CS 2710 Foundations of AI
Lecture 23

Supervised learning.
Multilayer neural networks.

Milos Hauskrecht
milos@cs.pitt.edu
5329 Sennott Square

Announcements

Homework 10:

- **due on Wednesday, November 30, 2005**

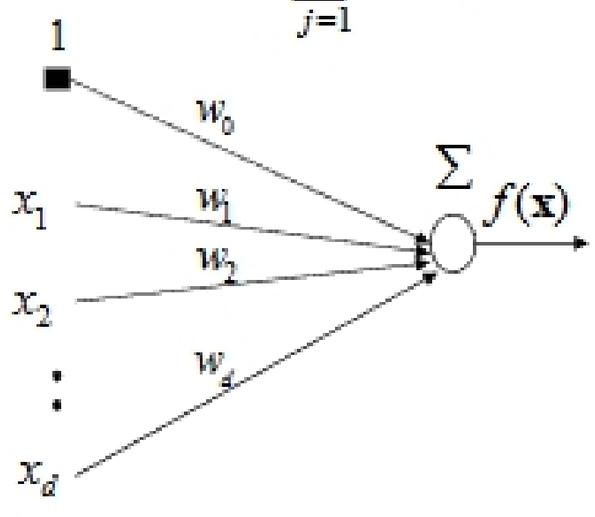
Final exam:

- **December 14, 2004 at 11:00am-1:00pm**
- Location: TBA
- Closed book
- Cumulative
- **AI prelim exam**

Linear units

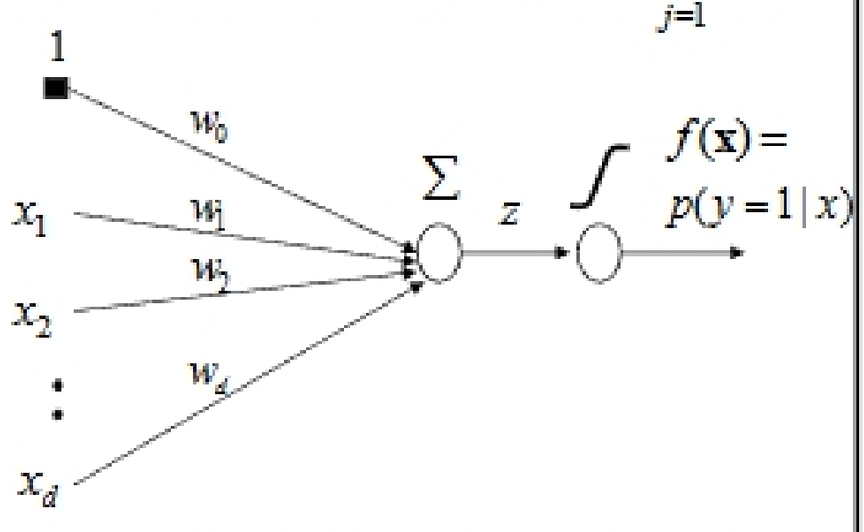
Linear regression

$$f(\mathbf{x}) = w_0 + \sum_{j=1}^{d} w_j x_j$$



Logistic regression

$$f(\mathbf{x}) = p(y=1 \mid \mathbf{x}, \mathbf{w}) = g(w_0 + \sum_{j=1}^{d} w_j x_j)$$



On-line gradient update:

$$w_0 \leftarrow w_0 + \alpha(y - f(\mathbf{x}))$$
$$\vdots$$
$$w_j \leftarrow w_j + \alpha(y - f(\mathbf{x}))x_j$$

The same

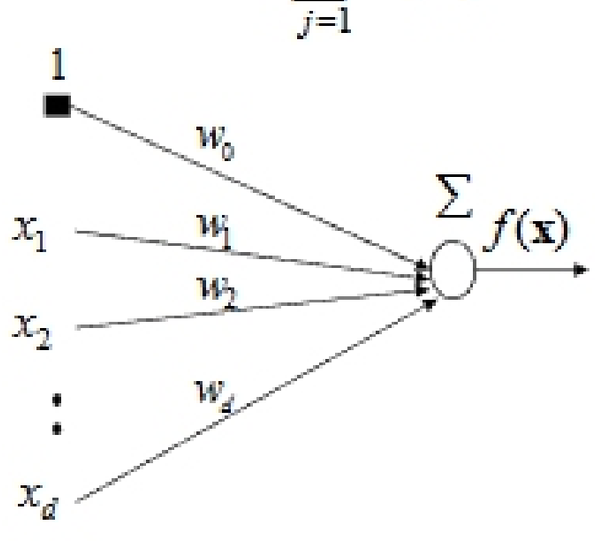
On-line gradient update:

$$w_0 \leftarrow w_0 + \alpha(y - f(\mathbf{x}))$$
$$\vdots$$
$$w_j \leftarrow w_j + \alpha(y - f(\mathbf{x}))x_j$$

