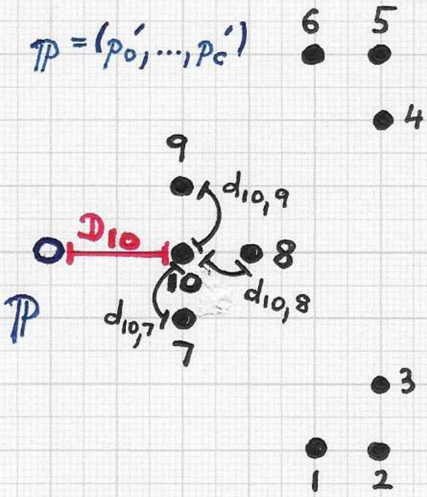
Bernd Hamann

Stratovan

10/19/2022

■ OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

• Laplacian eigenfunctions and neural networks:...



Abstract illustration of ten class signatures $S_1, \ldots, S_{10}$ and an unclassified $\mathbb{P}$-tuple. This geometrical sketch is intended to portrait distances between $\mathbb{P}$ and the signatures, with D_{10} being the minimal distance between $\mathbb{P}$ and the ten signatures.

Further, since S_{10} is the closest signature, from $\mathbb{P}$'s perspective, we also high-light the three closest signatures of S_{10}; they are signatures S_7, S_8 and S_9, from S_{10}'s perspective. The associated distance values are $d_{10,7}$, $d_{10,8}$ and $d_{10,9}$.

By considering the value of D_{10} in conjunction with the values $d_{10,7}$, $d_{10,8}$ and $d_{10,9}$ – AND D_7, D_8 and D_9 – one can evaluate "classification quality."

• Note. We briefly describe one possibility to use the distances D_{cl} and $d_{cl1,cl2}$ to "evaluate the reliability" of classification: An unclassified tuple $\mathbb{P} = (p_0', p_1', \ldots, p_c')$ must be classified. Ten material classes, class 1 ... class 10, are represented in the material sample database, together with their signatures, $S_1 \ldots S_{10}$. First, we compute the ten values of D_{cl}, $cl = 1 \ldots 10$, i.e., the distances from $\mathbb{P}$ to the ten class signatures. For example, D_{10} has the smallest value and (initially) $\mathbb{P}$ is classified as "10." The figure (left, top) sketches this example in an abstract geometrical way. Second, we now calculate/look up the ordered signature distances $d_{10,cl2}$. The ordered set of indices $\{cl2\}$ is $\{7, 8, 9, 1, 3, 4, 6, 2, 5\}$. Thus, the signatures S_7, S_8 and S_9 are the three closest signatures of S_{10}.

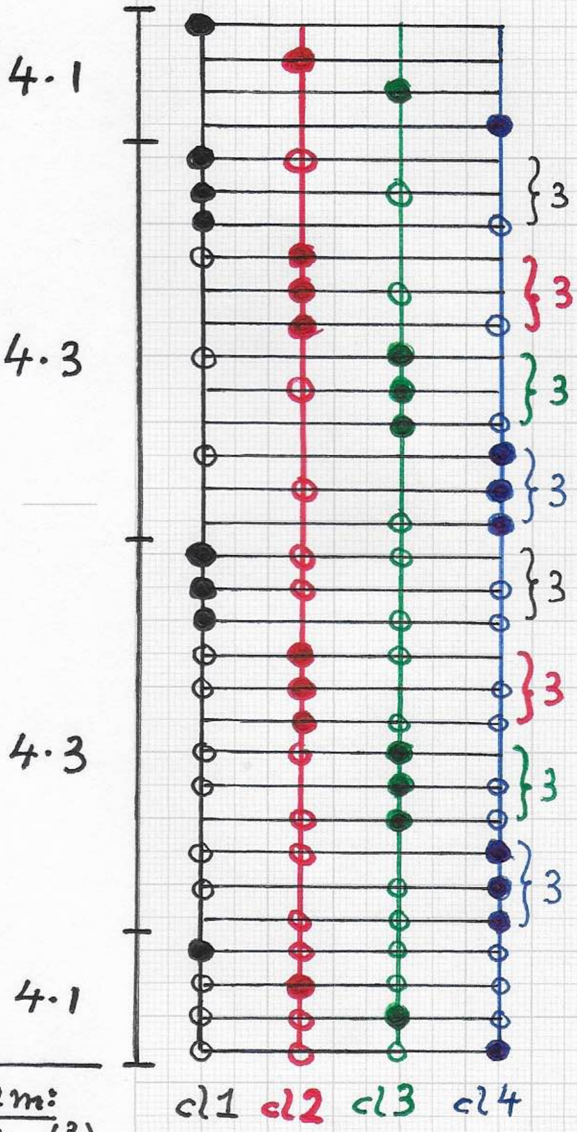
...

