

Back propagation — (11.46)

$$W^m(k+1) = W^m(k) - \alpha \Delta^m(k) \cdot \Delta^{m-1}(k)^T$$

↑
sensitivity

$0 < \alpha < 1$; convergence factor

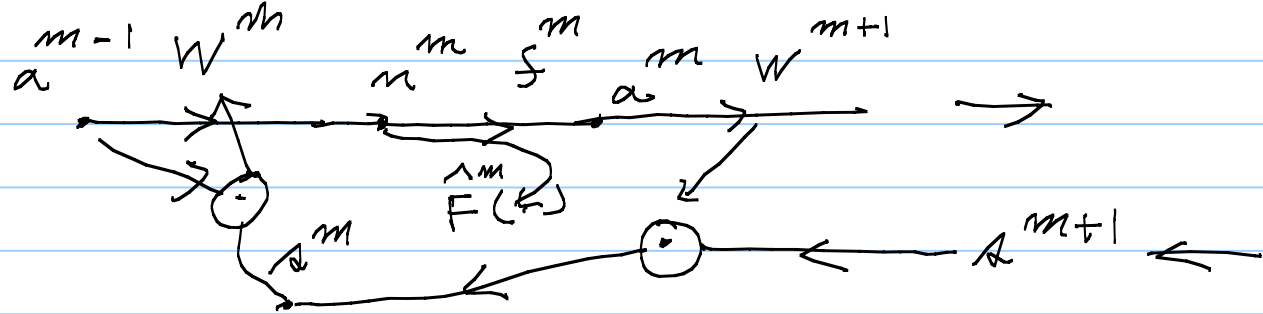
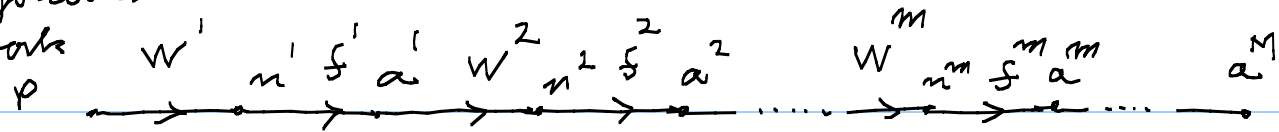
$m = \text{layer \#}; 1 \leq m \leq M$

for Δ^m use (11.35)

$$\Delta^m = \hat{F}^m(c^m) (W^{m+1})^T \Delta^{m+1} \left. \vphantom{\Delta^m} \right\} \text{ goes backwards}$$

↑
(11.34)

Feedforward
networks



need a^M to get started