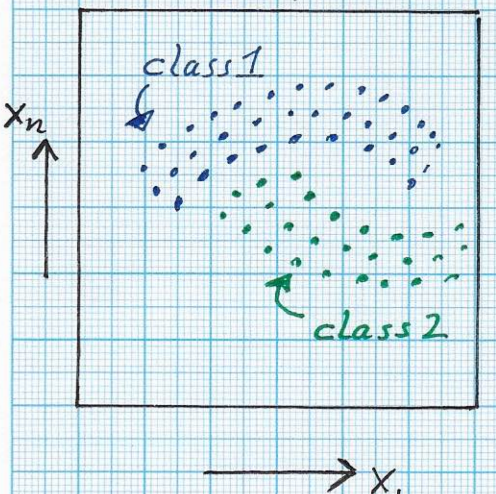
Bernd Hamann

Stratovan

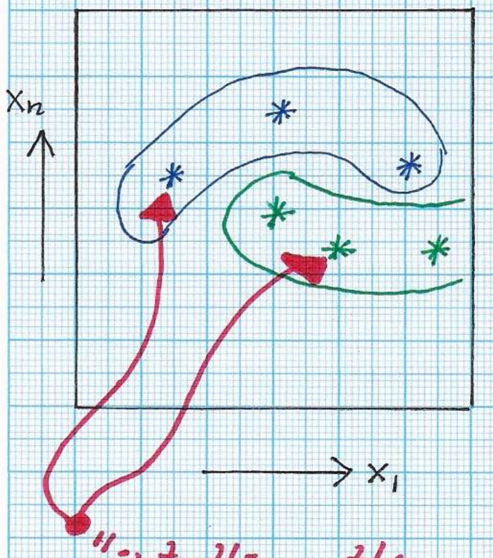
8/10/2020

RADIAL BASIS FUNCTIONS (RBFs) AND CLASSIFICATION

Feature space



RBF centers



"intelligently placed" centers of RBFs $r_l(\mathbb{X})$, $l = 1 \dots L$, with $L \ll K$ (much fewer RBFs than feature points!)

- **Assumptions:**

→ n-dimensional feature space with dimensions $x_1, \dots, x_n$

→ m classes of objects/materials, class 1, class 2, ...

→ "example" feature points

$$\overline{\mathbb{P}}_k = (x_1^k, \dots x_n^k, c_k)^T = \begin{pmatrix} \mathbb{X}_k \\ c_k \end{pmatrix}, \quad k = 1 \dots K, \quad c_k \in \{1, 2, \dots, m\},$$

where c_k is the class label of $\mathbb{X}_k$

→ m classifier functions

$$f_1(\mathbb{X}), \dots, f_m(\mathbb{X}),$$

where $f_j(\mathbb{X}) = f_j(x_1, \dots, x_n)$

→ m-dimensional classifier space with dimensions $f_1, \dots, f_m$, defining the multi-valued classifier function $\mathbb{F} = (f_1, \dots, f_m)^T$

■ RBFs AND CLASSIFICATION — Cont'd.

• **Assumptions:**

→ Each classifier function $f_j(\mathbf{x})$ is an expansion in the RBF basis $\{r_l(\mathbf{x})\}$:

$$f_j(\mathbf{x}) = \sum_{l=1}^{L} w_l^j \, r_l(\mathbf{x}) .$$

(The coefficients w_l^j are computed via a best approximation approach, using the normal equations resulting from a least squares method.)

→ Ideally, $f_j(\mathbf{x})$ ("the neuron returning 1 for values of $\mathbf{x}$ representing class j") returns 1 for all feature points $\mathbf{x}_k$ with class label j and 0 for all feature points with class labels different from j:

$$f_j(\mathbf{x}_k) = \begin{cases} 1, & \text{if } c_k = j \\ 0, & \text{otherwise} \end{cases} = \delta_{j,k}$$

Kronecker symbol

→ The conditions for f_j in "matrix form" are:

$$\boxed{\begin{aligned} f_j(\mathbf{x}_1) &= \sum_{l=1}^{L} w_l^j \, r_l(\mathbf{x}_1) = \delta_{j,1} \\ &\vdots \\ f_j(\mathbf{x}_K) &= \sum_{l=1}^{L} w_l^j \, r_l(\mathbf{x}_K) = \delta_{j,K} \end{aligned}}$$

!

RBFs AND CLASSIFICATION – Cont'd.

