Previously: given a square $n \times n$ symmetric matrix A

can write $A = PDP^T$ P = orthonormal
(so rows/columns form orthonormal basis)

$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ diagonal

Anton 9.4 Singular Value Decomposition

Let A be an $\underbrace{m \times n}$ matrix.

m & n could be different so A need not be square

$\underset{n \times m}{A^T} \underset{m \times n}{A}$ is symmetric, so

$A^T A = n \times n$ matrix

- eigenvalues are non-negative, so let's arrange them largest to smallest

$$\underbrace{\lambda_1 \geq \lambda_2 \cdots \geq \lambda_k}_{\text{positive}} > \underset{\lambda_{k+1}}{0} = \underset{\lambda_{k+2}}{0} \cdots = \underset{\lambda_n}{0}$$

(if some eigenvalues are 0)

- has orthonormal basis of eigenvectors $v_1, \ldots, v_n$ correspond to eigenvalues $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_k > 0 = 0 \cdots = 0$

Singular Value Decomposition says we can write.

$$\underset{m\times n}{A} = \underset{\substack{m\times m \\ \text{orthogonal}}}{U} \quad \underset{\substack{m\times n \\ \text{(main)} \\ \text{diagonal}}}{\Sigma} \quad \underset{\substack{n\times n \\ \text{orthogonal}}}{V^T}$$

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \sigma_2 & & 0 \\ & & \ddots\ \sigma_k & \\ \hline & 0 & & 0 \end{array}\right]$$

(generalizing $A = PDP^T$ for symmetric square matrices)

What must U, Σ, V^T be?

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

Then $\underbrace{A^TA}_{\text{symmetric matrix}} = (U\Sigma V^T)^T(U\Sigma V^T)$

$$= (V^T)^T \Sigma^T \underbrace{U^TU}_{I_m} \Sigma V^T$$

$$A^TA = V\underbrace{(\Sigma^T\Sigma)}_{n\times n \text{ diagonal matrix.}}V^T$$

So if $A^TA = PDP^T$ from the orthogonal diagonalization of A^TA, then

$V = P$

$\Sigma^T\Sigma = D$, so σ_i^2 are the eigenvalues of A^TA

for $i \leq \min(m,n)$

$$\left[\begin{array}{c|c} \sigma_1 \ \ddots \ \sigma_k & 0 \\ \hline 0 & 0 \end{array}\right]$$

Singular Value Decomposition of an $m \times n$ matrix A

$$\underset{m\times n}{A} = \underset{\substack{m\times m \\ \text{orthogonal}}}{U} \; \underset{\substack{m\times n \\ \text{(main)} \\ \text{diagonal}}}{\Sigma} \; \underset{\substack{n\times n \\ \text{orthonormal}}}{V^T}$$

$$= \begin{bmatrix} | & & | & | & & | \\ u_1 & \cdots & u_k & u_{k+1} & \cdots & u_m \\ | & & | & | & & | \end{bmatrix} \left[\begin{array}{ccc|c} \sqrt{\lambda_1} & & & \\ & \ddots & & \\ & & \sqrt{\lambda_k} & \\ \hline & & & \end{array}\right] \begin{bmatrix} - v_1^T - \\ \vdots \\ v_n^T \end{bmatrix}$$

- $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$ are positive eigenvalues for $\underline{A^T A}$
- $v_1, \ldots v_n$ orthonormal basis of eigenvectors for $\underline{A^T A}$ with v_i having eigenvalue λ_i
- $u_i = \dfrac{Av_i}{\|Av_i\|} = \dfrac{1}{\sqrt{\lambda_i}} A v_i$ for $i = 1, 2, \ldots, k$
- $\{u_1, \ldots u_k\}$ is orthonormal basis for column space of A
- $\{u_1, \ldots u_k, u_{k+1}, \ldots u_m\}$ is extension of $\{u_1, \ldots u_k\}$ to an orthonormal basis for $\mathbb{R}^m$

Proof is easy: (details later)

$U\Sigma = AV$ from $u_i \sqrt{\lambda_i} = A v_i$

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

$U \Sigma V^T = A$ since V is orthogonal

Terminology:

$v_1, \ldots v_n$ are called the <u>right singular vectors</u> of A

$u_1, \ldots, u_k$ are called the <u>left singular vectors</u> of A

$\sigma_1 = \sqrt{\lambda_1}, \ldots \sigma_k = \sqrt{\lambda_k}$ are <u>singular values</u> of A

Anton 9.4 Example 2.

