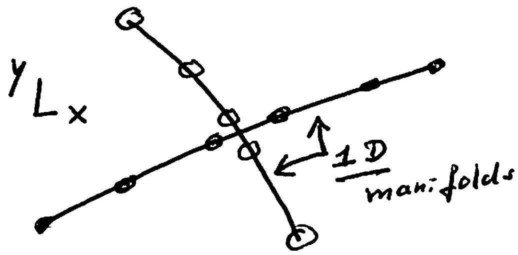
Bernd Hamann
Stratovan

9 Jan 2017

More on Sub-space Clustering etc.

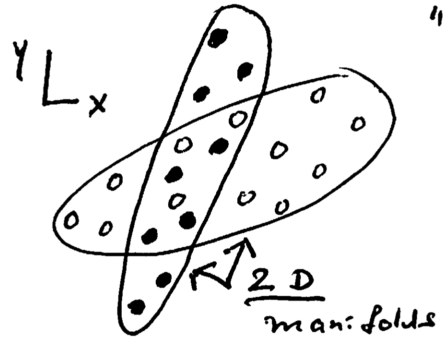
→ Examples of distributions:



{•} cluster C1
{o} cluster C2

"Points lie on 1D curves in 2D space."

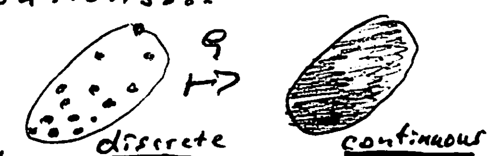
⇒ Can define density functions ρ along/on the 2 curves...



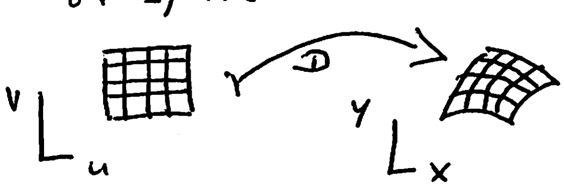
"C1 and C2 define 2D regions in 2D space."

⇓ Can define density functions ρ for each of the 2 point distributions...

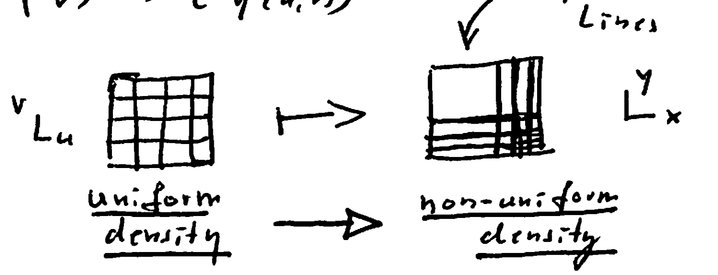
- Construct a smooth density function ρ: $^{y}\llcorner_{x}$



- **OR**: Think of a point distribution as a "deformation" of space:



$\begin{pmatrix} u \\ v \end{pmatrix} \overset{D}{\mapsto} \begin{pmatrix} x(u,v) \\ y(u,v) \end{pmatrix}$



⇒ Is it useful to interpret a point distribution as a "deformation of (uniform) space"?

⇒ Can consider a set $S = \{p_i\}$ of points and define a model/density function ρ. Can then determine to which of multiple models/density functions ρ is "closest to."

⇒ To compare 2 deformations – and define a distance metric – one can consider various derivatives: $\begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} x(u,v) \\ y(u,v) \end{pmatrix}$

1st deriv: $\begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix}$; 2nd deriv: x_{uu}, x_{uv}, x_{vv}, y_{uu}, y_{uv}, y_{vv} ; ...

Bernd Hamann
Stratovan

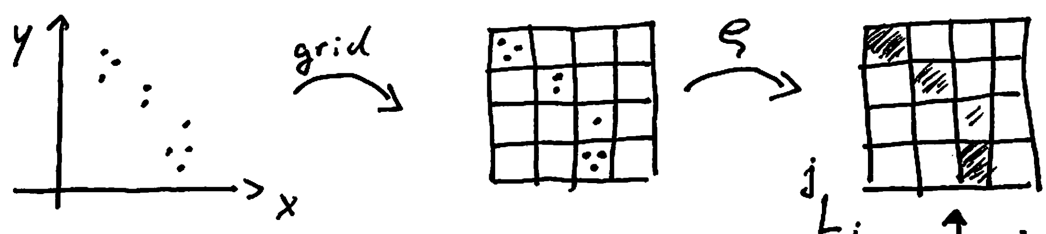
23 Jan 2017

... Cont'd: "Distance between 2 distributions / density functions" ...

