Fixed point representation

v: 没有生物学对应.

比较难算：网络很大时运算量很大，n^3的运算量（n: 神经元个数）

$$\frac{dr}{dt} = F(r, w, x, y) \qquad r \in R^n, w \in R^m$$

$r^*(w)$: r^*是关于w的函数

$$r_\infty = r^* \quad \frac{dr}{dt}=0 \quad 0 = F(r^*, w, x, y)$$

稳态

Gradient based learning:

欧式距离

$$loss\ function = l(r, y) = \frac{1}{2}(r-y)^2$$

$= \frac{1}{2}\|r-y\|^2$ 矢量

$$\frac{dl}{dw} = \frac{\partial l}{\partial r}\Big|_{r^*}\frac{dr^*}{dw}$$

链式法则：r^*已知，$\frac{\partial l}{\partial r}\big|_{r^*}$已知

如何求

两边同时求导

$$\frac{d0}{dw} = \frac{dF(r^*, w, x, y)}{dw} = \underbrace{\frac{\partial F}{\partial r}\Big|_{r^*}}_{\text{已知}}\frac{dr^*}{dw} + \frac{\partial F}{\partial w}$$

复合求导

example

$F = r_1^2 w$

$\frac{\partial F}{\partial w} = r_1^2$

$$J(r^*) = \frac{\partial F}{\partial r}\Big|_{r^*} \qquad \frac{\partial l}{\partial r^*} = \frac{\partial l}{\partial r}\Big|_{r^*}$$

雅可比矩阵

$$\frac{dr^*}{dw} = -J^{-1}\frac{\partial F}{\partial w}$$

$$\frac{dl}{dw} = -\frac{\partial l}{\partial r^*}J^{-1}\frac{\partial F}{\partial w}$$

BP

重新构建一个动力学

$$\frac{dv}{dt} = vJ + \frac{\partial l}{\partial r^*} \Rightarrow \text{稳态是}\ \frac{dv}{dt}=0$$

$$v^* = -\frac{\partial l}{\partial r^*}J^{-1}$$

$$\frac{dl}{dw} = v^*\frac{\partial F}{\partial w}$$

BP算法.

设计J使得J sparse

才能降低复杂度

Energy-base Model（一无法对）对v的理解（是用）

r^*是关于λ的函数

$$\frac{dr}{dt} = F(r, w, x, y) + \lambda\left(\frac{\partial l}{\partial r}\right)^T$$

条件：r^*，$\frac{dr}{dt}=0$

关于λ求导

$$\frac{d0}{d\lambda} = \frac{dF(r^*, w, x, y)}{d\lambda} + \left(\frac{\partial l}{\partial r}\right)^T = J\frac{dr^*}{d\lambda} + \left(\frac{\partial l}{\partial r}\right)^T\Big|_{r^*}$$

$$\left(\frac{dr^*}{d\lambda}\right)^T = -\frac{\partial l}{\partial r}\Big|_{r^*}J^{-T}$$

Energy与最终输出无关

If $J^{-T} = J^{-1} \Leftrightarrow$ exist E, s.t. $F = \frac{\partial E}{\partial r}$

$J^T = J$ J是对称矩阵，F两等式

$$\frac{dl}{dw} = \left(\frac{dr^*}{d\lambda}\right)^T\Big|_{\lambda=0}\frac{\partial F}{\partial w}$$

$r^*(\lambda)$

- Energy-based model

整体思路：

loss function =0 最优

通常loss function只有一个最小值

如何计算一个方程的零点（最小值）

连续函数也就是导数为0的地方

所以要对loss function求导

可以用实验模拟

$$\frac{dr^*(\lambda)}{d\lambda} = \frac{r^*(0+\lambda)-r^*(0)}{\lambda}$$

从而得到$\frac{dr^*(\lambda)}{d\lambda}$

$\lambda=0$，得到r^*

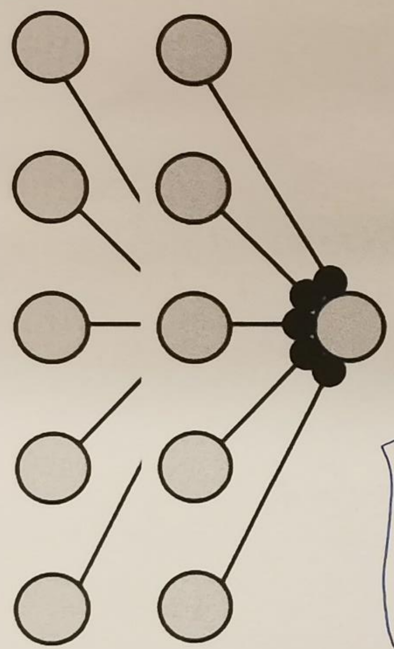
$\lambda=0.1$，得到r_1

BP算法计算过程.

