Bernd Hamann 8/24/2020

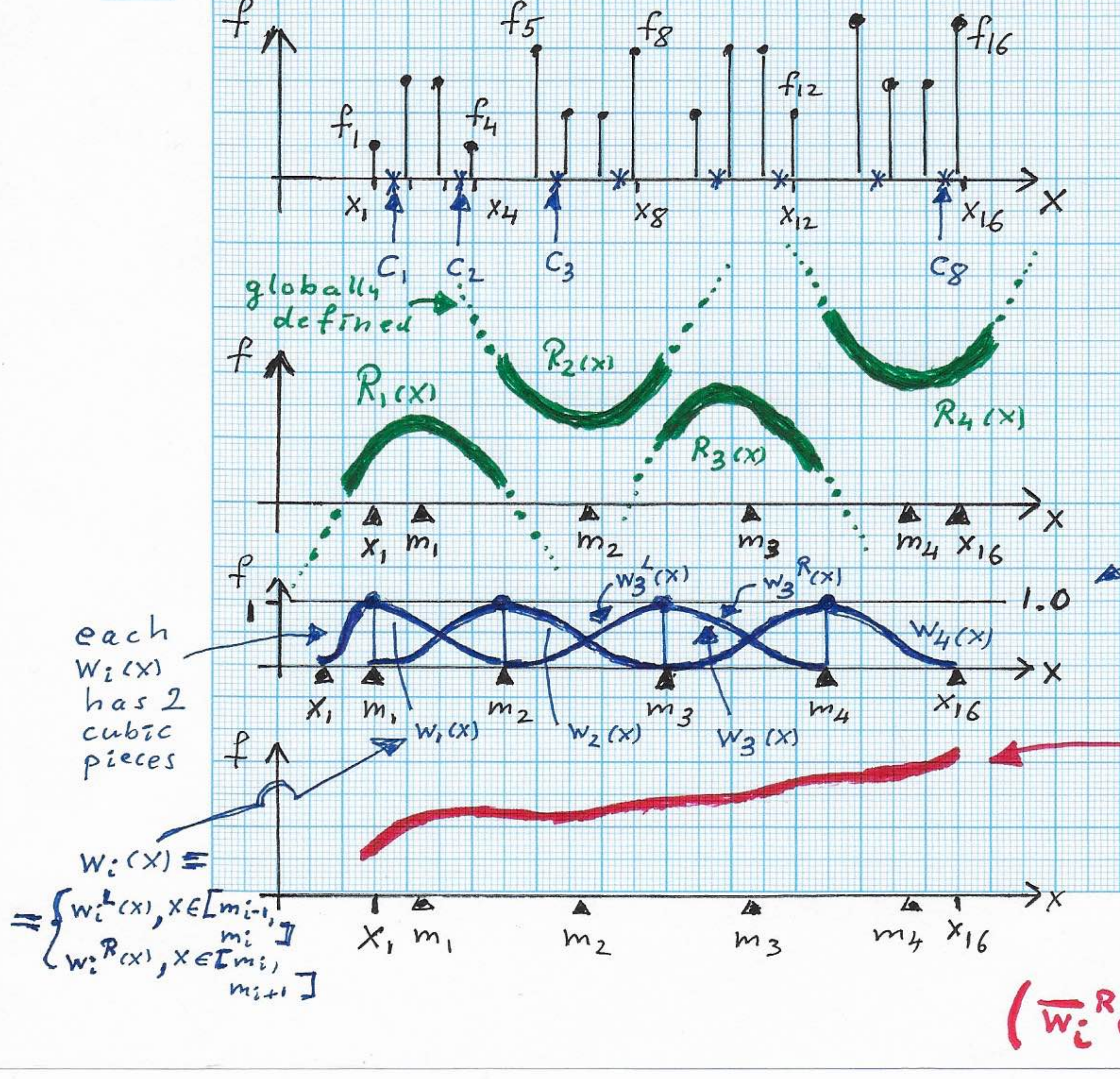
Stratovan

RBFs AND CLASSIFICATION - Cont'd.

