Fourier Analysis

(I) Fourier Analysis

- It is always possible to analyze „complex“ periodic waveforms into a set of sinusoidal waveforms
- Any periodic waveform can be approximated by adding together a number of sinusoidal waveforms
- Fourier analysis tells us what particular set of sinusoids go together to make up a particular complex waveform

- The period is the duration of one cycle of an event and is the reciprocal of the frequency *f*. For example, if we count 40 events in two seconds, the frequency is

$$\frac{40}{2\ s} = \frac{20}{1\ s} = 20\ \frac{1}{s} = 20\ hertz$$

- period is

$$T = p = \frac{1}{20} s.$$

- The frequency f is the inverse of the period

$$f = \frac{1}{T} = \frac{1}{p}$$

- If something changes rapidly, then we say that it has a high frequency.
- If it does not change rapidly, i.e., it changes smoothly, we say that it has a low frequency.

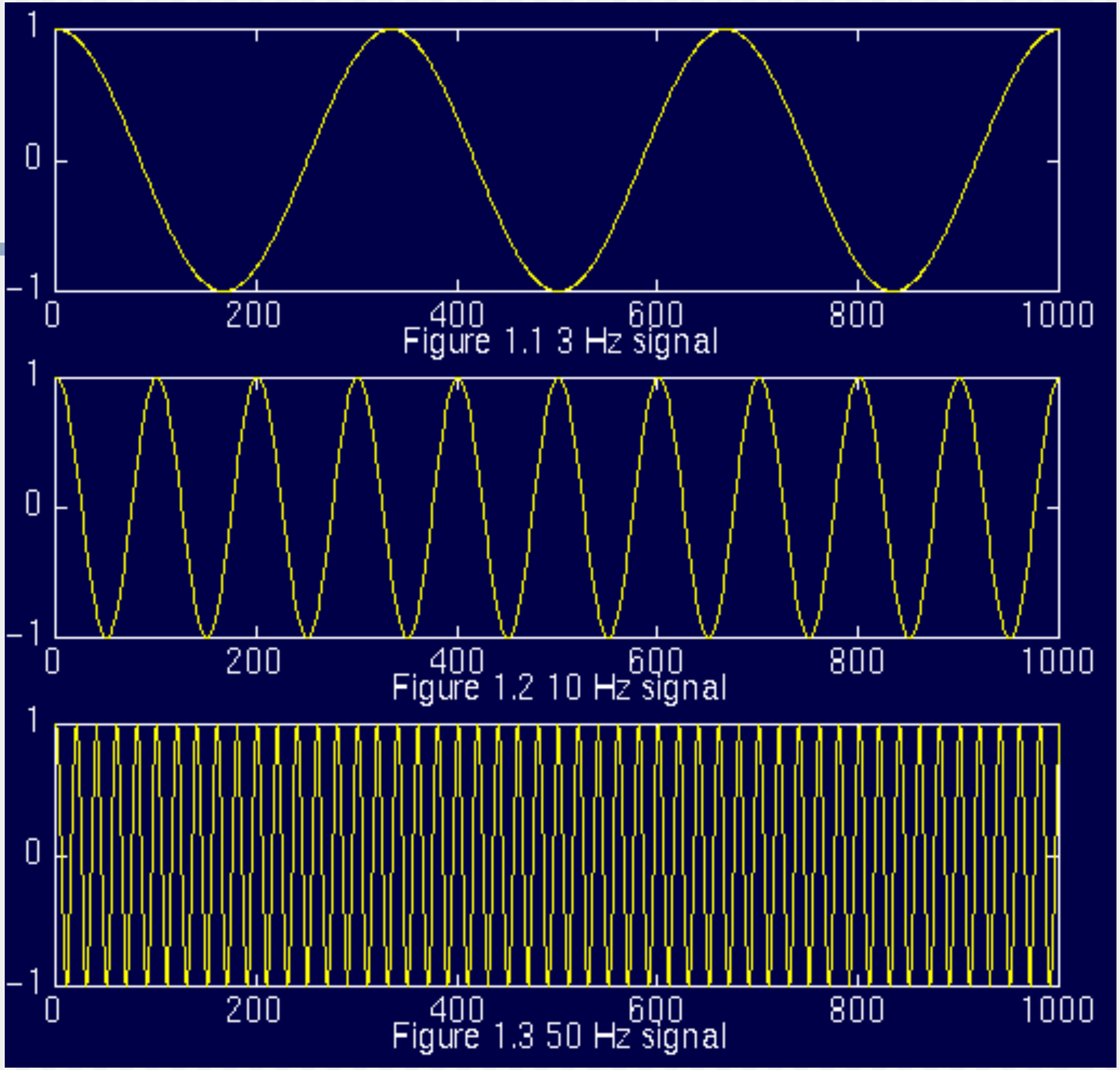


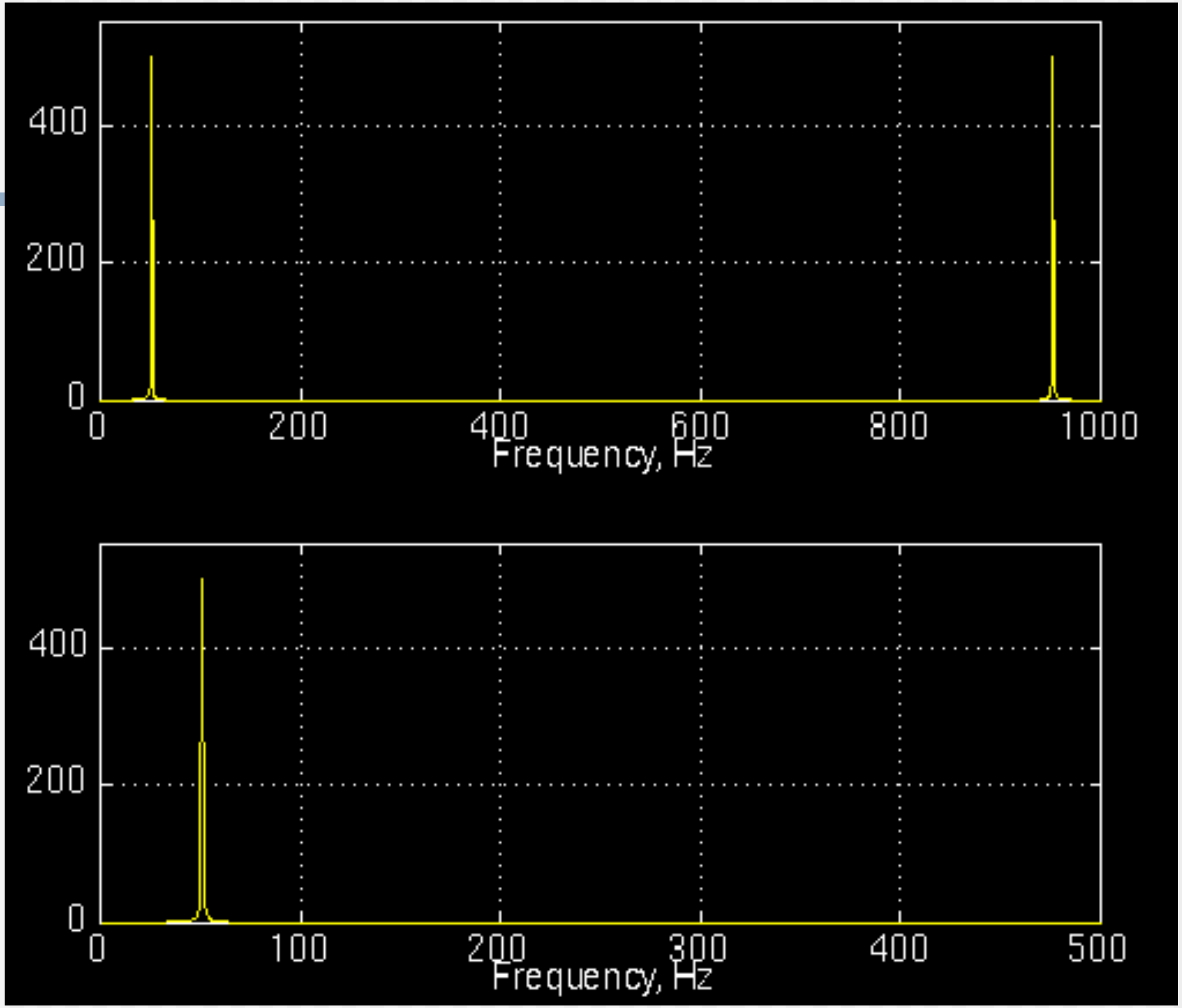
Figure 1.1 3 Hz signal

Figure 1.2 10 Hz signal

Figure 1.3 50 Hz signal

FOURIER TRANSFORM

- For example, if we take the FT of the electric current that we use in our houses,
- We will have one spike at 50 Hz
- Nothing elsewhere, since that signal has only 50 Hz frequency component


