Non-Euclidean Representations

Causes, Corrections, Informativeness

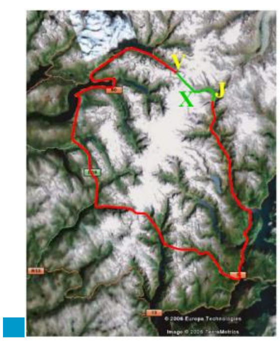
Non-Euclidean Representations: Causes

Computational Noise

Even for Euclidean distance matrices zero eigenvalues may show negative, e.g:

- X = N(50,20) : 50 points in 20 dimensions
- D = Dist(X): 50 x 50 distance matrix
- Expected: 49-20 = 29 zero eigenvalues
- Found: 15 negative eigenvalues

Lack of information

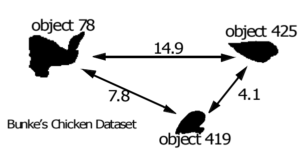


1800:
Crossing the Jostedalsbreen was impossible.
Travelling around (200 km) lasted 5 days.
Untill the shared point X was found.
People could visit each other in 8 hours.

D(V,J) = 5 days
D(V,X) = 4 hours
D(X,J) = 4 hours

Computational Problems

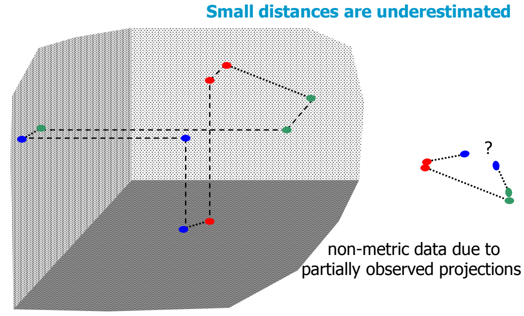
Large distances are overestimated due to computational problems



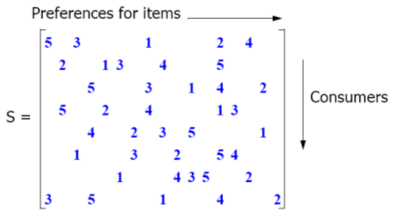
Weighted edit distance for strings

Projections - Occlusions

Small distances are underestimated



Projections - Occlusions

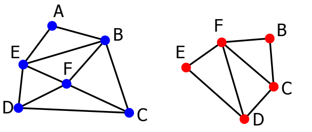


Example: consumer preferences for recommendation systems

TUDelft

Graph Matching → Dissimilarities

Representation by Connected Graphs



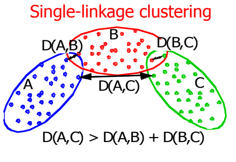
Graph (Nodes, Edges, Attributes)

Distance (Graph_1, Graph_2)

TUDelft

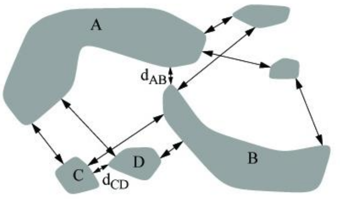
Intrinsicly Non-Euclidean Dissimilarity Measures Single Linkage

Distance(Table,Book) = 0
Distance(Table,Cup) = 0
Distance(Book,Cup) = 1



TUDelft

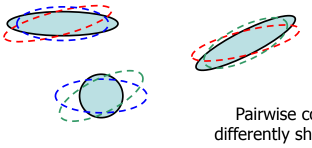
Boundary distances



A set of boundary distances may characterize sets of datapoints:
Distances → features

TUDelft

Intrinsicly Non-Euclidean Dissimilarity Measures Mahalanobis

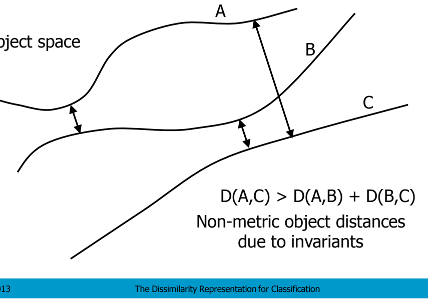


Pairwise comparison between
differently shaped data distributions

Different pairs → different comparison frameworks
→ non-Euclidean

TUDelft

Intrinsicly Non-Euclidean Dissimilarity Measures Invariants

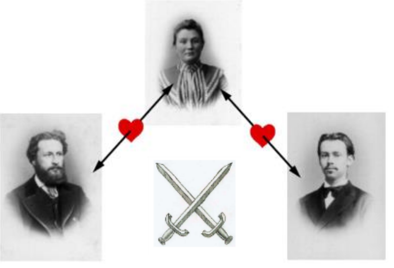


D(A,C) > D(A,B) + D(B,C)

Non-metric object distances
due to invariants

TUDelft

Intrinsicly Non-Euclidean Dissimilarity Measures



Non-Euclidean human relations

Objects may have an 'inner life'

In dissimilarity measures the 'inner life' of objects may be taken into account (e.g. invariants).

