

Искусственные нейронные сети

Высшая школа экономики, Нижегородский
филиал
Кафедра прикладной математики и информатики

ст.преп. Бабкина Татьяна Сергеевна

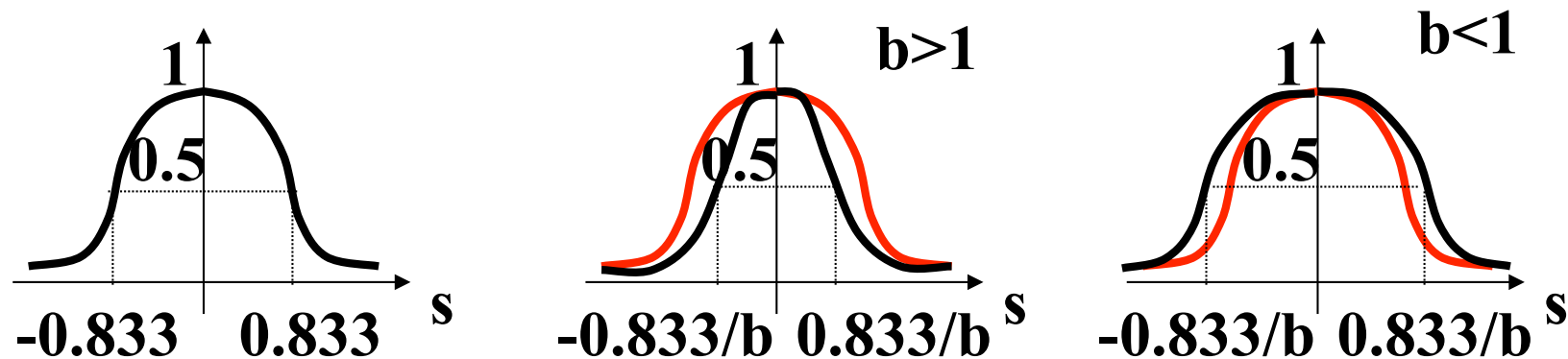
tbabkina@hse.ru

taty-bab@yandex.ru

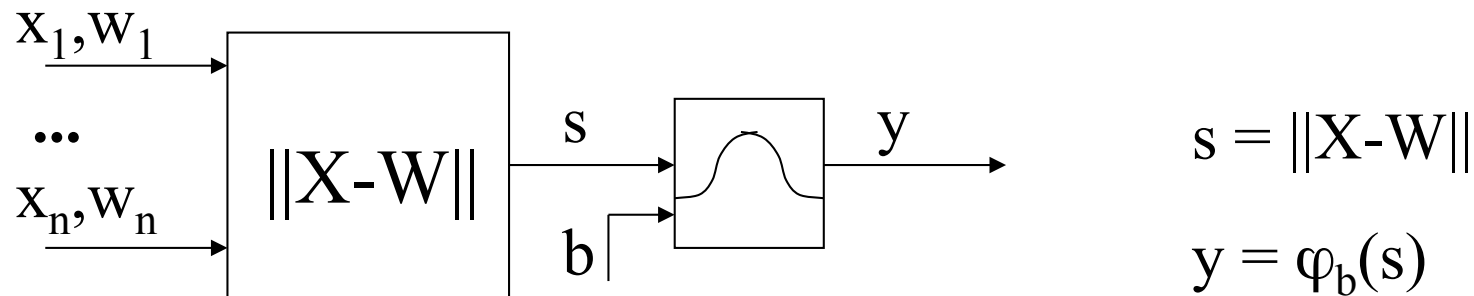
2015

Радиальная базисная функция

$$\varphi_b(s) = \exp(-(bs)^2)$$



Радиальный базисный нейрон



Детектор близости X и W

W – «центр» нейрона; b – «радиус действия» нейрона

«Зона действия» нейрона: $\varphi_b(s) \geq 0.5$