400
200
0
0
200
400
600
800
1000
Frequency, Hz
400
200
0
0
100
200
300
400
500
Frequency, Hz

- The frequency spectrum of a real valued signal is always symmetric. The top plot illustrates this point
- However, since the symmetric part is exactly a mirror image of the first part
- This symmetric second part is usually not shown

the Fourier transform of x(t)

$$X(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-2\pi i t f} dt$$

the inverse Fourier transform of X(f)

$$x(t) = \int_{-\infty}^{\infty} X(f) \cdot e^{2\pi i t f} df$$

- *t* stands for time, *f* stands for frequency, and *x* denotes the signal
- *x* denotes the signal in time domain and the *X* denotes the signal in frequency domain

- The signal *x(t),* is multiplied with an exponential term, at some certain frequency "*f*" , and then integrated over **ALL TIMES** !

Discrete Fourier Transform

- Operates on discrete complex-valued function
  - Given a function *a* :

$$a : [0,1,\ldots,N-1] \rightarrow C$$

  - The discrete Fourier transform produces a function *A* :

$$A : [0,1,\ldots,N-1] \rightarrow C$$

$$A(x) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a(k) \cdot e^{2\pi i \cdot \frac{kx}{N}}$$

- DFT can be seen as a linear transform talking the column vector ***a*** to a column vector ***A***

$$\begin{pmatrix} A(0) \\ A(1) \\ \vdots \\ A(N-1) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{N}} \cdot e^{2\pi i \frac{0\cdot 0}{N}} & \frac{1}{\sqrt{N}} \cdot e^{2\pi i \frac{0\cdot 1}{N}} & \cdots & \frac{1}{\sqrt{N}} \cdot e^{2\pi i \frac{0\cdot N-1}{N}} \\ \frac{1}{\sqrt{N}} \cdot e^{2\pi i \frac{0\cdot 1}{N}} & \frac{1}{\sqrt{N}} \cdot e^{2\pi i \frac{1\cdot 1}{N}} & \cdots & \frac{1}{\sqrt{N}} \cdot e^{2\pi i \frac{(N-1)\cdot 1}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sqrt{N}} \cdot e^{2\pi i \frac{0\cdot (N-1)}{N}} & \frac{1}{\sqrt{N}} \cdot e^{2\pi i \frac{1\cdot (N-1)}{N}} & \cdots & \frac{1}{\sqrt{N}} \cdot e^{2\pi i \frac{(N-1)\cdot (N-1)}{N}} \end{pmatrix} \cdot \begin{pmatrix} a(0) \\ a(1) \\ \vdots \\ a(N-1) \end{pmatrix}$$

- Simplification

$$\begin{pmatrix} A(0) \\ A(1) \\ \vdots \\ A(N-1) \end{pmatrix} = \frac{1}{\sqrt{N}} \cdot \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{2\pi i \frac{1\cdot 1}{N}} & \cdots & e^{2\pi i \frac{(N-1)\cdot 1}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{2\pi i \frac{1\cdot (N-1)}{N}} & \cdots & e^{2\pi i \frac{(N-1)\cdot (N-1)}{N}} \end{pmatrix} \cdot \begin{pmatrix} a(0) \\ a(1) \\ \vdots \\ a(N-1) \end{pmatrix}$$

- Example, N=4

$$\begin{pmatrix} A(0) \\ A(1) \\ \vdots \\ A(N-1) \end{pmatrix} = \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix} \cdot \begin{pmatrix} a(0) \\ a(1) \\ \vdots \\ a(N-1) \end{pmatrix}$$

- Let $a:[0,1,\ldots,N-1]\rightarrow C$ be a **periodic function**

```
In[30]:= TrigToExp[Cos[u] + i * Sin[u]]

         e^(i u)

In[49]:= ExpToTrig[Exp[-2 * π * i * u * x / N]]

Out[49]= Cos[2 π u (pi : 0.1 : pi) / N] + i Sin[2 π u (pi : 0.1 : pi) / N]

In[50]:=

         ExpToTrig[Exp[2 * π * i * u * x / N]]

Out[50]= Cos[2 π u (pi : 0.1 : pi) / N] - i Sin[2 π u (pi : 0.1 : pi) / N]
```

$$a(x) = e^{-2\pi i \frac{ux}{N}}$$

$$a(x) = \cos(2\pi \frac{ux}{N}) + i \cdot \sin(2\pi \frac{ux}{N})$$

$$e^{iu} = \cos(u) + i \cdot \sin(u)$$

- A complex root of unity is a complex number $\omega^N = 1$
- There are exactly *n*th roots of unity:

$$e^{2\pi i \frac{k}{N}} \qquad for \quad k = 0,1,\ldots,N-1$$

- We define $\omega_N = e^{2\pi i \frac{1}{N}}$

$$e^{iu} = \cos(u) + i \cdot \sin(u)$$

$$\begin{pmatrix} A(0) \\ A(1) \\ \vdots \\ A(N-1) \end{pmatrix} = \frac{1}{\sqrt{N}} \cdot \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{2\pi i \frac{1\cdot 1}{N}} & \cdots & e^{2\pi i \frac{(N-1)\cdot 1}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{2\pi i \frac{1\cdot (N-1)}{N}} & \cdots & e^{2\pi i \frac{(N-1)\cdot (N-1)}{N}} \end{pmatrix} \cdot \begin{pmatrix} a(0) \\ a(1) \\ \vdots \\ a(N-1) \end{pmatrix}$$

$$\begin{pmatrix} A(0) \\ A(1) \\ \vdots \\ A(N-1) \end{pmatrix} = \frac{1}{\sqrt{N}} \cdot \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega_N^{1\cdot 1} & \cdots & \omega_N^{(N-1)\cdot 1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_N^{1(N-1)} & \cdots & \omega_N^{(N-1)(N-1)} \end{pmatrix} \cdot \begin{pmatrix} a(0) \\ a(1) \\ \vdots \\ a(N-1) \end{pmatrix}$$

Remarks

$$\begin{pmatrix} A(0) \\ A(1) \\ \vdots \\ A(N-1) \end{pmatrix} = \frac{1}{\sqrt{N}} \cdot \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega_N^{1\cdot 1} & \cdots & \omega_N^{(N-1)\cdot 1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_N^{1(N-1)} & \cdots & \omega_N^{(N-1)(N-1)} \end{pmatrix} \cdot \begin{pmatrix} a(0) \\ a(1) \\ \vdots \\ a(N-1) \end{pmatrix}$$

$$\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-1} \end{pmatrix} = \frac{1}{\sqrt{N}} \cdot \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega_N^{1\cdot 1} & \cdots & \omega_N^{(N-1)\cdot 1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_N^{1(N-1)} & \cdots & \omega_N^{(N-1)(N-1)} \end{pmatrix} \cdot \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{pmatrix}$$

- Input vector of complex numbers of length *N*

$$x_0, x_2, \ldots\ldots x_{N-1}$$

$$y_0, y_2, \ldots\ldots y_{N-1}$$

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N} kj} \qquad k \in \{0,1,\ldots,N-1\}$$

$$inverse:$$

$$x_k = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} y_k e^{\frac{2\pi i}{N} kj} \qquad j \in \{0,1,\ldots,N-1\}$$

