Spatialisation

PART ONE

Learning Objectives

- What is Spatialisation?
- How can we visualise high-dimension data?

Spatialisation

- find 2D or 3D position of data entries
- typically used to visualise similarities
  - clustering

Measures of observation

- Classificatory concept
- Comparative concept
- Quantitative concept

Carnap, R. (1966), Philosophical Foundations of Physics, New York: Basic Books

Classificatory concept

- category
  - just to distinguish one from others

Comparative concept – Part I

- provides “logical/ordered structure of relationships”

Comparative concept – Part II

- Measurement, M(a)
  - Evaluation of a value (a).

- Empirical relations
  - Equivalence, EM(a,b)
    - EM(a, b) : M(a) = M(b)

  - Magnitude relationship, LM(a,b)

Comparative concept – Part III

- Equivalence, EM(a,b)
  - Symmetric
    - EM(a, b) == EM(b, a)
  - Transitive
    - E(a,b) & E(b,c) => E(a, c)

- Magnitude relationship, LM(a,b)
  - Asymmetric
    - LM(a, b) != LM(b, a)
  - Transitive
    - LM(a,b) & LM(b,c) => LM(a, c)

Quantitative concept

- Subjects will have numerical attributes
  - Distance between value a and b : d(a, b)
- Equivalence
  - EM(a, b) : M(a) = M(b)
- Magnitude relationships
  - LM(a, b) : M(a) < M(b)
- Assign a special value “s” to a particular state “p” :
  - M(p) = s
- Assign another special value “t” to another particular state “q” :
  - M(q) = t
- Equivalence in difference
  - EDM(a,b,c,d) : M(a) - M(b) == M(c) - M(d)

Quantitative concept (Cont.)

- EDM(a,b,c,d) : M(a) - M(b) == M(c) - M(d)
- EDM(a,b,c,d) : Empirical Condition
- EDM(a,b,c,d) : d(a,b) == d(c,d)

Multidimensional Scaling (MDS)

Psychophysical Measure

- Sensory Scaling
  - map physically measurable to psychological continuum

  - old techniques are statistic-based : map psychological observation to unidimensional continuum using probability density function.

Multidimensional Scaling (MDS)

- including both numerical and categorical observations
- number/detail of dimensions might be unknown
- Map data/states to multidimensional continuum

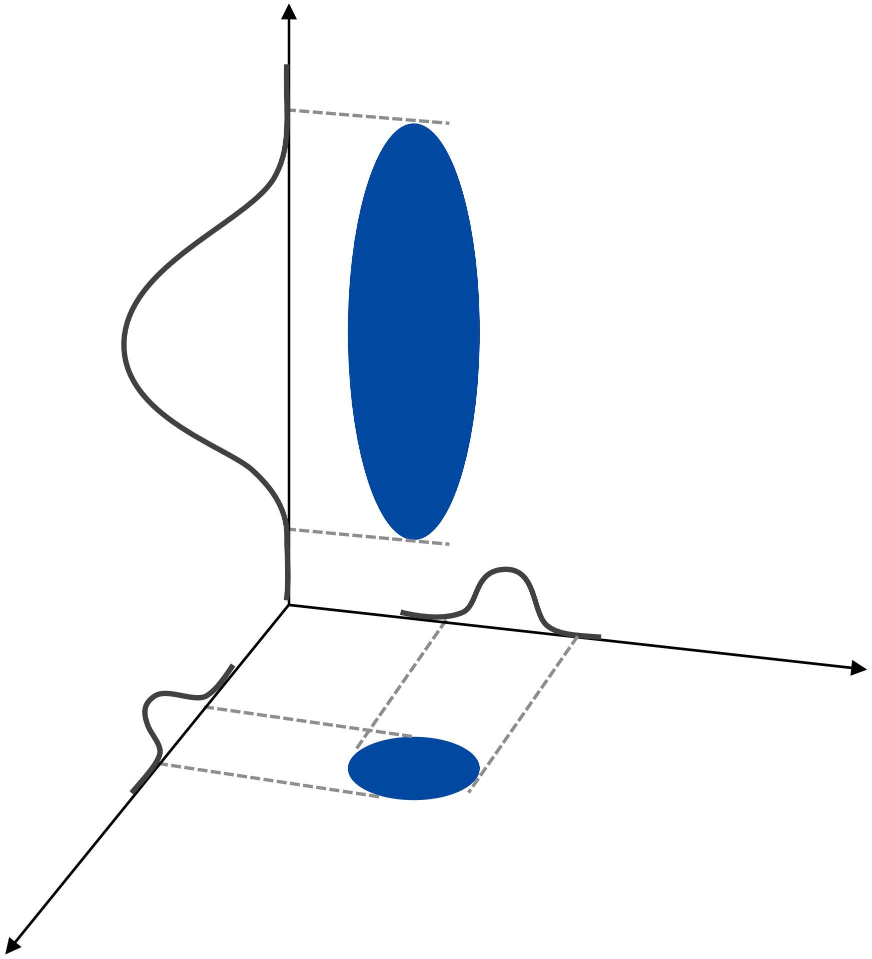
Multidimensional Scaling (MDS) (cont.)

- Given some similarity measures,
- Describe an object as a point in the multidimensional space,
- calculate the placement of data points,
- so that distance between points corresponds to the observed similarity

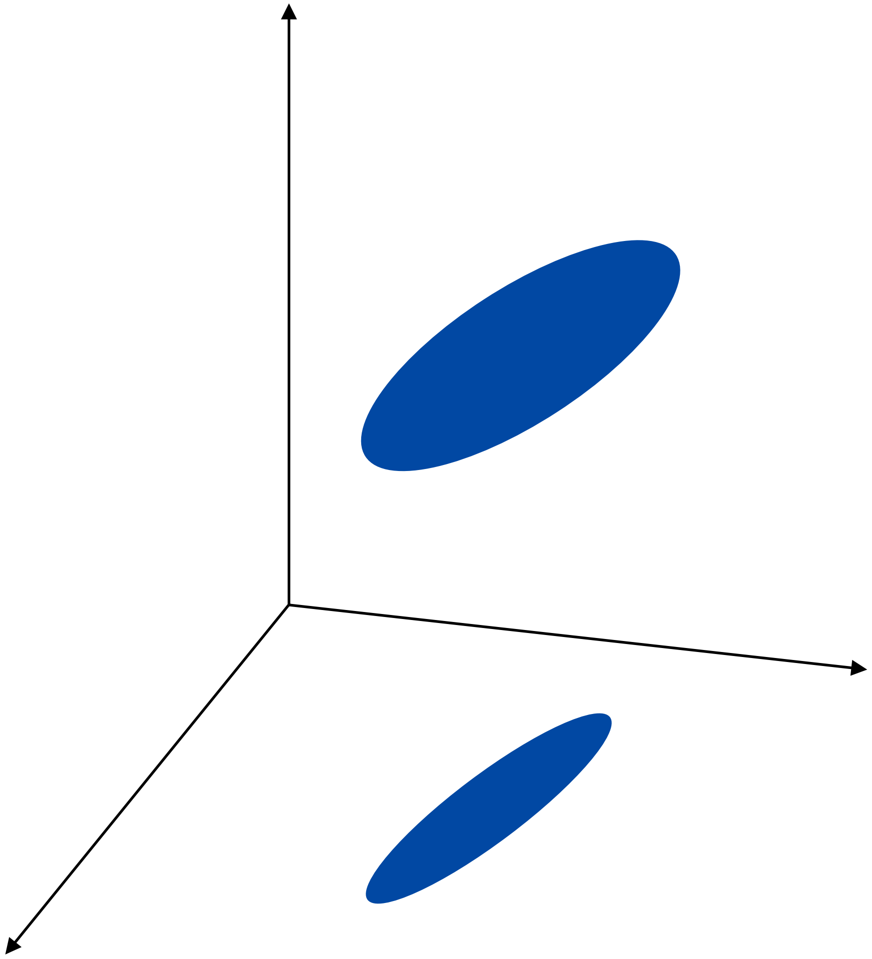
Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika*, *17*(4), 401-419.
https://doi.org/10.1007/BF02288916

