Sammon Mapping

- original data $X = \{x_1, \ldots, x_n\} \subset {\rm I\!R}^p$
- distance matrix $D^x_{ij} = \|x_i - x_j\|$
- transformed data $Y = \{y_1, \ldots, y_n\} \subset {\rm I\!R}^q$
- distance matrix $D^y_{ij} = \|y_i - y_j\|$

- wanted:

  (nonlinear) mapping $X \to Y$,

  so that $D^x_{ij} \approx D^y_{ij} \quad \forall i, j = 1, \ldots, n$

- or:

  given relational data D^x,

  wanted corresponding real valued object data Y,

  so that $D^x_{ij} \approx D^y_{ij} \quad \forall i, j = 1, \ldots, n$

Sammon Mapping

- error measures:

$$E_1 = \frac{1}{\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(D_{ij}^x\right)^2} \sum_{i=1}^{n}\sum_{j=i+1}^{n} \left(D_{ij}^y - D_{ij}^x\right)^2$$

$$E_2 = \sum_{i=1}^{n}\sum_{j=i+1}^{n} \left(\frac{D_{ij}^y - D_{ij}^x}{D_{ij}^x}\right)^2$$

$$E_3 = \frac{1}{\sum_{i=1}^{n}\sum_{j=i+1}^{n} D_{ij}^x} \sum_{i=1}^{n}\sum_{j=i+1}^{n} \frac{\left(D_{ij}^y - D_{ij}^x\right)^2}{D_{ij}^x}$$

- E_1 minimizes the global absolute error
- E_2 minimizes the relative local error
- E_3 is a compromise between E_1 and E_2 (best choice)

Numerical Optimization

- gradient descent
  1. input X, α, initialize Y
  2. repeat

$$y_k := y_k - \alpha \frac{\partial E}{\partial y_k}, \quad k = 1, \ldots, n$$

     until termination
  3. output Y
- Newton's method
  1. input X, initialize Y
  2. repeat

$$y_k := y_k - \left(\frac{\partial E}{\partial y_k} \right) / \left(\frac{\partial^2 E}{\partial y_k^2} \right), \quad k = 1, \ldots, n$$

     until termination
  3. output Y

Derivatives of the Sammon Function

- preliminary note

$$\frac{\partial D_{ij}^y}{\partial y_k} = \frac{\partial}{\partial y_k} \|y_i - y_j\| = \begin{cases} \frac{y_k - y_j}{D_{kj}^y} & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$$

- first derivative

$$\frac{\partial E_3}{\partial y_k} = \frac{2}{\sum\limits_{i=1}^{n} \sum\limits_{j=i+1}^{n} d_{ij}^x} \sum_{\substack{j=1 \\ j \neq k}}^{n} \left(\frac{1}{d_{kj}^x} - \frac{1}{d_{kj}^y} \right) \left(y_k - y_j \right)$$

- second derivative

$$\frac{\partial^2 E_3}{\partial y_k^2} = \frac{2}{\sum\limits_{i=1}^{n} \sum\limits_{j=i+1}^{n} d_{ij}^x} \sum_{\substack{j=1 \\ j \neq k}}^{n} \left(\frac{1}{d_{kj}^x} - \frac{1}{d_{kj}^y} - \frac{(y_k - y_j)^2}{(d_{kj}^y)^3} \right)$$

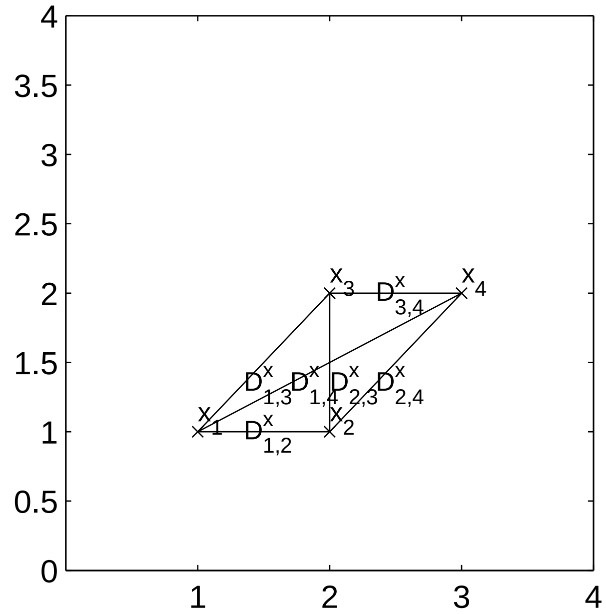
Example Sammon Mapping

$$X = \{(1,1),(2,1),(2,2),(3,2)\}$$

$$D^x = \begin{pmatrix} 0 & 1 & \sqrt{2} & \sqrt{5} \\ 1 & 0 & 1 & \sqrt{2} \\ \sqrt{2} & 1 & 0 & 1 \\ \sqrt{5} & \sqrt{2} & 1 & 0 \end{pmatrix}$$