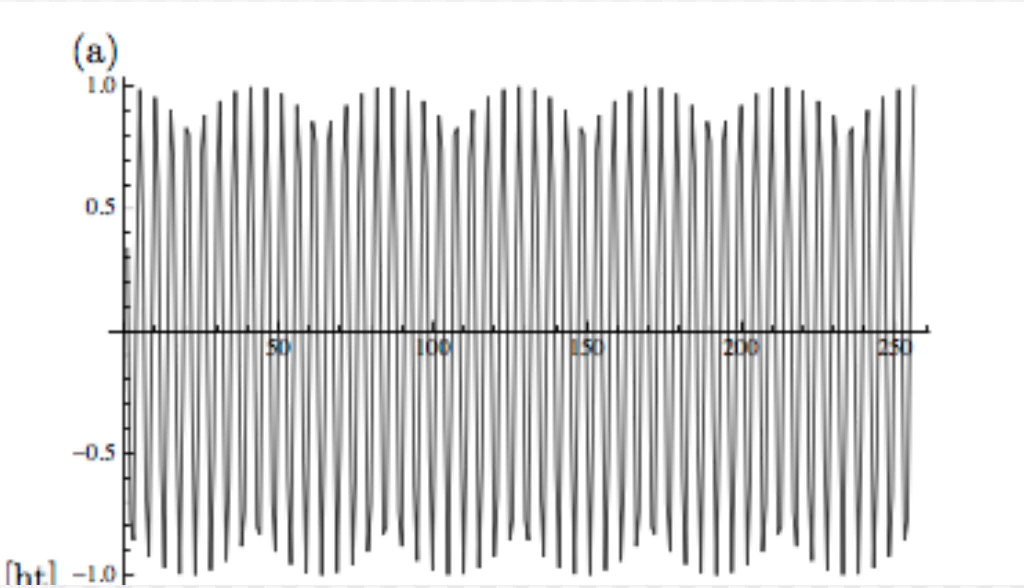
Noise reduction

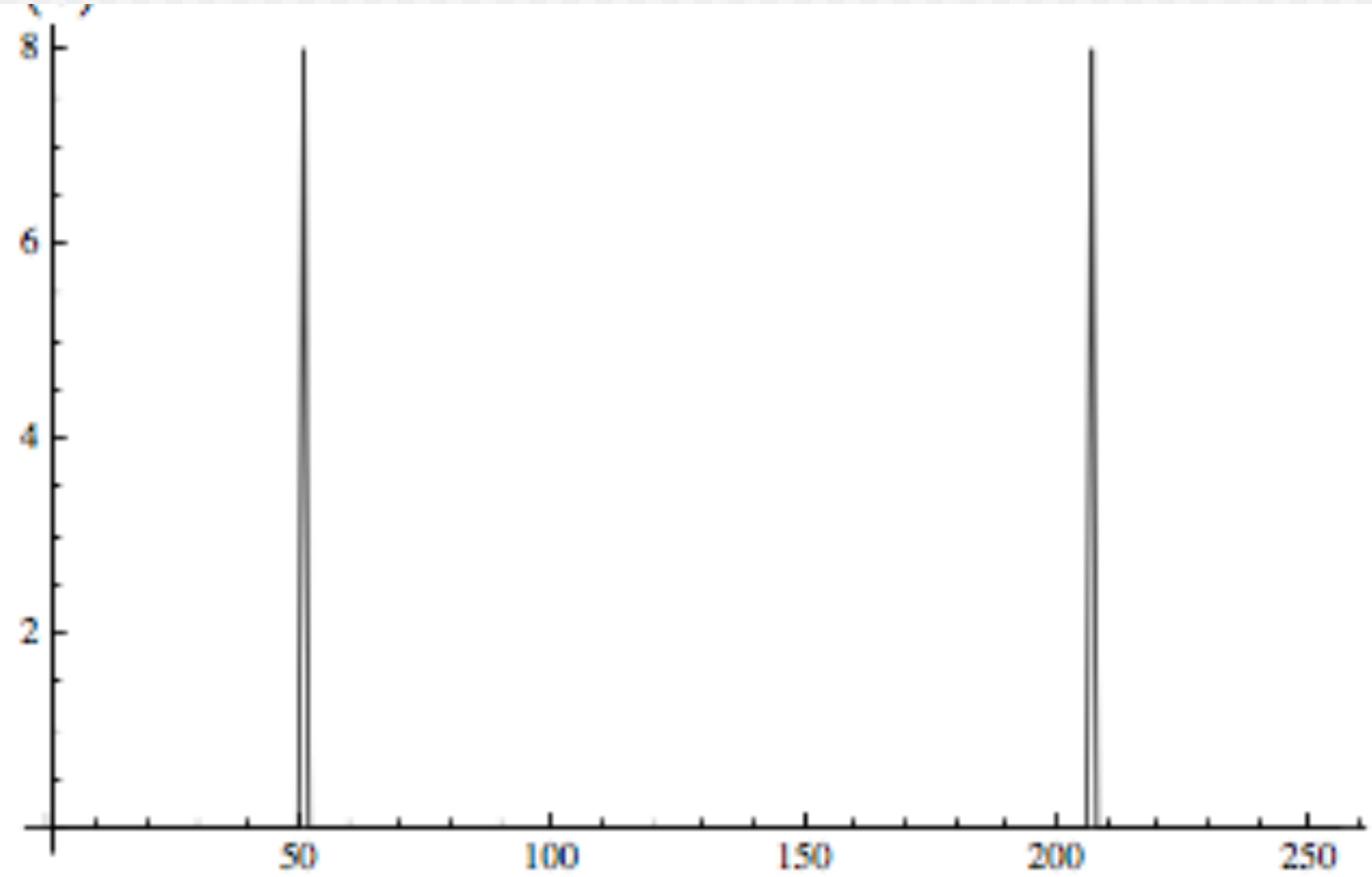
- It is difficult to identify the frequency components by looking at the original signal
- Converting to the frequency domain
- If dimension reduction, store only a fraction of frequencies (with high amplitude)
- If noise reduction
  - (remove high frequencies, fast change, smoothing)
  - (remove low frequencies, slow change, remove global trends)
  - Inverse discrete Fourier transform