■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:...

Considering the example described on the previous page, how can the three smallest values $d_{10,cl}$, $cl \in \{7, 8, 9\}$, be used to further assess the (initial) classification result for the tuple $\mathbb{P}$? At this point, $\mathbb{P}$ is viewed as a class-10 tuple, as D_{10} has the smallest value of all D_{cl} distances, $cl = 1 \ldots 10$. Let us consider the following cases, for example: (i) The values of distances D_7, D_8 and D_9 are all "much larger" than D_{10}; thus, the (initial) classification of $\mathbb{P}$ as "class 10" can be assumed to be of very high probability / certainty.
(ii) The values of D_{10} and D_7 are "nearly identical," and the values of D_8 and D_9 are "much larger" than D_{10}; thus, the (initial) classification of $\mathbb{P}$ as "class 10" can be assumed to be less certain, and "class 10" and "class 7" might be equally probable classifications of $\mathbb{P}$.

...



⇒ sum:

$$4 \cdot \binom{3}{0} + 4 \cdot \binom{3}{1} + 4 \cdot \binom{3}{2} + 4 \cdot \binom{3}{3} = 4 \cdot (1+3+3+1) = 4 \cdot 8 = \underline{\underline{32}}$$

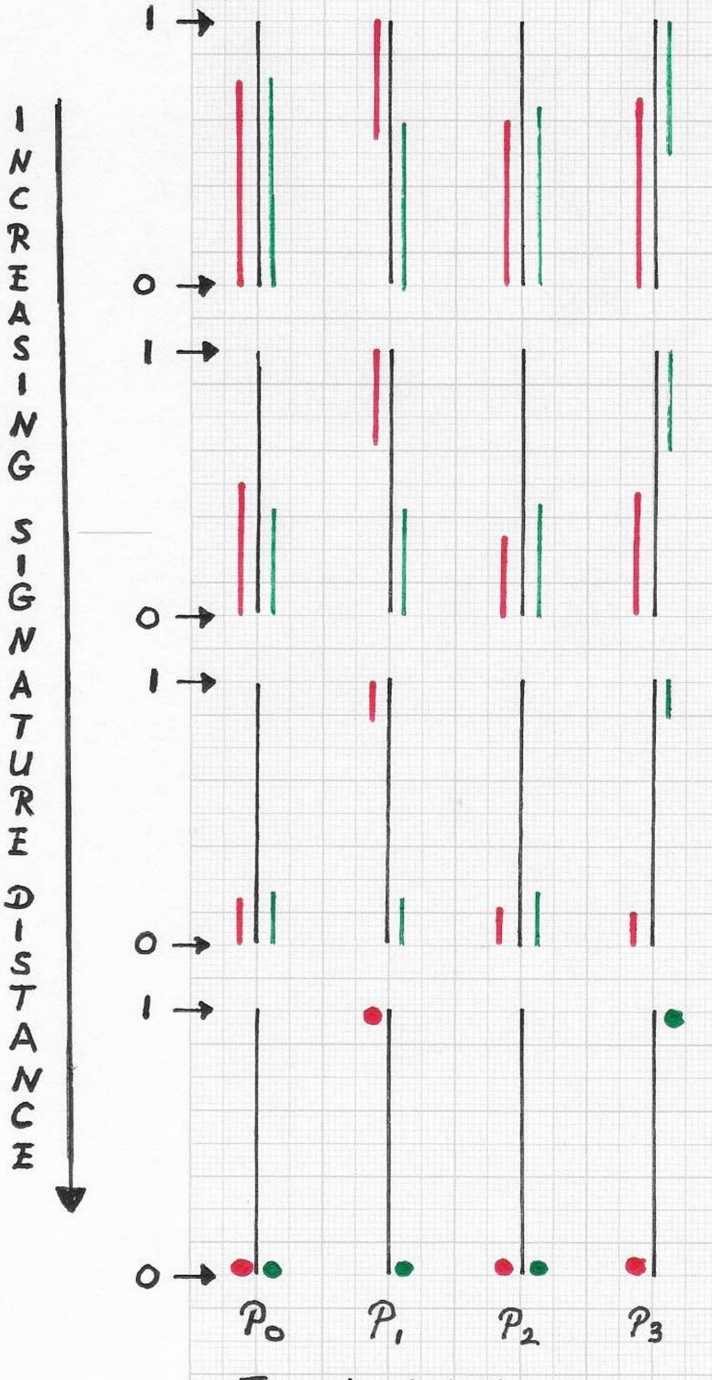
Four-class classification.