$$Y = \{1,2,3,4\}$$

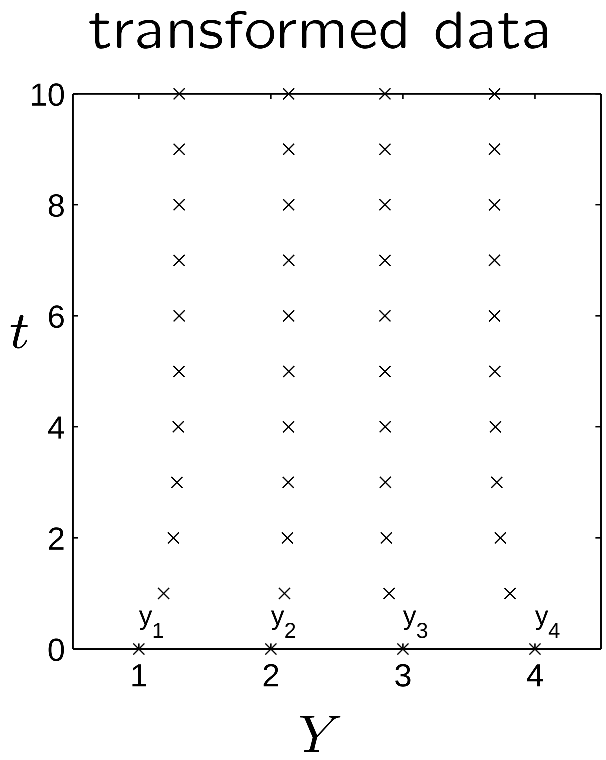
$$D^y = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{pmatrix}$$