Сети с RBE

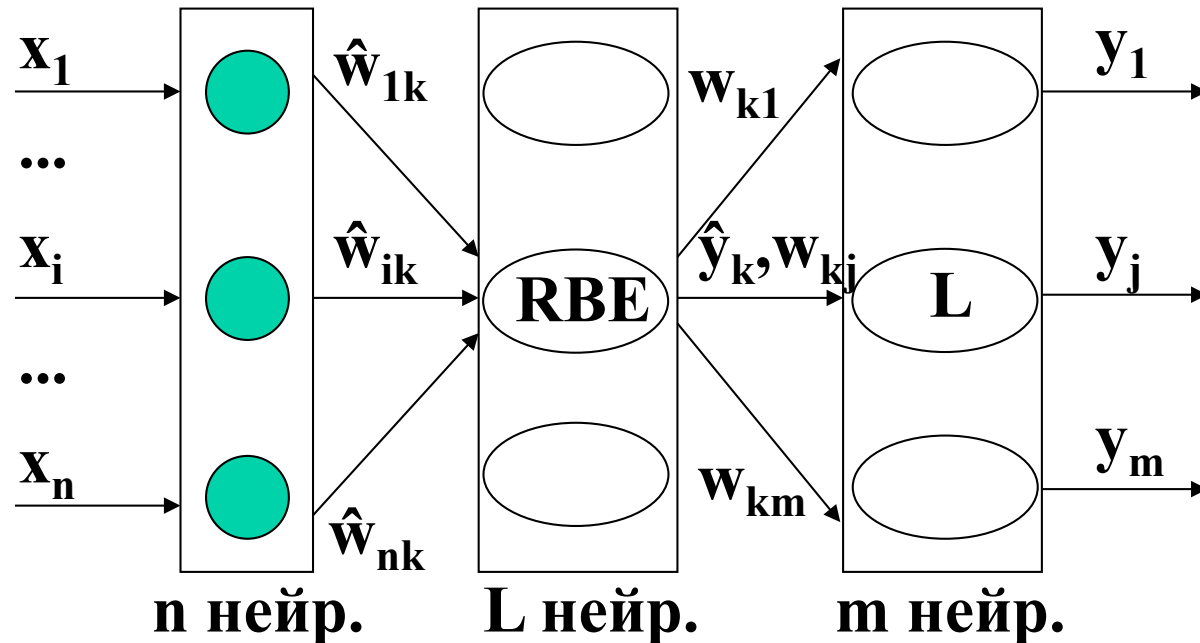
1988 Д. Брумхед, Д. Лоув

Многослойные сети с прямым распространением сигнала

По сравнению многослойным персептроном

- имеют один скрытый слой
 - обучаются быстрее (настройка при создании)
 - громоздкие (больше нейронов в скрытом слое)
 - не способны экстраполировать данные (плохо?)
- Радиальная базисная сеть
 - Обобщенно-регрессионная сеть
- Аппроксимация функций, анализ временных рядов*
- Вероятностная сеть
- Задачи классификации*

