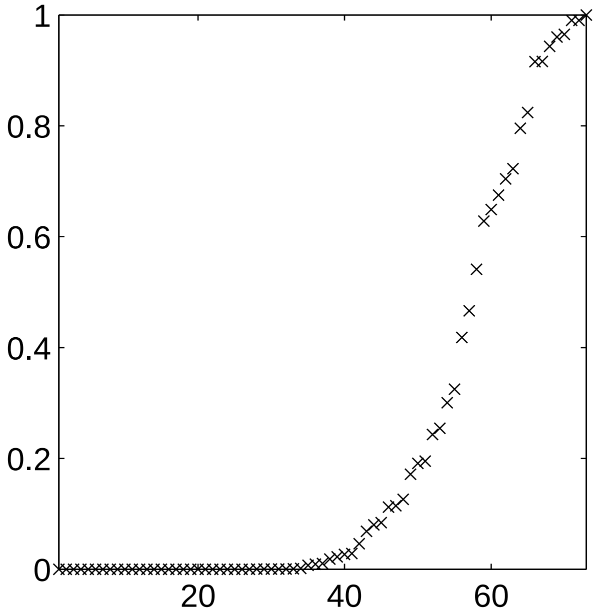
Chapter 4: Visualization

1. Diagrams
2. Principal Component Analysis
3. Multi Dimensional Scaling
4. Sammon Mapping
5. Auto–Encoder
6. Histograms
7. Spectral Analysis

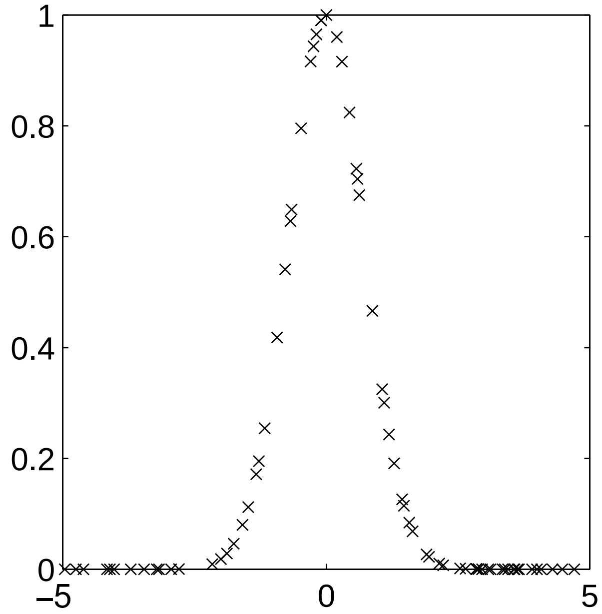
Prof. Dr. Thomas A. Runkler

Diagrams

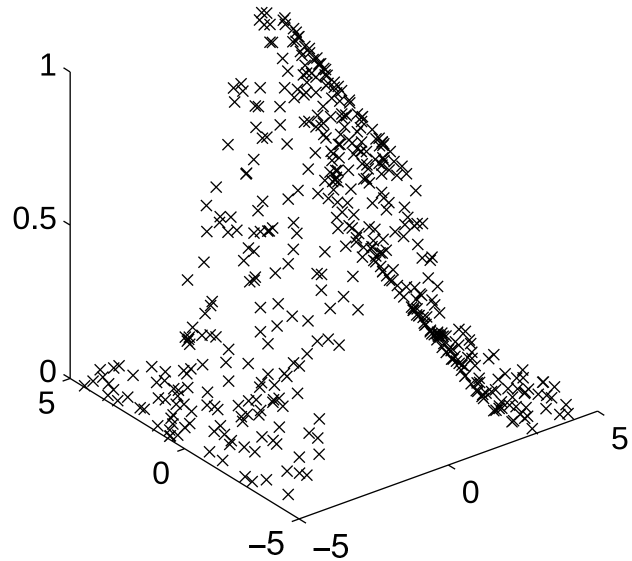
dimension $p = 1$
$X = \{x_k\}$
diagram $x_1, \ldots, x_n$