Principal Component Analysis (PCA)

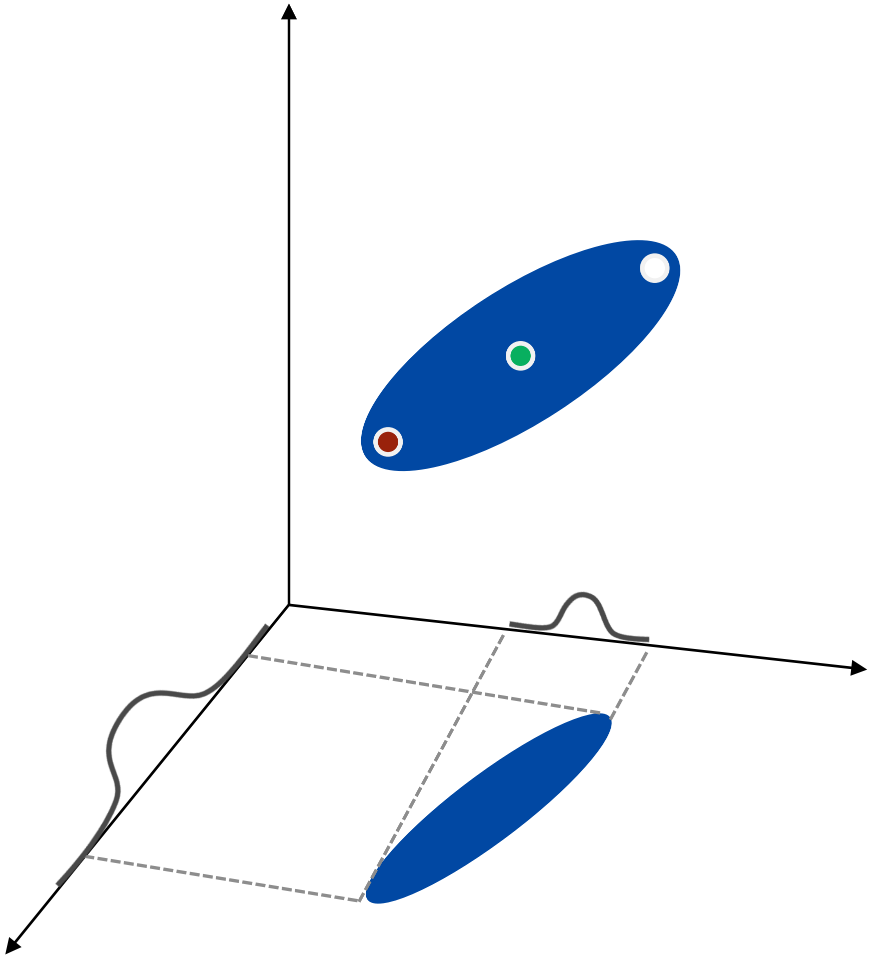
PCA (Principal Component Analysis) – Part I



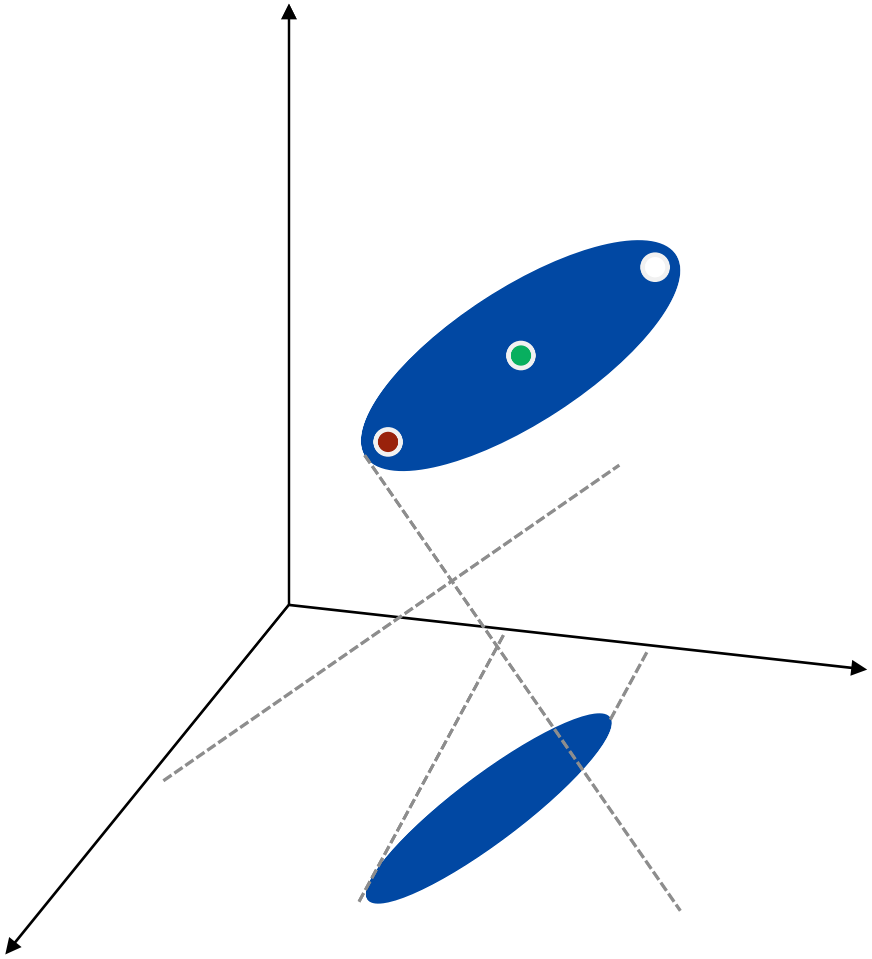
PCA (Principal Component Analysis) – Part I



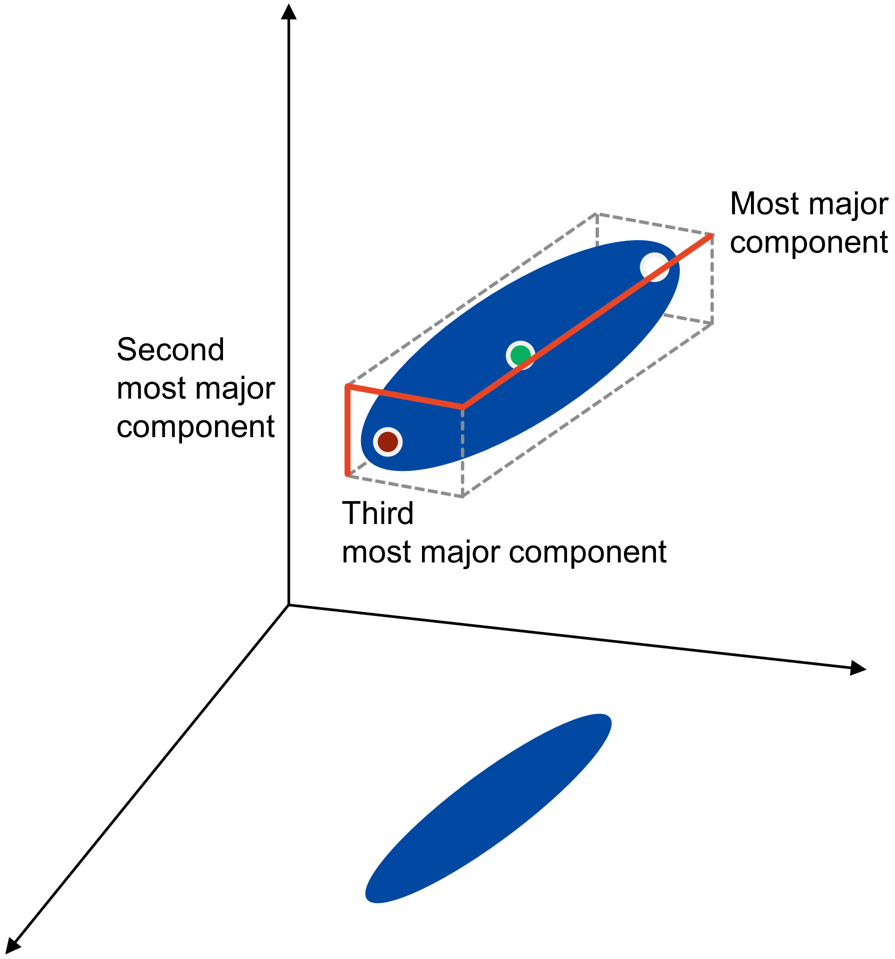
PCA (Principal Component Analysis) – Part I