Радиальная базисная сеть



$$s_k = \|X - \hat{W}_k\|$$

$$\hat{y}_k = \varphi_b(s_k)$$

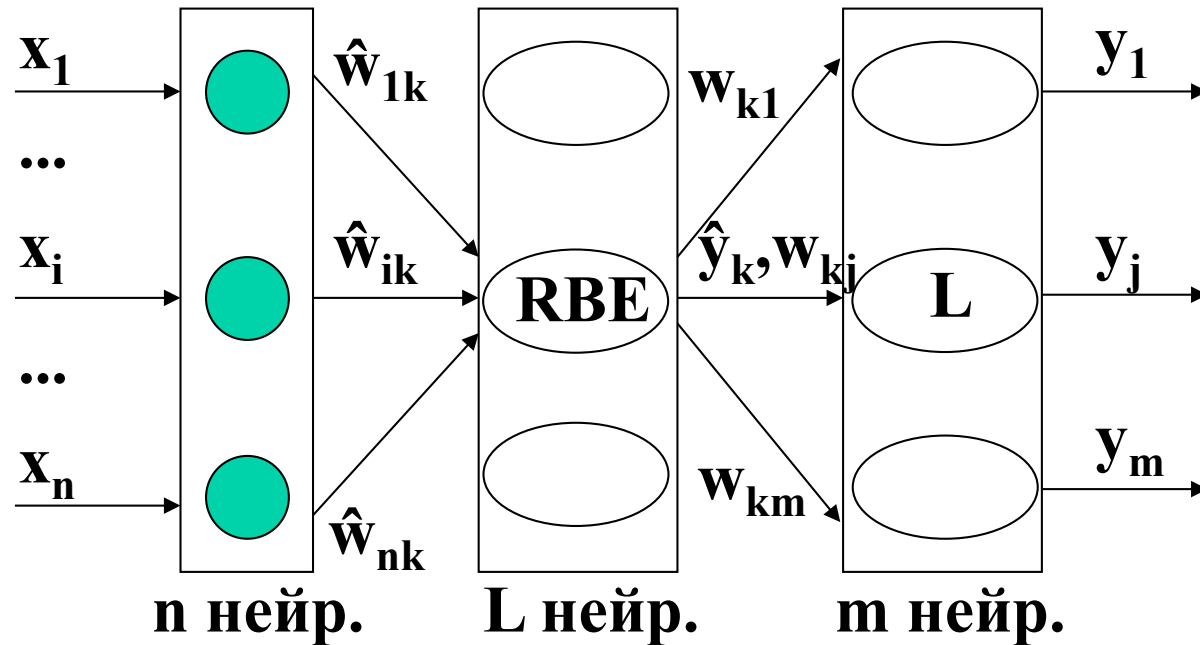
$$y_j = \sum_{k=1}^L \hat{y}_k w_{kj}$$

$$Q = \{(X^1, Y^1) \dots (X^N, Y^N)\}, \quad L \ll N$$

Если X близок к \hat{W}_r и далек от остальных \hat{W}_k , $k \neq r$, то $\hat{y}_r \approx 1$ и $\hat{y}_k \approx 0$; тогда $Y \approx (w_{r1} \dots w_{rm})$

В общем случае: несколько RBE возбуждается в разной степени, и выход сети = суперпозиции их выходов