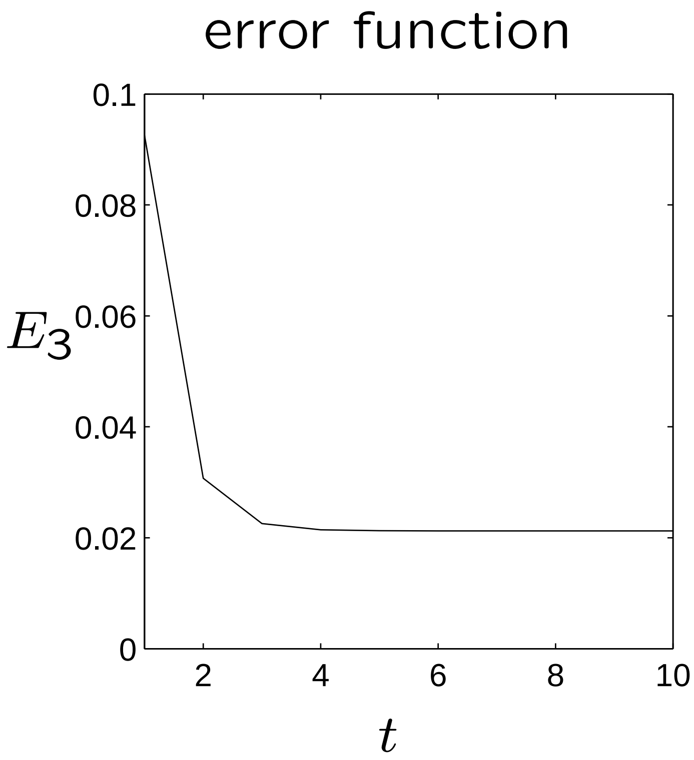


Example Sammon Mapping

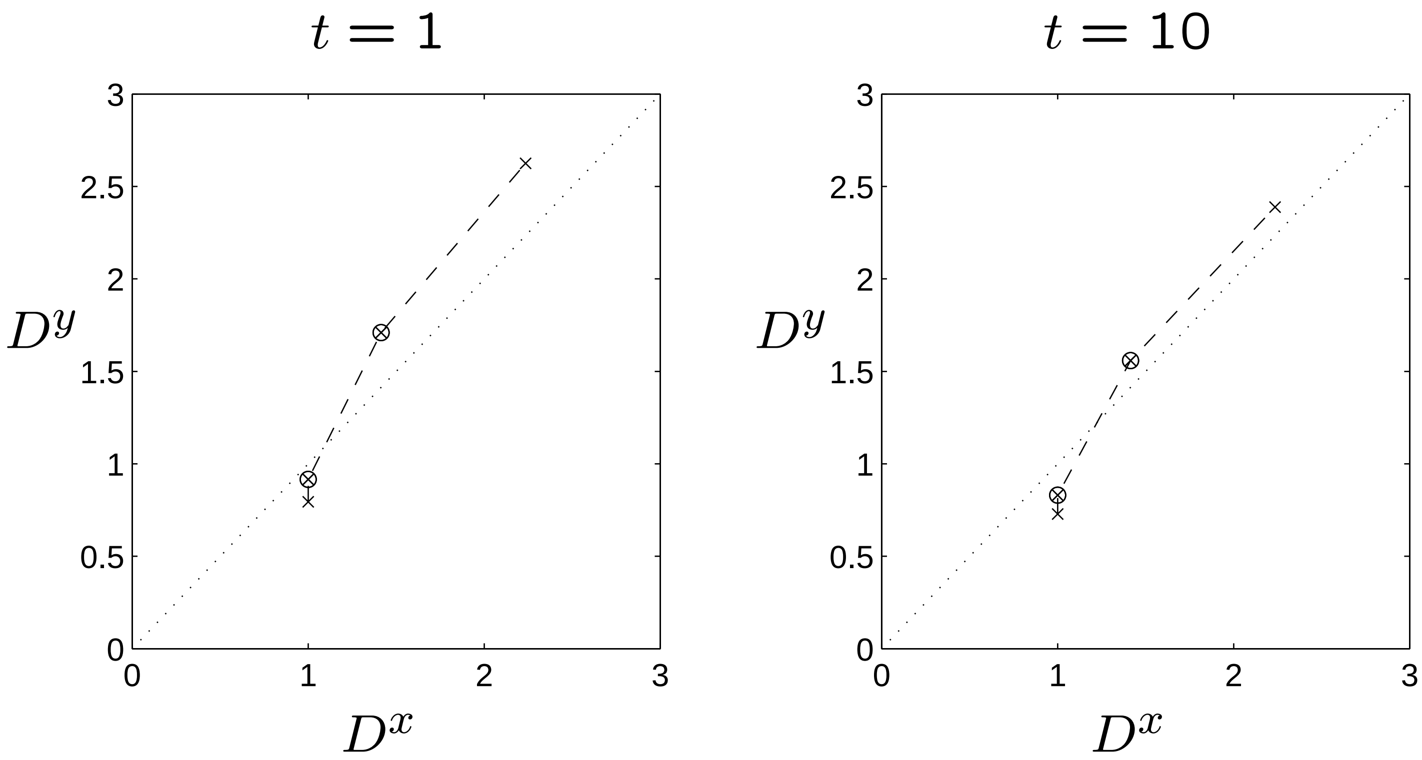
$$
\begin{aligned}
E_3 &= \frac{1}{3 \cdot 1 + 2 \cdot \sqrt{2} + \sqrt{5}} \cdot \left(2 \cdot \frac{\left(2 - \sqrt{2}\right)^2}{\sqrt{2}} + \frac{\left(3 - \sqrt{5}\right)^2}{\sqrt{5}} \right) = 0.0925 \\
\frac{\partial E_3}{\partial y_1} &= \frac{2}{3 \cdot 1 + 2 \cdot \sqrt{2} + \sqrt{5}} \cdot \left(-\frac{2 - \sqrt{2}}{\sqrt{2}} - \frac{3 - \sqrt{5}}{\sqrt{5}} \right) = -0.1875 \\
\frac{\partial E_3}{\partial y_2} &= \frac{2}{3 \cdot 1 + 2 \cdot \sqrt{2} + \sqrt{5}} \cdot \left(-\frac{2 - \sqrt{2}}{\sqrt{2}} \right) = -0.1027 \\
\frac{\partial E_3}{\partial y_3} &= \frac{2}{3 \cdot 1 + 2 \cdot \sqrt{2} + \sqrt{5}} \cdot \left(\frac{2 - \sqrt{2}}{\sqrt{2}} \right) = 0.1027 \\
\frac{\partial E_3}{\partial y_4} &= \frac{2}{3 \cdot 1 + 2 \cdot \sqrt{2} + \sqrt{5}} \cdot \left(\frac{3 - \sqrt{5}}{\sqrt{5}} + \frac{2 - \sqrt{2}}{\sqrt{2}} \right) = 0.1875 \\
y &\leftarrow (1.1875,\ 2.1027,\ 2.8973,\ 3.8125)
\end{aligned}
$$

