- IF LINEAR REGRESSION ($\sum e_n^2$)

$$\nabla E(\underline{w}) = \sum_{n=1}^{N} (\underline{w}^T \underline{\phi}_n - t_n)\underline{\phi}_n = \underline{\underline{\Phi}}^T \underline{\underline{\Phi}}\, \underline{w} - \underline{\underline{\Phi}}^T \underline{t}$$

$$\underline{\underline{H}} = \nabla\nabla E(\underline{w}) = \sum_{n=1}^{N} \underline{\phi}_n \underline{\phi}_n^T = \underline{\underline{\Phi}}^T \underline{\underline{\Phi}}$$

$$\underline{w}^{(new)} = (\underline{\underline{\Phi}}^T \underline{\underline{\Phi}})^{-1} \underline{\underline{\Phi}}^T \underline{t}$$

LEAST SQUARES

- FOR LOGISTIC REGRESSION (σ)

$$\nabla E(\underline{w}) = \sum_{n=1}^{N} (y_n - t_n)\phi_n = \underline{\underline{\Phi}}^T (\underline{y} - \underline{t})$$

$$\underline{\underline{H}} = \sum_{n=1}^{N} y_n(1-y_n)\phi_n \phi_n^T = \underline{\underline{\Phi}}^T \underset{\text{diag}[y_n(1-y_n)]}{\underline{\underline{R}}}\, \underline{\underline{\Phi}}$$

$$\underline{w}^{(new)} = \underline{w}^{(old)} - (\underline{\underline{\Phi}}^T \underline{\underline{R}}\, \underline{\underline{\Phi}})^{-1} \underline{\underline{\Phi}}^T (\underline{y} - \underline{t}) =$$

$$= (\underline{\underline{\Phi}}^T \underline{\underline{R}}\, \underline{\underline{\Phi}})^{-1} \left\{ \underline{\underline{\Phi}}^T \underline{\underline{R}}\, \underline{\underline{\Phi}}\, \underline{w}^{(old)} - \underline{\underline{\Phi}}^T (\underline{y} - \underline{t}) \right\}$$

$$= (\underline{\underline{\Phi}}^T \underline{\underline{R}}\, \underline{\underline{\Phi}})^{-1} \underline{\underline{\Phi}}^T \underline{\underline{R}}\, \underline{z} \qquad \underline{z} = \underline{\underline{\Phi}}\, \underline{w}^{(old)} - \underline{\underline{R}}^{-1} (\underline{y} - \underline{t})$$

<u>IRLS</u> - <u>ITERATIVE REWEIGHTED L S</u> ($\underline{\underline{R}}$ like variance; $\underline{z}$ like target $z_n \simeq a_n(\underline{w})$)

L MULTICLASS LOGISTIC REGRESSION (CH. 4.3.4)

$$E(\underline{w}_1, \dots, \underline{w}_K) \longrightarrow \nabla_{\underline{w}_j} E(\underline{w}_1, \dots, \underline{w}_K)$$

$$\nabla_{\underline{w}_k} \nabla_{\underline{w}_j} E(\underline{w}_1, \dots, \underline{w}_K)$$

ETC.

L CLASS CONDITIONAL DISTRIBUTION EXPONENTIAL $\longrightarrow$ EXTENSION TO OTHER MODELS
$\downarrow$
POSTERIOR CLASS PROBABILITY $\longleftarrow$ LOGISTIC TRANSFORMATION ON LINEAR FUNCTION OF THE FEATURES

L LAPLACE-APPROXIMATION: (CH. 4.4) $\quad q(z) \sim \frac{1}{Z} p(z)$ $\quad$ SECOND ORDER CENTERED ON THE MODE

$$q(z) = \left(\frac{A}{2\pi}\right)^{1/2} \exp\left\{ -\frac{A}{2}(z - z_0)^2 \right\} \qquad A = -\frac{d^2}{dz^2} \ln p(z)$$

$$\left.\frac{dp(z)}{dz}\right|_{z=z_0} = \phi$$