PCA (Principal Component Analysis) – Part I



PCA (Principal Component Analysis) – Part I



Principal Component Analysis (PCA) – Part II

- Use variance
- find the main axis, which has the largest variance

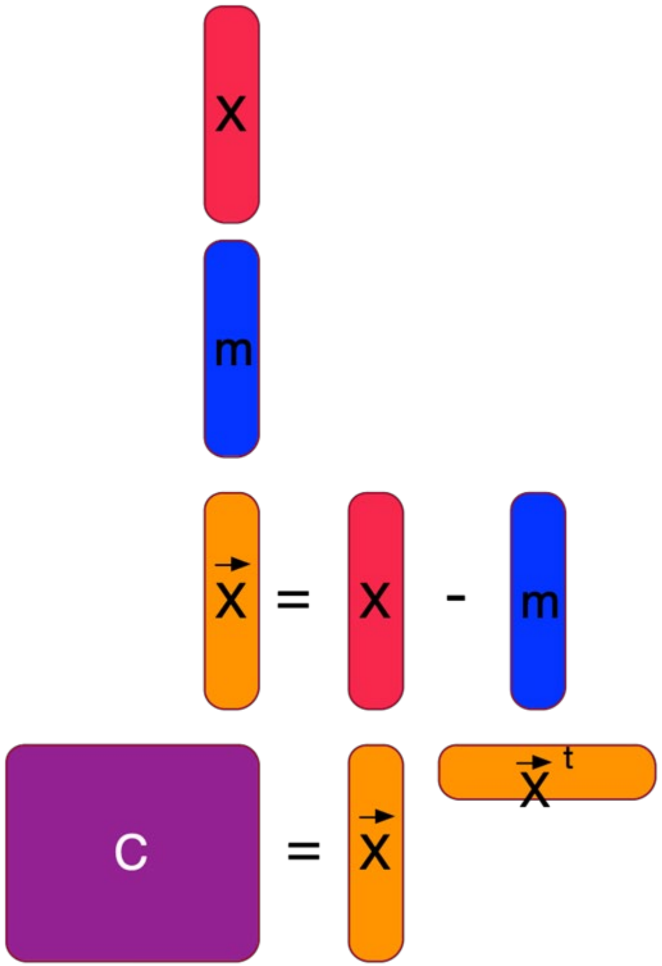
Principal Component Analysis (PCA) – Part III

$$m = \sum_{i=1}^{n} X_i$$

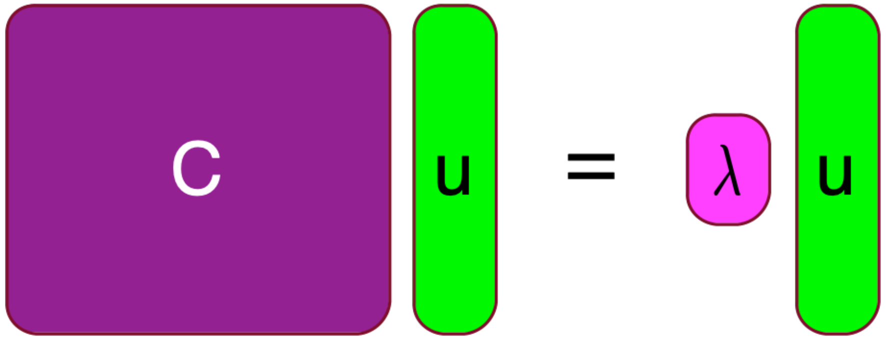
$$\vec{X} = X_i - m$$

$$\vec{C} = \frac{1}{n}\sum_{i=i}^{n} \vec{x}_i \vec{x}_i^t$$

$$\vec{C} = \frac{1}{n}\sum_{i=i}^{n} (\vec{x}_i - \vec{m})(\vec{x}_i - \vec{m})^t$$



Principal Component Analysis (PCA) – Part IV



$$\vec{C}\vec{u} = \lambda\vec{u}$$

$$where\ \vec{u} = \text{eigenvector}, \lambda = eigenvalue$$

Principal Component Analysis (PCA) – Part V

- Use the newly found principle axes to map nD data to 2D /3D space
- any problem??

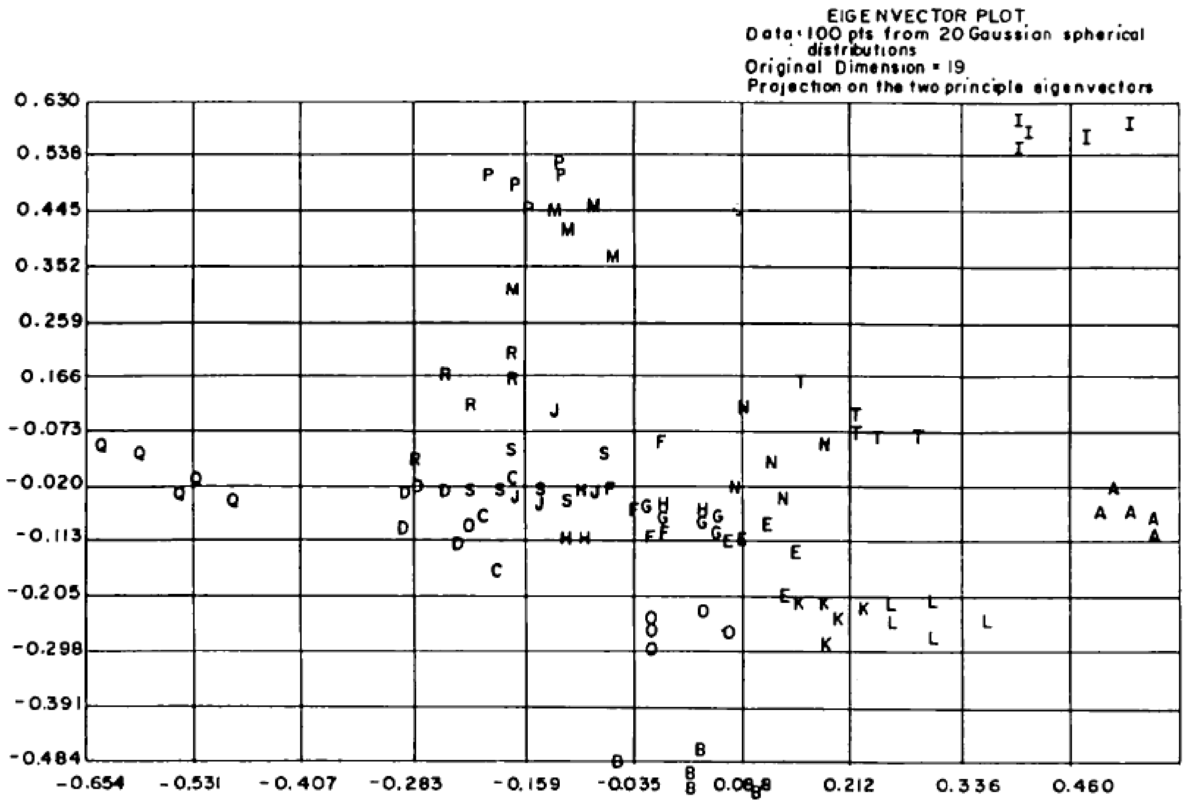


Fig. 7.