Радиальная базисная сеть



$$s_k = \|X - \hat{W}_k\|$$

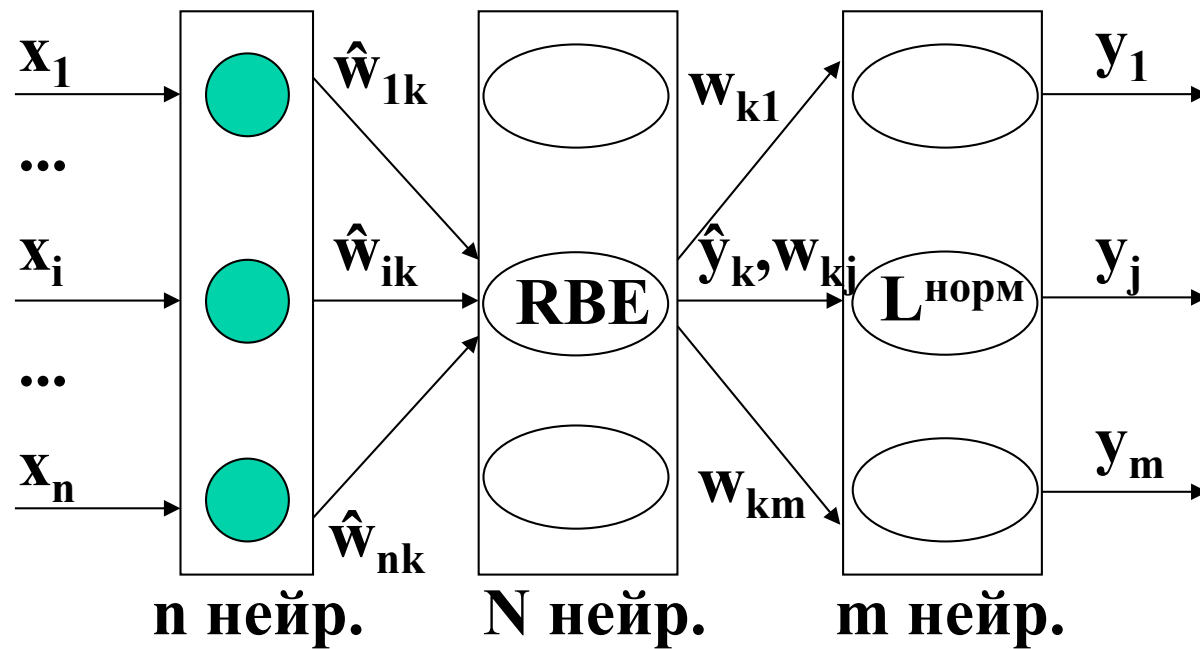
$$\hat{y}_k = \varphi_b(s_k)$$

$$y_j = \sum_{k=1}^L \hat{y}_k w_{kj}$$

Обучение:

- 1) Определить центры и радиусы RBE (центры – случайная выборка, метод k -средних, правило Кохонена; радиусы выбираются так, чтобы зоны действия нейронов слегка перекрывались)
- 2) Настроить веса линейного слоя (решение системы линейных уравнений)