→ Objects cannot be represented anymore as points

→ Non-Euclidean dissimilarities

Causes of Non-Euclidean Dissimilarities

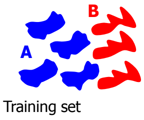
- **Computational / Observational Limitations**
  - numerical accuracy problems
  - overestimated large distances (too difficult to compute)
  - underestimated small distances (one-sided view of objects)
- **Essential non-Euclidean distance definitions**
  - the human distance concept differs from the mathematical one
  - no global framework
  - invariants

Euclidean corrections for non-Euclidean dissimilarities

SSSPR 2008

R.P.W. Duin, E. Pekalska, A. Harol, W.J. Lee and H. Bunke

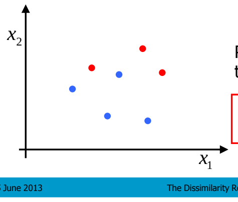
Alternative 2: Embedding



Training set

→ Dissimilarity matrix D → X

Is there a feature space for which Dist(X,X) = D ?



Position points in a vector space such that their Euclidean distances → D

However, many popular dissimilarity measures are not Euclidean!!

(Pseudo-)Euclidean Embedding

$m \times m$ D is a given, imperfect dissimilarity matrix of training objects.

Construct inner-product matrix: $B = -\frac{1}{2} J D^{(2)} J \quad J = I - \frac{1}{m}\mathbf{11}$

Eigenvalue Decomposition , $B = Q \Lambda Q^T$

Select k eigenvectors: $X = Q_k \Lambda_k^{\frac{1}{2}}$ (problem: $\Lambda_k < 0$)

Let $\Im_k$ be a k x k diag. matrix, $\Im_k(i,i)$ = sign($\Lambda_k(i,i)$)

$\Lambda_k(i,i) < 0 \rightarrow$ Pseudo-Euclidean

$n \times m$ D_z is the dissimilarity matrix between new objects and the training set.

The inner-product matrix: $B_z = -\frac{1}{2}(D_z^{(2)} J - \frac{1}{n}\mathbf{11}^T D^{(2)} J)$

The embedded objects: $Z = B_z Q_k |\Lambda_k|^{-\frac{1}{2}} \Im_k$

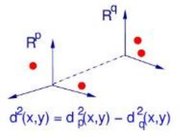
PES: Pseudo-Euclidean Space (Krein Space)

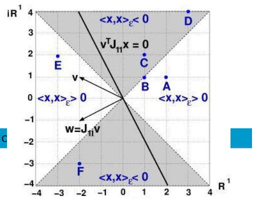
If D is non-Euclidean, B has p positive and q negative eigenvalues.

A pseudo-Euclidean space $\mathcal{E}$ with signature (p,q), $k = p+q$, is a non-degenerate inner product space $\Re_k = \Re_p \oplus \Re_q$ such that:

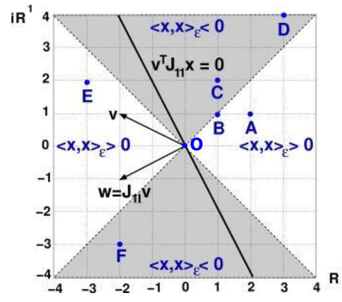
$$\langle x, y\rangle_\varepsilon = x^T \Im_{pq} y = \sum_{i=1}^{p} x_i y_i - \sum_{j=p+1}^{q} x_j y_j \qquad \Im_{pq} = \begin{bmatrix} I_{p\times p} & 0 \\ 0 & -I_{q\times q} \end{bmatrix}$$

$$d_\varepsilon^2(x,y) = \langle x-y, x-y\rangle_\varepsilon = d_p^2(x,y) - d_q^2(x,y)$$





Distances in PES



$d^2(O,A) > 0$
$d^2(O,E) > 0$
$d^2(O,B) = 0$
$d^2(O,D) < 0$

All points in the grey area are closer to O than O itself !?