v: 没有生物学对应，由脑科学中无法解释

北京大学 PEKING UNIVERSITY

Feedforward model



$$x_1 = f(w_1x_0)$$
$$x_2 = f(w_2x_1)$$

$$\frac{dr}{dt} = F(r,w,x,y)$$

$$\frac{dx_1}{dt} = -x_1 + f(w_1x_0)$$

$$\frac{dx_2}{dt} = -x_2 + f(w_2x_1)$$

$x_0 \searrow w_1 \nearrow x_1 \searrow w_2 \nearrow x_2$

前馈神经网络

$$F = \begin{pmatrix} f(w_1x_0)-x_1 \\ f(w_2x_1)-x_2 \end{pmatrix}$$

$$J = \frac{\partial F}{\partial x}$$

训练方法

BP算法

$$\frac{dl}{dw} = -\frac{\partial l}{\partial r^*}J^{-1}\frac{\partial F}{\partial w}$$

$$\frac{dv}{dt} = vJ + \frac{\partial l}{\partial r^*}$$

$$v^* = -\frac{\partial l}{\partial r^*}J^{-1}$$

$$\frac{dl}{dw} = v^*\frac{\partial F}{\partial w}$$

$$\frac{dl}{d(w_1,w_2)} = -\frac{\partial l}{\partial(x_1,x_2)}J^{-1}\frac{\partial F}{\partial(w_1,w_2)}$$

$$J = \begin{pmatrix} -1 & 0 \\ w_2f'(w_2x_1) & -1 \end{pmatrix}$$

$$\frac{d(v_1,v_2)}{dt} = (v_1,v_2)\begin{pmatrix} -1 & 0 \\ w_2f'(w_2x_1) & -1 \end{pmatrix} + \frac{\partial l}{\partial(x_1,x_2}$$

$$\begin{cases} -v_1 + v_2w_2f'(w_2x_2) + 0 = 0 \\ -v_2 + \frac{\partial l}{\partial x_2} = 0 \end{cases}$$

$$v_2^* = \frac{\partial l}{\partial x_2}, v_1^* = \frac{\partial l}{\partial x_2}w_2f'(w_2x_2)$$

$$\frac{dl}{d(w_1,w_2)} = \begin{pmatrix} \overset{[a_1]}{\frac{\partial l}{\partial x_2}w_2f'(w_2x_2)} \\ \underset{[a_2]}{\frac{\partial l}{\partial x_2}} \end{pmatrix} \begin{pmatrix} \overset{[b_1]}{\frac{\partial f(w_1x_0)}{\partial w_1}} & 0 \\ 0 & \underset{[b_2]?}{\frac{\partial f(w_2x_1}{\partial w_2}} \end{pmatrix}$$

$$\frac{dl_1}{dw_1} = a_1b_1 \quad \frac{dl_2}{dw_2} = a_2b_2$$

~~不能引入回转变量v~~

Trajectory representation

北京大学 PEKING UNIVERSITY

$$\frac{dr}{dt} = F(r, w, x, y) \qquad r \in R^n, w \in R^m$$

$=\frac{1}{2}(r_t - y_t)^2$

$$loss\ function:\ l = \int \alpha_t l_t(r_t, y_t)dt$$

$\alpha_t = 0$，对t时刻输出没要求.

$$\frac{dl_t(r_t, y_t)}{dw} = \underbrace{\frac{\partial l_t}{\partial r_t}}_{易} \underbrace{\frac{dr_t}{dw}}_{难}$$

w对每一时刻的r都有影响

$$\frac{dr_t}{dw} = \frac{d}{dw}\int_0^t dr_\tau = \frac{d}{dw}\int_0^t \boxed{\frac{dr_\tau}{d\tau}} d\tau$$

$$= \frac{d}{dw}\int_0^t F(r, w, x, y)d\tau$$

$$= \frac{d}{dw}\int_0^t F(r, w, x, y)d\tau$$

$$= \int_0^t \frac{dF(r, w, x, y)}{dw} d\tau$$

设 $P_t = \int_0^t \frac{dF(r,w,x,y)}{dw} d\tau$

Online

$$p_t = \frac{dr_t}{dw}$$

$$\frac{dp_t}{dt} = \frac{dF(r, w, x, y)}{dw} = \underbrace{\boxed{J(r_t)}}_{n*n} p_t + \frac{\partial F}{\partial w}$$

Real time recurrent learnin

Time: $O(n^2 m * T)$

Space: $O(mn + n^2)$

$m \approx n^2$

$$\frac{dp_t}{dt} = J(r_t)p_t + \frac{\partial F}{\partial w}$$

$$p_t = [J(r_{t-1})\Delta t + 1]p_{t-1} + \frac{\partial F(r_{t-1})}{\partial w}\Delta t \quad 欧拉展开$$

$r_t = F(r_{t-1})\Delta t + r_{t-1}$

$$\frac{\partial r_t}{\partial r_{t-1}} = \frac{\partial F}{\partial r}\Big|_{r_{t-1}}$$

$(J \cdot \Delta t + 1)$

$$p_t = \frac{\partial r_t}{\partial r_{t-1}} p_{t-1} + \frac{\partial F(r_{t-1})}{\partial w}\Delta t$$

$$p_t = \frac{\partial r_t}{\partial r_{t-1}}\frac{\partial r_{t-1}}{\partial r_{t-2}} p_{t-1} + \frac{\partial r_t}{\partial r_{t-1}}\frac{\partial F(r_{t-1})}{\partial w}\Delta t + \frac{\partial F(r_{t-1})}{\partial w}\Delta t$$

$$\frac{dr_t}{dw} = p_t = \int_0^t \frac{\partial r_t}{\partial r_\tau}\frac{\partial F(r_\tau, w, x, y)}{\partial w} d\tau$$

$$\frac{dl_t(r_t, y_t)}{dw} = \int_0^t \underbrace{\left(\frac{\partial l_t}{\partial r_t}\right)}_{1*n} \underbrace{\left(\frac{\partial r_t}{\partial r_\tau}\right)}_{n*n} \underbrace{\left(\frac{\partial F(r_\tau, w, x, y)}{\partial w}\right)}_{n*m} d\tau$$

BPTT

Time: $O(n^2 T + nmT)$

Space: $O(mn + n^2)$