... 5) COMACT SUPPORT (more detail):

→ Goal is computation of multiple local RBF-based approximations that are "blended" with simple blending functions to establish a final approximation.

→ Ex.: 1D case - Blending local RBF-based approximations with CUBIC blending functions



16 Locations x_i with values f_i

8 chosen RBF centers c_i

4 local RBF-based best approx. computed $= \{R_i(x)\}$, with "midpoints" m_i

4 weight functions / blending functions ($w_i \geq 0$, $\sum w_i = 1$)

final approx.:

$$f(x) = \sum_{i=1}^{4} \overline{w}_i^R(x) R_i(x) + \overline{w}_{i+1}^L(x) R_{i+1}(x)$$

(⇒ EFFICIENCY!)

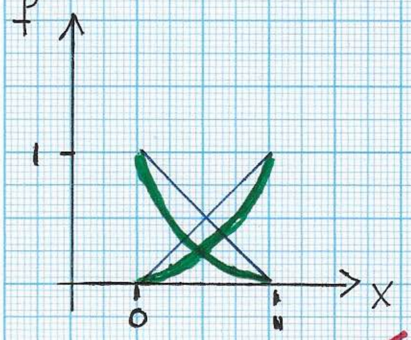
($\overline{w}_i^R(x)$, $\overline{w}_{i+1}^L(x)$ "weights"!)

RBFs AND CLASSIFICATION – Cont'd.

... 5) COMPACT SUPPORT:

→ If one computes local RBF-based, same-class best approximation, then one will have to "blend", combine them to establish an overall class-specific approximation. LINEAR or CUBIC splines enable simple, efficient blending.

Ex: • Blending 2 functions linearly



$w_1(x) = 1 - x$

$w_2(x) = x$

$R_1(x) = (1-x)^2$

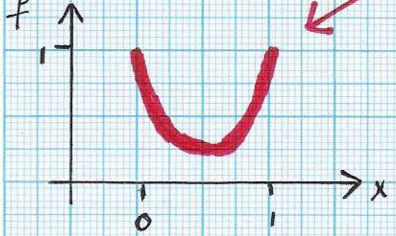
$R_2(x) = x^2$

$f(x) = \sum_{i=1}^{2} w_i(x) R_i(x) = (1-x)^3 + x^3$

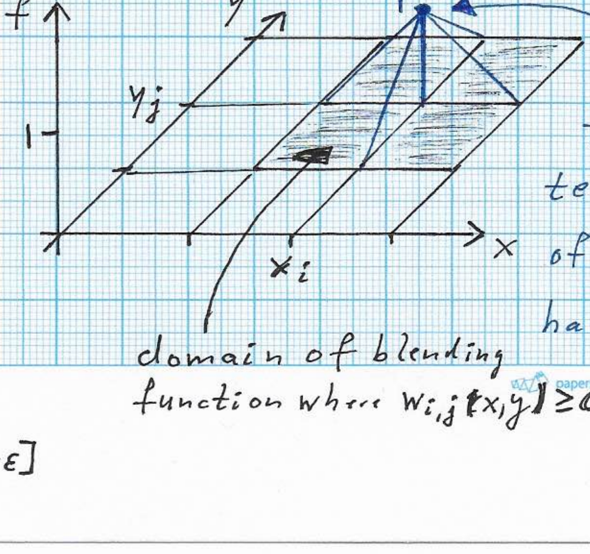
• properties of blending functions:

$w_i(x) \geq 0$ "positivity"

$\sum w_i(x) = 1$ "partition of 1"



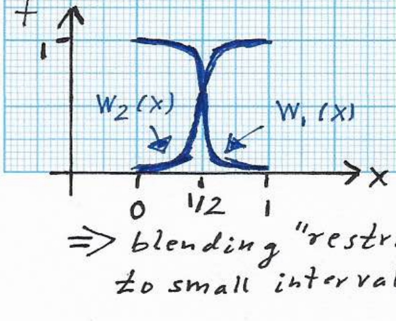
• Implied blending functions (hat functions) for 2D domain:



$w_{i,j}(x,y)$ = blending function is tensor product of univariate hat functions

domain of blending function where $w_{i,j}(x,y) \geq 0$

• Potentially "better" blending functions:

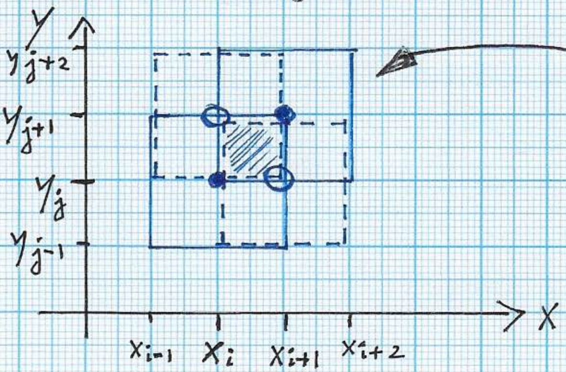


⇒ blending "restricted" to small interval $[\frac{1}{2}-\epsilon, \frac{1}{2}+\epsilon]$

■ RBFs AND CLASSIFICATION - Cont'd.

... 5) COMPACT SUPPORT:

- Blending local RBF approximations with tensor product blending functions serving as weights: [2D feature space]



The 4 blending functions $w_{i,j}, w_{i+1,j}, w_{i,j+1}, w_{i+1,j+1}$ have the "overlap domain" $[x_i, x_{i+1}] \times [y_j, y_{j+1}]$, shown as  .

In the overlap region the blending result of the 4 corresponding local RBF approximations $R_{i,j}, R_{i+1,j}, R_{i,j+1}, R_{i+1,j+1}$ is the linear and CONVEX combination

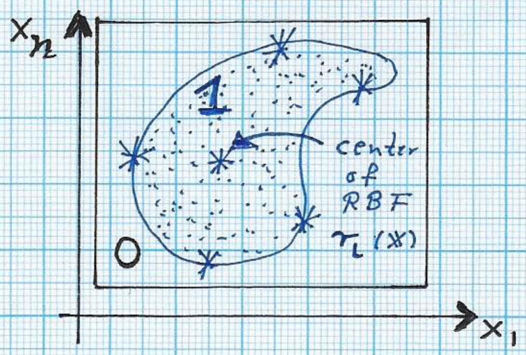
$$f(x,y) = w_{i,j}(x,y)\, R_{i,j}(x,y) + \ldots$$
$$+ w_{i+1,j+1}(x,y)\, R_{i+1,j+1}(x,y).$$

⇒ For an n-dim. feature space, the overlap region of these tensor product blending functions is an n-dim. hypercube with 2^n corners! Using 2^n blending functions is inefficient for large values of n.

⇒ More efficient blending highly desirable!

▪ RBFs AND CLASSIFICATION – Cont'd.

... 3) **RBF CENTERS**: The center locations of the RBFs should be "near-optimal" to support precision of classification. Consider the one-class (binary) classification problem:



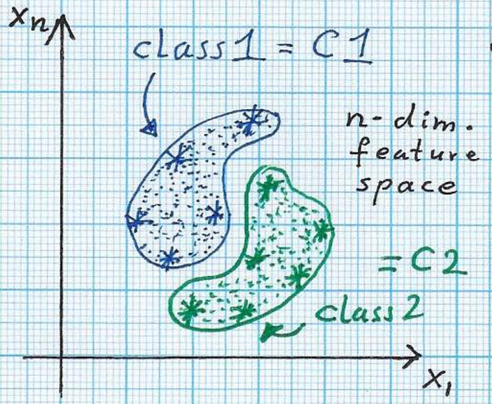
$$f(\mathbf{x}) = \begin{cases} 1 & (\text{"close to 1"}), \text{ if } \mathbf{x} \text{ of class } 1 \\ 0 & , \text{ otherwise} \end{cases}$$

