19.2.1 A mathematical formulation of Hebb's rule

In order to find a mathematically formulated learning rule based on Hebb's postulate we focus on a single synapse with efficacy w_{ij} that transmits signals from a presynaptic neuron j to a postsynaptic neuron i. For the time being we content ourselves with a description in terms of mean firing rates. In what follows, the activity of the presynaptic neuron is denoted by v_j and that of the postsynaptic neuron by v_i.

There are two aspects of Hebb's postulate that are particularly important: *locality* and *joint activity*. Locality means that the change of the synaptic efficacy can depend only on local variables, i.e., on information that is available at the site of the synapse, such as pre- and postsynaptic firing rate, and the actual value of the synaptic efficacy, but not on the activity of other neurons. Based on the locality of Hebbian plasticity we can write down a rather general formula for the change of the synaptic efficacy,

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = F(w_{ij}; v_i, v_j). \tag{19.1}$$

Here, $\mathrm{d}w_{ij}/\mathrm{d}t$ is the rate of change of the synaptic coupling strength and F is a so-far-undetermined function (Sejnowski and Tesauro, 1989). We may wonder whether there are other local variables (e.g., the input potential h_i; see Chapter 15) that should be included as additional arguments of the function F. It turns out that in standard rate models this is not necessary, since the input potential h_i is uniquely determined by the postsynaptic firing rate, $v_i = g(h_i)$, with a monotone gain function g.

The second important aspect of Hebb's postulate is the notion of "joint activity" which implies that pre- and postsynaptic neurons have to be active *simultaneously* for a synaptic weight change to occur. We can use this property to learn something about the function F. If F is sufficiently well behaved, we can expand F in a Taylor series about $v_i = v_j = 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = c_0(w_{ij}) + c_1^{\mathrm{pre}}(w_{ij}) v_j + c_1^{\mathrm{post}}(w_{ij}) v_i + c_2^{\mathrm{pre}}(w_{ij}) v_j^2$$
$$+ c_2^{\mathrm{post}}(w_{ij}) v_i^2 + c_{11}^{\mathrm{corr}}(w_{ij}) v_i v_j + \mathcal{O}(v^3). \tag{19.2}$$

The term containing c_{11}^{corr} on the right-hand side of (19.2) is bilinear in pre- and postsynaptic activity. This term implements the AND condition for joint activity. If the Taylor expansion had been stopped before the bilinear term, the learning rule would be called "non-Hebbian," because pre- or postsynaptic activity alone induces a change of the synaptic efficacy, and joint activity is irrelevant. Thus a Hebbian learning rule needs either the bilinear term $c_{11}^{\mathrm{corr}}(w_{ij}) v_i v_j$ with $c_{11}^{\mathrm{corr}} > 0$ or a higher-order term (such as $c_{21}(w_{ij}) v_i^2 v_j$) that involves the activity of both pre- and postsynaptic neurons.

Example: Hebb rules, saturation, and LTD

The simplest choice for a Hebbian learning rule within the Taylor expansion of Eq. (19.2) is to fix c_{11}^{corr} at a positive constant and to set all other terms in the Taylor expansion to zero. The result is the prototype of Hebbian learning,

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = c_{11}^{\text{corr}} \, v_i \, v_j. \tag{19.3}$$

We note in passing that a learning rule with $c_{11}^{\text{corr}} < 0$ is usually called anti-Hebbian because it weakens the synapse if pre- and postsynaptic neuron are active simultaneously, a behavior that is just contrary to that postulated by Hebb.

Note that, in general, the coefficient c_{11}^{corr} may depend on the current value of the weight w_{ij}. This dependence can be used to limit the growth of weights at a maximum value w^{max}. The two standard choices of weight-dependence are called "hard bound" and "soft bound," respectively. Hard bound means that $c_{11}^{\text{corr}} = \gamma_2$ is constant in the range $0 < w_{ij} < w^{\text{max}}$ and zero otherwise. Thus, weight growth stops abruptly if w_{ij} reaches the upper bound w^{max}.

Table 19.1 *The change $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij}$ of a synapse from j to i for various Hebb rules as a function of pre- and postsynaptic activity. "ON" indicates a neuron firing at high rate ($v > 0$), whereas "OFF" means an inactive neuron ($v = 0$). From left to right: Standard Hebb rule, Hebb with decay, Hebb with postsynaptic or presynaptic LTP/LTD threshold, covariance rule. The parameters are $0 < v_\theta < v^{\text{max}}$ and $0 < c_0 < (v^{\text{max}})^2$.*

A soft bound for the growth of synaptic weights can be achieved if the parameter c_{11}^{corr} in Eq. (19.3) tends to zero as w_{ij} approaches its maximum value w^{max},

$$c_{11}^{\text{corr}}(w_{ij}) = \gamma_2 \, (w^{\text{max}} - w_{ij})^\beta, \tag{19.4}$$

with positive constants γ_2 and β. The typical value of the exponent is $\beta = 1$, but other choices are equally possible (Gütig *et al.*, 2003). For $\beta \to 0$, the soft-bound rule (19.4) converges to the hard-bound one.