Metric space

- To define similarities, you need a metric space
- typical one....Euclidean space
- Minkowski's r-metric (Nonmetric Multidimensional Scaling)

$$d_{jk} = \left[\sum_{a=1}^{t} \left| x_{ja} - x_{ka} \right|^{r} \right]^{\frac{1}{r}}, (r \geq 1.0)$$

- LP-norm

Metric space (cont.)

$$d_{jk} = \left[\sum_{a=1}^{t} \left| x_{ja} - x_{ka} \right|^{r} \right]^{\frac{1}{r}}, (r \geq 1.0)$$

- when $r = 1$, Manhattan metric. (1-norm)
- $r = 2$, 2d Euclidean metric

  - EDM(a,b,c,d) : M(a) - M(b) == M(c) - M(d)

Sammon Mapping – Part I

- N L-dimensional data points :
  $X_1, X_2, \dots, X_N$
  $$d_{ij}^* = dist[X_i, X_j]$$

$$E = \frac{1}{\sum_{i<j}[d_{ij}^*]} \sum_{i<j}^{N} \frac{[d_{ij}^* - d_{ij}]^2}{d_{ij}^*}$$

- N d-dimensional data points :
  $Y_1, Y_2, \dots, Y_N$
  $$d_{ij} = dist[Y_i, Y_j]$$

Sammon, J. W. (1969). A Nonlinear Mapping for Data Structure Analysis. *IEEE Transactions on Computers*, *C-18*(5), 401-409. https://doi.org/10.1109/T-C.1969.222678

Sammon Mapping – Part II

$$E = \frac{1}{\sum_{i<j}[d_{ij}^*]} \sum_{i<j}^{N} \frac{[d_{ij}^* - d_{ij}]^2}{d_{ij}^*}$$

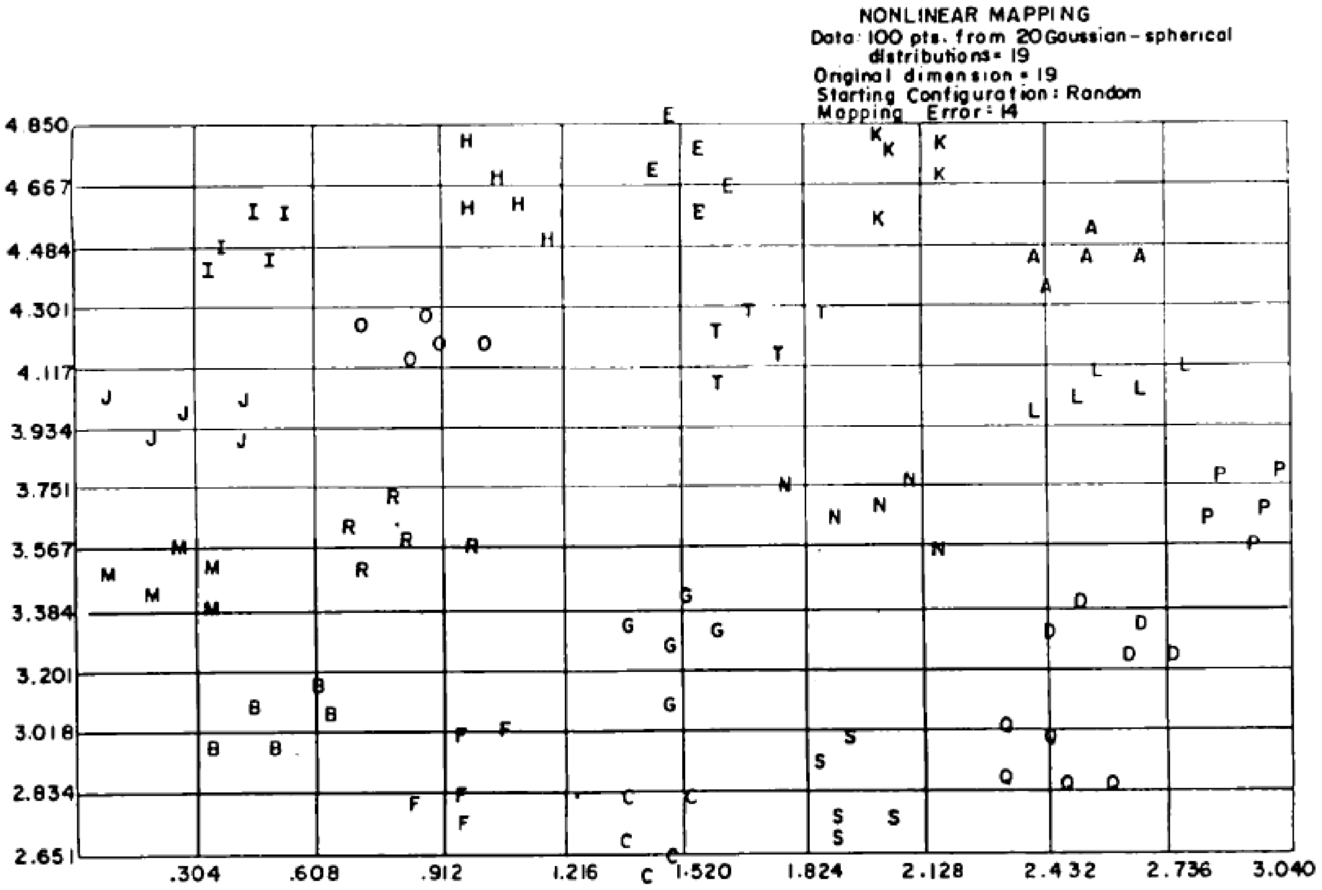


Fig. 6.

Sammon, J. W. (1969). A Nonlinear Mapping for Data Structure Analysis. *IEEE Transactions on Computers*, *C-18*(5), 401-409. https://doi.org/10.1109/T-C.1969.222678

