# Eigen Decomposition and Singular Value Decomposition

Based on the slides by Mani Thomas
Modified and extended by Longin Jan Latecki

# Introduction

- Eigenvalue decomposition
  - Spectral decomposition theorem
- Physical interpretation of eigenvalue/eigenvectors
- Singular Value Decomposition

# What are eigenvalues?

- Given a matrix, $\mathbf{A}$, $\mathbf{x}$ is the eigenvector and $\lambda$ is the corresponding eigenvalue if $\mathbf{Ax} = \lambda\mathbf{x}$
  - A must be square and the determinant of $\mathbf{A} - \lambda\, \mathbf{I}$ must be equal to zero

$$\mathbf{Ax} - \lambda\mathbf{x} = 0 \Rightarrow (\mathbf{A} - \lambda\mathbf{I})\, \mathbf{x} = 0$$

    - Trivial solution is if $\mathbf{x} = 0$
    - The non trivial solution occurs when $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$
- Are eigenvectors are unique?
  - If $\mathbf{x}$ is an eigenvector, then $\beta\mathbf{x}$ is also an eigenvector and $\beta\lambda$ is an eigenvalue

$$\mathbf{A}(\beta\mathbf{x}) = \beta(\mathbf{Ax}) = \beta(\lambda\mathbf{x}) = \lambda(\beta\mathbf{x})$$

# Calculating the Eigenvectors/values

- Expand the det(**A** - λ**I**) = 0 for a 2 **£** 2 matrix

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0$$

$$\det\begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} = 0 \Rightarrow (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0$$

$$\lambda^2 - \lambda(a_{11} + a_{22}) + (a_{11}a_{22} - a_{12}a_{21}) = 0$$

- For a 2 **£** 2 matrix, this is a simple quadratic equation with two solutions (maybe complex)

$$\lambda = \frac{(a_{11} + a_{22}) \pm \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}}{2}$$

- This "characteristic equation" can be used to solve for **x**

# Eigenvalue example

- Consider,

$$\mathbf{A}=\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}\Rightarrow\begin{cases}\lambda^2-(a_{11}+a_{22})\lambda+(a_{11}a_{22}-a_{12}a_{21})=0\\ \lambda^2-(1+4)\lambda+(1\cdot 4-2\cdot 2)=0\\ \lambda^2=(1+4)\lambda\Rightarrow\lambda=0,\lambda=5\end{cases}$$

- The corresponding eigenvectors can be computed as

$$\lambda=0\Rightarrow\left[\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}-\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}\right]\cdot\begin{bmatrix}x\\ y\end{bmatrix}=0\Rightarrow\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}\cdot\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}1x+2y\\ 2x+4y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\lambda=5\Rightarrow\left[\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}-\begin{bmatrix}5 & 0\\ 0 & 5\end{bmatrix}\right]\cdot\begin{bmatrix}x\\ y\end{bmatrix}=0\Rightarrow\begin{bmatrix}-4 & 2\\ 2 & -1\end{bmatrix}\cdot\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}-4x+2y\\ 2x-1y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