Example

We generate a list with $256 = 2^8$ elements containing a periodic signal α_t

$$\alpha_t = cos\left(\frac{50 \cdot t \cdot 2 \cdot \cdot \pi}{256}\right),$$



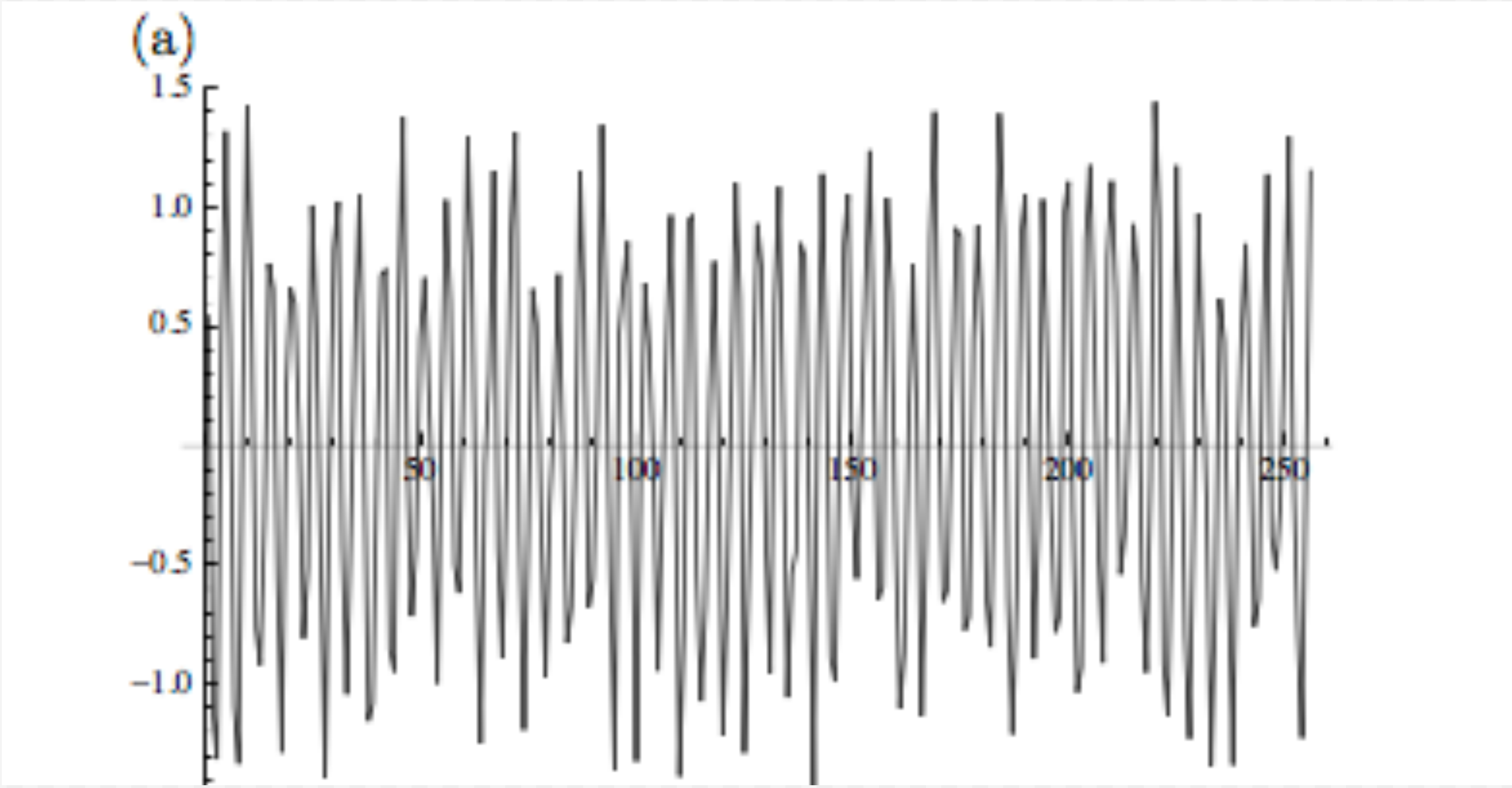

8
6
4
2
50
100
150
200
250

- The discrete Fourier transform ω_f of the real valued signal αt is symmetric. It shows a strong peak at *50 + 1* and a symmetric peak at *256 – 50 + 1* representing the frequency component of the signal

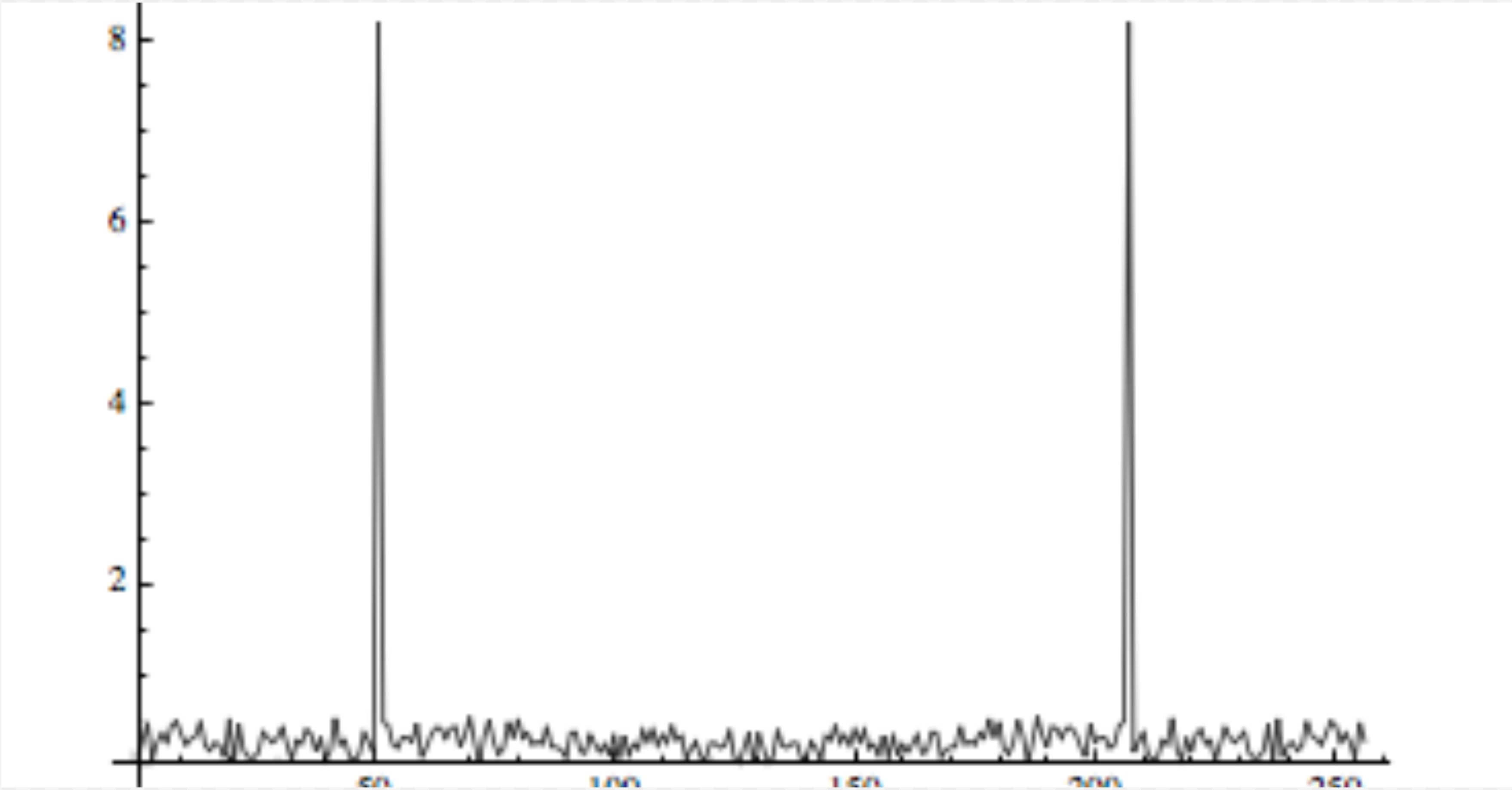
- We add to the periodic signal αt Gaussian random noise from the interval [*–0.5, 0.5*].

$$\alpha_t^* = cos\left(\frac{50 \cdot t \cdot 2 \cdot \cdot \pi}{256}\right) + noise.$$