Find a singular value decomposition of $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$

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

i.e Find the 3 matrices

Solution First find orthogonal diagonalization of $\underline{A^T A}$

$$B = A^T A = \underset{2\times 3}{\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}} \underset{3\times 2}{\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Find eigenvalues of B via characteristic polynomial

$$0 = \det(B - \lambda I) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)(2-\lambda) - 1\cdot 1$$

$$0 = 4 - 4\lambda + \lambda^2 - 1$$
$$0 = \lambda^2 - 4\lambda + 3$$

$$0 = (\lambda - 1)(\lambda - 3)$$ so eigenvalues are $\lambda_1 = 3,\ \lambda_2 = 1$

3 is biggest eigenvalue so let it be λ_1, not λ_2

$B = A^T A$

$Bv = \lambda_1 v$

$Bv - \lambda_1 v = 0$

$(B - \lambda_1 I)v = 0$

Find $\lambda_1 = 3$ eigenvector as null space of

$$B - \lambda_1 I = \begin{bmatrix} 2-3 & 1 \\ 1 & 2-3 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = (B - \lambda I)\, v = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

row reduce $B-3I$

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \xrightarrow{(-1)R_1} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \overset{v}{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$c_1 - c_2 = 0$ $\quad$ $c_1 = c_2$

c_2 is free $\quad$ $c_2 = c_2$ $\quad = \begin{bmatrix} 1 \\ 1 \end{bmatrix} c_2$

so $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector, but not a unit eigenvector.
It has $\| \quad \|^2 = 1^2 + 1^2 = 2$
so we scale it by $\frac{1}{\sqrt{2}}$

$$\boxed{v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

is $\underset{\text{unit}}{\wedge}$ eigenvector for A

Similarly, for $\lambda_2 = 1$: $\quad B = A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ $\quad B - 1 \cdot I = \begin{bmatrix} 2-1 & 1 \\ 1 & 2-1 \end{bmatrix}$

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

row reducing $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

$c_1 + c_2 = 0$ $\quad$ $c_1 = -c_2$

c_2 is free $\quad$ $c_2 = c_2$ $\quad = \begin{bmatrix} -1 \\ 1 \end{bmatrix} c_2$

so $\boxed{v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}}$

$\sqrt{2}$ is norm of $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

(note $-v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is also an acceptable choice)

What have we found so far?

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

$$\underset{3\times 2}{A} = \underset{3\times 3}{U}\ \underset{3\times 2}{\Sigma}\ \underset{2\times 2}{V^T} = U \begin{bmatrix} \sqrt{\lambda_1} & \\ & \sqrt{\lambda_2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} - v_1^T - \\ - v_2^T - \end{bmatrix}$$

$$= U \begin{bmatrix} \sqrt{3} & \\ & \sqrt{1} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}}[1, 1] \\ \frac{1}{\sqrt{2}}[-1, 1] \end{bmatrix}$$

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

So it remains to find $\underset{3\times 3}{U} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$

From the theorem,

$$u_1 = \frac{1}{\sqrt{\lambda_1}} A v_1 \qquad\qquad u_2 = \frac{1}{\sqrt{\lambda_2}} A v_2$$

$$= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \qquad\qquad u_2 = \frac{1}{\sqrt{1}} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$$

$$u_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \qquad\qquad u_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$\{u_1, u_2\}$ is an orthonormal set, and we need to complete it to an orthonormal basis $\{u_1, u_2, u_3\}$ of $\mathbb{R}^m = \mathbb{R}^3$

If $u_3 = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ then $0 = \langle u_1, u_3 \rangle = \langle u_2, u_3 \rangle$

implies

$$\begin{array}{l} \text{scalar mult of } u_1 \to \\ \text{Scalar mult of } u_2 \to \end{array} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

i.e we need to find (an orthonormal basis for) the null space of $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$

row reduce it:

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix} \xrightarrow{R1 \to R1 - R2} \begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & -1 \end{bmatrix} \to \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}}_{\text{rref}}$$

So $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

x, y pivot vars
z is free

$$\begin{aligned} x + 0y + z &= 0 \\ y - z &= 0 \end{aligned} \Rightarrow \begin{aligned} x &= -z \\ y &= z \\ z &= z \end{aligned} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} z$$

So basis is $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, has norm $\sqrt{(-1)^2 + 1^2 + 1^2} = \sqrt{3}$

So $\quad u_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

So collecting everything:

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

$$\underset{3\times 2}{\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}} = \underset{3\times 3}{\begin{bmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{bmatrix}} \underset{3\times 2}{\begin{bmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\lambda_1} \\ 0 & 0 \end{bmatrix}} \begin{bmatrix} - & v_1^T & - \\ - & v_2^T & - \end{bmatrix}$$

$$\boxed{\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \underset{U}{\begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}} \underset{\Sigma}{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{1} \\ 0 & 0 \end{bmatrix}} \underset{V^T}{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}}$$

is SVD of A

What is it good for? Spectral decomposition

$A_1 = \sigma_1 u_1 v_1^T$ is best rank 1 approx of A

$$= \sqrt{3} \cdot \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \frac{1}{\sqrt{2}} [1, 1] = \frac{1}{2} \begin{bmatrix} 2 & 2 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}}_{\text{rank 1 matrix}}$$

has entries pretty ~~close~~ to A

SVD in SageMath

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{1} \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{V^T}$$

is SVD of A

```
A=matrix(RDF, 3, 2, [1, 1, 0, 1, 1, 0])
A
```

```
[1.0 1.0]
[0.0 1.0]
[1.0 0.0]
```

```
A.SVD()
```

```
([    -0.8164965809277257 -1.8557752066599326e-16      -0.5773502691896257]
[    -0.4082482904638629      -0.7071067811865475       0.5773502691896257]
[     -0.408248290463863       0.7071067811865474       0.5773502691896257], [1.7320508075688772                0.0]
[               0.0                1.0]
[               0.0                0.0], [-0.7071067811865477  0.7071067811865474]
[-0.7071067811865474 -0.7071067811865477])
```

```
typeset_mode()
```

```
A.SVD()
```

$\sqrt{3}$

$$\left(\begin{pmatrix} -0.8164965809277257 & -1.8557752066599326\times 10^{-16} & -0.5773502691896257 \\ -0.4082482904638629 & -0.7071067811865475 & 0.5773502691896257 \\ -0.408248290463863 & 0.7071067811865474 & 0.5773502691896257 \end{pmatrix}, \begin{pmatrix} 1.7320508075688772 & 0.0 \\ 0.0 & 1.0 \\ 0.0 & 0.0 \end{pmatrix}, \underbrace{\begin{pmatrix} -0.7071067811865477 & 0.7071067811865474 \\ -0.7071067811865474 & -0.7071067811865477 \end{pmatrix}}\right)$$

SageMath returns V not V^T

```
U, S, V=A.SVD()
```

```
(U*S*V).round()
```

not A! $\begin{pmatrix} 1.0 & -1.0 \\ 1.0 & 0.0 \\ 0.0 & -1.0 \end{pmatrix}$

V^T

```
(U*S*(V.T)).round()
```

$A = \begin{pmatrix} 1.0 & 1.0 \\ 0.0 & 1.0 \\ 1.0 & -0.0 \end{pmatrix}$

SVD in Python

In [25]:
```python
A=[[1,1], [0,1], [1,0]]
```

In [26]:
```python
import numpy as np
```

In [27]:
```python
np.linalg.svd(A)
```

Out[27]:
```
(array([[-8.16496581e-01, -1.85577521e-16, -5.77350269e-01],
        [-4.08248290e-01, -7.07106781e-01,  5.77350269e-01],
        [-4.08248290e-01,  7.07106781e-01,  5.77350269e-01]]),
 array([1.73205081, 1.        ]),
 array([[-0.70710678, -0.70710678],
        [ 0.70710678, -0.70710678]]))
```

← Python returns V^T not V

In [32]:
```python
U, s, V = np.linalg.svd(A)
```

In [29]:
```python
np.diag(s)
```

Out[29]:
```
array([[1.73205081, 0.        ],
       [0.        , 1.        ]])
```

In [30]:
```python
np.vstack((np.diag(s), [0,0]))
```

Out[30]:
```
array([[1.73205081, 0.        ],
       [0.        , 1.        ],
       [0.        , 0.        ]])
```

```python
U@(np.vstack((np.diag(s), [0,0])))@V
```

```
array([[ 1.00000000e+00,  1.00000000e+00],
       [ 5.61334798e-17,  1.00000000e+00],
       [ 1.00000000e+00, -1.56386917e-16]])
```

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

Spectral decomposition just as for symmetric matrices

$$\underset{m\times n}{A} = \underset{\substack{m\times m\\ \text{orthogonal}}}{U} \; \underset{\substack{m\times n\\ \text{(main)}\\ \text{diagonal}}}{\Sigma} \; \underset{\substack{n\times n\\ \text{orthonormal}}}{V^T}$$

$$= \begin{bmatrix} | & & | & | & & | \\ u_1 & \cdots & u_k & u_{k+1} & \cdots & u_n \\ | & & | & | & & | \end{bmatrix} \left[\begin{array}{c|c} \begin{matrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_k} \end{matrix} & \\ \hline & \end{array}\right] \begin{bmatrix} - v_1^T - \\ \vdots \\ v_n^T \end{bmatrix}$$

Let $\sigma_i = \sqrt{\lambda_i}$

For any $r = 1, 2, \ldots, k$

let $A_r = \sigma_1 \underbrace{u_1}_{m\times 1} \underbrace{v_1^T}_{1\times n} + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$

$= m\times n$

is a rank r matrix closely approximating A.

Application: Image compression