Example: Sammon Mapping, Part III

Sammon Mapping

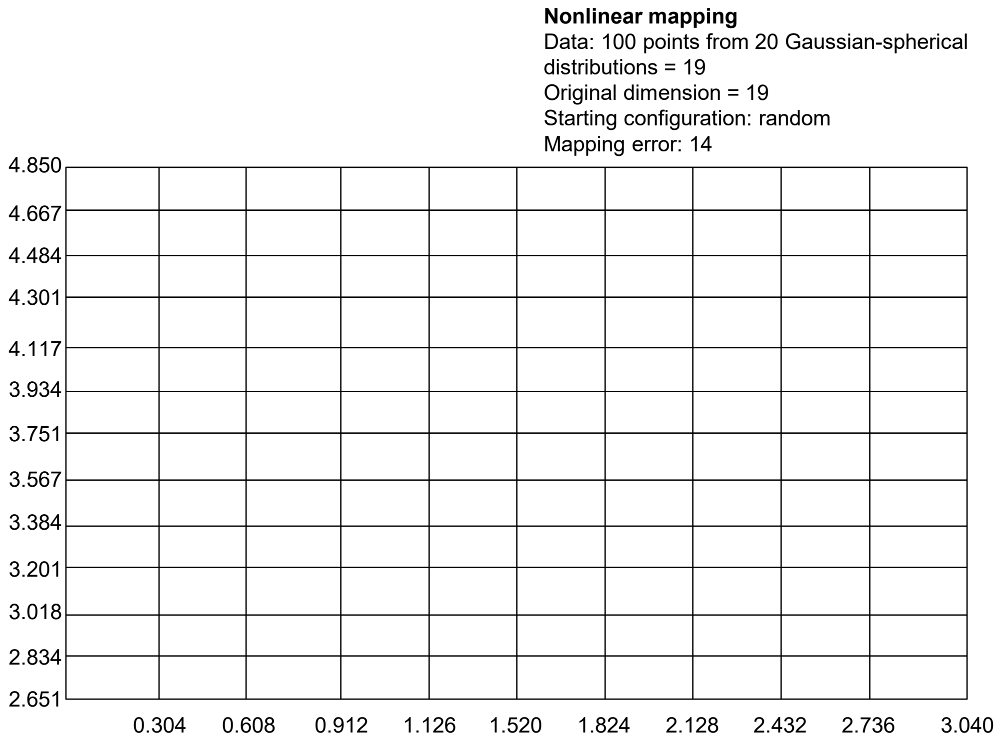


Figure 6

PCA

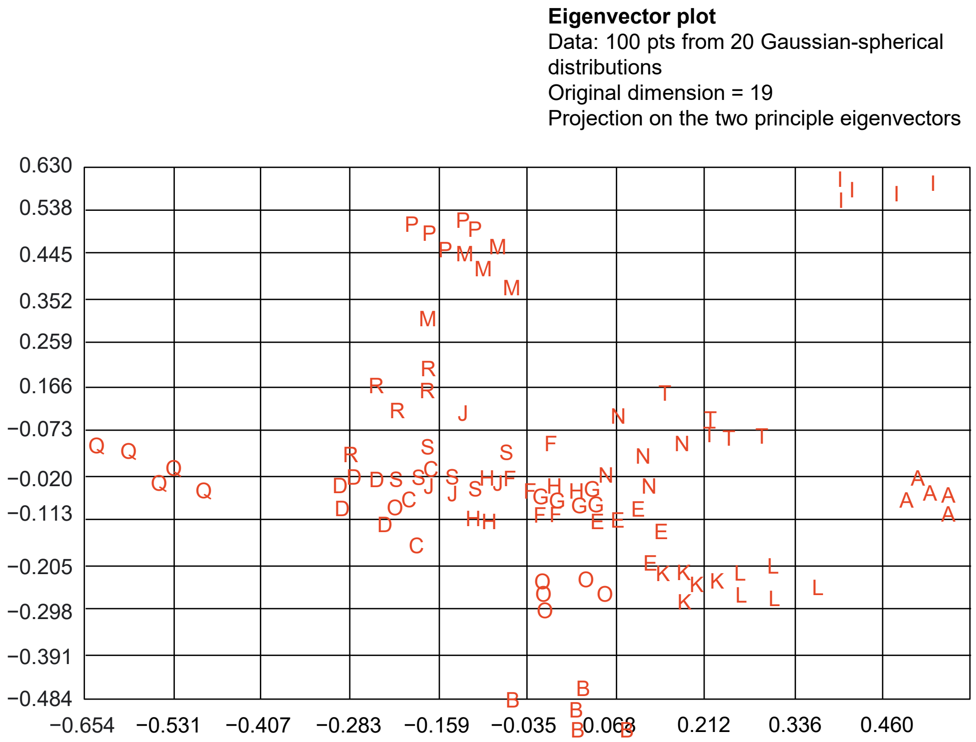


Figure 7

Example: Sammon Mapping, Part III

Sammon Mapping

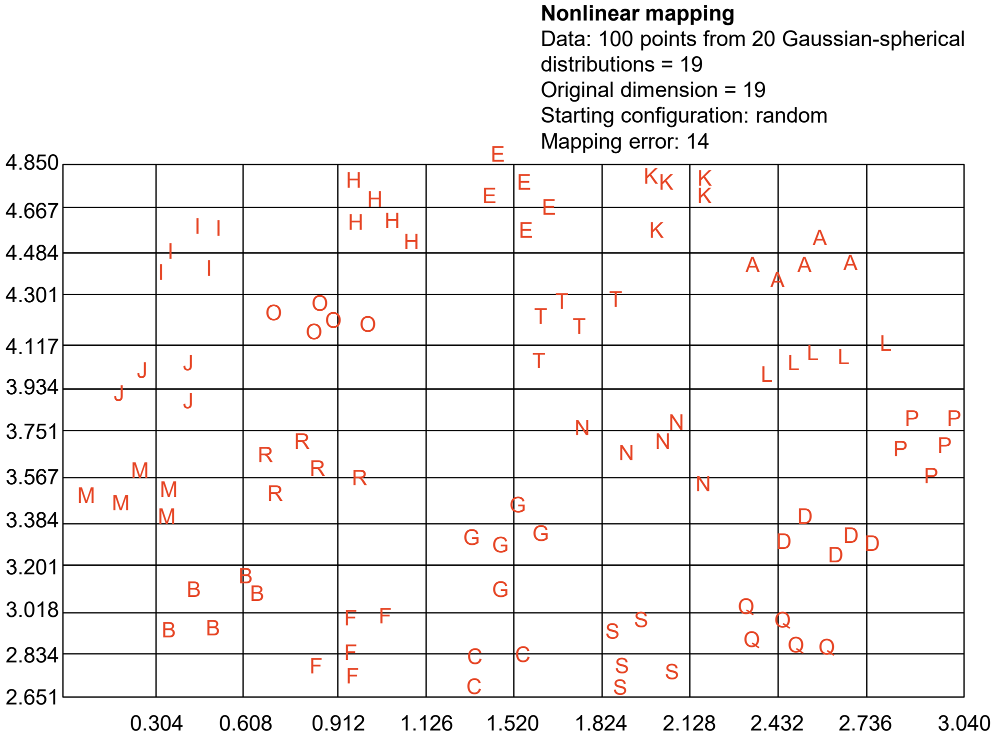