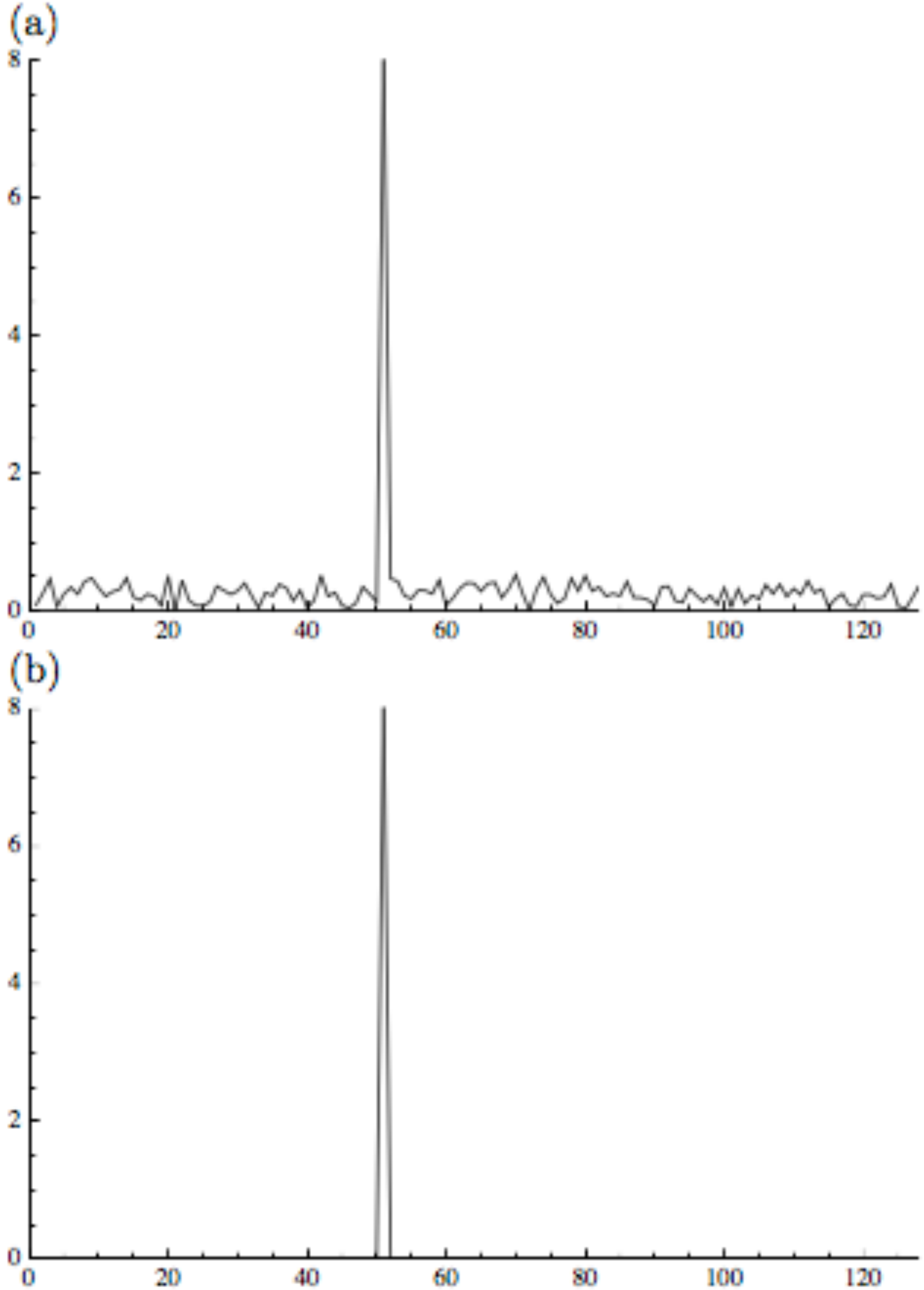
- The represented data looks random



The frequency component



- A filter that reduces Gaussian noise based on DFT removes frequencies with low amplitude of ω_f and performs the inverse discrete Fourier transform
- For dimension reduction of the signal, only a fraction of frequencies with high amplitude are represented.


(a)
8
6
4
2
0
0
20
40
60
80
100
120
(b)
8
6
4
2
0
0
20
40
60
80
100
120

Feature space

- Sample $\left\{\vec{x}^{(1)},\vec{x}^{(2)},\ldots,\vec{x}^{(k)},\ldots,\vec{x}^{(n)}\right\}$

$$\vec{x} = \begin{cases} x_1 \\ x_2 \\ .. \\ .. \\ x_d \end{cases} \in \Re^d \qquad \left\|\vec{x} - \vec{y}\right\| = \sqrt{\sum_{i=1}^{d}(x_i - y_i)^2}$$

Scaling

- A well-known scaling method consists of performing some scaling operations
  - subtracting the mean and dividing the standard deviation

$$y_i = \frac{(x_i - m_i)}{s_i}$$

- m_i sample mean
- s_i sample standard deviation

- According to the scaled metric the **scaled** feature vector is expressed as

$$\| \vec{y} \|_s = \sqrt{\sum_{i=1}^{n} \frac{(x_i - m_i)^2}{s_i^2}}$$

- shrinking **large** variance values
  - $s_i > 1$
- stretching **low** variance values
  - $s_i < 1$

- Fails to preserve distances when *general linear transformation* is applied!

Covariance

- Covariance
  - Measuring the tendency two features x_i and x_j varying in the **same direction**
  - The covariance between features x_i and x_j is estimated for *n* patterns

$$c_{ij} = \frac{\sum_{k=1}^{n} \left(x_i^{(k)} - m_i\right)\left(x_j^{(k)} - m_j\right)}{n-1}$$

$$C = \begin{bmatrix} c_{11} & c_{12} & .. & c_{1d} \\ c_{21} & c_{22} & .. & c_{2d} \\ .. & .. & .. & .. \\ c_{d1} & c_{d2} & .. & c_{dd} \end{bmatrix}$$

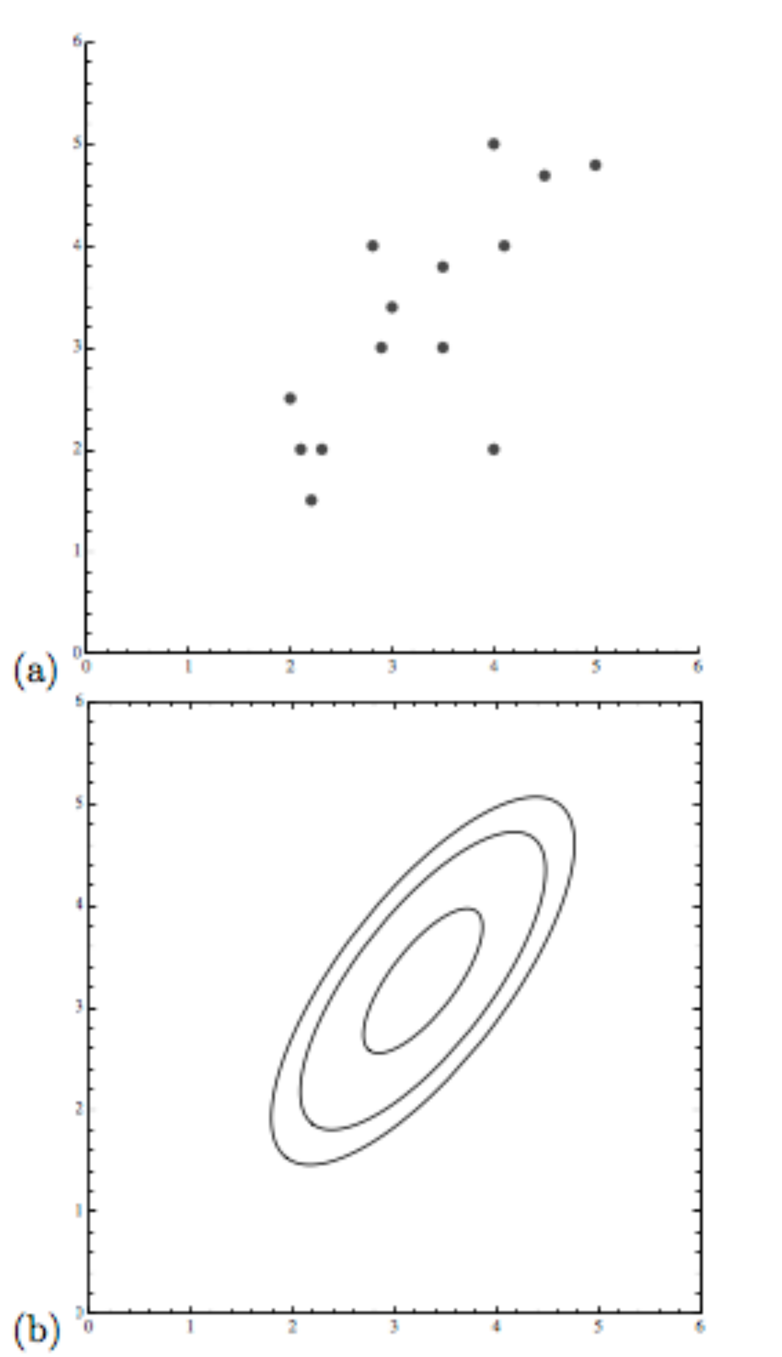
Correlation

- Covariances are symmetric $c_{ij}=c_{ji}$
- Covariance is related to correlation

$$r_{ij} = \frac{\sum_{k=1}^{n}\left(x_i^{(k)} - m_i\right)\left(x_j^{(k)} - m_j\right)}{(n-1)s_i s_j} = \frac{c_{ij}}{s_i s_j} \in [-1,1]$$

$$\Sigma = \{(2.1, 2), (2.3, 2), (2.9, 3), (4.1, 4), (5, 4.8), (2, 2.5), (2.2, 1.5), (4, 5), (4, 2), (2.8, 4), (3, 3.4), (3.5, 3.8), (4.5, 4.7), (3.5, 3)\}$$