dimension $p = 2$
$X = \{(x_k, y_k)\}$
scatter diagram $x(y)$

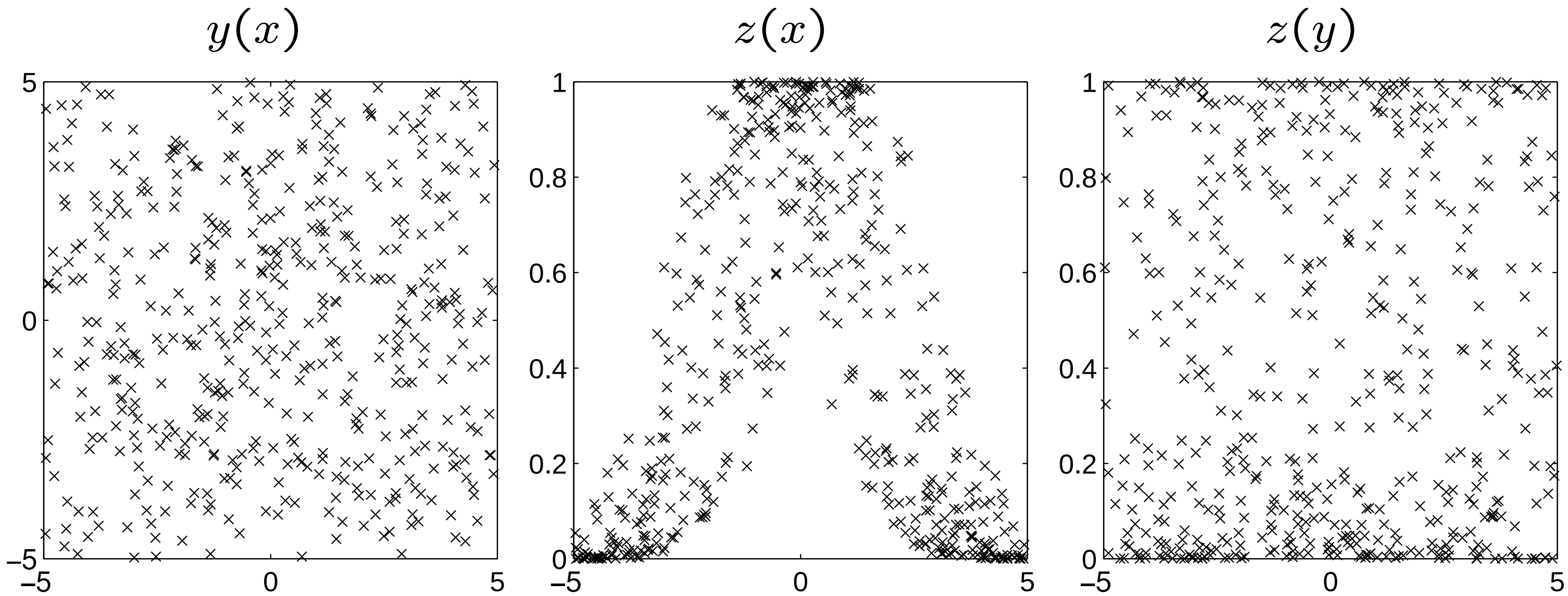


dimension $p = 3$
$X = \{(x_k, y_k, z_k)\}$
3D scatter diagram $z(x, y)$



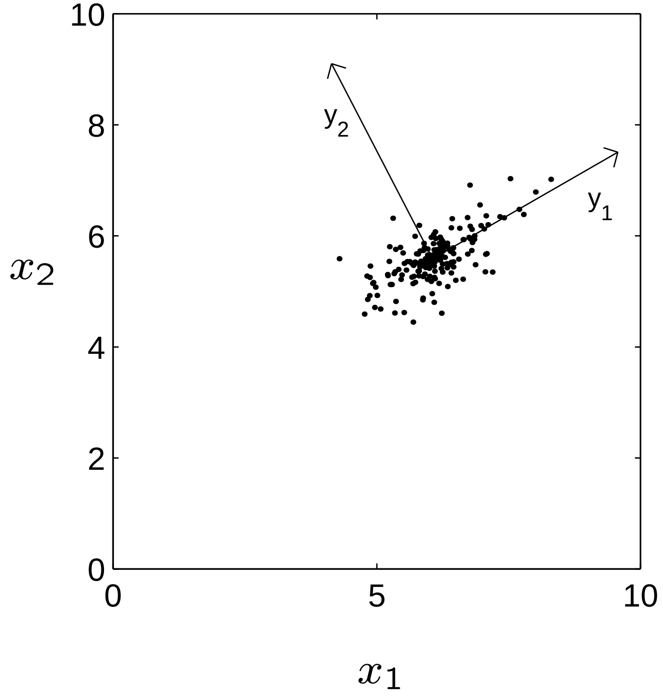
Prof. Dr. Thomas A. Runkler

Projection



Principal Component Analysis

- principal component analysis (PCA)
- Karhunen-Loève transform
- singular value decomposition (SVD)
- eigenvector projection
- Hotelling transform
- proper orthogonal decomposition