Figure 6

PCA

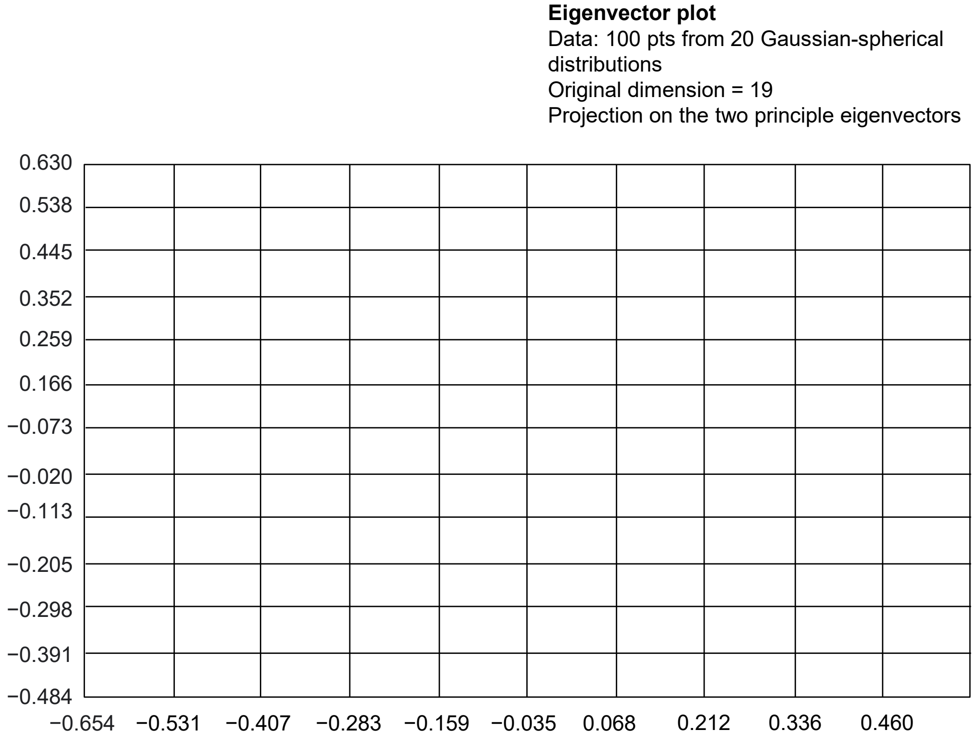


Figure 7

Example: Sammon Mapping, Part III

Sammon Mapping

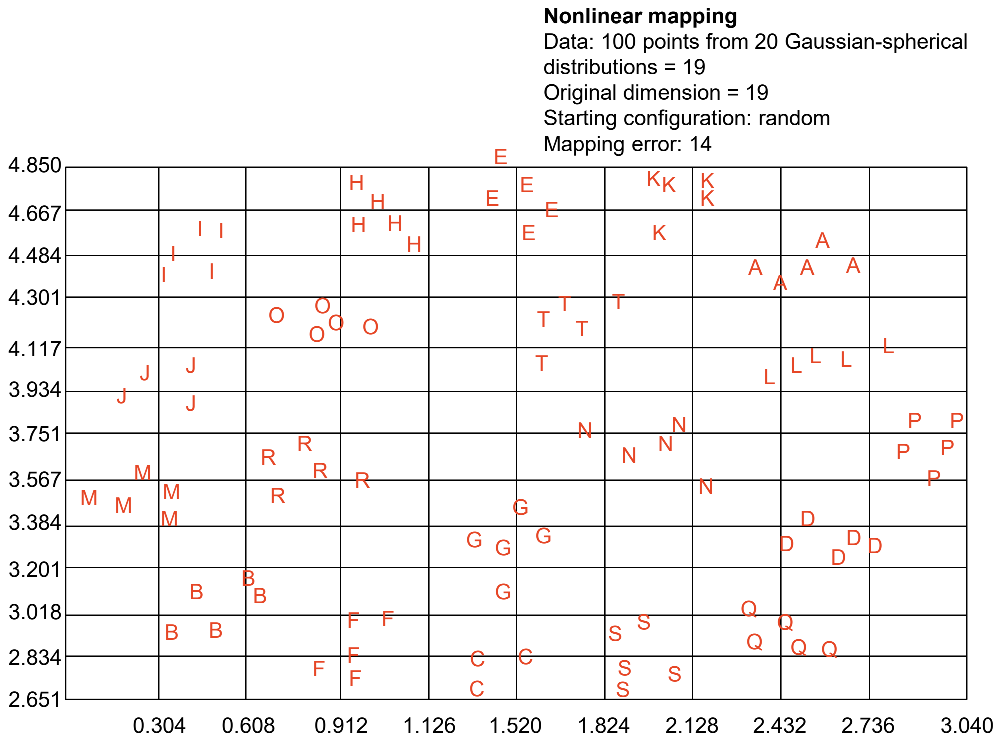


Figure 6

PCA

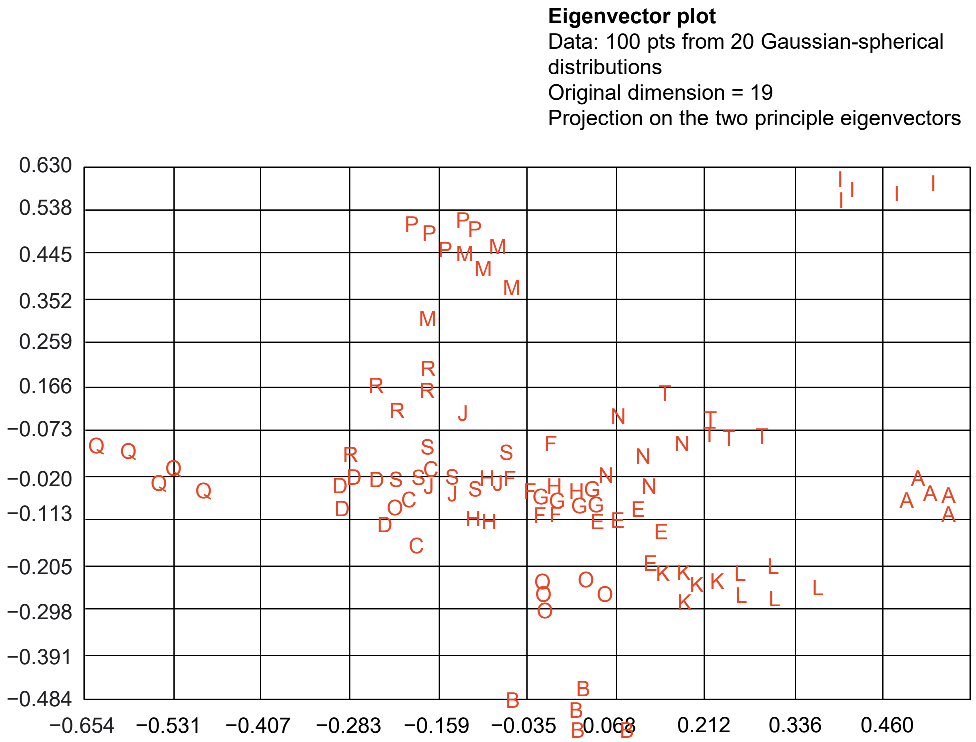
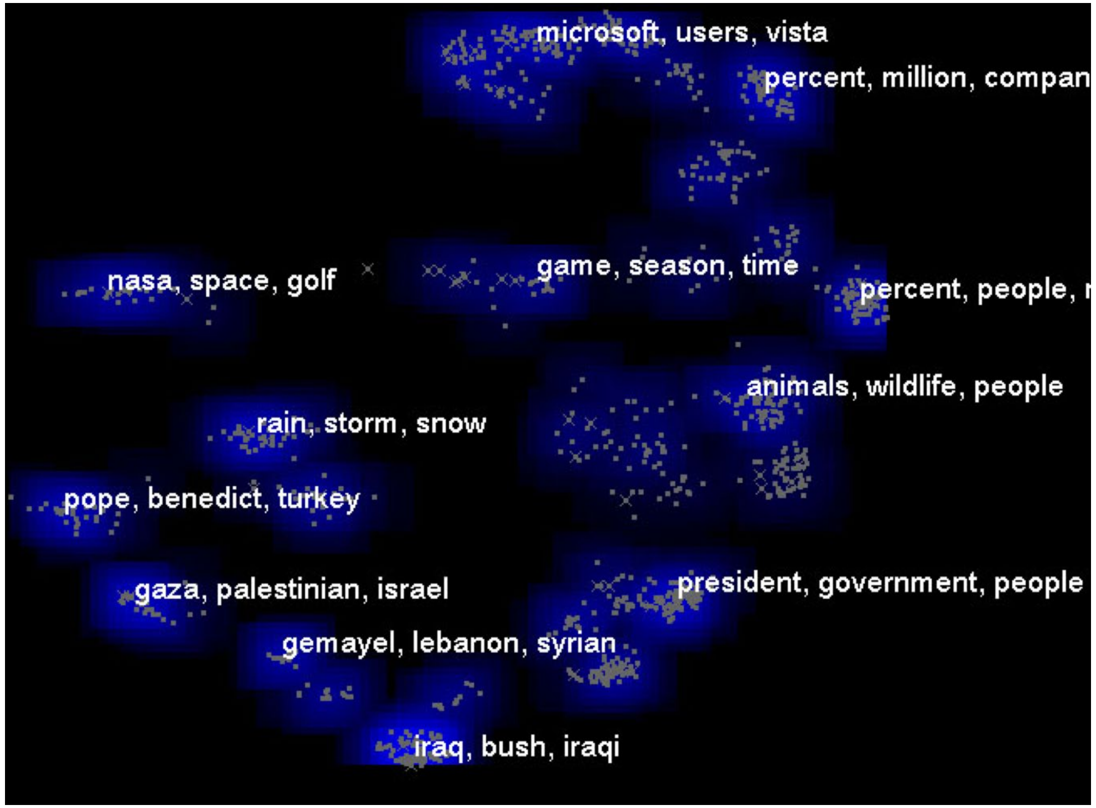