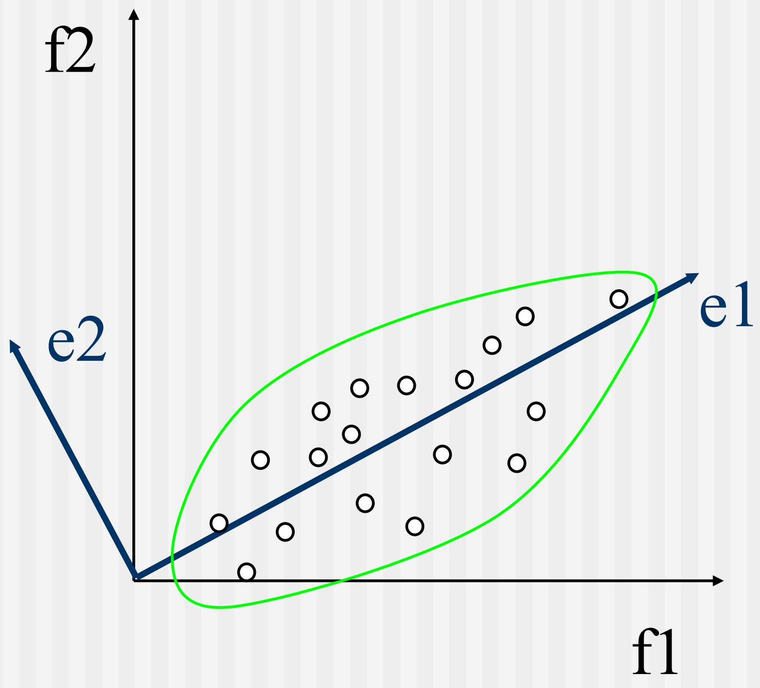
$$c_{ij} = \frac{\sum_{k=1}^{n} (x_{k,i} - \overline{x_i}) \cdot (y_{k,j} - \overline{y_j})}{n-1}$$



$$C = \begin{pmatrix} 0.912582 & 0.82456 \\ 0.82456 & 1.34247 \end{pmatrix}$$

Principal Component Analysis

- Intuition: find the axis that shows the greatest variation, and project all points into this axis



(II) Karhunen-Loève Transformation

- Covariance matrix *C* of (a $d \times d$ matrix)
  - Symmetric and positive definite

$$U^T C U = \Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_d)$$

$$(\lambda I - C)u = 0$$

  - There are *d* eigenvalues and eigenvectors

$$C\vec{u}_i = \lambda_i \vec{u}_i$$

  - is the λ_i ith eigenvalue of *C* and $\boldsymbol{u}_i$ the ith column of ***U***, the ith eigenvectors

- Eigenvectors are always **orthogonal**
- *U* is an orthonormal matrix $UU^T = U^T U = I$
- *U* defines the K-L transformation
- The transformed features by the K-L transformation are given by

$$\vec{y} = U^T \vec{x} \qquad \text{(linear Transformation)}$$

- K-L transformation rotates the feature space into alignment with **uncorrelated** features

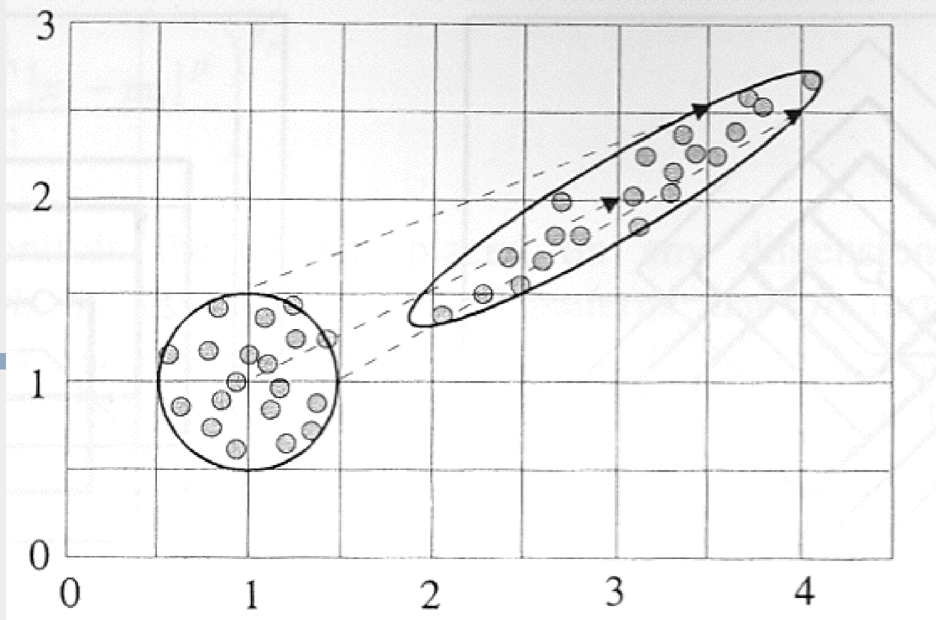
Example

$$C = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \qquad |\lambda I - C| = 0 \qquad \lambda^2 - 3\lambda + 1 = 0$$

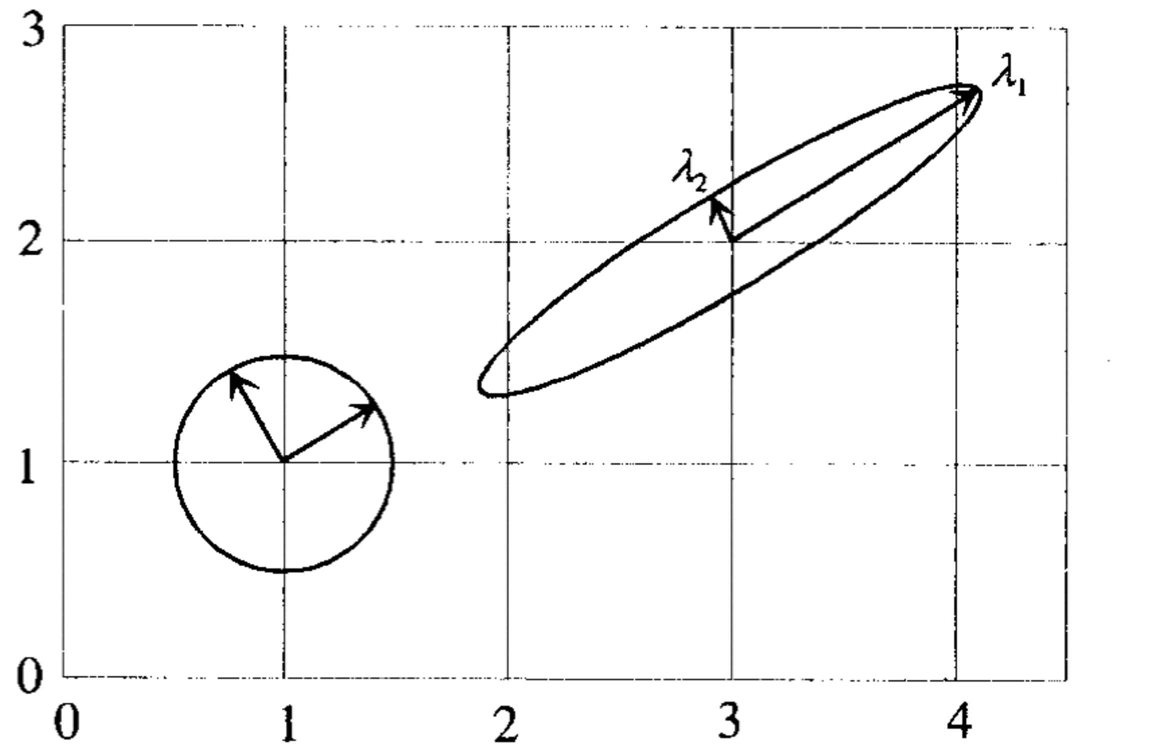
- λ_1=2.618 λ_2=0.382

$$\begin{bmatrix} -0.618 & -1 \\ -1 & 1.618 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