Note that neither Hebb's original proposal nor the simple rule (19.3) contains a possibility for a decrease of synaptic weights. However, in a system where synapses can only be strengthened, all efficacies will eventually saturate at their upper maximum value. Our formulation (19.2) is sufficiently general to allow for a combination of synaptic potentiation and depression. For example, if we set $w^{\text{max}} = \beta = 1$ in (19.4) and combine it with a choice $c_0(w_{ij}) = -\gamma_0 \, w_{ij}$, we obtain a learning rule

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = \gamma_2 \, (1 - w_{ij}) \, v_i \, v_j - \gamma_0 \, w_{ij}, \tag{19.5}$$

where, in the absence of stimulation, synapses spontaneously decay back to zero. Many other combinations of the parameters $c_0, \ldots, c_{11}^{\text{corr}}$ in Eq. (19.2) exist. They all give rise to valid Hebbian learning rules that exhibit both potentiation and depression; see Table 19.1.

Example: Covariance rule

Sejnowski (1977) has suggested a learning rule of the form

$$\frac{d}{dt} w_{ij} = \gamma \left(v_i - \langle v_i \rangle \right) \left(v_j - \langle v_j \rangle \right), \tag{19.6}$$

called the covariance rule. This rule is based on the idea that the rates $v_i(t)$ and $v_j(t)$ fluctuate around mean values $\langle v_i \rangle, \langle v_j \rangle$ that are taken as running averages over the recent

firing history. To allow a mapping of the covariance rule to the general framework of Eq. (19.2), the mean firing rates $\langle v_i \rangle$ and $\langle v_j \rangle$ have to be constant in time.

Example: Oja's rule

All of the above learning rules had $c_2^{\text{pre}} = c_2^{\text{post}} = 0$. Let us now consider a nonzero quadratic term $c_2^{\text{post}} = -\gamma w_{ij}$. We take $c_{11}^{\text{corr}} = \gamma > 0$ and set all other parameters to zero. The learning rule

$$\frac{d}{dt} w_{ij} = \gamma [v_i v_j - w_{ij} v_i^2] \tag{19.7}$$

is called Oja's rule (Oja, 1982). Under some general conditions Oja's rule converges asymptotically to synaptic weights that are normalized to $\sum_j w_{ij}^2 = 1$ while keeping the essential Hebbian properties of the standard rule of Eq. (19.3); see Exercises. We note that normalization of $\sum_j w_{ij}^2$ implies competition between the synapses that make connections to the same postsynaptic neuron, i.e., if some weights grow, others must decrease.

Example: Bienenstock–Cooper–Munro rule

Higher-order terms in the expansion on the right-hand side of Eq. (19.2) lead to more intricate plasticity schemes. Let us consider

$$\frac{d}{dt} w_{ij} = \phi \left(v_i - v_\theta \right) v_j \tag{19.8}$$

with a nonlinear function ϕ and a reference rate v_θ. If we take v_θ to be a function $f(\langle v_i \rangle)$ of the average output rate $\langle v_i \rangle$, then we obtain the so-called Bienenstock–Cooper–Munro (BCM) rule (Bienenstock *et al.*, 1982).

The basic structure of the function ϕ is sketched in Fig. 19.5. If presynaptic activity is combined with moderate levels of postsynaptic excitation, the efficacy of synapses activated by presynaptic input is *decreased*. Weights are *increased* only if the level of postsynaptic activity exceeds a threshold, v_θ. The change of weights is restricted to those synapses which are activated by presynaptic input. A common choice for the function ϕ is

$$\frac{d}{dt} w_{ij} = \eta \, v_i \left(v_i - v_\theta \right) v_j = c_{21} v_i^2 v_j - c_{11}^{corr} v_i v_j, \tag{19.9}$$

which can be mapped to the Taylor expansion of Eq. (19.2) with $c_{21} = \eta$ and $c_{11}^{corr} = -\eta v_\theta$.

For stationary input, it can be shown that the postsynaptic rate v_i under the BCM-rule (19.9) has a fixed point at v_θ which is unstable (see Exercises). To avoid the postsynaptic firing rate blowing up or decaying to zero, it is therefore necessary to turn v_θ into an adaptive variable which depends on the average rate $\langle v_i \rangle$. The BCM rule leads to input

selectivity (see Exercises) and has been successfully used to describe the development of receptive fields (Bienenstock *et al.*, 1982).