第 l 層のユニット j の出力を作る際に、活性化関数に与える入力を $a_j^{(l)}$ とします。つまり第 l 層のユニット j の出力は $\sigma_l(a_j^{(l)})$ です。また

$$a_j^{(l)} = \sum_i w_{ji}^{(l-1)} \sigma_{l-1}(a_i^{(l-1)}) + b_j^{(l-1)}$$

の関係があるので、合成関数の微分を使うと以下が成立します。

$$\frac{\partial E_k}{\partial w_{ji}^{(l-1)}} = \frac{\partial E_k}{\partial a_j^{(l)}} \frac{\partial a_j^{(l)}}{\partial w_{ji}^{(l-1)}} = \frac{\partial E_k}{\partial a_j^{(l)}} \sigma_{l-1}(a_i^{(l-1)})$$

$$\frac{\partial E_k}{\partial b_j^{(l-1)}} = \frac{\partial E_k}{\partial a_j^{(l)}} \frac{\partial a_j^{(l)}}{\partial b_j^{(l-1)}} = \frac{\partial E_k}{\partial a_j^{(l)}}$$

つまり、第 $l-1$ 層と第 l 層の間に存在するパラメータは $\frac{\partial E_k}{\partial a_j^{(l)}}$ を計算することで求めることができます。

　更に多変数関数の合成関数の微分を使うと以下が成立します。

$$\frac{\partial E_k}{\partial a_j^{(l)}} = \sum_h \frac{\partial E_k}{\partial a_h^{(l+1)}} \frac{\partial a_h^{(l+1)}}{\partial a_j^{(l)}}$$

ここで

$$a_h^{(l+1)} = \sum_j w_{hj} \sigma_l(a_j^{(l)}) + b_k^{(l)}$$

なので

$$\frac{\partial a_h^{(l+1)}}{\partial a_j^{(l)}} = \sum_j w_{hj} \sigma_l'(a_j^{(l)})$$

が成立し、結局、

$$\frac{\partial E_k}{\partial a_j^{(l)}} = \sum_h \frac{\partial E_k}{\partial a_h^{(l+1)}} \sum_j w_{hj} \sigma_l'(a_j^{(l)})$$

となっています。つまり $\frac{\partial E_k}{\partial a_j^{(l)}}$ を計算するには、１つ上の層の $\frac{\partial E_k}{\partial a_h^{(l+1)}}$ を計算できればよいことがわかります。$\frac{\partial E_k}{\partial a_j^{(l)}}$ は第 l 層のユニット j の誤差を表しています。なので出力層の誤差から入力層に向かって、つまり逆向きに、誤差を伝播させてゆくことでパラメータを求める形になっているため、この手法は誤差逆伝播法と呼ばれます。

　なお最も上位の層となる出力層におけるユニット j の出力が $\sigma_2(a_j^{(3)})$ であり、しかもそれは f_j であったので