**CN510: Principles and Methods of Cognitive and Neural Modeling**

# Learning Rules for Continuous Models

## Lecture 12

Instructor: Anatoli Gorchetchnikov <anatoli@bu.edu>

# Hebbian Postulate

*When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased*

Usually written as: $\dot{w} = \eta x_{pre} x_{post}$

Minimalistic correlation based rule often voiced as “neurons that fire together wire together”

Due to this interpretation correlational learning is often called Hebbian learning

# Levy and Desmond's Postulated Learning Rules

Based on results of neurophysiological studies of neural plasticity *Levy and Desmond (1985)* suggested several different learning processes

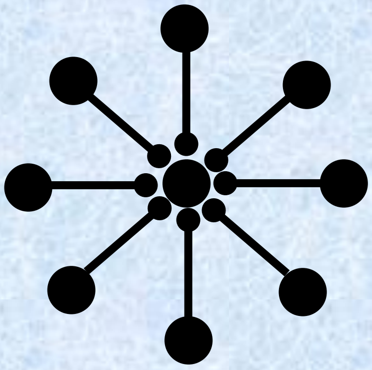
Based on LTP in various areas of the brain:

Rule 1 (Hebb/Anti-Hebb):

$$\dot{w}_{ij} = c_1 y_j \left( x_i - c_2 w_{ij} \right)$$

Rearranging the terms shows that this is what in modeling literature is called Hebbian rule with postsynaptically gated decay (aka instar)

$$\dot{w}_{ij} = \eta x_i y_j - \alpha y_j w_{ij}$$



# Post-synaptically Gated Decay

$$\dot{w}_{ij} = \eta x_i y_j - \alpha y_j w_{ij}$$

The rules like that has been reported in various parts of the nervous system: most notably hippocampus and visual cortex

It consists of Hebbian part and anti-Hebbian part:

- When presynaptic cell fires and postsynaptic cell fires – the weight increases
- When presynaptic cell is silent but postsynaptic cell fires – the weight decreases

Since decrease of weight happens driven by postsynaptic process this rule is local to the synapse

# Levy and Desmond's Postulated LTM Rules

Based on synaptic proliferation following postsynaptic inactivity

Rule 2:

$$\sum_i w_{ij} = \frac{c_1}{c_2 + E(y_j^p)}$$

A total strength of incoming synaptic weights is fixed for a given firing rate

Increase in firing rate leads to decrease in weights and vice versa

Weight normalization based on neurophysiological results

Couple issues:

- Need to know all incoming weights to set each of them (non-local)
- No mechanism to set relative magnitude of individual weight: needs another component to do this, Hebb will do

# BCM Family of Learning Rules

Also includes the use of averaged firing rates:

$$\dot{w}_{ij} = \eta x_i \left( y_j - \theta_j \right) y_j$$

$$\dot{\theta}_j = \varepsilon \left( y_j^2 - \theta_j \right)$$

If the current firing rate is higher than average then the weight increases in presence of presynaptic firing

But the running average will increase faster than the weight and eventually overshoot the firing rate, so the weight will start to decrease

Net effect is similar: weights are adjusted to keep constant postsynaptic firing rate

This rule is local, but also does not pay precise attention to individual input components

# Levy and Desmond's Postulated LTM Rules

Based on axonal growth properties

Rule 3:

$$\sum_j w_{ij} = cE(x_i)$$

Total outgoing weights are proportional to the expected cell activity (note that this is presynaptic)

Can be rewritten in BCM-like formalism; gives one of the covariance rules

$$\dot{w}_{ij} = \eta\left(x_i - \theta_i\right) y_j$$

$$\dot{\theta}_i = \varepsilon\left(x_i - \theta_i\right)$$

There is no fixed point here (therefore the rule is unstable) and postsynaptic activity does not matter

*Levy and Desmond* postulated also some rules for inhibitory synapses, but these were not based on neurophysiology

There are many other rules that we postulated over years from mathematical as well as biological considerations

# Variations of Classic Hebb

Simple $\dot{w}_{ij} = \eta x_i y_j$

Passive decay $\dot{w}_{ij} = \eta x_i y_j - \alpha w_{ij}$

Postsynaptically gated decay $\dot{w}_{ij} = \eta x_i y_j - \alpha y_j w_{ij}$

Oja rule $\dot{w}_{ij} = \eta x_i y_j - \alpha y_j^2 w_{ij}$

Presynaptically gated decay $\dot{w}_{ij} = \eta x_i y_j - \alpha x_i w_{ij}$

Dual gating of decay: additive (OR)

$$\dot{w}_{ij} = \eta x_i y_j - \alpha \left( x_i + y_j \right) w_{ij}$$

Dual gating of decay: multiplicative (AND)

$$\dot{w}_{ij} = \eta x_i y_j - \alpha x_i y_j w_{ij}$$

Note: the latter is not used in continuous firing rate models and is provided for completeness only