Example Sammon Mapping





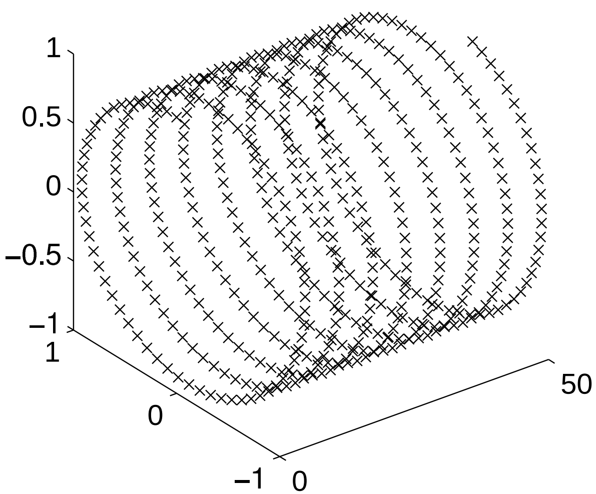
Shepard Diagrams



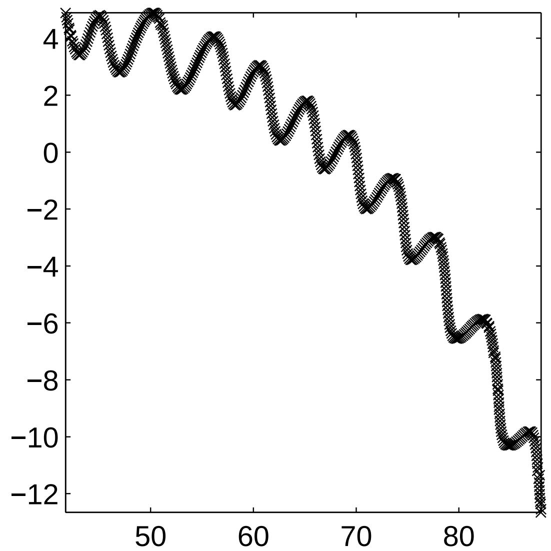
Example Sammon Mapping: Helix

$$X = \{(t, \sin t, \cos t)^T \mid t \in \{0, 0.1, 0.2, \ldots, 50\}\}$$

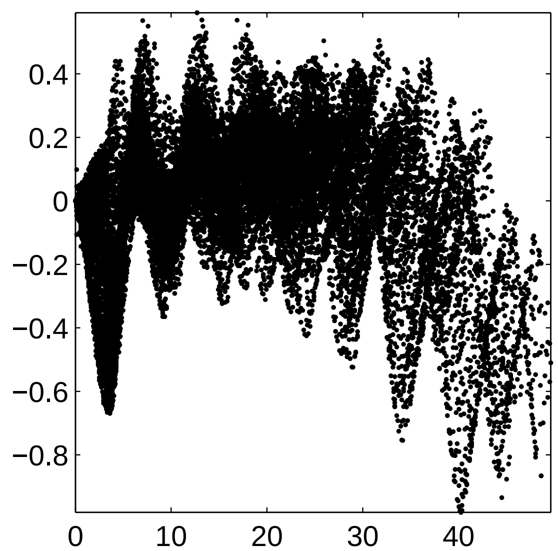
original data



Sammon projection



projection error

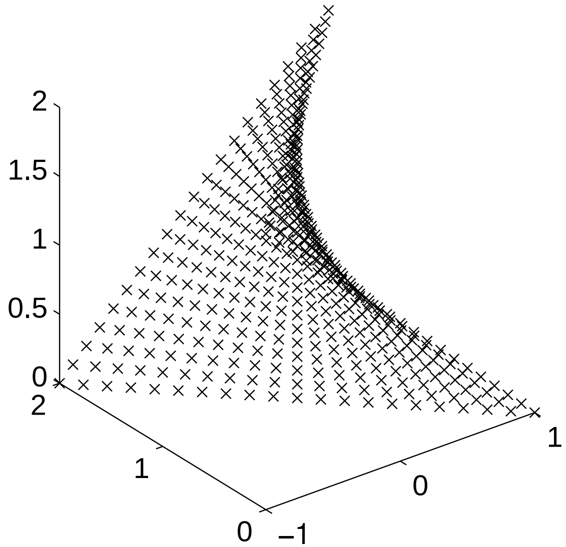


Prof. Dr. Thomas A. Runkler

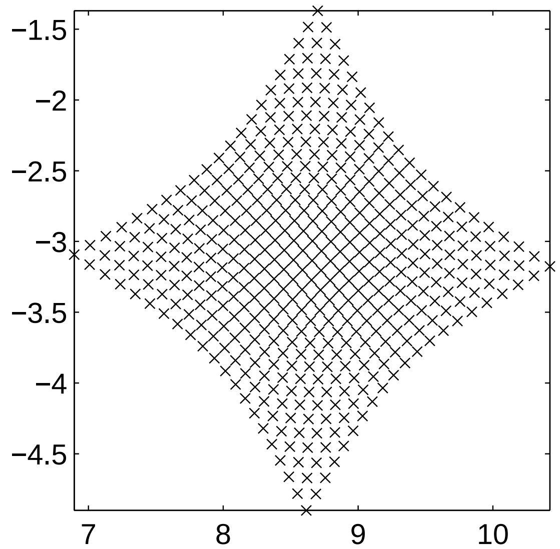
Example Sammon Mapping: Bent Square

$$X = \{((t_1 - 1) \cdot (t_2 - 1), t_1, t_2)^T \mid t_1, t_2 \in \{0, 0.1, 0.2, \ldots, 2\}\}$$