- View the problem as a COMBINATORIAL optimization problem:
  - K feature points: $(x_1^k, \ldots, x_n^k)^T$, $k = 1 \ldots K$
  - L RBFs : $r_L(\mathbf{x})$, $L = 1 \ldots L$
  - "Select L feature points from set of all feature points." $\Rightarrow \binom{K}{L}$ possibilities
  - Expansion of RBFs: $f(\mathbf{x}) = \sum_{L=1}^{L} w_L r_L(\mathbf{x})$
  - Solve $\sum_{L=1}^{L} w_L r_L(\mathbf{x}^k) = 1$, $k = 1 \ldots K$, via least squares / normal equations.
  - Compute error measure $E = \left(\frac{1}{K} \sum_{k=1}^{K} (1 - f(\mathbf{x}^k))^2 \right)^{1/2}$
  - "Select a new set of L feature points and compute a new $f(\mathbf{x})$ and E value".
  - Employ an efficient optimization method!

BH

■ RBFs AND CLASSIFICATION – Cont'd.

7) MULTI-CLASS CLASSIFICATION:



$*$ opt. centers of L_1 RBFs for $f_1(\mathbf{x})$

$*$ opt. centers of L_2 RBFs for $f_2(\mathbf{x})$ $\Rightarrow$ $\mathbb{F} = \left(f_1(\mathbf{x}), f_2(\mathbf{x}) \right)^T$

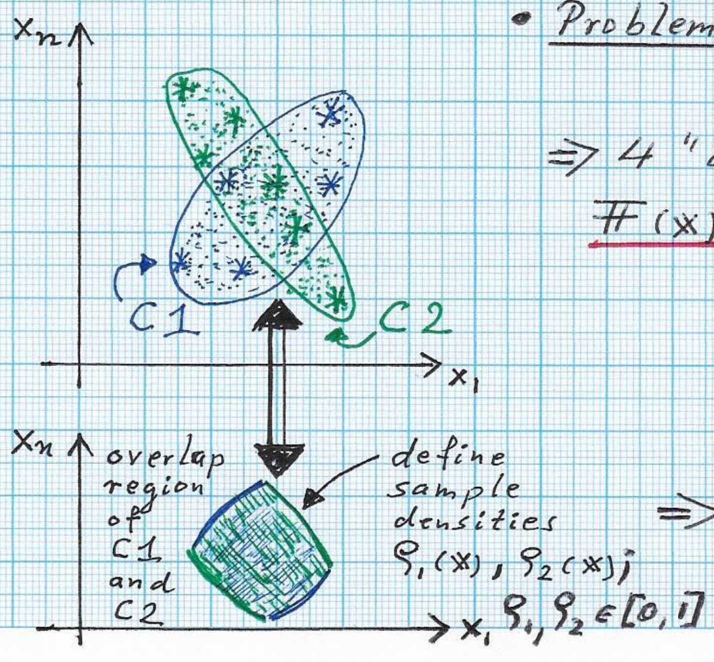
- Ex: 2-class case

$\Rightarrow$ construct 2 best RBFs $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ that "fire" (return value 1) when $\mathbf{x}$ represents C1 or C2

$\Rightarrow$ $f_1(\mathbf{x}) = \sum_{l=1}^{L_1} w_l^1 r_l^1(\mathbf{x})$,

$f_2(\mathbf{x}) = \sum_{l=1}^{L_2} w_l^2 r_l^2(\mathbf{x})$

$\Rightarrow$ IDEAL: $f_1(\mathbf{x}) \approx 1$ $\left(f_2(\mathbf{x}) \approx 1 \right)$ $\Leftrightarrow$ $\mathbf{x}$ is of C1(C2).



- Problem: value of $\mathbf{x}$ can indicate C1 or C2

$\Rightarrow$ 4 "idealized cases":

$$\mathbb{F}(\mathbf{x}) = \begin{cases} \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{– "nothing"} \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{– C1} \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{– C2} \\ \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \text{– C1 or C2} \end{cases}$$

$\Rightarrow$ possibility: compute sample densities $\rho_1(\mathbf{x})$ and $\rho_2(\mathbf{x})$ in overlap region of C1 and C2

$\Rightarrow$ weighted classification: $\mathbb{F}(\mathbf{x}) = \left(\rho_1(\mathbf{x}) f_1(\mathbf{x}), \rho_2(\mathbf{x}) f_2(\mathbf{x}) \right)^T$