \sigma_{\min\{m,n\}}^2$ are the eigenvalues of A^TA and AA^T.
- The columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V are the eigenvectors of A^TA. They are also called the right singular vectors of A.
- The columns $\mathbf{u}_1, \ldots, \mathbf{u}_m$ of U are the eigenvectors of AA^T. They are also called the left singular vectors of A.
- $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ forms an orthonormal basis for $\mathrm{Col}\, A^T$.
- $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ forms an orthonormal basis for $\mathrm{Nul}\, A$.
- $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ forms an orthonormal basis for $\mathrm{Col}\, A$.
- $\{\mathbf{u}_{r+1}, \ldots, \mathbf{v}_m\}$ forms an orthonormal basis for $\mathrm{Nul}\, A^T$.

Singular Value Decomposition [Compact form]

Let A be an $m \times n$ rectangular matrix and $d = \min\{m, n\}$. Then there exist

- an $m \times d$ matrix $U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_d \end{bmatrix}$ with $U^T U = I_d$,
- an $n \times d$ matrix $V = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_d \end{bmatrix}$ with $V^T V = I_d$,
- a $d \times d$ diagonal matrix Σ with diagonal entries $\sigma_1 \geq \cdots \geq \sigma_d$,

such that

$$A = U \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_d \end{bmatrix} V^T = \sum_{j=1}^{d} \sigma_j \mathbf{u}_j \mathbf{v}_j^T.$$

Furthermore, suppose $r = \text{Rank } A$. Then there exist

- a $m \times r$ matrix $U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_r \end{bmatrix}$ with $U^T U = I_r$,
- a $n \times r$ matrix $V = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_r \end{bmatrix}$ with $V^T V = I_r$,
- a $r \times r$ diagonal matrix Σ with diagonal entries $\sigma_1 \geq \cdots \geq \sigma_r > 0$,

such that

$$A = U \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix} V^T = \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^T.$$

Denote $\#(X)$ by the number of entries required to storage a matrix X.

Consider the Singular Value Decomposition $A = U\Sigma V^T$ with Σ being a $r \times r$ diagonal matrix. Then

- $\#(A) = mn$
- $\#(U) = mr$
- $\#(\Sigma) = r$
- $\#(V) = nr$

Then

$$\#(U\Sigma V^T) = \#(U) + \#(\Sigma) + \#(V) = mr + r + nr = r(m + n + 1).$$

If the rank of A is much smaller than m and n, then $\#(U\Sigma V^T)$ will be much smaller than $\#(A)$.

For example, if $m = n = 1{,}000$ and $r = 50$, then

$$\#(A) = 1{,}000{,}000 \quad \text{and} \quad \#(U\Sigma V^T) = 100{,}050.$$

Singular Value Decomposition

Problem: Given an $m \times n$ matrix A with rank r and an integer $0 < k < r$.. Find the $m \times n$ matrix $\hat{A}$ such that

$$\|A - \hat{A}\|_F = \min_{\text{Rank } X = k} \|A - X\|_F.$$

This problem is well known as the low-rank approximation problem.

Low-rank Approximation [Eckart-Young Theorem]

Suppose A is an $m \times n$ matrix with rank r with the singular value decomposition

$$A = \sum_{j=1}^{d} \sigma_j \mathbf{u}_j \mathbf{v}_j^T \quad \text{with} \quad \sigma_1 \geq \cdots \geq \sigma_d.$$

Given $0 < k < r \leq d = \min\{m, n\}$ and define

$$\hat{A} = \sum_{j=1}^{k} \sigma_j \mathbf{u}_j \mathbf{v}_j^T.$$

Notice that $\hat{A}$ is an $m \times n$ rank-k matrix. Furthermore,

$$\|A - \hat{A}\|_F \leq \|A - X\|_F \quad \text{for any } m \times n \text{ rank-}k \text{ matrix } X.$$

In other words, the optimal rank-k approximation to A is $\hat{A}$.