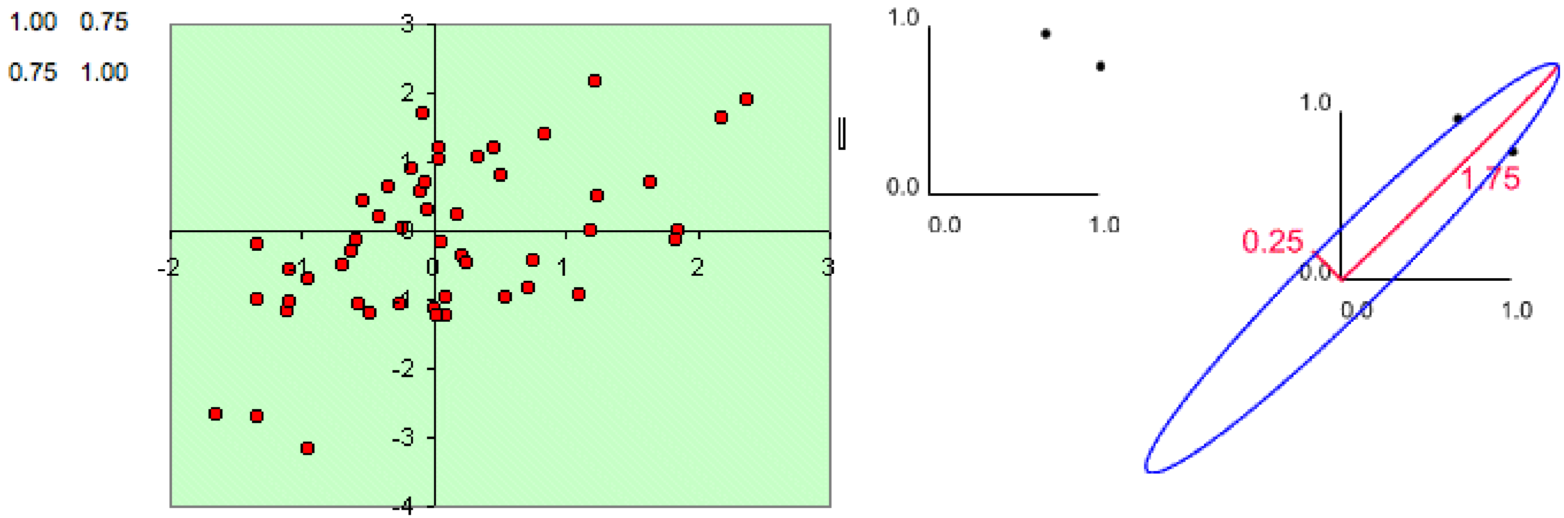
  - For $\lambda = 0$, one possible solution is $\mathbf{x} = (2, -1)$
  - For $\lambda = 5$, one possible solution is $\mathbf{x} = (1, 2)$

For more information: Demos in Linear algebra by G. Strang, http://web.mit.edu/18.06/www/

# Physical interpretation

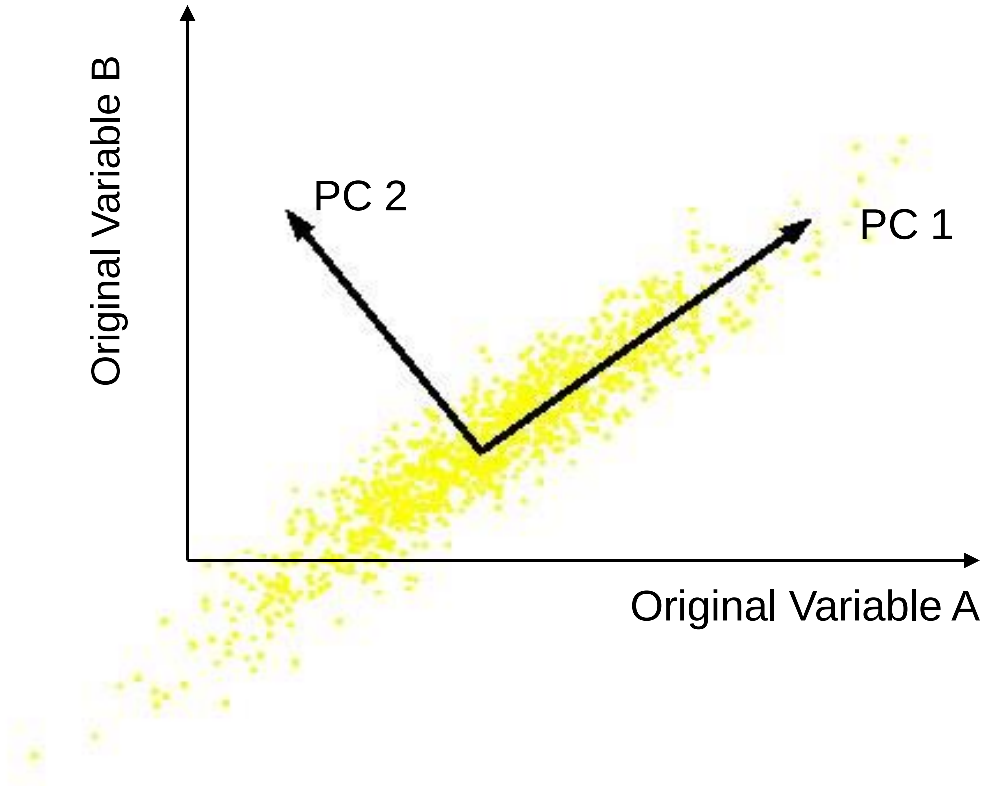
- Consider a covariance matrix, **A**, i.e., A = 1/n S $S^T$ for some S