Principal Component Analysis

- coordinate transformation: translation and rotation

  transformation $\mathrm{IR}^p \to \mathrm{IR}^q$: $\quad y_k = (x_k - \bar{x}) \cdot E$

  inverse transformation $\mathrm{IR}^q \to \mathrm{IR}^p$: $\quad x_k = y_k \cdot E^T + \bar{x}$

- mean

$$\bar{x} = \frac{1}{n} \sum_{k=1}^{n} x_k$$

- determine rotation matrix E by maximizing the variance of Y

Principal Component Analysis

- variance in Y

$$
\begin{aligned}
v_y &= \frac{1}{n-1}\sum_{k=1}^{n} y_k^T y_k \\
&= \frac{1}{n-1}\sum_{k=1}^{n} ((x_k-\bar{x})\cdot E)^T \cdot ((x_k-\bar{x})\cdot E) \\
&= \frac{1}{n-1}\sum_{k=1}^{n} E^T \cdot (x_k-\bar{x})^T \cdot (x_k-\bar{x})\cdot E \\
&= E^T \left(\frac{1}{n-1}\sum_{k=1}^{n} (x_k-\bar{x})^T \cdot (x_k-\bar{x}) \right) \cdot E \\
&= E^T \cdot C \cdot E
\end{aligned}
$$

- covariance matrix of X

$$
c_{ij} = \frac{1}{n-1}\sum_{k=1}^{n} (x_k^{(i)} - \bar{x}^{(i)})(x_k^{(j)} - \bar{x}^{(j)})
$$

Principal Component Analysis

- constraint: rotation, no dilation

$$E^T \cdot E = I$$

- Lagrange optimization

$$\begin{aligned} L &= E^T C E - \lambda(E^T E - I) \\ \frac{\partial L}{\partial E} = 0 &\Rightarrow CE - \lambda E + E^T C - \lambda E^T = 0 \\ &\Rightarrow CE = \lambda E \text{ (eigenproblem)} \\ \text{solution} & \quad (v_1, \ldots, v_p, \lambda_1, \ldots, \lambda_p) = \text{eig}(C) \end{aligned}$$

- solution using homogeneous equation system

$$(C - \lambda I) \cdot E = 0$$

Prof. Dr. Thomas A. Runkler

Principal Component Analysis

- E is matrix of eigenvectors of C

$$E = (v_1, \ldots, v_q)$$

- variances in Y are eigenvalues of C

$$CE = \lambda E \quad \Leftrightarrow \quad \lambda = E^T C E = v_y$$

- suitable dimensionality q

$$\sum_{i=1}^{q} \lambda_i \Big/ \sum_{i=1}^{p} \lambda_i \geq 95\%$$

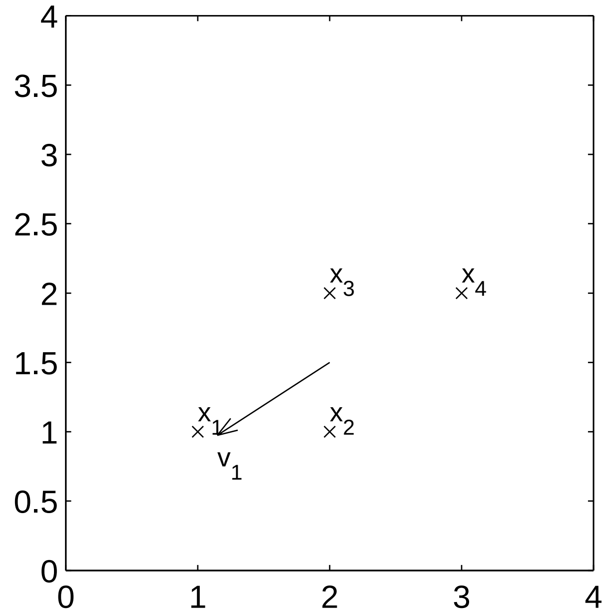
- transformation error

$$e = \frac{1}{n} \sum_{k=1}^{n} (x_k - x'_k)^2 = \frac{n-1}{n} \sum_{i=q+1}^{p} \lambda_i$$