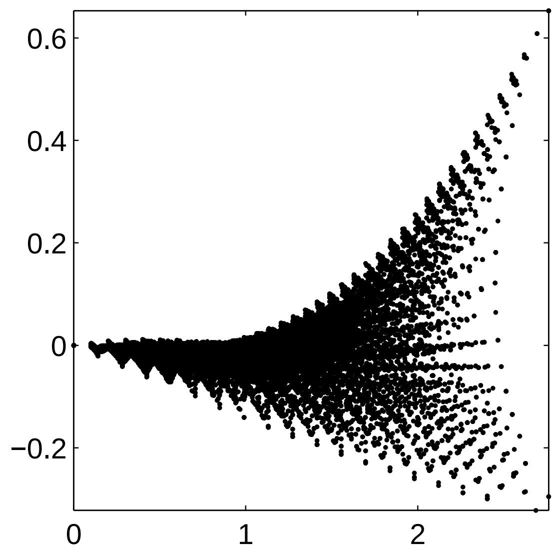
original data



Sammon projection



projection error



Auto–Encoder

- find $f : R^p \rightarrow R^q$ and $g : R^q \rightarrow R^p$

$$\begin{aligned} y_k &= f(x_k) \\ x_k &\approx g(y_k) \end{aligned}$$

- find $g \circ f$ by regression

$$x_k \approx g \circ f(x_k) = g(f(x_k))$$

- use f for mapping

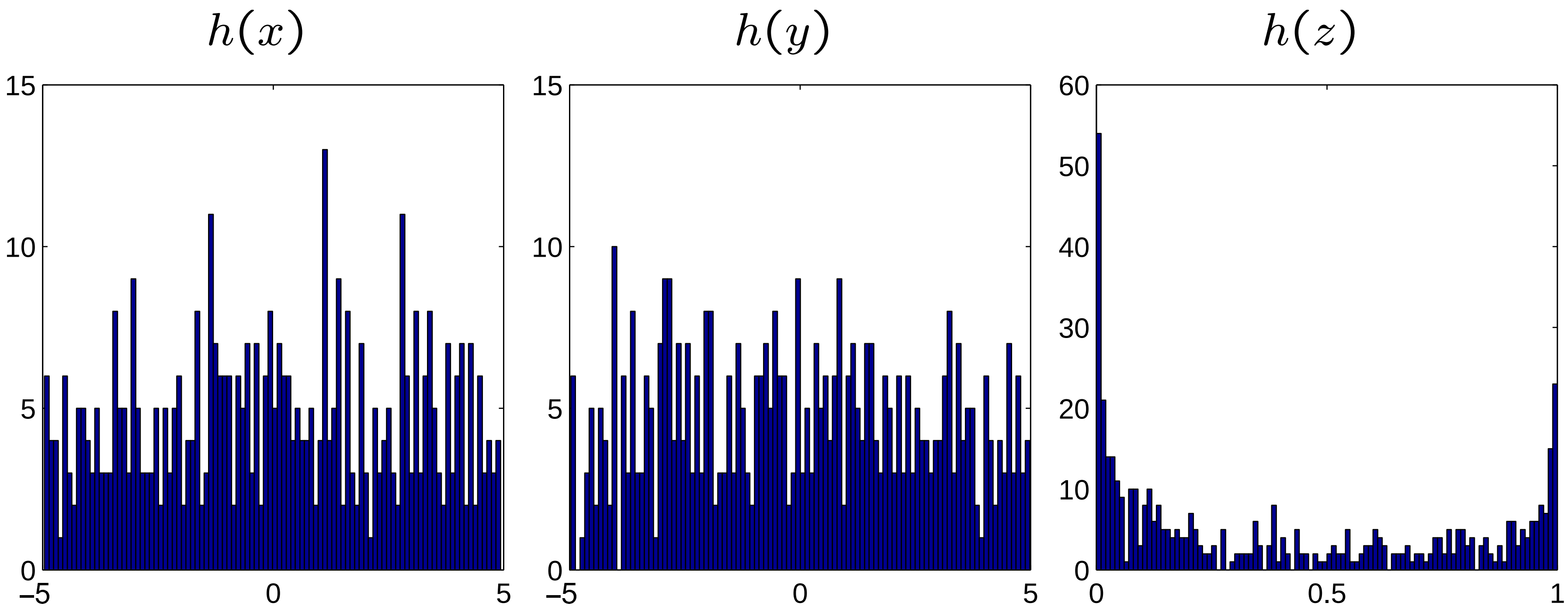
$$y_k = f(x_k)$$

- suitable regression models will be presented in chapter 6
- neural network auto–encoders are widely used as layers of deep learning architectures

Histogram

$$h_k(X) = |\{\xi \in X \mid \xi_k \leq \xi < \xi_{k+1}\}|, \quad k = 1, \ldots, m$$

$$\Delta x = (\max X - \min X)/m, \quad \xi_k = \min X + (k-1) \cdot \Delta x$$



Prof. Dr. Thomas A. Runkler