$$\mathbf{A} = \begin{bmatrix} 1 & .75 \\ .75 & 1 \end{bmatrix} \Rightarrow \lambda_1 = 1.75, \lambda_2 = 0.25$$




- Error ellipse with the major axis as the larger eigenvalue and the minor axis as the smaller eigenvalue

# Physical interpretation




- Orthogonal directions of greatest variance in data
- Projections along PC1 (Principal Component) discriminate the data most along any one axis

# Physical interpretation

- First principal component is the direction of greatest variability (covariance) in the data
- Second is the next orthogonal (uncorrelated) direction of greatest variability
  - So first remove all the variability along the first component, and then find the next direction of greatest variability
- And so on …
- Thus each eigenvectors provides the directions of data variances in decreasing order of eigenvalues

For more information: See Gram-Schmidt Orthogonalization in G. Strang's lectures

# Multivariate Gaussian

- Multivariate Normal (MVN)

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \Sigma) \stackrel{\text{def}}{=} \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp[-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})]$$

- Exponent is the Mahalanobis distance between x and μ

$$\Delta = (\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})$$

Σ is the covariance matrix (positive definite)

$$\mathbf{x}^T\Sigma\mathbf{x} > 0 \; \forall \mathbf{x}$$

# Bivariate Gaussian

- Covariance matrix is

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$

where the correlation coefficient is

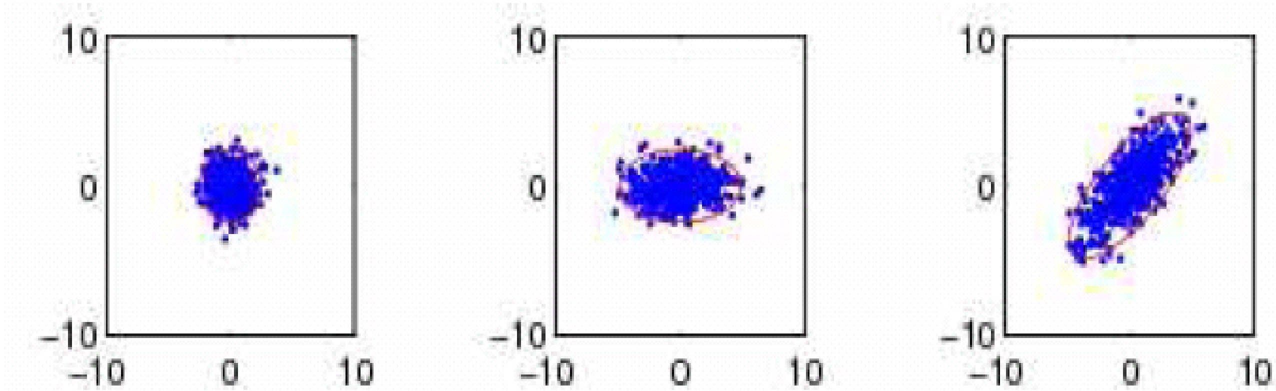
$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

and satisfies $-1 \leq \rho \leq 1$

- Density is

$$p(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} - \frac{2\rho xy}{(\sigma_x\sigma_y)}\right)\right)$$