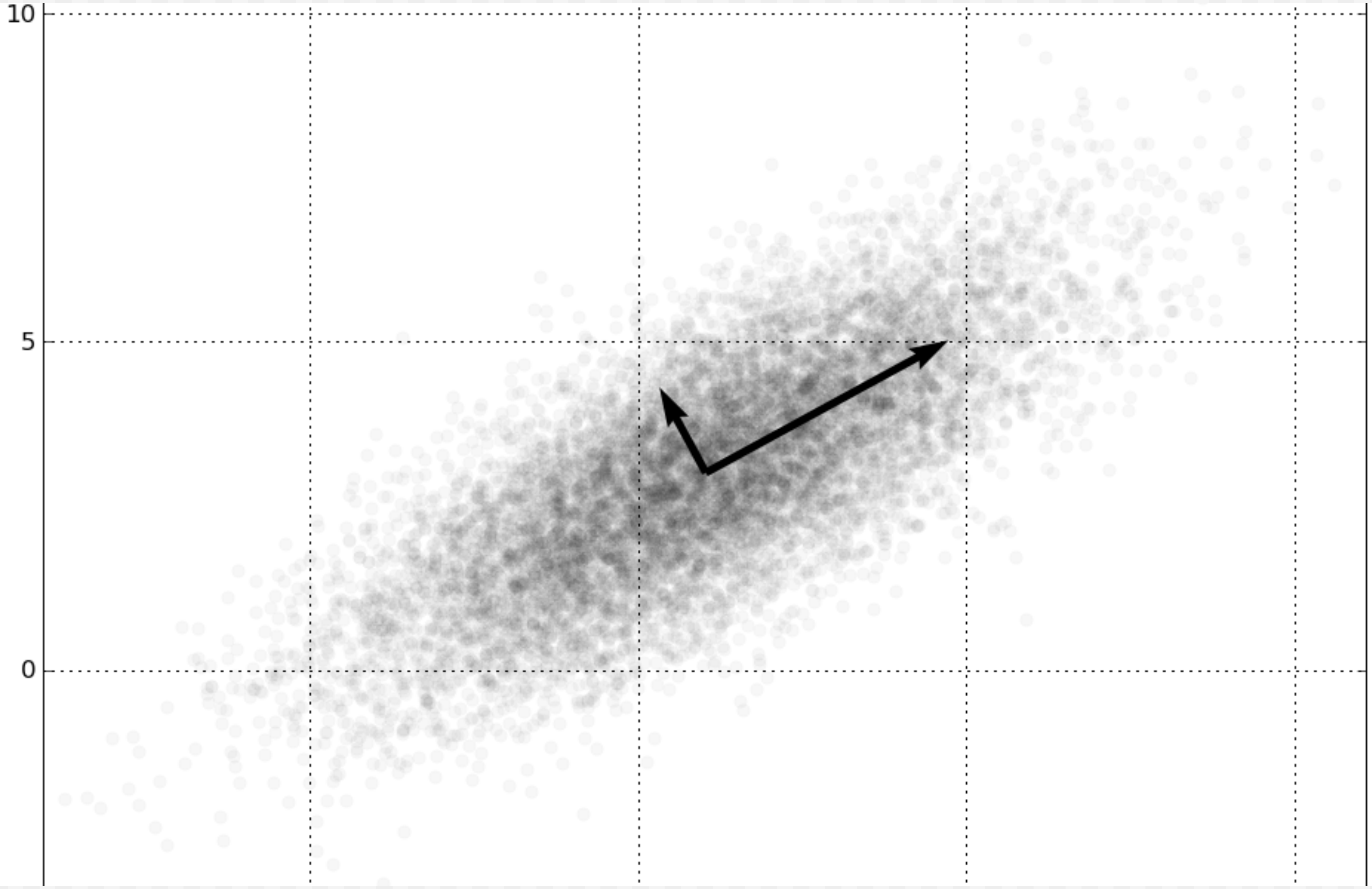
- $u^{(1)}$=[1 0.618] $u^{(2)}$=[-1 1.618]
- **U**=[$u^{(1)}$,$u^{(2)}$]

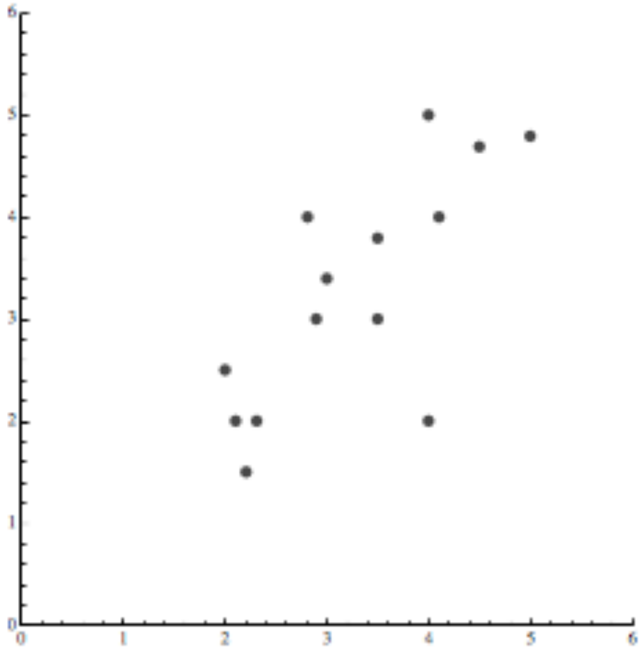


3
2
1
0
0
1
2
3
4



3
2
1
0
0
1
2
3
4
λ_1
λ_2


10
5
0



(a)

The squares of the eigenvalues represent the variances along the eigenvectors. The eigenvalues corresponding to the covariance matrix of the data set Σ are

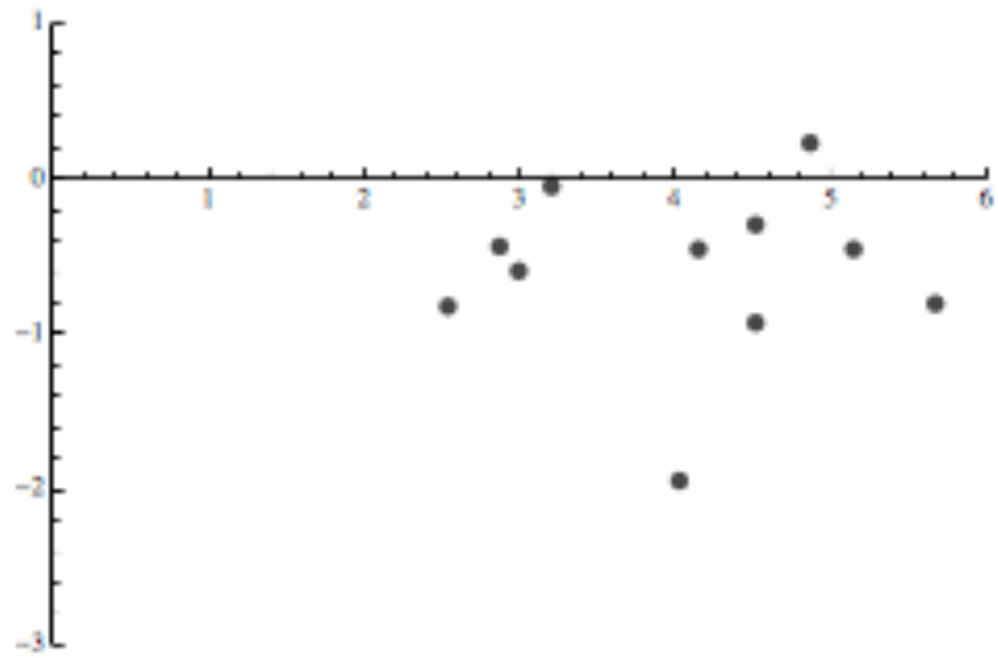
$$\lambda_1 = 1.97964, \quad \lambda_2 = 0.275412$$

and the corresponding normalized eigenvectors are

$$\mathbf{u}_1 = \begin{pmatrix} 0.611454 \\ 0.79128 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -0.79128 \\ 0.611454 \end{pmatrix}.$$

$$U = \begin{pmatrix} 0.611454 & -0.79128 \\ 0.79128 & 0.611454 \end{pmatrix}.$$

$$\mathbf{y} = U^{\top} \cdot \mathbf{x} = \begin{pmatrix} 0.611454 & 0.79128 \\ -0.79128 & 0.611454 \end{pmatrix} \cdot \mathbf{x},$$



