5.26 Proving 5.201:

Using 5.128,

$$\Omega_n = \frac{1}{2}\sum_k \left(\left.\frac{\partial y_{nk}}{\partial \xi}\right|_{\xi=0} \right)^2$$

From 5.126, we know that

$$\left.\frac{\partial y_{nk}}{\partial \xi}\right|_{\xi=0} = \sum_{i=1}^{D} \frac{\partial y_{nk}}{\partial x_i} \left.\frac{\partial x_i}{\partial \xi}\right|_{\xi=0} \qquad \text{and} \qquad \frac{\partial x_i}{\partial \xi} = \tau_i$$

Substituting above, we get:

$$\Omega_n = \frac{1}{2}\sum_k \left(\sum_{i=1}^{D} \tau_i \left.\frac{\partial y_{nk}}{\partial x_i}\right|_{\xi=0} \right)^2$$

Substituting for $\mathcal{G}$, where

$$\mathcal{G} \equiv \sum_i \tau_i \frac{\partial}{\partial x_i}$$

we get

$$\Omega_n = \frac{1}{2}\sum_k \left(\mathcal{G} y_{nk}\right)^2.$$

Proving 5.203:

$$\begin{aligned}
\mathcal{G} z_j &= \sum_i \tau_i \frac{\partial z_j}{\partial x_i} \\
&= \sum_i \tau_i \frac{\partial h(a_j)}{\partial x_i} \\
&= \sum_i \tau_i \frac{\partial h(a_j)}{\partial a_j} \frac{\partial a_j}{\partial x_i}
\end{aligned}$$

$$= \frac{\partial h(a_j)}{\partial a_j} \sum_i \tau_i \frac{\partial a_j}{\partial x_i}$$

$$= h'(a_j) \mathcal{G} a_j$$

$$= h'(a_j) \beta_j$$

$$\mathcal{G} a_j = \sum_i \tau_i \frac{\partial a_j}{\partial x_i}$$

$$= \sum_i \tau_i \frac{\partial \left(\sum_{i'} w_{ji'} z_{i'}\right)}{\partial x_i}$$

$$= \sum_i \tau_i \left(\sum_{i'} w_{ji'} \frac{\partial z_{i'}}{\partial x_i} \right)$$

$$= \sum_i \sum_{i'} \tau_i w_{ji'} \frac{\partial z_{i'}}{\partial x_i}$$

$$= \sum_{i'} w_{ji'} \left(\sum_i \tau_i \frac{\partial z_{i'}}{\partial x_i} \right)$$

$$= \sum_{i'} w_{ji'} \mathcal{G} z_{i'}$$

$$= \sum_{i'} w_{ji'} \alpha_{i'}.$$

Proving 5.206:

$$\Omega_n = \frac{1}{2} \sum_k \left(\sum_i J_{nki} \tau_{ni} \right)^2$$

$$\frac{\partial \Omega_n}{\partial w_{rs}} = \frac{\partial}{\partial w_{rs}} \left(\frac{1}{2} \sum_k \left(\sum_i J_{nki} \tau_{ni} \right)^2 \right)$$

$$= \frac{1}{2}\sum_k \frac{\partial}{\partial w_{rs}} \left(\sum_i J_{nki}\tau_{ni}\right)^2$$

$$= \frac{1}{2}\sum_k 2\left(\sum_i J_{nki}\tau_{ni}\right)\left(\frac{\partial}{\partial w_{rs}}\left(\sum_i J_{nki}\tau_{ni}\right)\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\sum_i \frac{\partial J_{nki}\tau_{ni}}{\partial w_{rs}}\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\sum_i \tau_{ni}\frac{\partial J_{nki}}{\partial w_{rs}}\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\sum_i \tau_{ni}\frac{\partial^2 y_k}{\partial x_i \partial w_{rs}}\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\sum_i \tau_{ni}\frac{\partial}{\partial x_i}\left(\frac{\partial y_k}{\partial w_{rs}}\right)\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\sum_i \tau_{ni}\frac{\partial}{\partial x_i}\left(\frac{\partial y_k}{\partial a_r}\frac{\partial a_r}{\partial w_{rs}}\right)\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\sum_i \tau_{ni}\frac{\partial}{\partial x_i}\left(\delta_{kr} z_s\right)\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\tau_{ni}\left(\delta_{kr}\sum_i \frac{\partial z_s}{\partial x_i} + z_s \sum_i \frac{\partial \delta_{kr}}{\partial x_i}\right)\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\delta_{kr}\sum_i \tau_{ni}\frac{\partial z_s}{\partial x_i} + z_s \sum_i \tau_{ni}\frac{\partial \delta_{kr}}{\partial x_i}\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\delta_{kr}\mathcal{G}z_s + z_s\mathcal{G}\delta_{kr}\right)$$

$$= \sum_k \left(\sum_i J_{nki}\tau_{ni}\right)\left(\delta_{kr}\alpha_s + z_s\phi_{kr}\right)$$

Backpropagation equations for δ_{kr}:

$$\delta_{kr} = \frac{\partial y_k}{\partial a_r}$$

$$= \sum_s \frac{\partial y_k}{\partial a_s} \frac{\partial a_s}{\partial a_r}$$

where the sum runs over all units s to which unit r sends connections.

$$= \sum_s \delta_{ks} \frac{\partial a_s}{\partial a_r}$$

$$= \sum_s \delta_{ks} \frac{\partial a_s}{\partial z_r} \frac{\partial z_r}{a_r}$$

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

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

Backpropagation equations for ϕ_{kr}:

$$\phi_{kr} = \mathcal{G}\delta_{kr}$$

$$= \mathcal{G} \left(h'(a_r) \sum_s \delta_{ks} \, w_{sr} \right)$$

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

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

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