Number of Histogram Intervals

- Sturgess: number of data

$$m = 1 + \log_2 n$$

- Scott: Gaussian distribution

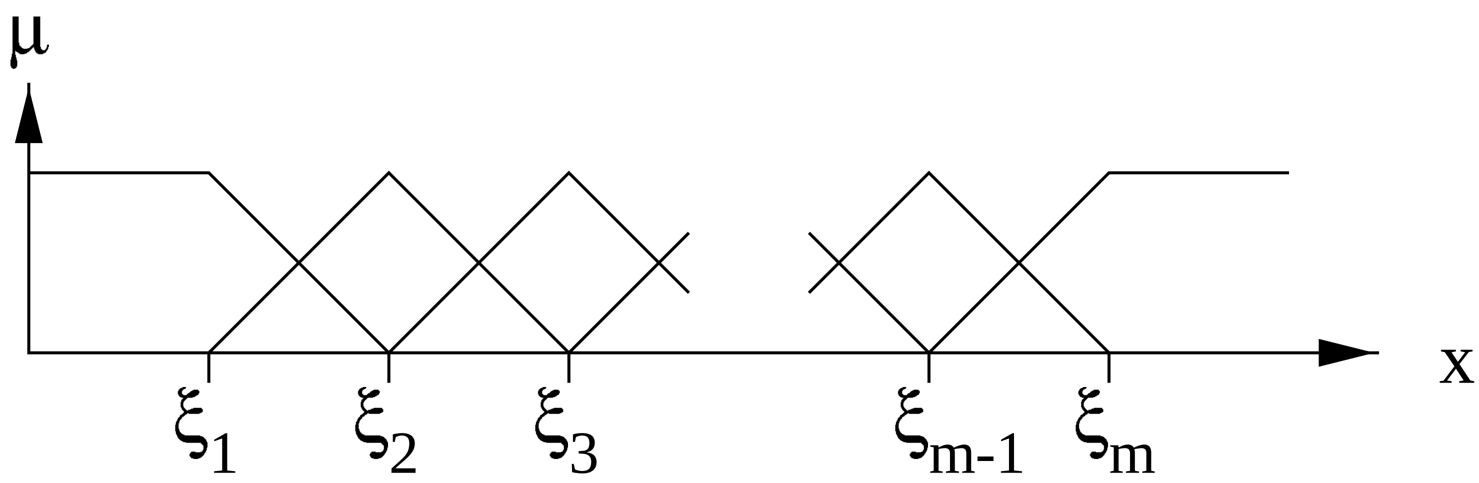
$$m = \frac{3.49 \cdot s}{\sqrt[3]{n}}$$

- Freedman and Diaconis: middle 50% quantile

$$m = \frac{2 \cdot (Q_{75\%} - Q_{25\%})}{\sqrt[3]{n}}$$

$$|\{x \in X \mid x \leq Q_\varphi\}| = \varphi \cdot n$$

Fuzzy Histogram



$$\tilde{h}_k(X) = \sum_{x \in X} \mu_k(x)$$

Spectral Analysis

- Fourier theorem

$$
\begin{aligned}
f(x) &= \int\limits_0^\infty \left(a(y)\,\mathsf{cos}\,xy + b(y)\,\mathsf{sin}\,xy\right)\,dy \quad \text{with} \\
a(y) &= \frac{1}{\pi}\int\limits_{-\infty}^{\infty} f(u)\,\mathsf{cos}\,yu\,du \\
b(y) &= \frac{1}{\pi}\int\limits_{-\infty}^{\infty} f(u)\,\mathsf{sin}\,yu\,du
\end{aligned}
$$