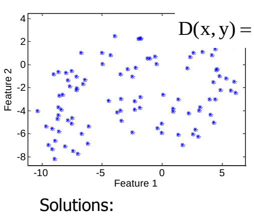
Any point has a negative square distance to some points on the line $v^T J x = 0$.
Can it be used as a classifier?
Can we define a margin as in the SVM?

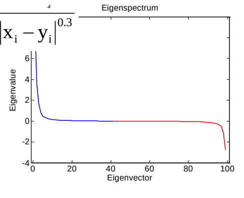
TUDelft

Pseudo-Euclidean Embedding

If D is non-Euclidean then B has p positive and q negative eigenvalues

$$D(x,y) = \sqrt[0.3]{\sum |x_i - y_i|^{0.3}}$$





Solutions:

- Remove all eigenvectors with small and negative eigenvalues
- or, take absolute values of eigenvalues and proceed
- or, construct a pseudo-Euclidean space

TUDelft

Correction Procedures PES ←→ ES

- **PES**: Pseudo Euclidean Space

  $$d_\varepsilon^2(x,y) = d_p^2(x,y) - d_q^2(x,y)$$

- **PES+**: Positive contributions only

  $$d_\varepsilon^2(x,y) = d_p^2(x,y)$$

- **AES**: Treat entire space as Euclidean

  $$d_\varepsilon^2(x,y) = d_p^2(x,y) + d_q^2(x,y)$$

TUDelft

Correction Procedures PES ←→ ES (2)

- **DEC**: Enlarging dissimilarities

  $$d_\varepsilon^2(x,y) \Leftarrow d_\varepsilon^2(x,y) + c, x \neq y$$

  use smallest c such that $D \succ 0$

- **Relax**: Relaxing dissimilarity measure

  $$d_\varepsilon(x,y) \Leftarrow d_\varepsilon(x,y)^{1/c}, c \geq 1$$

  use smallest c such that $D \succ 0$

TUDelft

Correction Procedures PES ←→ ES (3)

- **Laplace**:
  Adaptation of Laplace correction from spectral graph theory:

  $$d_\varepsilon^2(x,y) \Leftarrow \mathrm{Norm}(d_\varepsilon^2(x,y))$$
  $$d_\varepsilon^2(x,y) \Leftarrow 1 - \delta(x,y) + d_\varepsilon^2(x,y)$$
  $$d_\varepsilon^2(x,y) \Leftarrow d_\varepsilon^2(x,y) + c, x \neq y \quad D \succ 0$$

- (unpublished result: a 'minimum' Laplace correction can be obtained by normalizing the dissimilarity matrix, followed by the DEC correction)

TUDelft

Example: Chickenpieces (H. Bunke, Bern)

BACK

BREAST

DRUMSTICK

THIGH-AND-BACK

WING

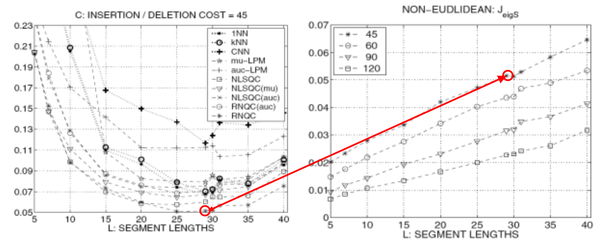
446 binary images, varying size, e.g.: 100 x 130

Andreu, G., Crespo, A., Valiente, J.M.: Selecting the toroidal self-organizing feature maps (TSOFM) best organized to object recogn. In: ICNN. (1997) 1341–1346.

Shape classification by weighted-edit distances (Bunke)

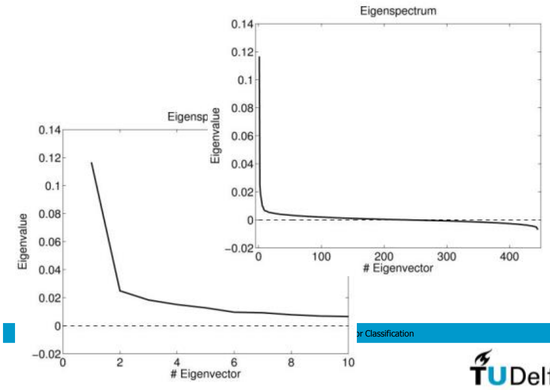
*Bunke, H., Buhler, U.: Applications of approximate string matching to 2D shape recognition. Pattern recognition **26** (1993) 1797–1812*