Bullets (●, ●, ●, ●) indicate initial classification based on D_{cl} values.

Circles (○, ○, ○, ○) indicate potential "secondary classifications" of equal probability.

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

- Laplacian eigenfunctions and neural networks: ...



From top to bottom: sequence of confidence intervals of two material classes, 1 and 3. The lengths of the intervals decrease, thereby increasing "discriminative power."

The example illustrated in the figure on the previous page used $C=4$ classes. Therefore, potential "secondary classifications" – in addition to the "primary classification" resulting from the D_{cl}-values – must be drawn from the other $(C-1)=3$ classes. Thus, considering all combinatorial possibilities, 0, 1, 2 or 3 "secondary classifications" can be made. Concerning the four-class classification table on the previous page, the number of rows / possibilities is 32, which is

$$4\cdot\binom{3}{0}+4\binom{3}{1}+4\cdot\binom{3}{2}+4\binom{3}{3}$$

$$=4\cdot\sum_{i=0}^{3}\binom{3}{i}=4\cdot 2^3=32.$$

The general case including C classes therefore has

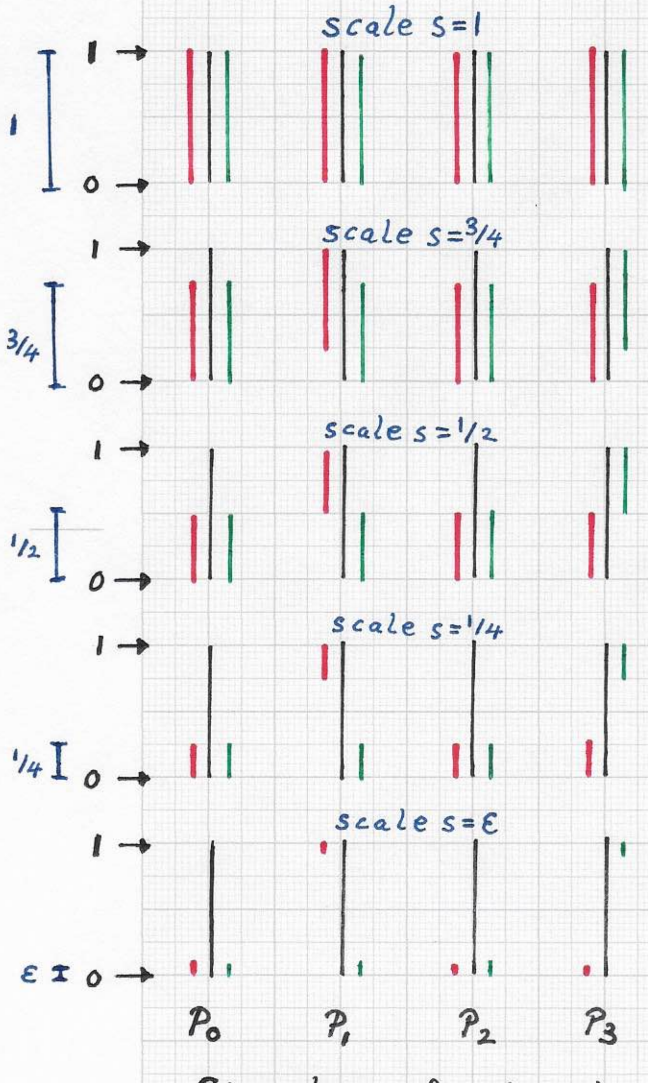
$$C\cdot\sum_{i=0}^{C-1}\binom{C-1}{i}=C\cdot 2^{C-1}$$

combinatorially viable possibilities. In order to keep computational complexity reasonably low, the number of classes used as "secondary classes" must be a small number.