Prof. Dr. Thomas A. Runkler

Spectral Analysis

- Fourier cosine transform

$$F_c(y) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(x) \cos xy \, dx$$

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty F_c(y) \cos xy \, dy$$

- Fourier sine transform

$$F_s(y) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(x) \sin xy \, dx$$

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty F_s(y) \sin xy \, dy$$

Spectral Analysis

- discretization: $x = k \cdot T$, $y = l \cdot \omega$
- discrete Fourier cosine transform

$$y_l^c = \frac{2}{n} \sum_{k=1}^{n} x_k \cos kl\omega T$$

$$x'_k = \frac{n\omega T}{\pi} \sum_{l=1}^{m} y_l^c \cos kl\omega T$$

- discrete Fourier sine transform

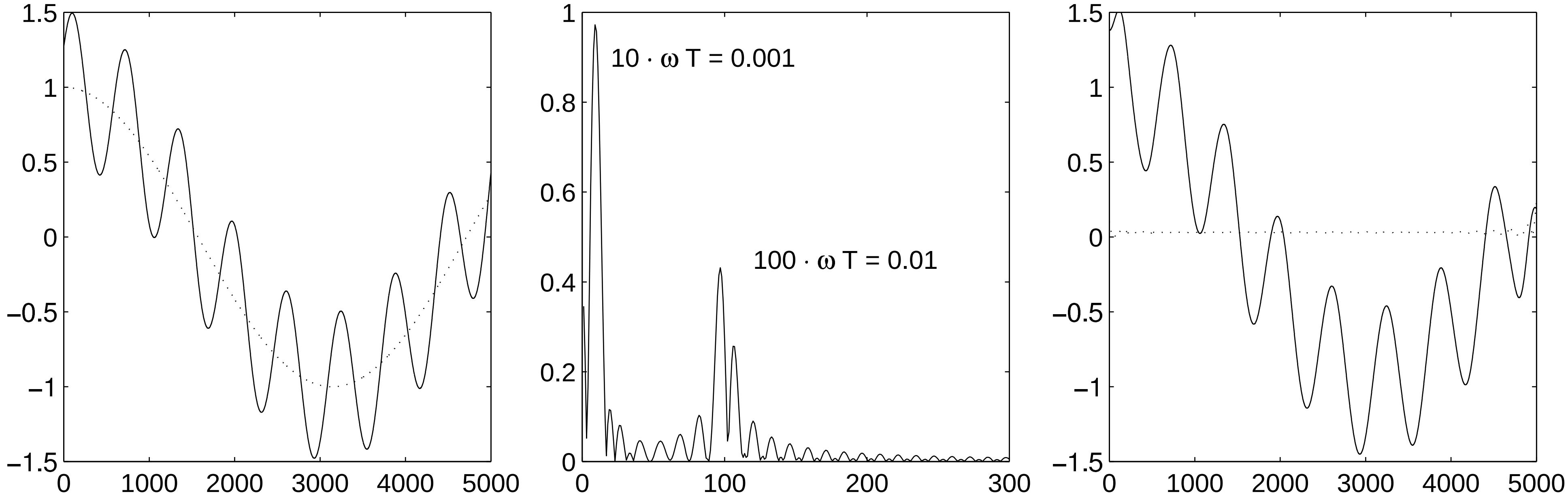
$$y_l^s = \frac{2}{n} \sum_{k=1}^{n} x_k \sin kl\omega T$$

$$x'_k = \frac{n\omega T}{\pi} \sum_{l=1}^{m} y_l^s \sin kl\omega T$$

Example Cosine Transform

$X = \{(i, \cos(0.001 \cdot i) + 0.5 \cdot \cos(0.01 \cdot i - 1)) \mid i \in \{1, \ldots, 5000\}\}$
$m = 300$ spectral values, $\omega T = 10^{-4}$

original data $X = \{x_1, \ldots, x_{5000}\}$ | cosine spectrum $|Y^c| = \{|y^c_1|, \ldots, |y^c_{300}|\}$ | inverse transform $X' = \{x'_1, \ldots, x'_{5000}\}$

