11 Apr 1990 Gradient Descent — Backward Error Propagation (BEP) — Notation

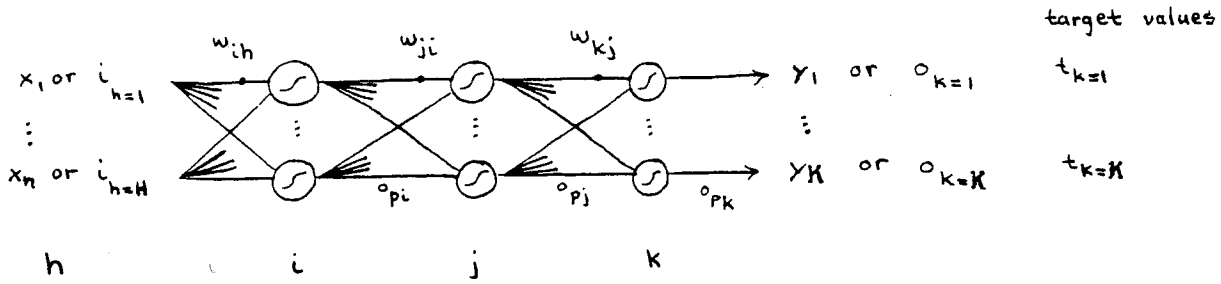
Neil E Cotter

modified

20 Apr 1990

Neil E Cotter

Network



Each layer has as many neurons as desired.

Notation: PDP book uses notation as follows

i_{Ph} ≡ input # h for pattern p

o_{pi} ≡ output of ~~layer~~ neuron i (input to ^{j} layer

i_{pj} ≡ input? No! Not defined.

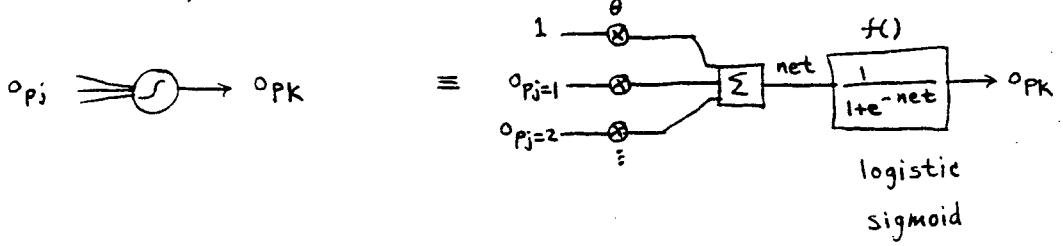
o_{pj} ≡ output of neuron j (from j layer, is input to k layer)

p ≡ training pattern #

w_{kj} ≡ synapse from neuron in j layer to " " k "

destination ↗ ↖ source

Note: I will drop the p subscripts. They are confusing. I follow PDP notation where possible. However, I eliminate the "input layer" that just distributes in



$$f(x) \equiv \frac{1}{1+e^{-x}}$$ is logistic sigmoid

$$f' = f(1-f) \equiv f(net)\left[1-f(net)\right] = \left.\frac{\partial f(net)}{\partial net}\right|_{net}$$