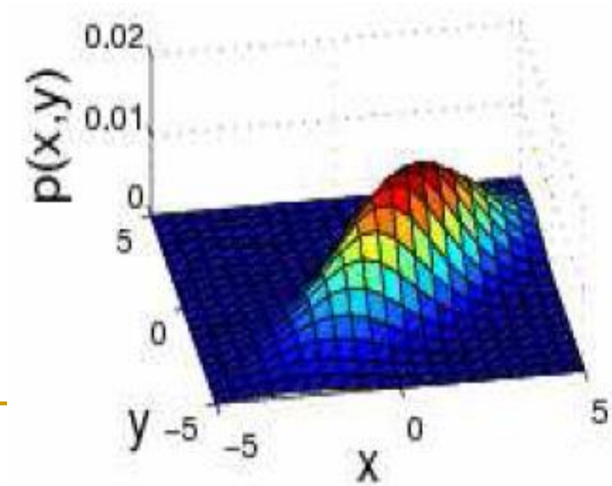
# Spherical, diagonal, full covariance

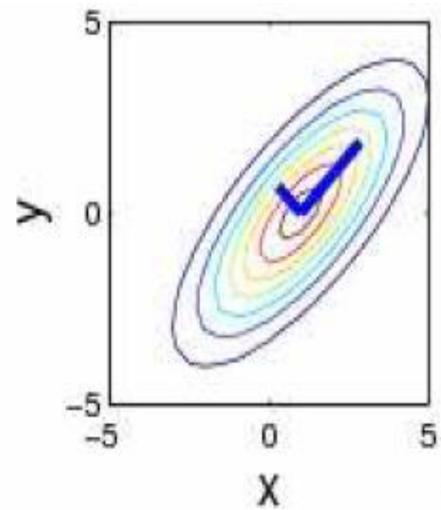



$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$



# Eigen/diagonal Decomposition

- Let $\mathbf{S} \in \mathbb{R}^{m \times m}$ be a **square** matrix with ***m* linearly independent eigenvectors** (a "non-defective" matrix)
- **Theorem**: Exists an **eigen decomposition**

$$\mathbf{S} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$$

*diagonal*

Unique for distinct eigen-values

  - (cf. matrix diagonalization theorem)
- Columns of ***U*** are **eigenvectors** of ***S***
- Diagonal elements of $\mathbf{\Lambda}$ are **eigenvalues** of $\mathbf{S}$

$$\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_m), \quad \lambda_i \geq \lambda_{i+1}$$

# Diagonal decomposition: why/how

Let ***U*** have the eigenvectors as columns: $U = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}$

Then, ***SU*** can be written

$$SU = S\begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \ldots & \lambda_n v_n \end{bmatrix} = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}\begin{bmatrix} \lambda_1 & & \\ & \ldots & \\ & & \lambda_n \end{bmatrix}$$

Thus $\boldsymbol{SU=U\Lambda}$, or $\boldsymbol{U^{-1}SU=\Lambda}$

And $\boldsymbol{S=U\Lambda U^{-1}}$.

# Diagonal decomposition - example

Recall $S = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}; \lambda_1 = 1, \lambda_2 = 3.$

The eigenvectors $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ form $U = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

Inverting, we have $U^{-1} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$

Recall $UU^{-1} = 1$.

Then, ***S=UΛU⁻¹ =*** $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$

# Example continued

Let's divide ***U*** (and multiply ***U***$^{-1}$) by $\sqrt{2}$

Then, ***S***= $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

$\boldsymbol{Q}$ $\qquad$ $\boldsymbol{\Lambda}$ $\qquad$ $(\boldsymbol{Q}^{-1} = \boldsymbol{Q}^{T})$

Why? Stay tuned …

# Symmetric Eigen Decomposition

- If $\mathbf{S} \in \mathbb{R}^{m \times m}$ is a **symmetric** matrix:
- **Theorem**: Exists a (unique) **eigen decomposition** $S = Q \Lambda Q^T$
- where ***Q*** is **orthogonal:**
  - $\boldsymbol{Q^{-1} = Q^T}$
  - Columns of ***Q*** are normalized eigenvectors
  - Columns are orthogonal.
  - (everything is real)