Classification Results for Various Dissimilarity Measures

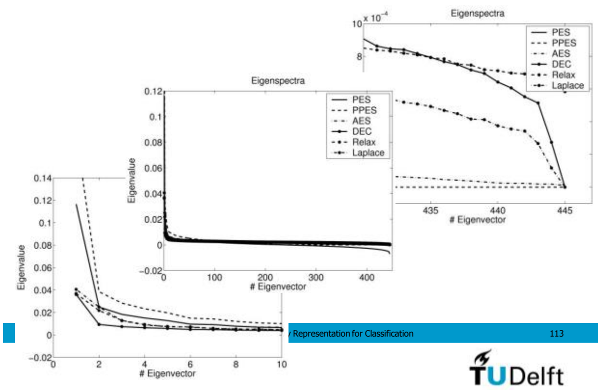


Best classification result is for a very non-Euclidean dissimilarity measure !

Eigenspectrum original data (PES)



Eigenspectra corrected data

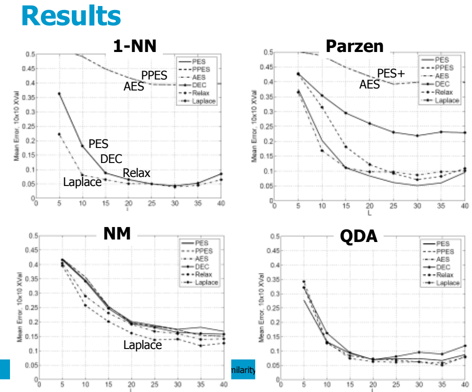


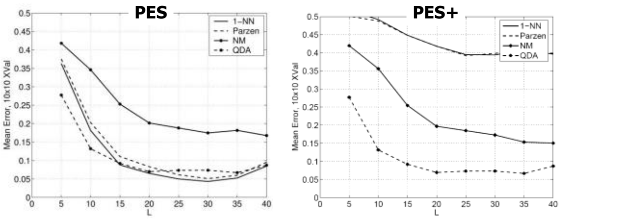
Classifiers:

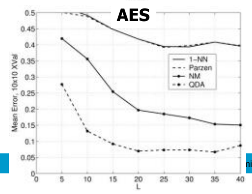
- **1-NN**: 1-Nearest Neighbor (local distances)
- **Parzen**: non-parametric density estimation based on given dissimilarities (local densities)
- **NM**: Nearest Mean Classifier (global distances)
- **QDA**: Quadratic Discriminant Analysis (global densities)

All four classifiers can be computed in ES as well as in PES

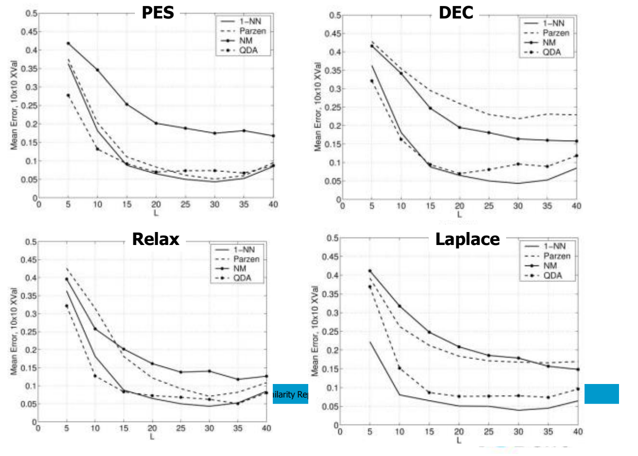
Results







Neglecting or sign-changing "negative" directions in the Pseudo-Euclidean Space deteriorates local classifiers (1-NN and Parzen) significantly.



Conclusion w.r.t. Euclidean corrections

- Globally sensitive classifiers are hardly affected by corrections.
- 1-NN is insensitive to monotonic corrections, but really suffers from crisp corrections in the PES.
- Parzen is always disturbed by corrections, so they damage the local structure in the data.
- Corrections should be studied in relation with the classifier.

TUDelft

Non-Euclidean Representations: Informativeness

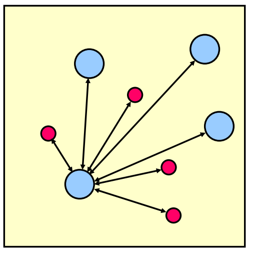
Negative Euclidean Fraction

$$\mathrm{NEF} = \frac{\sum_{\lambda_i < 0} |\lambda_i|}{\sum_{\forall \lambda_i} |\lambda_i|}$$

$$0 \leq \mathrm{NEF} \leq 1$$

TUDelft

Artificial Example ,Ball Distances



- Generate sets of balls (classes) uniformly, in a (hyper)cube; not intersecting.
- Balls of the same class have the same size.
- Compute all distances between the ball surfaces.