$\Rightarrow$ PCA yields projection with minimal quadratic error

Example Principal Component Analysis

$$
\begin{aligned}
X &= \{(1,1),(2,1),(2,2),(3,2)\} \\
\bar{x} &= \frac{1}{2}\cdot(4,3) \\
C &= \frac{1}{3}\cdot\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \\
\lambda_1 &= 0.8727 \\
\lambda_2 &= 0.1273 \\
v_1 &= \begin{pmatrix} -0.85065 \\ -0.52573 \end{pmatrix} \\
v_2 &= \begin{pmatrix} 0.52573 \\ -0.85065 \end{pmatrix}
\end{aligned}
$$



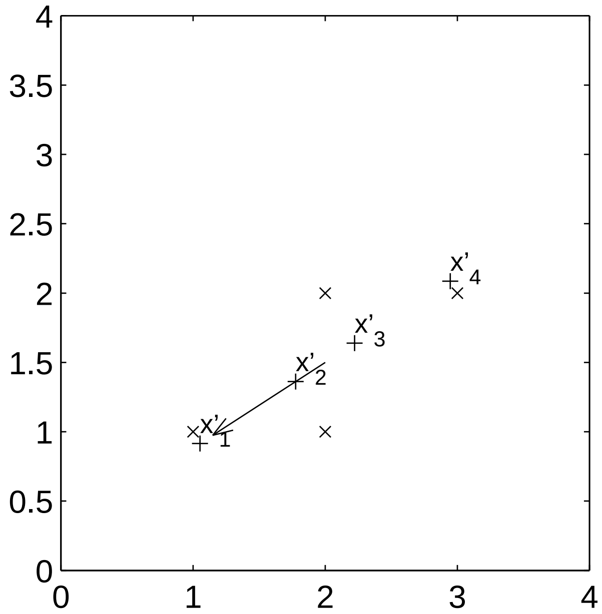
Example Principal Component Analysis

projection onto first axis:

$$E = v_1 = \begin{pmatrix} -0.85065 \\ -0.52573 \end{pmatrix}$$

$$Y = (1.1135, 0.2629, -0.2629, -1.1135)$$

$$X' = \{(1.0528, 0.91459), (1.7764, 1.3618), (2.2236, 1.6382), (2.9472, 2.0854)\}$$



Prof. Dr. Thomas A. Runkler

Multi Dimensional Scaling

- eigendecomposition of the positive semi–definite matrix XX^T

$$XX^T = Q\Lambda Q^T = (Q\sqrt{\Lambda}^T) \cdot (\sqrt{\Lambda}Q^T) = (Q\sqrt{\Lambda}^T) \cdot (Q\sqrt{\Lambda}^T)^T$$

  $Q = (v_1, \ldots, v_p)$: matrix of eigenvectors of XX^T, usually $p < n$
  Λ: diagonal matrix of eigenvalues of XX^T
- estimate for X

$$Y = Q\sqrt{\Lambda}^T$$

- lower dimensional projections $Y \subset R^q$, $q < p$
  - use only first q dimensions
  - scale eigenvectors (square norms match eigenvalues)

MDS of a Distance Matrix D

- assume $\tilde{X}$ and choose anchor point $\tilde{x}_a$, $a \in \{1, \ldots, n\}$
- transform $\tilde{X}$ (origin 0) to X (origin $\tilde{x}_a$)

$$x_k = \tilde{x}_k - \tilde{x}_a$$

- difference vectors are invariant

$$\tilde{x}_i - \tilde{x}_j = x_i - x_j$$

- scalar product of each side with itself

$$(\tilde{x}_i - \tilde{x}_j)(\tilde{x}_i - \tilde{x}_j)^T = (x_i - x_j)(x_i - x_j)^T$$

$$\Rightarrow d_{ij}^2 = x_i x_i^T - 2 x_i x_j^T + x_j x_j^T = d_{ia}^2 - 2 x_i x_j^T + d_{ja}^2$$

$$\Rightarrow x_i x_j^T = (d_{ia}^2 + d_{ja}^2 - d_{ij}^2)/2$$

- XX^T can be computed from D

Example Multi Dimensional Scaling

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

$$X = \left\{(-1,-\frac{1}{2}),(0,-\frac{1}{2}),(0,\frac{1}{2}),(1,\frac{1}{2})\right\} \quad \text{(mean subtracted)}$$