PCA (Principal Components Analysis)

- New features ***y*** are uncorrelated with the covariance Matrix
- Each eigenvector $\boldsymbol{u}_i$ is **associated** with some variance associated by λ_i
- Uncorrelated features with **higher variance** (represented by λ_i) contain **more information**
- Idea:
  - Retain only the significant eigenvectors $\boldsymbol{u}_i$
  - Example
    - **U**=[$u^{(1)}$,$u^{(2)}$] λ_1=2.618 λ_2=0.382
    - **U***=[$u^{(1)}$]
    -

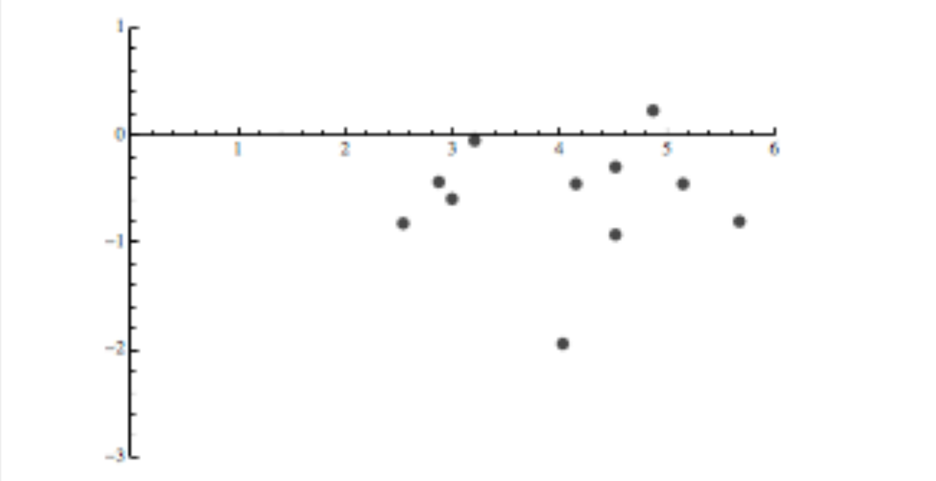
$$\vec{y} = U^{*T}\vec{x}$$

Dimension Reduction

- How many eigenvectors (and corresponding eigenvector) to retain

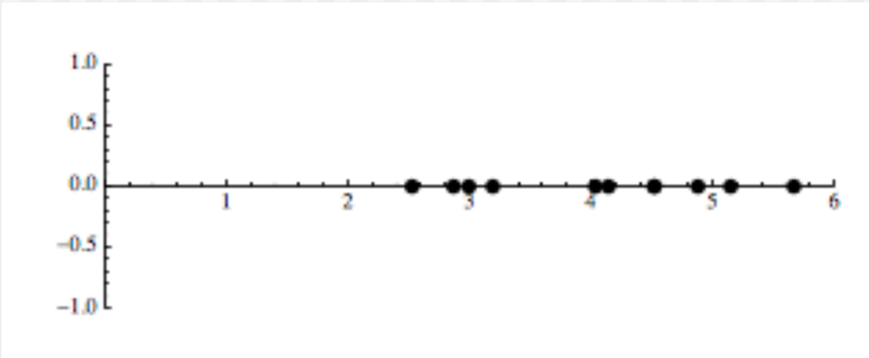
- Kaiser criterion
  - Discards eigenvectors whose eigenvalues are below 1

$$\mathbf{y} = U^\top \cdot \mathbf{x} = \begin{pmatrix} 0.611454 & 0.79128 \\ -0.79128 & 0.611454 \end{pmatrix} \cdot \mathbf{x},$$



$$\mathbf{z} = W^\top \cdot \mathbf{x}$$

$$\mathbf{z} = W^\top \cdot \mathbf{x} = \begin{pmatrix} 0.611454 & 0.79128 \end{pmatrix} \cdot \mathbf{x}$$



Problems

- Principal components are linear transformation of the original features
- It is difficult to attach any semantic meaning to principal components
- For new data which is added to the dataset, the PCA has to be recomputed

- Suppose we have a covariance matrix:

$$C = \begin{pmatrix} 3 & 1 \\ 1 & 21 \end{pmatrix}$$

- What is the matrix of the K-L transformation?

- First we have to compute the eiganvalues
- The system has to become linear depentable (singular)

$$\left|\lambda I - C\right| = 0$$

- The determinant has to become zero

$$\lambda^2 - 24\lambda + 62 = 0$$

- Solving it we get

- λ_1=2,94461
- λ_2=21,05538

Now, lets compute the two eigenvectors….

- To do it you have to solve two singular, dependent systems
- for the first eigenvalue λ_1=2,94461

$$(\lambda_1 I - C)\vec{u}_1 = 0$$

- Or if you prefer more ...

$$\lambda_1 \vec{u}_1 = C\vec{u}_1$$

$$C\vec{u}_1 = \lambda_1 \vec{u}_1$$

Now, lets compute the two eigenvectors….

- To do it you have to solve two singular, dependent systems
- And for the second eigenvalue λ_2=21,05538

$$(\lambda_2 I - C)\vec{u}_2 = 0$$

- Or if you prefer more ...

$$\lambda_2 \vec{u}_2 = C\vec{u}_2$$

$$C\vec{u}_2 = \lambda_2 \vec{u}_2$$

For λ_1=2,94461
$\mathbf{u}_1$=(u_1,u_2)

$$\left[\begin{bmatrix} 2{,}94461 & 0 \\ 0 & 2{,}94461 \end{bmatrix} - \begin{bmatrix} 3 & 1 \\ 1 & 21 \end{bmatrix}\right]\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

- We have to find a nontrivial solution!
  - Trivial solution is u=[0,0]...

$$\begin{bmatrix} -0{,}05539 & -1 \\ -1 & -18{,}055 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

- Because the system is linear dependable, the left column is multiple value of the right column
- There are infinity many solution!!!!

- We have only to determine the direction of the eigenvectors $\mathbf{u_1}$ and $\mathbf{u_2}$
- But be careful, the ***normalized*** vectors have to be orthogonal to each other
- $<\mathbf{u_1},\mathbf{u_2}>=0$

- Let be u_1=1 then we have to determine u_2

$$\begin{bmatrix} -0,05539 & -1 \\ -1 & -18,055 \end{bmatrix} \begin{bmatrix} 1 \\ u_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} -0,05539 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 18,055 \end{bmatrix} u_2$$

- $\mathbf{u_1}=[u_1,u_2]=[1,-0,05539]$

For λ_2= 21,05538
$\mathbf{u}_1$=(u_1,u_2)

$$\left[\begin{bmatrix} 21{,}05538 & 0 \\ 0 & 21{,}05538 \end{bmatrix} - \begin{bmatrix} 3 & 1 \\ 1 & 21 \end{bmatrix}\right]\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

- We have to find a nontrivial solution!
  - Trivial solution is u=[0,0]... Déjà vu?

$$\begin{bmatrix} 18{,}055 & -1 \\ -1 & 0{,}05538 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

- Let be u_1=1 then we have to determine u_2

$$\begin{bmatrix} 18{,}055 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0{,}05538 \end{bmatrix} u_2$$

- $\mathbf{u_2}=[u_1,u_2]=[1,\ 18{,}055]$

- $\mathbf{u_1}$=[u_1,u_2]=[1,-0,05539]
  - for λ_1=2,94461
- $\mathbf{u_2}$=[u_1,u_2]=[1,18,055]
  - λ_2=21,05538
- Orthogonal? Yes $<\mathbf{u_1},\mathbf{u_2}>=0$

- Which of the two eigenvectors is more significant?
- $\mathbf{u_2}$, because λ_1=2,94461 < λ_2=21,05538
- Remember, we have to normalize the Eigenvectors