Singular Value Decomposition

[1]:
```
import numpy as np
import matplotlib.pyplot as plt
from PIL import Image
import os
```

[2]:
```
# input image
IMG = Image.open('BlockZ.png')
```

[3]:
```
# File size
file_size = os.path.getsize('BlockZ.png'); file_size
```

[3]: 11041601

[4]:
```
# Plot the orignal image
plt.imshow(IMG)
file_size = os.path.getsize('BlockZ.png');
plt.title("Original / File Size = %s Bytes" %file_size)
plt.show()
```

THE HONG KONG POLYTECHNIC UNIVERSITY 香港理工大學

Singular Value Decomposition

[5]:
```
#convert image to matrix
IMG_mat = np.array(list(IMG.getdata(band=0)), float)
IMG_mat.shape = (IMG.size[1], IMG.size[0])
```

[6]:
```
IMG_mat
```

[6]:
```
array([[123., 123., 125., ...,  89.,  89.,  89.],
       [125., 125., 125., ...,  89.,  89.,  89.],
       [128., 127., 126., ...,  89.,  89.,  89.],
       ...,
       [  0.,   0.,   0., ...,  25.,  28.,  29.],
       [  0.,   0.,   0., ...,  25.,  28.,  29.],
       [  1.,   1.,   1., ...,  13.,  17.,  19.]])
```

[7]:
```
# Image Matrix Size
IMG_mat.shape
```

[7]:
```
(3270, 5252)
```

[8]:
```
#Singular Value Decomposition
U, S, V = np.linalg.svd(IMG_mat)
```

[9]:
```
S
```

[9]:
```
array([7.11692788e+05, 1.04473952e+05, 6.36477970e+04, ...,
       5.50141017e+00, 5.47601291e+00, 5.41345454e+00])
```


Singular Value Decomposition

[10]:
```
U
```

[10]:
```
array([[-2.11287306e-02, -1.73201650e-02, -1.63692991e-02, ...,
         1.55302105e-02,  1.81898752e-02, -3.13303540e-02],
       [-2.11286058e-02, -1.73409438e-02, -1.63650167e-02, ...,
        -2.07066166e-02, -1.09210090e-02,  5.51064864e-02],
       [-2.11291949e-02, -1.73622514e-02, -1.63397961e-02, ...,
        -4.23506390e-02, -3.01032125e-02, -2.08291708e-02],
       ...,
       [-8.86029103e-03, -2.14506373e-05,  1.76507318e-02, ...,
        -6.48376920e-05,  3.82014595e-04,  2.60432303e-04],
       [-8.89369969e-03,  1.34553943e-04,  1.78083342e-02, ...,
         8.74721643e-04, -2.47302936e-04, -1.40882566e-04],
       [-8.82726495e-03,  1.10209059e-07,  1.69906356e-02, ...,
        -2.36243194e-04,  4.85242251e-04, -2.26682713e-04]])
```

[11]:
```
V
```

[11]:
```
array([[-8.92943734e-03, -8.92023161e-03, -8.90544695e-03, ...,
        -1.13182873e-02, -1.13242451e-02, -1.13225670e-02],
       [ 1.90953546e-02,  1.90048136e-02,  1.88537246e-02, ...,
         2.93563447e-02,  2.93157073e-02,  2.93117420e-02],
       [-5.06685409e-03, -5.04020940e-03, -4.94616658e-03, ...,
         1.44143894e-02,  1.45876485e-02,  1.46057531e-02],
       ...,
       [ 7.60796818e-04, -3.32574554e-03,  2.08830540e-03, ...,
         5.16046490e-01, -1.28437911e-01, -4.04860630e-02],
       [-2.64949953e-03,  9.29913218e-03,  1.93017301e-04, ...,
        -1.31648042e-01,  5.47825936e-01, -2.19775330e-01],
       [-2.00165304e-02, -3.40311018e-03,  4.26387337e-03, ...,
        -3.00906022e-02, -2.18187562e-01,  3.51703033e-01]])
```