Figure 7

Multidimensional Scaling as Spatialization

- Observe target objects
- Observed data could be
  - Categorical
  - Ordinal
  - Numerical
- form a vector (n-dimension)
- use MDS to find 2D (or 3D) location



Self-Organising Map (SOM)

What is Self-Organising Map?

- The striate cortex of a cat
  Neurons respond to particular
  angles of light stripes

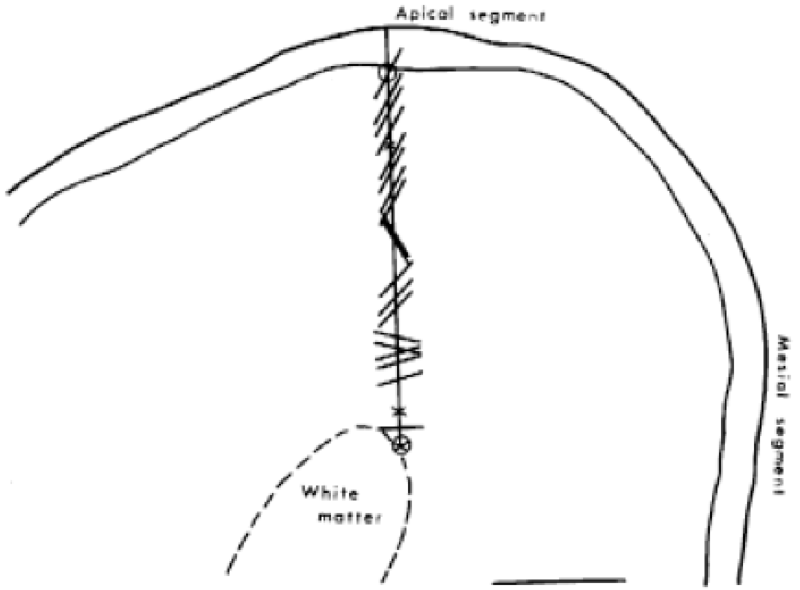


FIG. 4. Reconstruction of a microelectrode penetration through the postlateral gyrus of the left hemisphere. This 2½-month-old kitten had its right eye covered from birth by lid suture. Lines intersecting the electrode track represent cortical cells; directions of these lines indicate the receptive-field orientations. Crosses indicate cortical cells uninfluenced by light stimulation. Simultaneous recordings from two units, which occurred three times in this penetration, are each indicated by only one line or cross. A lesion was made while recording from the first unit, and another at the end of the penetration; these are marked by small circles. The ocular-dominance distribution of units recorded in this penetration is shown in Fig. 3. All fields positioned 5–6° to the left of the area centralis, slightly below the horizontal meridian. Scale, 0.5 mm.

Hubel and Wiesel 1963

Artificial Model of Self-Organization

- von der Malsburg (1973) simulated feature extracting cells organizing in a 2D space
- lateral excitatory/inhibitory connections.

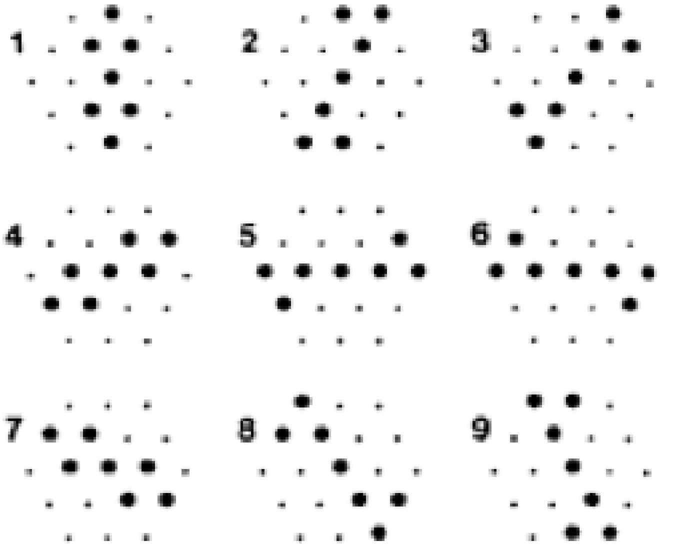


Fig. 5. The standard set of stimuli used on the model "retina". Large and small dots represent active and non-active fibres respectively

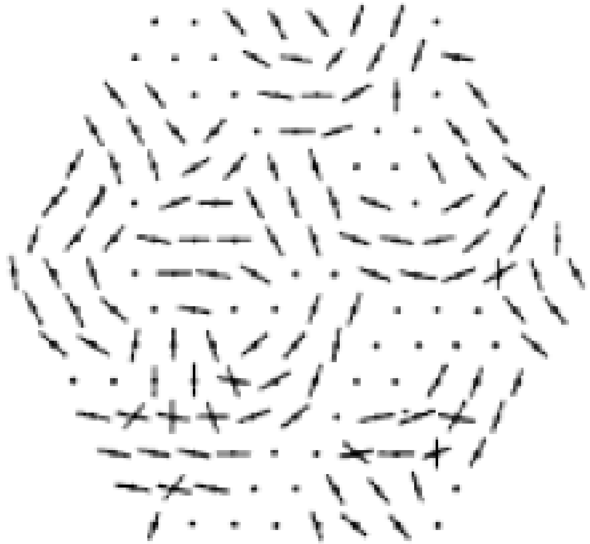


Fig. 12. View onto the cortex. Each bar indicates the optimal orientation of the *E*-cell (for definition see text). Dots without a bar are cells which never reacted to the standard set of stimuli. Two bars indicate two separate sensitive regions

Mathematical Model (for Biological SOM)

- Amari (1980)
  topographic organization of
  nerve fields
- Linsker (1986)
  Emergence of column orientation
- Tanaka (1990)
  Cortical Map formation model