Limitations of basic linear units

Linear regression

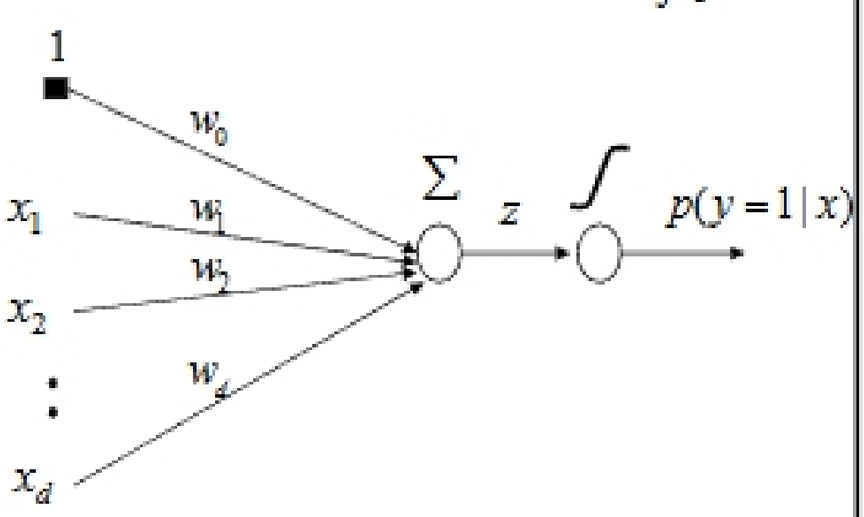
$$f(\mathbf{x}) = w_0 + \sum_{j=1}^{d} w_j x_j$$



Function linear in inputs !!

Logistic regression

$$f(\mathbf{x}) = p(y=1 \mid \mathbf{x}, \mathbf{w}) = g(w_0 + \sum_{j=1}^{d} w_j x_j)$$



Linear decision boundary!!

Extensions of simple linear units

- use **feature (basis) functions** to model **nonlinearities**

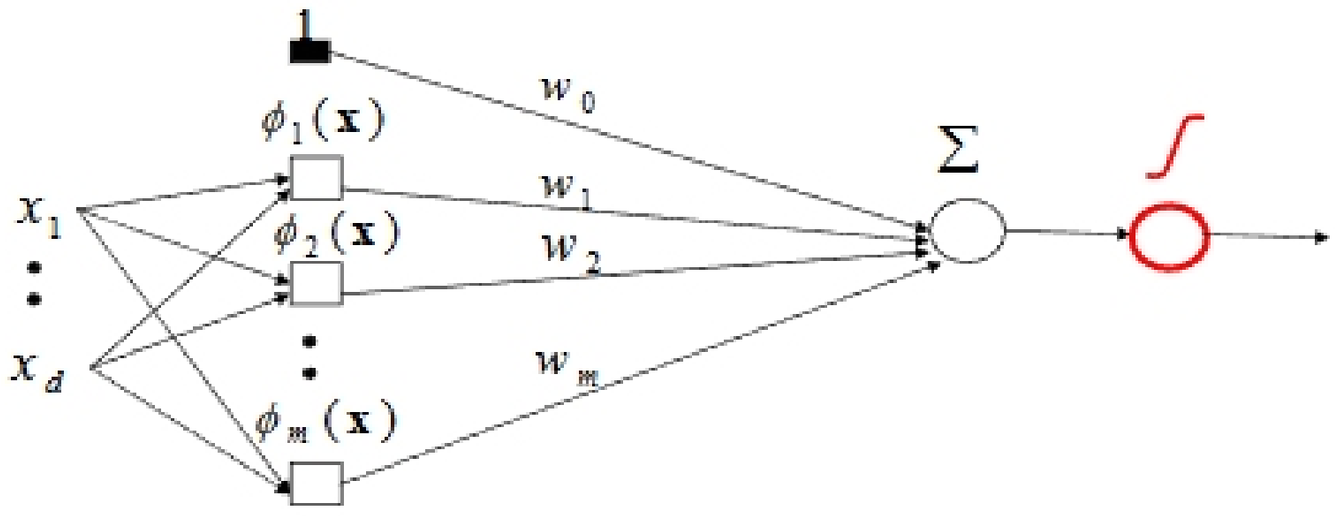
Linear regression

$$f(\mathbf{x}) = w_0 + \sum_{j=1}^{m} w_j \phi_j(\mathbf{x})$$

Logistic regression

$$f(\mathbf{x}) = g(w_0 + \sum_{j=1}^{m} w_j \phi_j(\mathbf{x}))$$

$\phi_j(\mathbf{x})$ - an arbitrary function of $\mathbf{x}$



Regression with the quadratic model.

Limitation: linear hyper-plane only

a non-linear surface can be better