... Conditions for classifier function $f_j(\mathbf{x})$:

$$\begin{bmatrix} r_1(\mathbf{x}_1) & \cdots & r_L(\mathbf{x}_1) \\ \vdots & & \vdots \\ r_1(\mathbf{x}_K) & \cdots & r_L(\mathbf{x}_K) \end{bmatrix} \cdot \begin{bmatrix} w_1^j \\ \vdots \\ w_L^j \end{bmatrix} = \begin{bmatrix} \delta_{j,1} \\ \vdots \\ \delta_{j,K} \end{bmatrix}$$

$$\Leftrightarrow \quad R \cdot \vec{w}_j = \vec{\delta}_j$$

→ Conditions "summarized" for all functions $f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})$:

$$\begin{bmatrix} R \end{bmatrix} \cdot \begin{bmatrix} \begin{bmatrix} \vec{w}_1 \end{bmatrix} \cdots \begin{bmatrix} \vec{w}_m \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \vec{\delta}_1 \end{bmatrix} \cdots \begin{bmatrix} \vec{\delta}_m \end{bmatrix} \end{bmatrix}$$

$$\Leftrightarrow \quad R \cdot W = \Delta$$

→ BEST approximation of $\mathbb{F} = (f_1, \ldots, f_m)^T$ obtained via least squares / normal equations:

$$W = \left(R^T \cdot R \right)^{-1} \cdot R^T \cdot \Delta$$

→ "IDEAL": Classification vector $\mathbb{F}$ has component $f_j(\mathbf{x})$ return 1 if $\mathbf{x}$ represents class j (and all other components return 0).

RBFs AND CLASSIFICATION – Cont'd

→ **ALTERNATIVE VIEW AND APPROACH:**
Instead of computing a classification vector, a multi-valued function $\mathbb{F} = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))^T$, compute a single scalar-valued function $f(\mathbf{x})$ that – ideally – returns a specific value representing the class associated with $\mathbf{x}$.

!

In other words, construct a function

$$\underline{f(\mathbf{x})} = \begin{cases} v_1, & \text{if } \mathbf{x} \text{ represents class } 1 \\ v_2, & \text{"} \ \mathbf{x} \ \text{"} \ \text{"} \ 2 \\ \ldots & \\ v_m, & \text{"} \ \mathbf{x} \ \text{"} \ \text{"} \ m \\ \underset{=v_0}{0}, & \text{otherwise; "class 0"} \end{cases}$$

For example, use "class values" $0 < v_1 < v_2 < \ldots < v_m$.

→ In this case, the conditions for $f(\mathbf{x})$ are:

$$\underline{\underline{f(\mathbf{x}_k) = \sum_{l=1}^{L} w_l \, r_l(\mathbf{x}_k) = \sum_{j=1(0)}^{m} v_j \cdot \delta_{c_k, j}}},$$

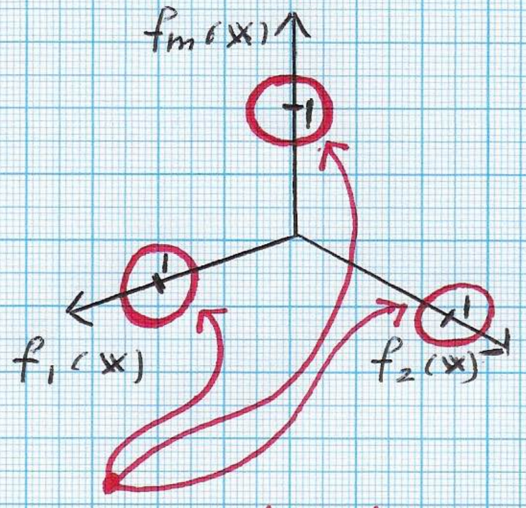
class label

$$k = 1 \ldots K, \quad c_k \in \{1, \ldots, m \ (,0)\}.$$

(**Note:** $\sum_{j=1(0)}^{m} v_j \cdot \delta_{c_k, j}$ returns value v_j when $\mathbf{x}_k$ has class label j.)

RBFs AND CLASSIFICATION - Cont'd

→ Use of multi-valued classification function $\mathbb{F} = (f_1, \dots, f_m)$ vs. single-/scalar-valued classification function f:



ε-neighborhoods of class tuples $\mathbb{F}_j$

• Ideal: $\mathbb{F} = (1, 0, 0, \dots, 0)^T$

$\Rightarrow$ class 1

...

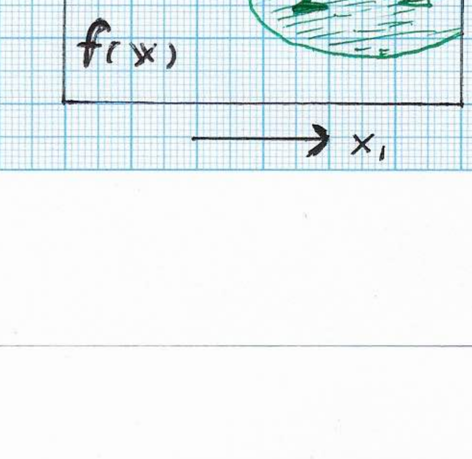
$\mathbb{F} = (0, \dots, 0, 0, 1)^T$

$\Rightarrow$ class m

$\mathbb{F}_j = (0, \dots, 0, 1, 0, \dots, 0)^T \Rightarrow$ class j

$\Rightarrow \|\mathbb{F}(x) - \mathbb{F}_j(x)\| < \varepsilon$ identifies $\begin{cases} j, & \text{if for } j \in \{1, \dots, m\} \text{ "class } j\text{"} \\ 0, & \text{otherwise} \end{cases}$

Feature space



class-1 region

$\Rightarrow |f(x) - v_1| < \varepsilon$

class-2 region

$\Rightarrow |f(x) - v_2| < \varepsilon$

(class) j identified if $|f(x) - v_j| < \varepsilon$

~BH