Обобщенно-регрессионная сеть



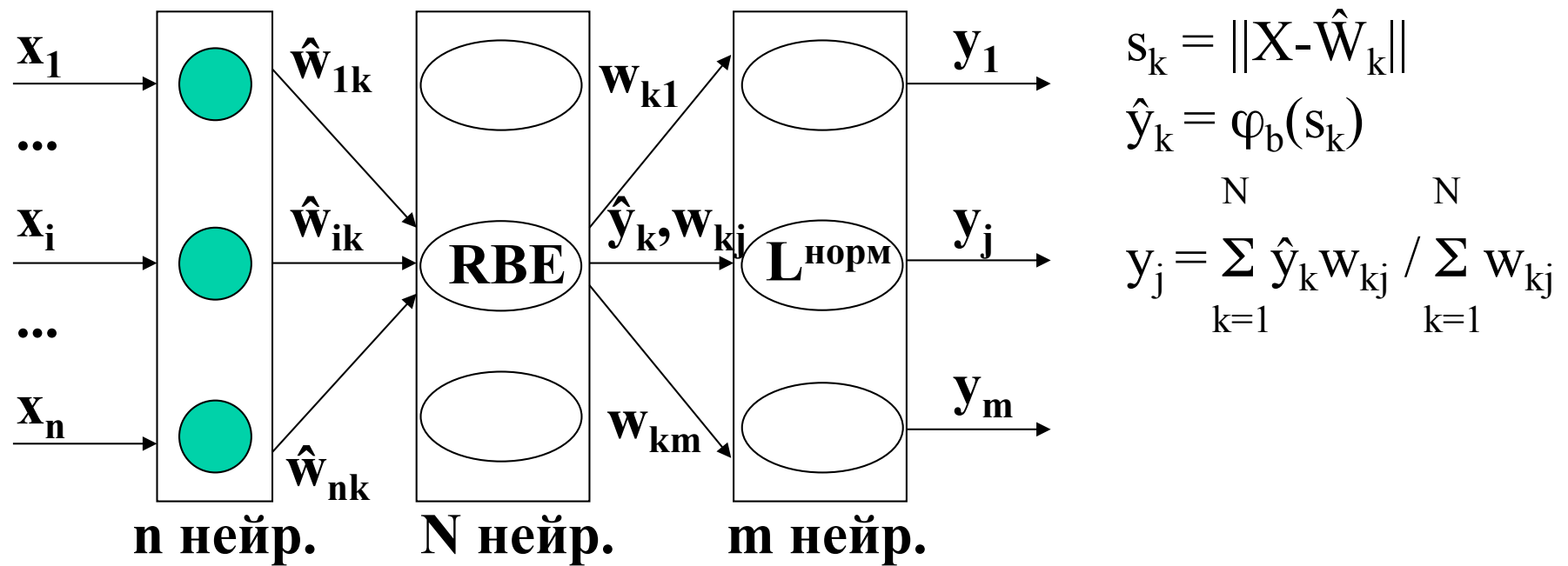
$$s_k = \|X - \hat{W}_k\|$$

$$\hat{y}_k = \varphi_b(s_k)$$

$$y_j = \frac{\sum_{k=1}^N \hat{y}_k w_{kj}}{\sum_{k=1}^N w_{kj}}$$

Выходной слой вычисляет взвешенное среднее выходов скрытого слоя

Обобщенно-регрессионная сеть

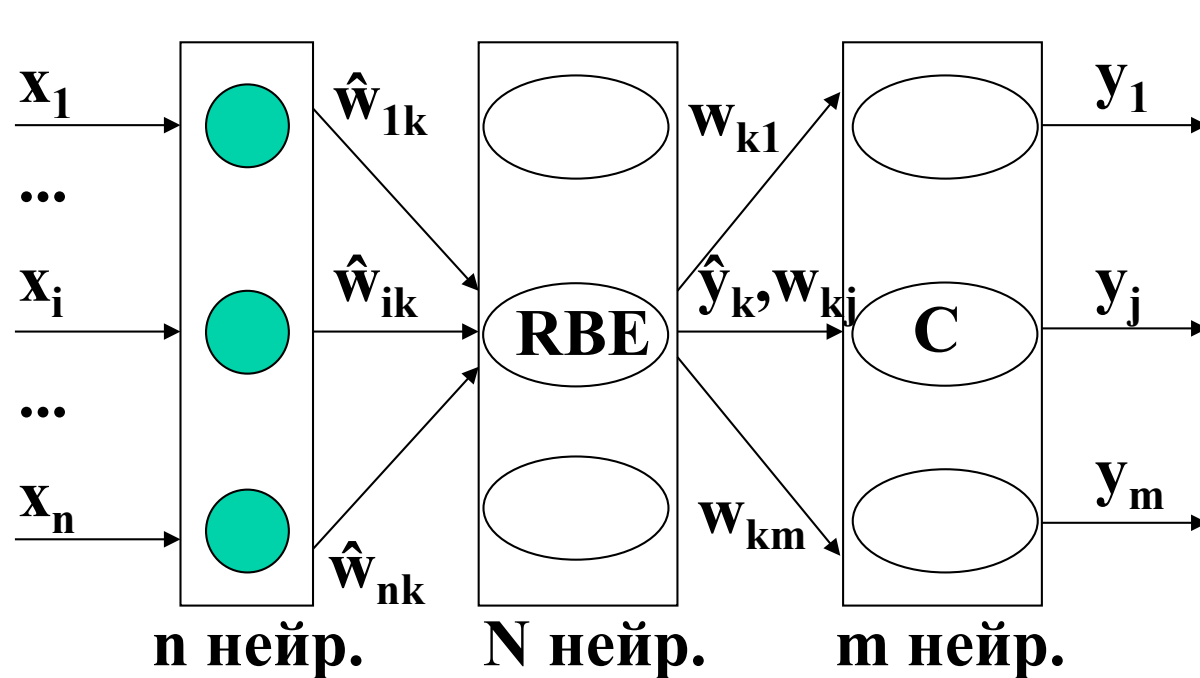


Обучение: настройка при создании

- 1) Центры RBE = обучающие точки $\{X^1, \dots, X^N\}$
- 2) Веса линейного слоя = ответы $\{Y^1, \dots, Y^N\}$