-> Dissimilarity matrix D

TUDelft

Balls3D

Classifier	PE Sp	Ass Sp	Pos Sp	Neg Sp	Cor Sp
1-NN	47.4 (2.0)	47.4 (2.0)	47.4 (2.0)	44.2 (1.5)	47.4 (2.0)
Parzen	45.7 (1.7)	45.5 (1.6)	45.6 (1.7)	35.5 (1.7)	45.7 (1.7)
NM	47.5 (2.0)	47.7 (2.0)	47.6 (1.9)	49.6 (0.2)	48.1 (1.8)
SVM-1	50.7 (2.2)	50.0 (2.7)	50.0 (2.5)	62.1 (1.7)	50.1 (2.0)

Classifier	PE Dis Sp	Ass Dis Sp	Pos Dis Sp	Neg Dis Sp	Cor Dis Sp
1-NN	49.8 (2.2)	49.8 (2.2)	49.8 (2.2)	5.1 (0.8)	49.7 (2.2)
Parzen	47.9 (2.2)	47.9 (2.2)	47.9 (2.2)	4.6 (0.5)	47.9 (2.2)
NM	49.8 (2.2)	49.8 (2.2)	49.8 (2.2)	5.0 (0.8)	49.9 (2.2)
SVM-1	50.2 (1.6)	50.8 (1.7)	50.7 (1.7)	1.9 (0.5)	49.8 (1.5)

10 x (2-fold crossvalidation of 50 objects per class)

TUDelft

Representation Strategies

Avoiding the PE space

Dissimilarity Space: $X = D$

Correcting

Associated space $X = \{[Xp, Xq], \varnothing\}$ $\tilde{d}_{ij}^2 = d_p^2(x_i, x_j) + d_q^2(x_i, x_j)$

Positive space $X = X_p$ $\tilde{d}_{ij}^2 = d_p^2(x_i, x_j)$

Negative space $X = X_q$ $\tilde{d}_{ij}^2 = d_q^2(x_i, x_j)$

Additive Correction $\tilde{d}_{ij}^2 = d_{ij}^2 + c, i \neq j$ $X = \text{Embedding}(\tilde{D})$

As it is

Pseudo Euclidean Space $X = \{Xp, Xq\}$ $d_{ij}^2 = d_p^2(x_i, x_j) - d_q^2(x_i, x_j)$

Classifiers to be developed further

TUDelft

	size	classes	Non-Metric	NEF	Rand Err	Original, D	Positive, D_p	Negative, D_q
Chickenpieces45	446	5	0	0.156	0.791	**0.022**	0.132	0.175
Chickenpieces60	446	5	0	0.162	0.791	**0.020**	0.067	0.173
Chickenpieces90	446	5	0	0.152	0.791	**0.022**	0.052	0.148
Chickenpieces120	446	5	0	0.130	0.791	**0.034**	0.108	0.148
FlowCyto	612	3	1e-5	0.244	0.598	0.103	**0.100**	0.327
WoodyPlants50	791	14	5e-4	0.229	0.928	**0.075**	0.076	0.442
CatCortex	65	4	2e-3	0.208	0.738	**0.046**	0.077	0.662
Protein	213	4	0	0.001	0.718	0.00[illegible]	[illegible]	[illegible]
Balls3D	200	2	3e-4	0.001	0.500	**0.470**	0.495	0.000
GaussM1	500	2	0	0.262	0.500	**0.202**	**0.202**	0.228
GaussM02	500	2	5e-4	0.393	0.500	0.204	**0.174**	0.252
CoilYork	288	4	8e-8	0.258	0.750	**0.267**	0.313	0.618
CoilDelftSame	288	4	0	0.027	0.750	**0.413**	0.417	0.597
CoilDelftDiff	288	4	8e-8	0.128	0.750	**0.34**[illegible]	[illegible]	[illegible]
NewsGroups	600	4	4e-5	0.202	0.733	**0.198**	0.213	0.435
BrainMRI	124	2	5e-5	0.112	0.499	0.226	**0.218**	0.556
Pedestrians	689	3	4e-8	0.111	0.348	**0.010**	0.015	0.030

Informative

Extremely Informative

Not Informative

Conclusions

- Pseudo Euclidean Space (PES) is sometimes informative (corrections are not helpful).
- The corresponding problems may be intrinsic non-Euclidean
- Classifiers for non-Euclidean data have to be studied further

TUDelft