# Stability: General Approach

Convert the network to matrix notation

$$\dot{w}_{ij} = \eta x_i y_j - \alpha f\left(x_i, y_j\right) w_{ij} \qquad Y = X^T W$$

$$\dot{w}_i = \eta x_i y - \alpha f\left(x_i, y\right) w_i \quad \rightarrow \quad A = \eta X X^T - \alpha f(X, Y)$$

$$y = \sum_i x_i w_i \qquad \dot{W} = AW$$

Evaluate eigenvalues of the matrix $A$ to analyze stability

We can also look at equilibrium solution for

$$\dot{w}_{ij} = \eta x_i y_j - \alpha f\left(x_i, y_j\right) w_{ij}$$

to analyze what the final weight distribution would be like

# Simple Hebb $\dot{w}_{ij} = \eta x_i y_j$

Obviously if $x_i$, $y_j$, and $\eta$ are $>0$ then $\lim_{t \to \infty} w_{ij} = \infty$

We can also recast it in matrix notation

$$\dot{W} = \eta X y_j = \eta X X^T W$$

or

$$\dot{W} = AW$$

$$A = \eta X X^T = \eta C$$

with the solution $W = W_0 e^{At} = P diag(e^{At}) P^{-1} W_0$

dominated by the largest eigenvalue of $A$

But matrix $A$ is positive semidefinite and thus has all eigenvalues positive

Therefore the weights will grow exponentially

# Hebb with Passive Decay $\dot{w}_{ij} = \eta x_i y_j - \alpha w_{ij}$

Here the equilibrium solution is $\lim_{t \to \infty} w_{ij} = \frac{\eta x_i y_j}{\alpha}$

That means that given the limited $x_i$ and $y_j$, the weights will also be limited

Matrix analysis adds to it

$$\dot{W} = \eta X X^T W - \alpha W = AW$$

$$A = \eta X X^T - \alpha I$$

The eigenvalues of matrix $A$

$$\lambda_1 = \eta\left(x_1^2 + x_2^2 + \ldots + x_m^2\right) - \alpha$$

$$\lambda_2 = \lambda_3 = \ldots = \lambda_m = -\alpha$$

So the condition for stability is $x_1^2 + x_2^2 + \ldots + x_m^2 \leq \frac{\alpha}{\eta}$

# Hebb with Passive Decay $\dot{w}_{ij} = \eta x_i y_j - \alpha w_{ij}$

Note that if $x_1^2 + x_2^2 + \ldots + x_m^2 < \dfrac{\alpha}{\eta}$

then all eigenvalues are negative, meaning that the solution will decay to zero

Basically the only good case is $x_1^2 + x_2^2 + \ldots + x_m^2 = \dfrac{\alpha}{\eta}$

which requires normalization of squares of activations – hard task to accomplish

Summary:

- with too much input it will go out of bounds,
- with too little – weights will go to 0
- having just the right input is non-trivial

# Pre-synaptically Gated Decay (Outstar)

$$\dot{w}_{ij} = \eta x_i y_j - \alpha x_i w_{ij}$$

Equilibrium solution

$$\lim_{t \to \infty} w_{ij} = \frac{\eta x_i y_j}{\alpha x_i} = \frac{\eta y_j}{\alpha} \sim y_j$$

Meaning that the weights will track the activity of the postsynaptic cell

Opposite of Levi's rule 2 since here the higher the postsynaptic activity *y* the stronger it will make the input

In matrix notation $A = \eta X X^T - \alpha X$

Diagonal can be factored out $A = diag(X)\left(\eta [X_i] - \alpha I\right)$

And eigenvalues are

$$\lambda_1 = \eta\left(x_1 + x_2 + \ldots + x_m\right) - \alpha$$

$$\lambda_2 = \lambda_3 = \ldots = \lambda_m = -\alpha$$

# Pre-synaptically Gated Decay (Outstar)

$$\dot{w}_{ij} = \eta x_i y_j - \alpha x_i w_{ij}$$

The network is stable if

$$x_1 + x_2 + \ldots + x_m \leq \frac{\alpha}{\eta}$$

Weights will not converge to zero if

$$x_1 + x_2 + \ldots + x_m = \frac{\alpha}{\eta}$$

Similar to passive decay, but no squares, thus this condition is easier to achieve

Given normalized inputs this rule

- is usable when inputs are all or none and response (*y*) magnitude is set elsewhere and
- is important: e.g. ART top down projections

# Post-synaptically Gated Decay (Instar)

$$\dot{w}_{ij} = \eta x_i y_j - \alpha y_j w_{ij}$$

Equilibrium solution

$$\lim_{t \to \infty} w_{ij} = \frac{\eta x_i y_j}{\alpha y_j} = \frac{\eta x_i}{\alpha} \sim x_i$$