# Spectral Decomposition theorem

- If **A** is a symmetric and positive definite k £ k matrix ($\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} > 0$) with $\lambda_{\mathbf{i}}$ ($\lambda_{\mathbf{i}} > 0$) and $\mathbf{e}_{\mathrm{i}}$, i = 1 ⋯ k being the k eigenvector and eigenvalue pairs, then

$$\underset{(k\times k)}{\mathbf{A}} = \lambda_1 \underset{(k\times 1)}{\mathbf{e}_1} \underset{(1\times k)}{\mathbf{e}_1^T} + \lambda_2 \underset{(k\times 1)}{\mathbf{e}_2} \underset{(1\times k)}{\mathbf{e}_2^T} \cdots + \lambda_k \underset{(k\times 1)}{\mathbf{e}_k} \underset{(1\times k)}{\mathbf{e}_k^T} \Rightarrow \underset{(k\times k)}{\mathbf{A}} = \sum_{i=1}^{k} \lambda_i \underset{(k\times 1)}{\mathbf{e}_i} \underset{(1\times k)}{\mathbf{e}_i^T} = \mathbf{P}\Lambda\mathbf{P}^T$$

$$\underset{(k\times k)}{\mathbf{P}} = [\mathbf{e}_1, \mathbf{e}_2 \cdots \mathbf{e}_k], \underset{(k\times k)}{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix}$$

  - This is also called the eigen decomposition theorem

- Any symmetric matrix can be reconstructed using its eigenvalues and eigenvectors

# Example for spectral decomposition

- Let **A** be a symmetric, positive definite matrix

$$A = \begin{bmatrix} 2.2 & 0.4 \\ 0.4 & 2.8 \end{bmatrix} \Rightarrow \det(A - \lambda I) = 0$$

$$\Rightarrow \lambda^2 - 5\lambda + (6.16 - 0.16) = (\lambda - 3)(\lambda - 2) = 0$$

- The eigenvectors for the corresponding eigenvalues are $\mathbf{e}_1^T = \left[ 1/\sqrt{5}, 2/\sqrt{5} \right], \mathbf{e}_2^T = \left[ 2/\sqrt{5}, -1/\sqrt{5} \right]$
- Consequently,

$$A = \begin{bmatrix} 2.2 & 0.4 \\ 0.4 & 2.8 \end{bmatrix} = 3 \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} + 2 \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}$$

$$= \begin{bmatrix} 0.6 & 1.2 \\ 1.2 & 2.4 \end{bmatrix} + \begin{bmatrix} 1.6 & -0.8 \\ -0.8 & 0.4 \end{bmatrix}$$

# Singular Value Decomposition

- If **A** is a rectangular m £ k matrix of real numbers, then there exists an m £ m orthogonal matrix **U** and a k £ k orthogonal matrix **V** such that

$$\underset{(m\times k)}{\mathbf{A}} = \underset{(m\times m)}{\mathbf{U}}\,\underset{(m\times k)}{\Lambda}\,\underset{(k\times k)}{\mathbf{V}^T} \qquad \mathbf{U}\mathbf{U}^T = \mathbf{V}\mathbf{V}^T = \mathbf{I}$$

  - Λ is an m £ k matrix where the (i, j)th entry $\lambda_i$ ¸ 0, i = 1 ⋯ min(m, k) and the other entries are zero
    - The positive constants $\lambda_i$ are the singular values of **A**
- If A has rank r, then there exists r positive constants $\lambda_1, \lambda_2, \cdots \lambda_r$, r orthogonal m £ 1 unit vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_r$ and r orthogonal k £ 1 unit vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r$ such that

$$\mathbf{A} = \sum_{i=1}^{r} \lambda_i \mathbf{u}_i \mathbf{v}_i^T$$

  - Similar to the spectral decomposition theorem

# Singular Value Decomposition (contd.)

- If **A** is a symmetric and positive definite then
  - SVD = Eigen decomposition
    - EIG($\lambda_i$) = SVD($\lambda_i^2$)
- Here $\mathbf{AA}^T$ has an eigenvalue-eigenvector pair ($\lambda_i^2$,$u_i$)

$$\mathbf{AA}^T = \left(\mathbf{U}\Lambda\mathbf{V}^T\right)\left(\mathbf{U}\Lambda\mathbf{V}^T\right)^T$$
$$= \mathbf{U}\Lambda\mathbf{V}^T\mathbf{V}\Lambda\mathbf{U}^T$$
$$= \mathbf{U}\Lambda^2\mathbf{U}^T$$