{ Would it be possible at all to determine for a single point p whether it belongs to one of two sub-space clusters when these 2 clusters overlap ??? }

- **The 2D Case**

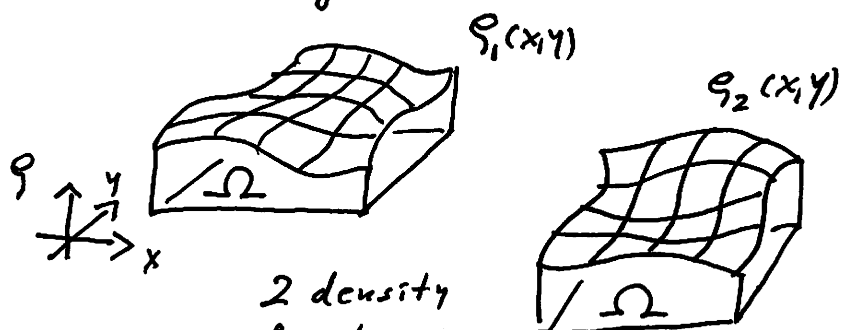
→ given: points in 2D plane, - wanted: density function ρ



$\rho_{i,j}$ = # of points in cell

↑ piecewise constant density function $\rho(x,y)$
↳ Can define a smoother density function from this!

⇒ Can compare 2 density functions by computing a "meaningful" distance measure:



2 density functions over same domain Ω

- distance of ρ_1 and ρ_2:

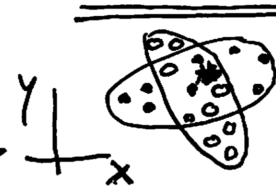
$$d = \int_\Omega (\rho_1 - \rho_2)^2 \, dx\, dy$$

or $$d = \sqrt{\frac{1}{\text{area}(\Omega)} \int_\Omega (\rho_1 - \rho_2)^2 \, dx\, dy}$$

or ...

- Practical problem: Given 1 point / feature vector (of an imaged object), does this point "belong to" a density ρ_1 or density ρ_2 (with ρ_1 and ρ_2 overlapping, i.e., being different from 0 in same region)?

⟺



$\{\circ\} \Rightarrow \rho_1(x,y)$
$\{\bullet\} \Rightarrow \rho_2(x,y)$
$*$ given single point

- TEST: 'INSERT' $*$ INTO SET $\{\circ\} \Rightarrow \bar{\rho}_1(x,y)$
  'INSERT' $*$ INTO SET $\{\bullet\} \Rightarrow \bar{\rho}_2(x,y)$
  → COMPUTE DISTANCES $d(\rho_1, \bar{\rho}_1)$ AND $d(\rho_2, \bar{\rho}_2)$
  → $*$ BELONGS TO THAT CLUSTER WHOSE DENSITY IS CHANGED LESS AFTER INSERTION OF $*$!

Bernd Hamann
Stratovan

23 Jan. 2017

■ **Support Vector Machines (SVMs) – Specific Issues: Dimensionality, Geometry, Feature Vector, Classification**

[Reference: ● "Geometry of Support Vector Machines";
Martin Jaggi
Seminar Talk of [redacted], 28 Feb. 2008
[+ Papers of Martin Jaggi, EPFL]]

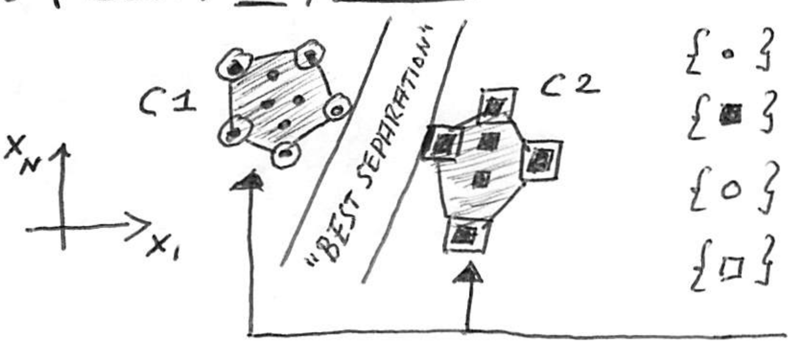
⇒ The theory of SVMs is related to the theory of CONVEX HULLS etc., studied in computational geometry.