Meaning that the weights will track the activity of the presynaptic cell

In matrix notation $A = \eta X X^T - \alpha X^T W = \eta C - \alpha X^T W$

which is weight (or time) dependent and this changes the solution of the equation $\dot{W} = AW$ to

$$W = \frac{\eta}{\alpha} X \frac{1}{I + e^{-\eta C t \left( \frac{\eta C}{\alpha X^T W_0} - I \right)}}$$

Here exponent goes to *0* with time and the weights are stable

# Oja Rule $\dot{w}_{ij} = \eta x_i y_j - \alpha y_j^2 w_{ij}$

Equilibrium solution

$$\lim_{t \to \infty} w_{ij} = \frac{\eta x_i y_j}{\alpha y_j^2} = \frac{\eta x_i}{\alpha y_j} \sim \frac{x_i}{y_j}$$

Meaning that the weights will track the ratio of pre- and postsynaptic activities

In matrix notation $A = \eta C - \alpha W^T CWI$ where $W^T CW$ is a scalar

The solution for $\dot{W} = AW$ is $W(t) = \sqrt{\frac{\eta}{\alpha}} \frac{e^{\eta C t} \cdot W_0}{\sqrt{\|e^{\eta C t} W_0\|^2 - \|W_0\|^2 + \frac{\eta}{\alpha}}}$

Which converges with time to a principal eigenvector of covariance matrix $C = XX^T$ corresponding to eigenvalue

$$\lambda = x_1^2 + x_2^2 + \ldots + x_m^2$$

# BCM Family of Rules

Based on two equations

$$\dot{w}_{ij} = \eta x_i \left( y_j - \theta_j \right) y_j$$

$$\dot{\theta}_j = \varepsilon \left( y_j^2 - \theta_j \right)$$

Variations of the first equation led to

IBCM $\dot{w}_{ij} = \eta x_i \left( y_j - \theta_j \right) y_j \sigma\left( y_j \right)$

LBCM $\dot{w}_{ij} = \eta x_i \left( y_j - \theta_j \right) \frac{y_j}{\theta_j}$

Original BCM also had passive decay term

In matrix form $\dot{W} = \eta X (y - \theta) y$

Multiplied from the left by $X^T$ : $X^T \dot{W} = \eta X^T X (y - \theta) y$

And substituting $X^T X = \|X\|$ : $\dot{y} = \eta \|X\| (y - \theta) y$

System

$$\dot{y} = \eta \|X\| (y - \theta) y$$

$$\dot{\theta}_j = \varepsilon \left( y_j^2 - \theta_j \right)$$

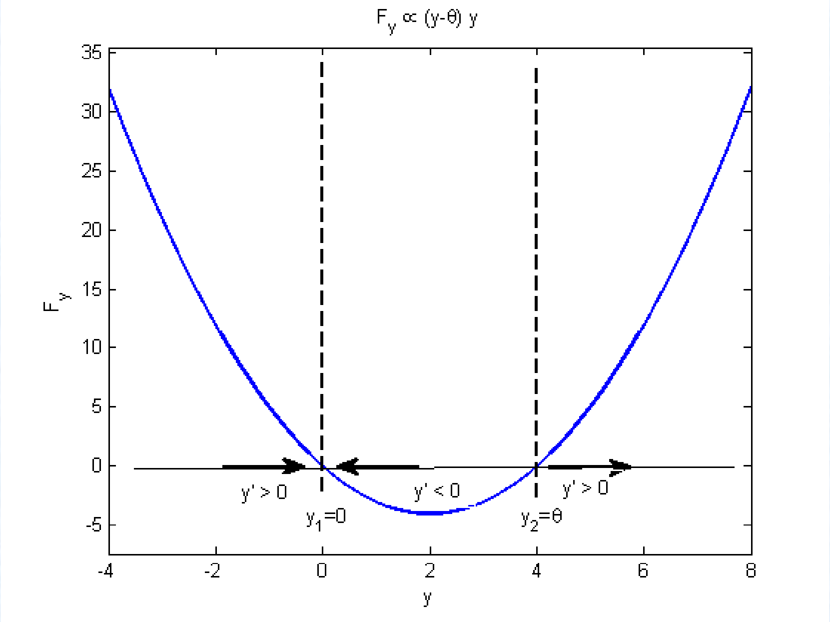
with fixed $\theta$ has two equilibrium points: $y=0$ and $y=\theta$ which can be analyzed further

# Stability with Fixed Threshold

If $y<0$ then it will increase and so will the weight

If $0<y<\theta$ then it will decrease and so will the weight

If $y>\theta$ then it will increase and so will the weight




This shows that stability of the rule indeed relies on sliding threshold

$$\dot{y} = \eta \|X\| (y - \theta) y$$

# Stability with Sliding Threshold

Substituting equilibrium point of the second equation $\theta=y^2$ in the first equation gives