- Alternatively, the $\mathbf{v}_i$ are the eigenvectors of $\mathbf{A}^T\mathbf{A}$ with the same non zero eigenvalue $\lambda_i^2$

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\Lambda^2\mathbf{V}^T$$

```
>> A = rand(10, 3);
>> B = A'*A;
>> [e, l] = eig(B)

e =

    0.8280    0.1313    0.5451
   -0.1508   -0.8842    0.4421
   -0.5400    0.4483    0.7123


l =

    0.4660         0         0
         0    0.9342         0
         0         0    9.3340

>> [u d v] = svd(B*B')

u =

   -0.5451    0.1313   -0.8280
   -0.4421   -0.8842    0.1508
   -0.7123    0.4483    0.5400


d =

   87.1229         0         0
         0    0.8728         0
         0         0    0.2171


v =

   -0.5451    0.1313   -0.8280
   -0.4421   -0.8842    0.1508
   -0.7123    0.4483    0.5400
```

# Example for SVD

- Let **A** be a symmetric, positive definite matrix
  - **U** can be computed as

$$\mathbf{A}=\begin{bmatrix}3 & 1 & 1\\ -1 & 3 & 1\end{bmatrix}\Rightarrow \mathbf{AA}^T=\begin{bmatrix}3 & 1 & 1\\ -1 & 3 & 1\end{bmatrix}\begin{bmatrix}3 & -1\\ 1 & 3\\ 1 & 1\end{bmatrix}=\begin{bmatrix}11 & 1\\ 1 & 11\end{bmatrix}$$

$$\det\left(\mathbf{AA}^T-\gamma\mathbf{I}\right)=0\Rightarrow \gamma_1=12,\gamma_2=10\Rightarrow \mathbf{u}_1^T=\left[\tfrac{1}{\sqrt{2}},\tfrac{1}{\sqrt{2}}\right],\mathbf{u}_2^T=\left[\tfrac{1}{\sqrt{2}},\tfrac{-1}{\sqrt{2}}\right]$$

  - **V** can be computed as

$$\mathbf{A}=\begin{bmatrix}3 & 1 & 1\\ -1 & 3 & 1\end{bmatrix}\Rightarrow \mathbf{A}^T\mathbf{A}=\begin{bmatrix}3 & -1\\ 1 & 3\\ 1 & 1\end{bmatrix}\begin{bmatrix}3 & 1 & 1\\ -1 & 3 & 1\end{bmatrix}=\begin{bmatrix}10 & 0 & 2\\ 0 & 10 & 4\\ 2 & 4 & 2\end{bmatrix}$$

$$\det\left(\mathbf{A}^T\mathbf{A}-\gamma\mathbf{I}\right)=0\Rightarrow \gamma_1=12,\gamma_2=10,\gamma_3=0$$

$$\Rightarrow \mathbf{v}_1^T=\left[\tfrac{1}{\sqrt{6}},\tfrac{2}{\sqrt{6}},\tfrac{1}{\sqrt{6}}\right],\mathbf{v}_2^T=\left[\tfrac{2}{\sqrt{5}},\tfrac{-1}{\sqrt{5}},0\right],\mathbf{v}_3^T=\left[\tfrac{1}{\sqrt{30}},\tfrac{2}{\sqrt{30}},\tfrac{-5}{\sqrt{30}}\right]$$

# Example for SVD

- Taking $\lambda^2{}_1=12$ and $\lambda^2{}_2=10$, the singular value decomposition of **A** is

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

$$= \sqrt{12}\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 1/\sqrt{6}, 2/\sqrt{6}, 1/\sqrt{6} \end{bmatrix} + \sqrt{10}\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 2/\sqrt{5}, -1/\sqrt{5}, 0 \end{bmatrix}$$

- Thus the U, V and $\Lambda$ are computed by performing eigen decomposition of $\mathbf{AA}^T$ and $\mathbf{A}^T\mathbf{A}$
- Any matrix has a singular value decomposition but only symmetric, positive definite matrices have an eigen decomposition