• **Key Problem(s):**

"What are the minimally necessary attributes/components of an N-dimensional feature vector (point) associated with 'parts'(*) of objects of different material classes allowing one to determine to what class a 'part' belongs?"

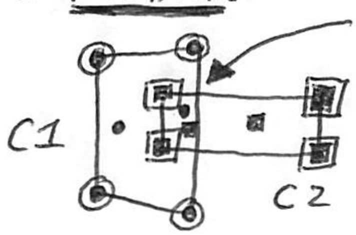
((*) 'Parts' can be the individual voxels of an object.)

• (Best) Separation Done in N-dim. Feature Space:



{•} cluster 1, C1
{■} cluster 2, C2
{○} support points of C1
{□} support points of C2
convex hulls of C1 & C2
⇒ only need representation of boundary of convex hulls!

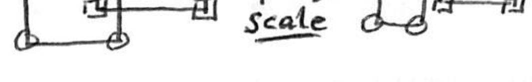
→ **Problem:**



overlapping convex hulls!

⇒ "Solution/idea":

Shrink (scale-down) convex hull (boundary), eliminating overlap:



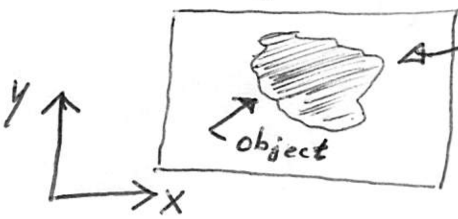
new problems...

Bernd Hamann
Stratovan

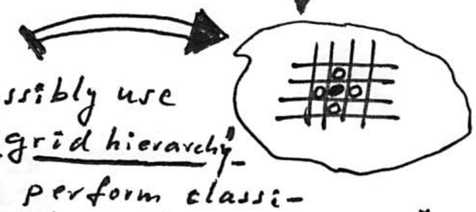
30 Jan. 2017 /
6 Feb. 2017

■ SVMs – cont'd. : Kernel Functions etc.

→ 2D image:



"Kernel function" – Define object in terms of a function / set of variables...



→ Possibly use a "grid hierarchy" to perform classification at multiple scales!

"Object," understood as a set of pixels (or voxels) can locally be represented via a function defined by a (small) neighborhood of pixels (voxels).

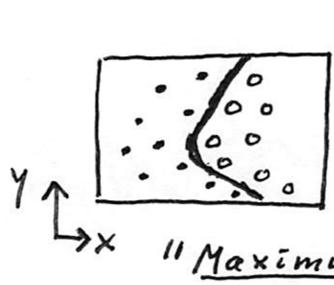
→ For example, (local)

$$f(x,y) = c_{00} + c_{10}x + c_{01}y + c_{11}xy + c_{20}x^2 + c_{02}y^2 + \ldots$$

⇒ Various differential properties of f can be used to define (local) "features", e.g., $f, f_x, f_y, f_{xx}, \ldots$

⇒ Thus, one can establish an N-dimensional "feature vector" (feature point) for each pixel (voxel) constituting an object. (Scale and rotation invariance?)

→ In N-dim. feature space, two objects / materials can possibly be separated via one separating hyper-plane – leading to an object boundary in 2D (or 3D) space that is 'curved':



separation (and classification) in image space

"Maximum Margin Classification"

→ "A crucial problem" (to be solved):

IS IT POSSIBLE TO DEFINE N AND THE TYPES OF THE N COMPONENTS OF A FEATURE VECTOR SUCH THAT UNIQUE CLASSIFICATION IS POSSIBLE?

⟺ DOES SEPARATING PLANE EXIST? ■ BH