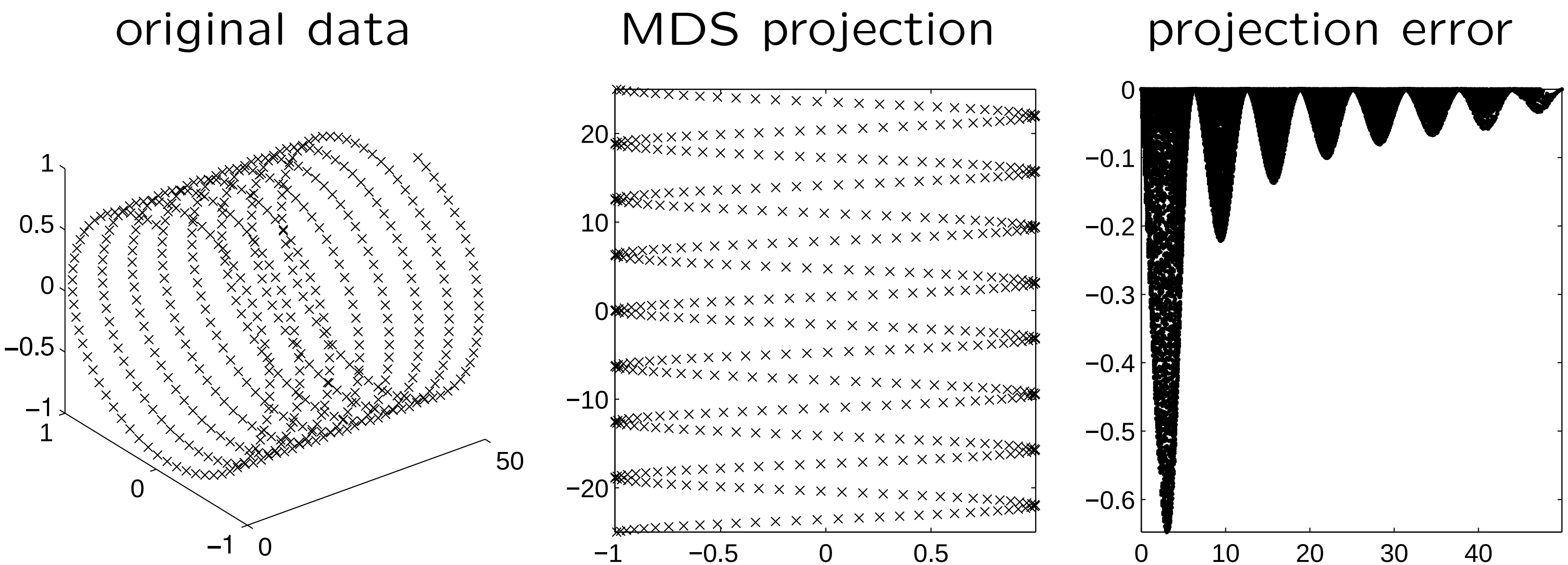
$$XX^T = \frac{1}{4}\begin{pmatrix} 5 & 1 & -1 & -5 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -5 & -1 & 1 & 5 \end{pmatrix}$$

$$v_1 \approx \begin{pmatrix} -0.6882 \\ -0.1625 \\ 0.1625 \\ 0.6882 \end{pmatrix}, \quad \lambda_1 \approx 2.618$$

$$\Rightarrow Y \approx \begin{pmatrix} -0.6882 \\ -0.1625 \\ 0.1625 \\ 0.6882 \end{pmatrix}\sqrt{2.618} \approx \begin{pmatrix} -1.1135 \\ -0.2629 \\ 0.2629 \\ 1.1135 \end{pmatrix}$$