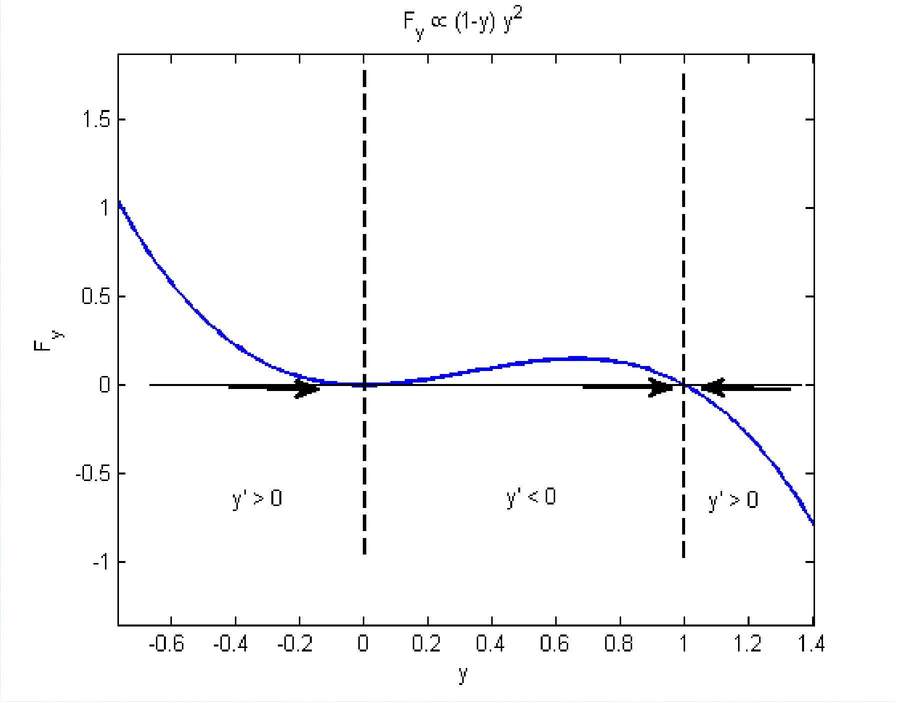
$$\dot{y}=\eta\|X\|\left(y-y^2\right)y$$

or

$$\dot{y}=\eta\|X\|(1-y)\,y^2$$

which has two equilibrium points: $y=0$ and $y=1$

But here the second one is stable




BCM ensures that postsynaptic rate settles to 1

Lets name $$F_y(y,\theta) = \dot{y} = \eta \left\| X \right\| \left( y - \theta \right) y$$
$$F_\theta(y,\theta) = \dot{\theta}_j = \varepsilon \left( y_j^2 - \theta_j \right)$$

The stability of the critical point is determined by the eigenvalues of the Jacobian matrix

$$\begin{pmatrix} \dfrac{\partial F_y(y,\theta)}{\partial y} & \dfrac{\partial F_y(y,\theta)}{\partial \theta} \\ \dfrac{\partial F_\theta(y,\theta)}{\partial y} & \dfrac{\partial F_\theta(y,\theta)}{\partial \theta} \end{pmatrix}$$

# Stability Conditions

$$F_y(y,\theta) = \eta\|X\|(y-\theta)\,y$$
$$F_\theta(y,\theta) = \varepsilon\left(y^2-\theta\right)$$

At the point *y=0*

$$\begin{pmatrix} -\eta\|X\|\theta & 0 \\ 0 & -\varepsilon \end{pmatrix}$$

For rates $\eta$ and $\varepsilon$ *>0* the stability condition is $\|X\| > 0$

At the point *y=1*

$$\begin{pmatrix} \eta\|X\|\theta & -\eta\|X\|\theta \\ 2\varepsilon\theta & -\varepsilon \end{pmatrix}$$

The stability condition is $\dfrac{\eta}{\varepsilon}\|X\| < 1$

Stability condition

$$\frac{\eta}{\varepsilon}\|X\|<1$$

means that given slow enough learning rate and fast enough threshold sliding rate the rule is stable for any input

Also means that if inputs are normalized the rule is stable if weight changes even slightly slower than threshold

Same condition applies to LBCM, IBCM is always stable

For all rules in BCM family

$$\lim_{t\to\infty} y_j = 1$$

# Next Week

No lecture on Tuesday

Next Thursday:

Spike-Timing-Dependent Plasticity

Readings:

Morrison, A, Diesmann, M and Gerstner, W. (2008). Phenomenological models of synaptic plasticity based on spike timing. *Biological Cybernetics 98*: 459-478.

Izhikevich E.M. and Desai N.S. (2003). Relating STDP to BCM. *Neural Computation 15*: 1511-1523