Hebbian learning & Oja's rule

A single linear neuron with a Hebbian-type adaptation rule for its synaptic weights can evolve into a filter for the first principal component of the input distribution.

- The model is linear in the sense that the model's output is a linear combination of its inputs.
- The neuron receives a set of m input signals x1, x2, ..., xm through a corresponding set of m synapses with weights w1, w2 , ..., wm, respectively

$$y = \sum_{i=1}^{m} w_i x_i \tag{8.36}$$

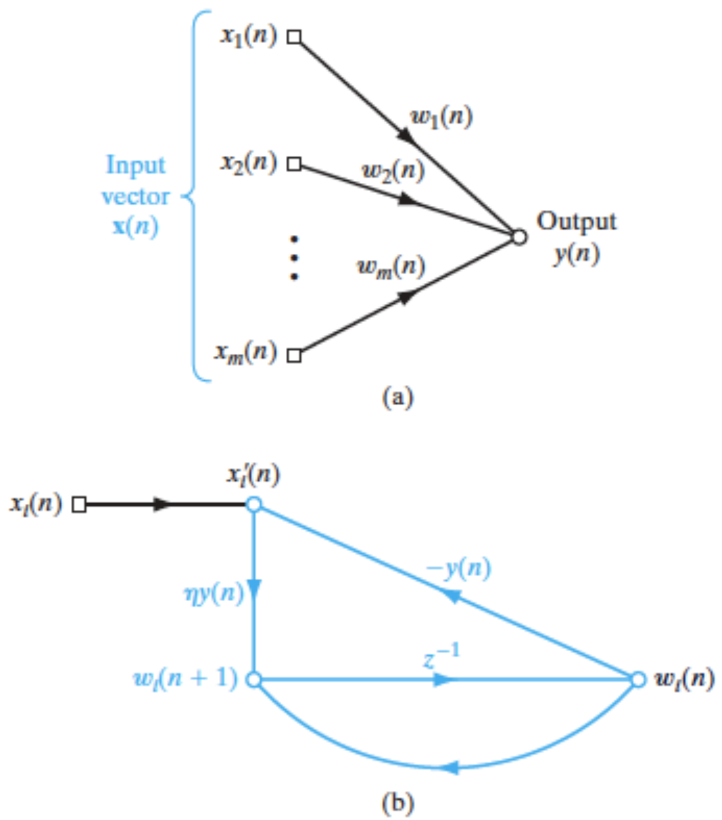

FIGURE 8.6 Signal-flow graph representation of maximum eigenfilter. (a) Graph of Eq. (8.36). (b) Graph of Eqs. (8.41) and (8.42).

Note: n is the discrete time index / iteration. In the slides this is denoted by t.

- Hebb: “Neurons that fire together, wire together”
  => synaptic weight varies with time, growing strong when the presynaptic signal xi and postsynaptic signal y coincide, hence:

$$w_i(n+1) = w_i(n) + \eta y(n) x_i(n), \qquad i = 1, 2, ..., m \qquad (8.37)$$

...where η is the learning rate.

Problem:
Hebb's rule has synaptic weights approaching infinity with a positive learning rate (→ unstable!)

Solution:
Normalizing the weights so that each weight's magnitude is restricted between 0, corresponding to no weight, and 1, corresponding to being the only input neuron with any weight. We do this by normalizing the weight vector to be of length one.

Oja’s rule:

- single-neuron special case of the Generalized Hebbian Algorithm
- stable, unlike Hebb's rule

$$w_i(n+1) = \frac{w_i(n) + \eta y(n) x_i(n)}{\left(\sum_{i=1}^{m} (w_i(n) + \eta y(n) x_i(n))^2\right)^{1/2}} \qquad (8.38)$$

Different notation in the **slides** (“Euclidean weights normalization”):

$$\underline{\mathbf{w}}(t+1) = \frac{\overbrace{\underline{\mathbf{w}}(t) + \varepsilon y^{(\alpha)} \underline{\mathbf{x}}^{(\alpha)}}^{\text{Hebbian learning}}}{\underbrace{\left| \underline{\mathbf{w}}(t) + \varepsilon y^{(\alpha)} \underline{\mathbf{x}}^{(\alpha)} \right|}_{\text{Euclidean weights normalization}}}$$

Via Taylor expansion (power series) we get to:

$$w_i(n+1) = w_i(n) + \eta y(n)(x_i(n) - y(n)w_i(n)) \tag{8.40}$$

Notation in the **slides**:

Oja's rule

$$\Delta \mathrm{w}_j = \varepsilon y_{\left(\underline{\mathbf{x}}^{(\alpha)};\underline{\mathbf{w}}\right)} \left\{ \underbrace{\mathrm{x}_j^{(\alpha)}}_{\substack{\text{Hebbian}\\\text{learning}}} - \underbrace{y_{\left(\underline{\mathbf{x}}^{(\alpha)};\underline{\mathbf{w}}\right)} \mathrm{w}_j}_{\text{decay term}} \right\}$$

Oja's rule = Hebbian Learning with weight normalization

Derivation of Oja's rule via Taylor expansion (question on exercise sheet!):

Let be

$$f(\varepsilon) := \frac{w_i + \varepsilon y x_i}{\left(\sum_{j=1}^{N} (w_j + \varepsilon y x_j)^2\right)^{1/2}}$$

Then it follows

$$\frac{\partial f(\varepsilon)}{\partial \varepsilon} = \frac{y x_i}{\left(\sum_{j=1}^{N} (w_j + \varepsilon y x_j)^2\right)^{1/2}} - \frac{(w_i + \varepsilon y x_i)\left(\sum_{j=1}^{N} (w_j + \varepsilon y x_j) y x_j\right)}{\left(\sum_{j=1}^{N} (w_j + \varepsilon y x_j)^2\right)^{3/2}}$$

With $\varepsilon_0 = 0$ this yields

$$\frac{\partial f(\varepsilon_0)}{\partial \varepsilon} = \frac{y x_i}{\left(\sum_{j=1}^{N} w_j^2\right)^{1/2}} - \frac{w_i\left(\sum_{j=1}^{N} w_j y x_j\right)}{\left(\sum_{j=1}^{N} w_j^2\right)^{3/2}} = \frac{y x_i - w_i\left(\sum_{j=1}^{N} w_j y x_j\right)\left(\sum_{j=1}^{N} w_j^2\right)^2}{\left(\sum_{j=1}^{N} w_j^2\right)^{1/2}} = \frac{y x_i - w_i\left(\sum_{j=1}^{N} w_j y x_j\right)||w||^2}{||w||}$$

Since $||w|| = 1$ and $\sum_{j=1}^{N} w_j x_j = y$ we have

$$\frac{\partial f(\varepsilon_0)}{\partial \varepsilon} = y x_i - w_i y\left(\sum_{j=1}^{N} w_j x_j\right) = y x_i - w_i y y$$

The Taylor expansion of f is

$$Tf(\varepsilon; \varepsilon_0) = f(\varepsilon_0) + \frac{\partial f(\varepsilon_0)}{\partial \varepsilon}(\varepsilon - \varepsilon_0)$$

With $\varepsilon_0 = 0$ and $||w|| = 1$ it follows

$$Tf(\varepsilon; \varepsilon_0) = w_i + (y x_i - w_i y y)\varepsilon = w_i + \varepsilon y(x_i - y w_i)$$

→ For other derivation see **Haykin p. 384-385**