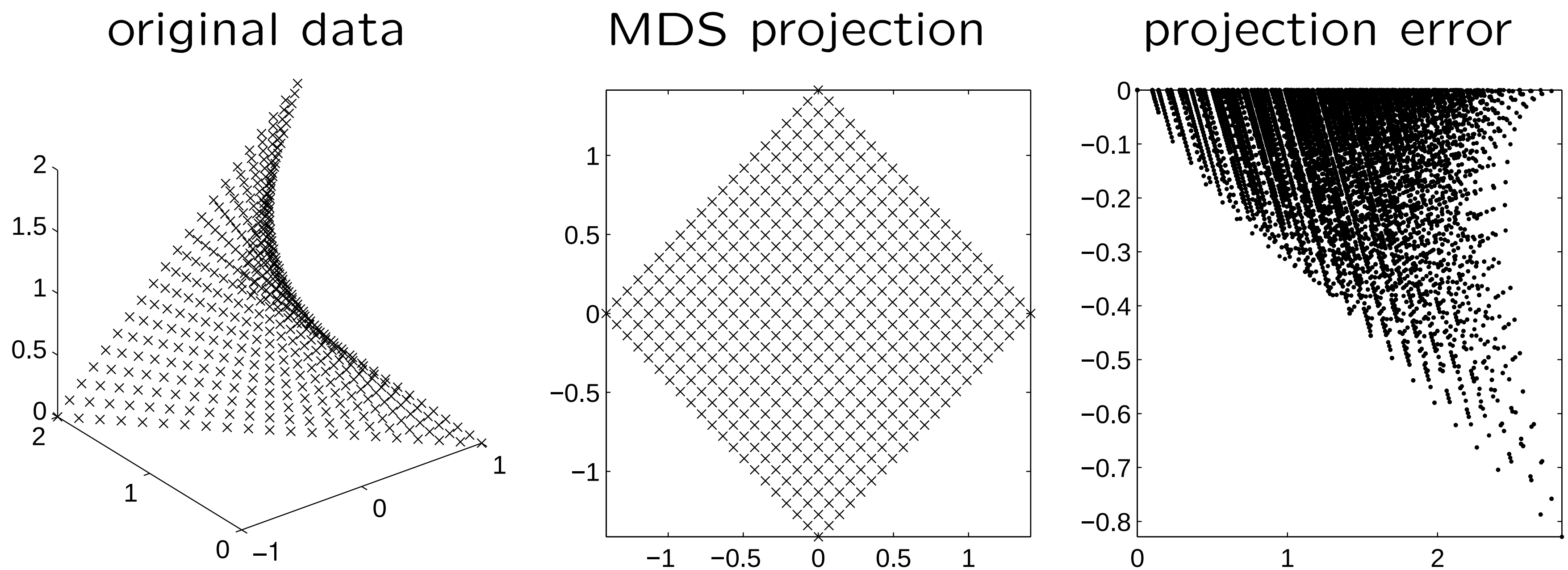
Example MDS: Helix

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



Example MDS: Bent Square

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

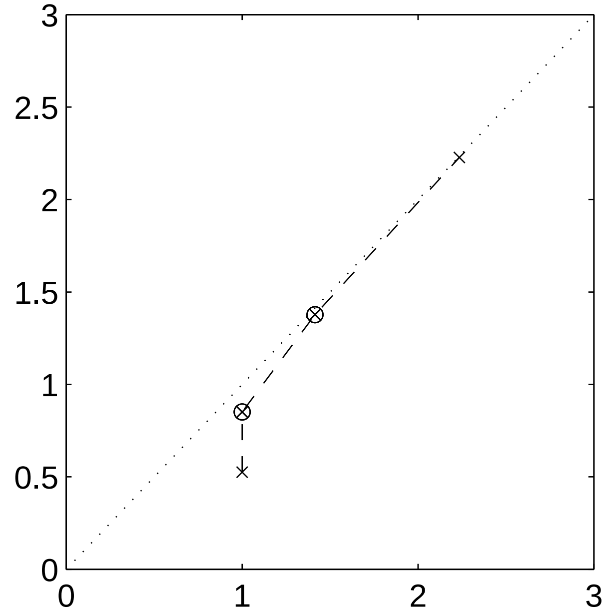


Prof. Dr. Thomas A. Runkler

Shepard Diagram

$$D^x \approx \begin{pmatrix} 0 & 1 & 1.4142 & 2.2361 \\ 1 & 0 & 1 & 1.4142 \\ 1.4142 & 1 & 0 & 1 \\ 2.2361 & 1.4142 & 1 & 0 \end{pmatrix}$$

$$D^y \approx \begin{pmatrix} 0 & 0.8507 & 1.3764 & 2.2270 \\ 0.8507 & 0 & 0.5257 & 1.3764 \\ 1.3764 & 0.5257 & 0 & 0.8507 \\ 2.2270 & 1.3764 & 0.8507 & 0 \end{pmatrix}$$



- alternative criteria for multidimensional scaling
  - strict monotonicity (Torgerson)
  - points close to main diagonal (Sammon)

