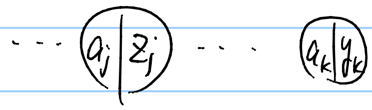
(目標) Ω_nを計算できる順方向の方程式を求める

$$\alpha_j \equiv G z_j, \quad \beta_j \equiv G a_j \quad \cdots (5.205)$$



なので

$$\alpha_k = G y_k$$

である。これより

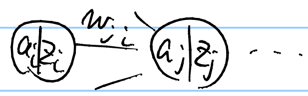
$$\Omega_n = \frac{1}{2} \sum_k (G y_k)^2 \quad \cdots (5.201)$$

は

$$\Omega_n = \frac{1}{2} \sum_k \alpha_k^2$$

と書ける。

順伝播の方程式の



$$a_j = \sum_i w_{ji} z_i, \quad z_j = h(a_j) \quad \cdots (5.204)$$

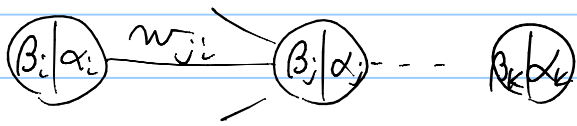
両辺にGを作用させて

$$G a_j = \sum_i w_{ji} G z_i, \quad G z_j = h'(a_j) G a_j$$

を得る。(5.205)を使って

$$\beta_j = \sum_i w_{ji} \alpha_i, \quad \alpha_j = h'(a_j) \beta_j \quad \cdots (5.204)$$

を得る。(5.204)は



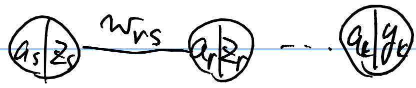
という感じの順方向の方程式になっている。

(目標) $\frac{\partial \Omega_n}{\partial w_{rs}}$ が (5.206) の形になることを示す。

(5.204) より

$$\frac{\partial \Omega_n}{\partial w_{rs}} = \frac{\partial}{\partial w_{rs}} \frac{1}{2}\sum_k (\mathcal{G} y_k)^2 = \frac{1}{2}\sum_k 2(\mathcal{G} y_k)\frac{\partial}{\partial w_{rs}}(\mathcal{G} y_k)$$

$$= \sum_k (\mathcal{G} y_k)\left(\mathcal{G}\frac{\partial y_k}{\partial w_{rs}}\right)$$



$$= \sum_k (\mathcal{G} y_k)\left(\mathcal{G}\frac{\partial a_r}{\partial w_{rs}}\frac{\partial y_k}{\partial a_r}\right)$$

$a_r = \sum_l w_{rl} z_l$ なので $\frac{\partial a_r}{\partial w_{rs}} = z_s$

$$= \sum_k (\mathcal{G} y_k)\left(\mathcal{G}\, z_s \frac{\partial y_k}{\partial a_r}\right)$$

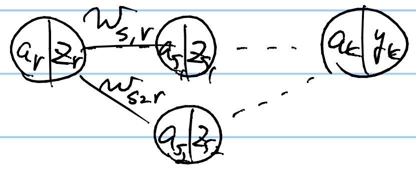
$$= \sum_k \underset{\alpha_k}{\underset{\shortparallel}{(\mathcal{G} y_k)}}\left\{\underset{\alpha_s}{\underset{\shortparallel}{(\mathcal{G} z_s)}}\underset{\delta_{kr}}{\underset{\shortparallel}{\frac{\partial y_k}{\partial a_r}}} + z_s \underbrace{\mathcal{G}\underset{\delta_{kr}}{\underset{\shortparallel}{\frac{\partial y_k}{\partial a_r}}}}_{\underset{\phi_{kr}}{\shortparallel}}\right\}$$

ここで $\alpha_i = \mathcal{G} z_i$, $\delta_{kr} = \frac{\partial y_k}{\partial a_r}$, $\phi_{kr} = \mathcal{G}\delta_{kr}$ とおくと

$$\frac{\partial \Omega_n}{\partial w_{rs}} = \sum_k \alpha_k \{\alpha_s \delta_{kr} + z_s \phi_{kr}\} \quad \cdots (5.206)$$

を得る。

(目標) $\overset{\phi_{kr}}{\delta_{kr}}$ を算出できる逆伝播方程式を導く

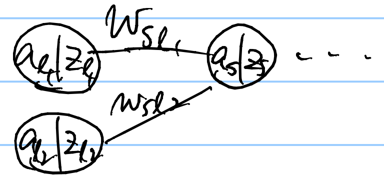


$$\delta_{kr} = \frac{\partial y_k}{\partial a_r} = \sum_s \frac{\partial a_s}{\partial a_r}\frac{\partial y_k}{\partial a_s}$$ ← チェインルール

$$= \sum_s \frac{\partial a_s}{\partial a_r}\delta_{ks} = \sum_s w_{sr} h'(a_r)\delta_{ks}$$

$$\therefore \delta_{kr} = \sum_s w_{sr} h'(a_r)\delta_{ks} = h'(a_r)\sum_s w_{sr}\delta_{ks}$$

↑
和が上流を指す逆伝播の式
になっている。



$$a_s = \sum_l w_{sl} z_l = \sum_l w_{sl} h(a_l)$$
なので
$$\frac{\partial a_s}{\partial a_r} = w_{sr} h'(a_r)$$

$$\delta_{kk} = \frac{\partial y_k}{\partial a_k} = \sigma'(a_k) \quad \cdots ②$$

①,②が δ_{kr} を求める逆伝播の式

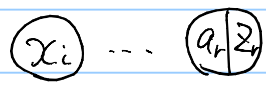
ϕ_{kr} については

$$\phi_{kr} = G\delta_{kr} = G\, h'(a_r)\sum_s w_{sr}\delta_{ks}$$

$$= (G h'(a_r))\sum_s w_{sr}\delta_{ks} + h'(a_r)\sum_s w_{sr} G\delta_{ks}$$

$$= (G a_r) h''(a_r)\sum_s w_{sr}\delta_{ks} + h'(a_r)\sum_s w_{sr}\phi_{ks}$$

$$= \beta_r h''(a_r)\sum_s w_{sr}\delta_{ks} + h'(a_r)\sum_s w_{sr}\phi_{ks}$$



$$G h'(a_r) = \sum_i \tau_i \frac{\partial}{\partial x_i} h'(a_r) = \sum_i \tau_i \frac{\partial a_r}{\partial x_i} h''(a_r) = (G a_r) h''(a_r)$$

を得る。これは δ_{ks} と上流の ϕ_{ks} から ϕ_{kr} を求める逆伝播の式になっている。