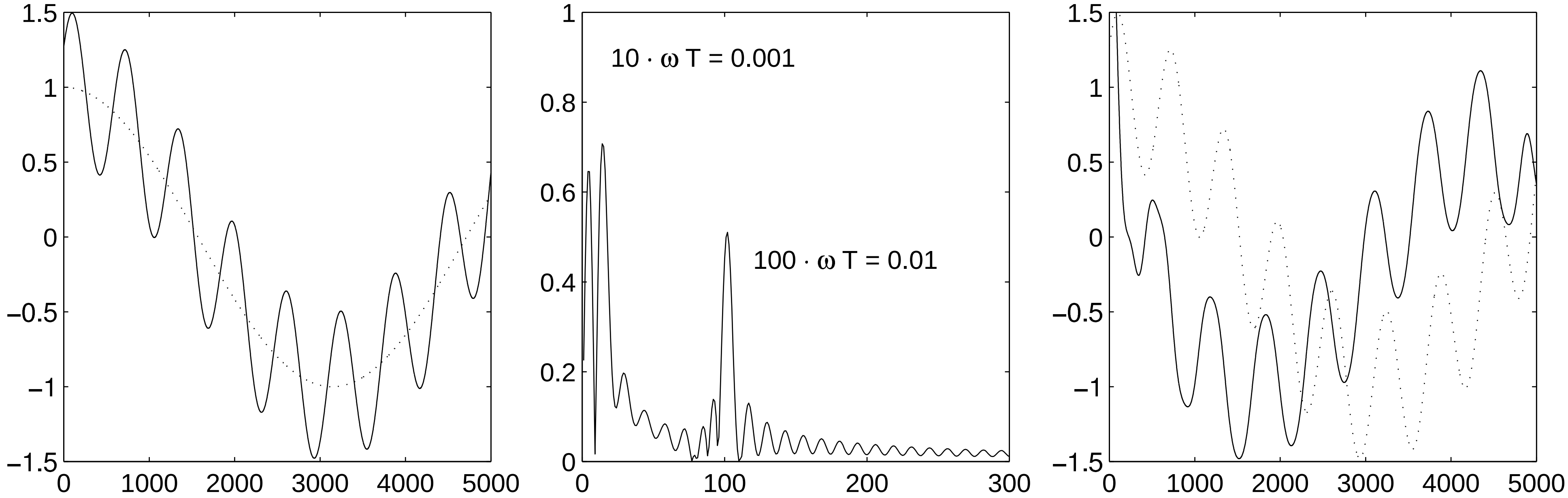


Prof. Dr. Thomas A. Runkler

Example Sine Transform

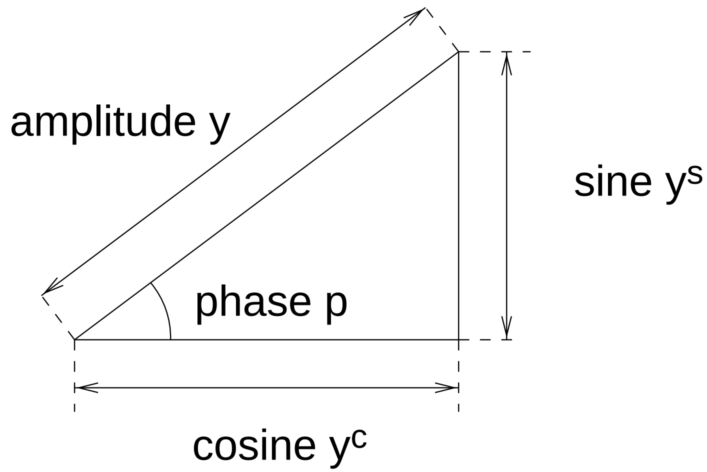
$X = \{(i, \cos(0.001 \cdot i) + 0.5 \cdot \cos(0.01 \cdot i - 1)) \mid i \in \{1, \ldots, 5000\}\}$
$m = 300$ spectral values, $\omega T = 10^{-4}$

original data $X = \{x_1, \ldots, x_{5000}\}$ | sine spectrum $|Y^s| = \{|y_1^s|, \ldots, |y_{300}^s|\}$ | inverse transform $X' = \{x'_1, \ldots, x'_{5000}\}$



Prof. Dr. Thomas A. Runkler

Spectral Analysis



amplitude spectrum $$y_l = \sqrt{(y_l^c)^2 + (y_l^s)^2}$$

phase spectrum $$p_l = \arctan \frac{y_l^s}{y_l^c}$$

Prof. Dr. Thomas A. Runkler

Example Amplitude and Phase Spectrum

$X = \{(i, \cos(0.001 \cdot i) + 0.5 \cdot \cos(0.01 \cdot i - 1)) \mid i \in \{1, \ldots, 5000\}\}$
$m = 300$ spectral values, $\omega T = 10^{-4}$

original data $X = \{x_1, \ldots, x_{5000}\}$ | amplitude spectrum $Y = \{y_1, \ldots, y_{300}\}$ | phase spectrum $P = \{p_1, \ldots, p_{300}\}$