Singular Value Decomposition

[12]:
```python
# The first 20 singular values
S[0:20]
```

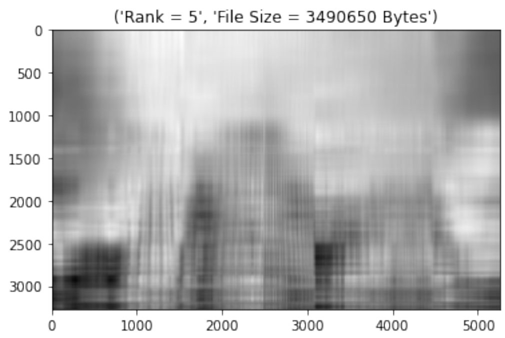
[12]:
```
array([711692.78787915, 104473.95173412,  63647.79699123,  47523.35334201,
        42185.00854629,  34022.4435739 ,  30527.23271654,  27093.10942512,
        26325.19804892,  25042.64360772,  24665.51132907,  19884.99161464,
        19317.54885838,  18471.46085386,  18261.03204542,  17491.65767212,
        16884.5122241 ,  16599.63778144,  16153.38345255,  15370.29097403])
```

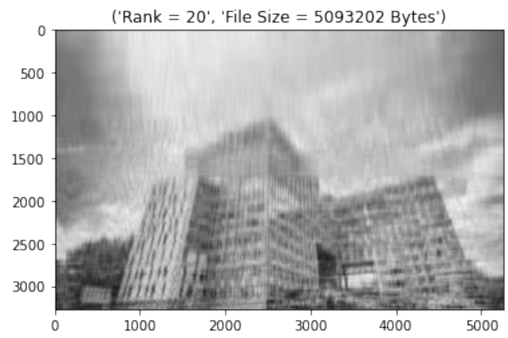
[13]:
```python
# The last 20 singular values
S[3250:3270]
```

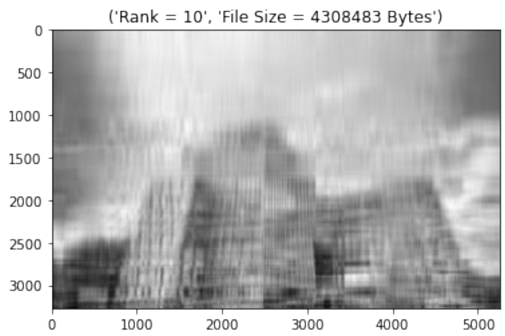
[13]:
```
array([5.98396226, 5.95490429, 5.94065185, 5.92239014, 5.9017396 ,
       5.87392038, 5.85189078, 5.82260538, 5.79400483, 5.76139741,
       5.74155833, 5.69073336, 5.66677736, 5.64142337, 5.60524532,
       5.57973692, 5.56430795, 5.50141017, 5.47601291, 5.41345454])
```

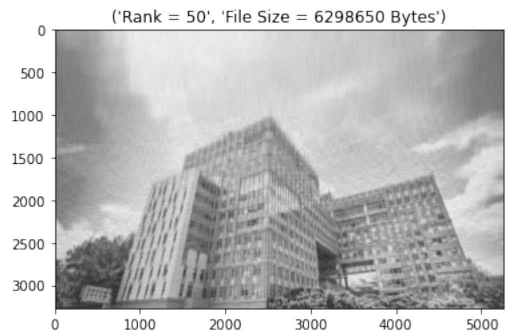
[14]:
```python
#Low-rank approximation for different rank
import matplotlib
list = [10,50,100,200,500,1000,3000]
for i in list:
    IMG_approx = np.matrix(U[:, :i])*np.diag(S[:i])*np.matrix(V[:i, :])
    filename = 'BlockZ_%s.png'%i
    matplotlib.image.imsave(filename, IMG_approx,cmap='gray')
    file_size = os.path.getsize(filename);
    plt.imshow(IMG_approx,cmap='gray')
    title = "Rank = %s" %i, "File Size = %s" %file_size
    plt.title(title)
    plt.show()
```

Singular Value Decomposition









Singular Value Decomposition

Singular Value Decomposition