сеть копирует в себя все обучающие данные

Вероятностная сеть



$$s_k = \|X - \hat{W}_k\|$$

$$\hat{y}_k = \varphi(bs_k)$$

$$\hat{s}_j = \sum_{k=1}^N \hat{y}_k w_{kj}$$

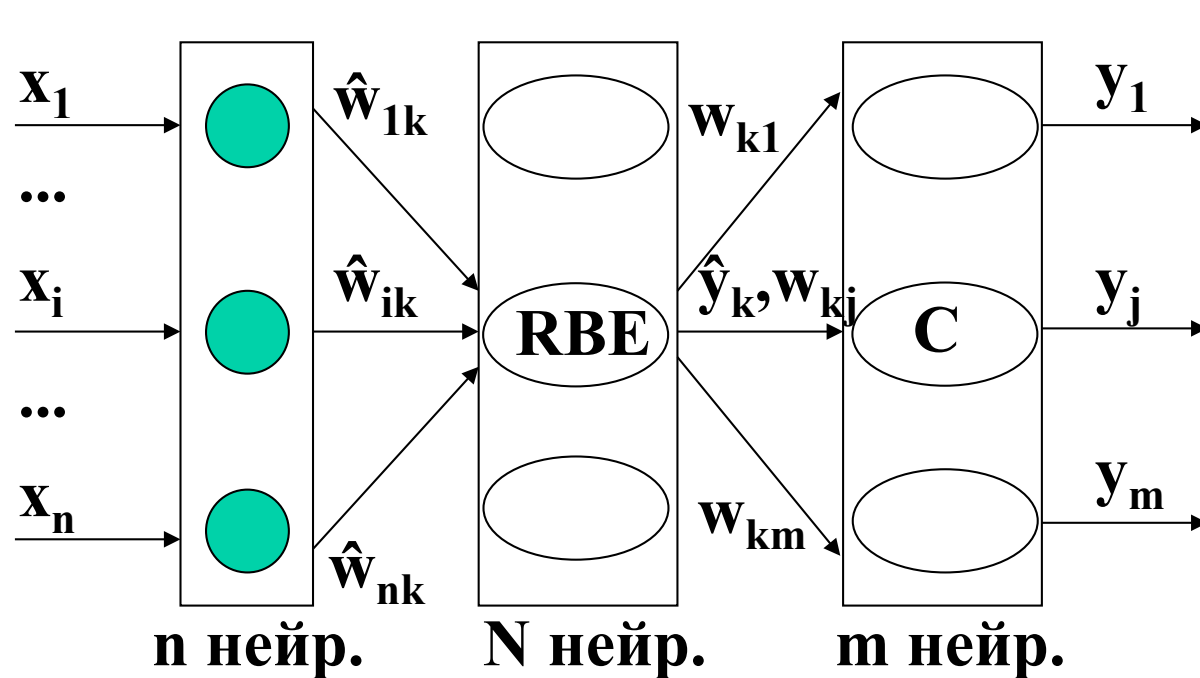
$$r = \operatorname{argmax}_j \hat{s}_j$$

$$y_r = 1$$

$$y_p = 0, p \neq r$$

Выходной слой вычисляет вероятность принадлежности X к каждому из m классов

Вероятностная сеть



$$s_k = \|X - \hat{W}_k\|$$

$$\hat{y}_k = \varphi(bs_k)$$

$$\hat{s}_j = \sum_{k=1}^N \hat{y}_k w_{kj}$$

$$r = \operatorname{argmax}_j \hat{s}_j$$

$$y_r = 1$$

$$y_p = 0, p \neq r$$

Обучение: настройка при создании

- 1) Центры RBE = обучающие точки $\{X^1, \dots, X^N\}$
- 2) Веса выходного слоя: если X^k принадлежит классу j , то $w_{kj} = 1$; иначе $w_{kj} = 0$.

сеть копирует в себя все обучающие данные