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

- Laplacian eigenfunctions and neural networks:...



Signatures of material classes 1 and 3. Top: signatures are identical and cover the entire interval [0,1]; they have no "discriminative power." Bottom: signatures clearly represent classes 1 and 3, as they nearly reflect the binary tuples (0,1,0,0) and (0,0,0,1), respectively.

For example, if one were to permit either 0 or 1 "secondary classifications" – in addition to the C possible "primary classifications" – the the total number of possibilities would be

$$C\cdot\left(\binom{C-1}{0}+\binom{C-1}{1}\right)=C\cdot(1+C-1)$$
$$=C\cdot C=C^2.$$

Nevertheless, in apractical application, an overall classification result would only be (i) a single ("primary") classification, or (ii) one "primary" and one "secondary classification," or (iii) one "primary" and two "secondary classifications etc.

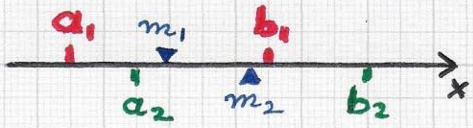
- Note. The figures on the previous and this page provide examples of signatures of two material classes, 1 and 3, with increasing "discriminative power," from top to bottom. For the five signature pairs shown in the figure on this page we compute $d_{1,3}$.

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:...

Toward a better, a more appropriate distance measure for two intervals in the classification context:



We consider two intervals, $[a_1, b_1]$ and $[a_2, b_2]$.

A simple distance is

$$\begin{aligned} d &= (a_2-a_1)+(a_2-b_1) \\ &\quad +(b_2-a_1)+(b_2-b_1) \\ &= 2(a_2-a_1+b_2-b_1) \\ &= 4\left(\frac{a_2+b_2}{2}-\frac{a_1+b_1}{2}\right) \\ &= 4(m_2-m_1). \end{aligned}$$

This distance measure has two main advantages: (i) Since d is simply the difference of the two interval midpoints, it can be quickly computed; (ii) since d is based on interval midpoints, it can be used for intervals of length 0, i.e., points.

The fact that this distance can be used for intervals of length 0 makes it possible to also consider it to define the distance D_{cl} of a tuple $(p_0', \dots, p_c')$.

Concerning the signature pairs illustrated on the previous page, we can associate "scaling factors" with the lengths of the "confidence intervals." We call this factor s, and its values are $1, 3/4, 1/2, 1/4$ and ε, from top to bottom. We use the distance formula provided on p. 25 (10/18/2022) to compute the five values for $d_{1,3}$. First, we must compute the values for A, B and C_n.

For $s \in [1/2, 1]$ we obtain

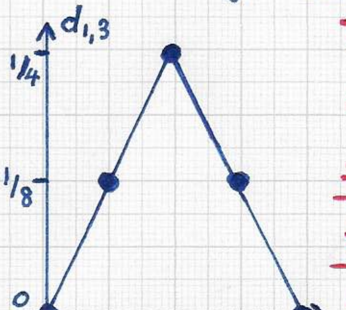
$$A = 4s\,,\ B = 4s\,,\ C_n = 6s-2\,,$$

and for $s \in [0, 1/2)$ we obtain

$$A = 4s\,,\ B = 4s\,,\ C_n = 2s\,.$$

Using the formula $d_{1,3} = \ldots = (A+B-2C_n)/8$, from p. 25, we obtain the following $d_{1,3}(s)$-values:

s	$d_{1,3}$
ε	$\varepsilon/2$
$1/4$	$1/8$
$1/2$	$1/4$
$3/4$	$1/8$
1	0



The values of $d_{1,3} = d_{1,3}(s)$ should monotonically increase when s changes from 1 to 0.

...