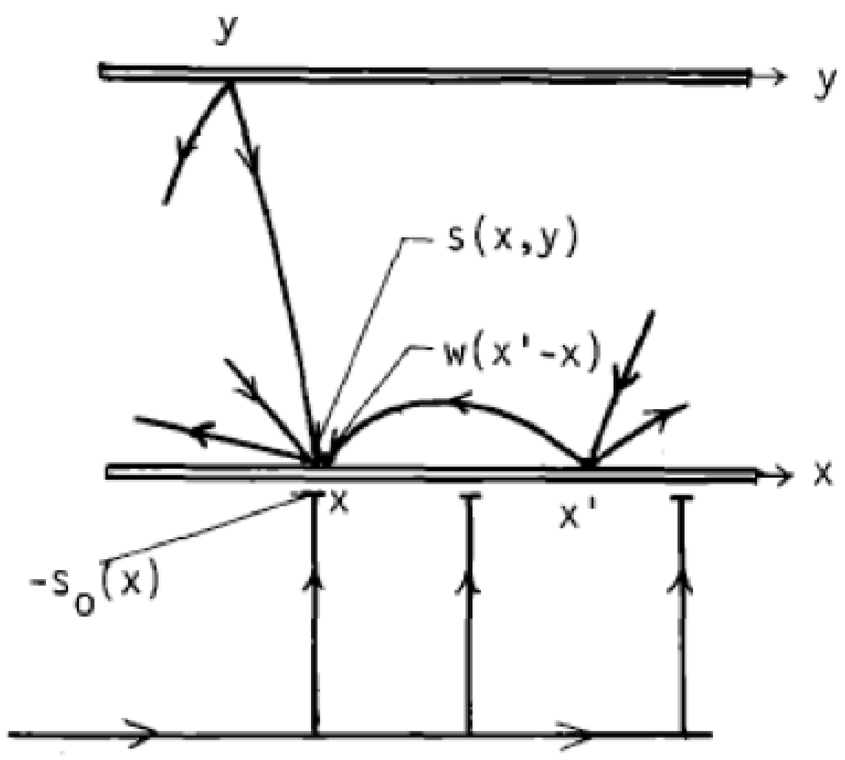


Figure 1. Connection of neural fields

Amari (1980)

Computational Model

- Kohonen (1982)
  Topological mapping
  - cf. Sammon mapping (1969)
    Nonlinear Mapping from n-dim to 2D (MDS).
- build a semantic map (Ritter/Kohonen:1989)
- concept building (Ichiki, et al.: 1991)
- information visualization (Takatsuka:00, Skupin:03)

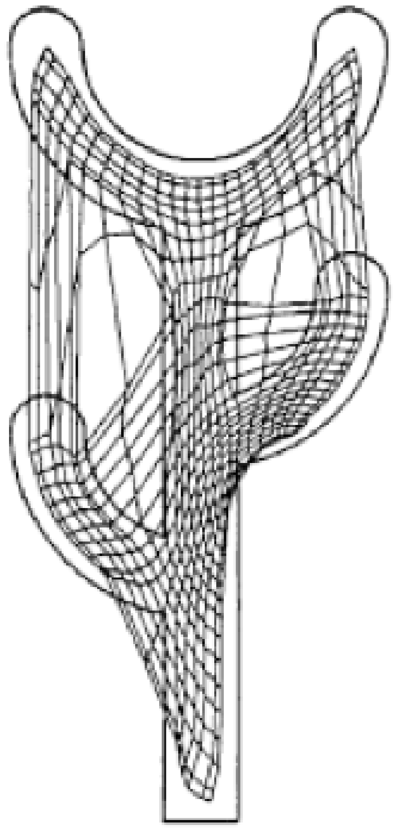
Kohonen's SOM

Kohonen’s Self-Organizing Map – Part I

- A type of Artificial Neural Network
- Self-Organizing Map
- Telvo Kohonen (1982)
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59-69. https://doi.org/10.1007/BF00337288
- unsupervised learning
- No biological equivalence
- Try to find an equal probability density function.

Kohonen’s Self-Organizing Map – Part II

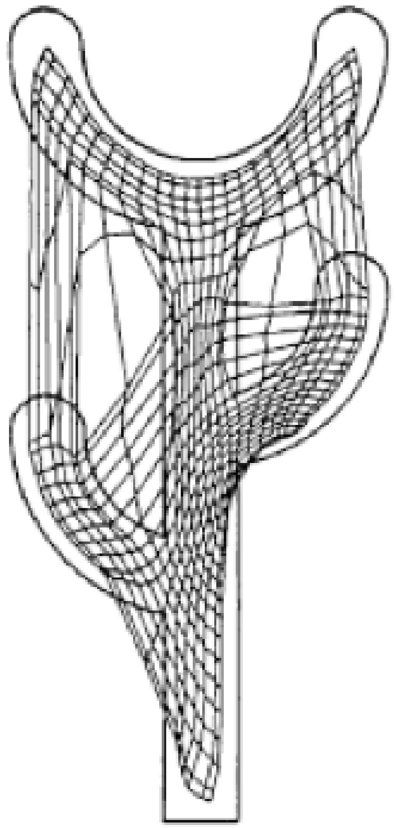
- Neurons are placed on a 2D grid.
- Each neuron has an n-dim. weight vector (representing a point in the n-dim space)
- Through competitive learning algorithm, neurons update their weight vectors.
- Use a neighbourhood function to preserve topology



Kohonen (1989)

Kohonen's Self-Organizing Map – Part III

- input vectors,
  $$x_i \; : \; i \; = \; 0, \ldots, N$$
- each neuron's connection weight,
  $$w_j : j \; = \; 0, \ldots, M$$
- Best Matching Unit, w_j^b
- $w_j^{new} = w_j + \alpha(r)(x_i - w_j)$
- Monitor Quantization error,
- $Q(r) = \sum_{i=0}^{N} \left\{ \left\| x_i - w_j^b \right\|^2 \right\}$



Kohonen (1989)

Example of Kohonen's SOM – Part I

- Input Data (multidimensional)

Animal\ Attribute		dove	hen	duck	goose	owl	hawk	eagle	fox	dog	wolf	cat	tiger	lion	horse	zebra	cow
is	small	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0
	medium	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
	big	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
has	2 legs	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
	4 legs	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	hair	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	hooves	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	mane	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0
	feathers	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
likes to	hunt	0	0	0	0	1	1	1	1	0	1	1	1	1	0	0	0
	run	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1	0
	fly	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0
	swim	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0

Example of Kohonen's SOM – Part II

- Result (13 D to 2D)

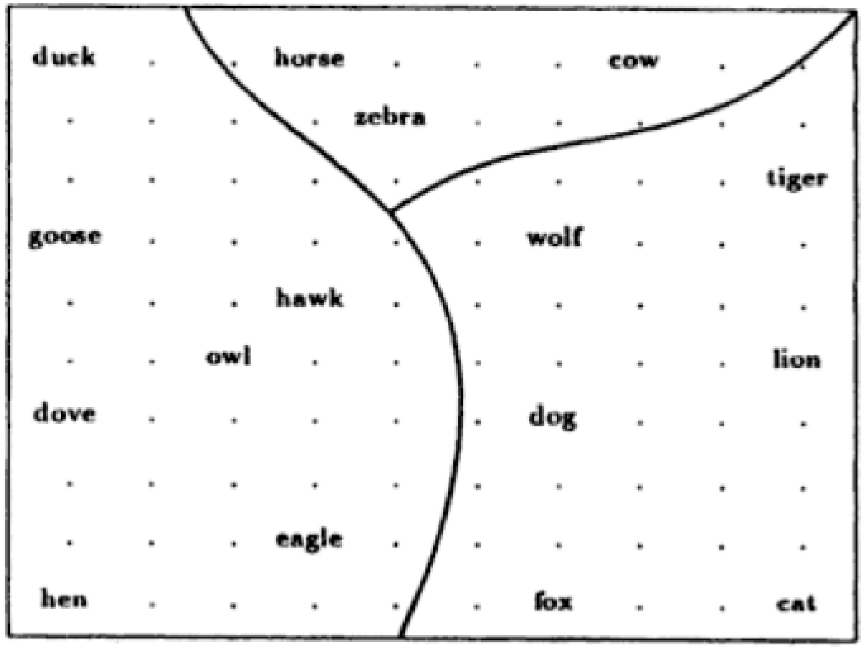


Fig. 3.29. After the network had been trained with inputs describing attribute sets from Table 3.6, the map was calibrated by the columns of Table 3.6 and labeled correspondingly. A grouping according to similarity has emerged

THE UNIVERSITY OF SYDNEY