WEIGTHED LS — DATA WITH LARGER ERROR COUNTS LESS!

(EXERCISE 3.3)

$$e_n \longleftrightarrow \sigma_n^2$$

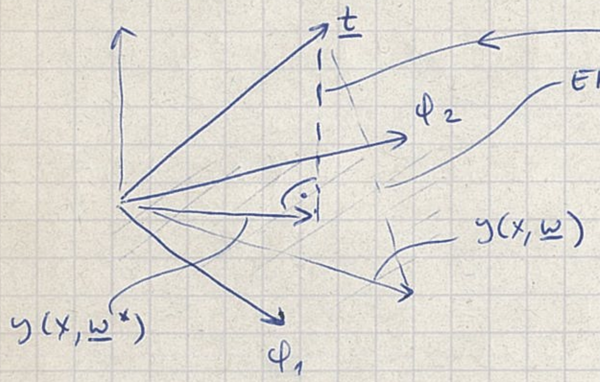
$$E(\underline{w}) = \frac{1}{2} \sum_{n=1}^{N} \frac{1}{\sigma_n^2} \{ \cdots \}^2 + \lambda \cdots$$

LET: $\underline{\underline{\Sigma}}^{-1} = \begin{bmatrix} 1/\sigma_1^2 & & \phi \\ & \ddots & \\ \phi & & 1/\sigma_N^2 \end{bmatrix}$ $\downarrow$ (CHECK IT)

$$\underline{w}_{LS}^{*} = \left(\underline{\underline{\Phi}}^T \underline{\underline{\Sigma}}^{-1} \underline{\underline{\Phi}} \right)^{-1} \underline{\underline{\Phi}}^T \underline{\underline{\Sigma}}^{-1} \underline{t}$$

GEOMETRICAL INTERPRETATION OF LS — ORTHOGONAL PROJECTION

(BISHOP 3.1.2)



FOR MIN ERROR:

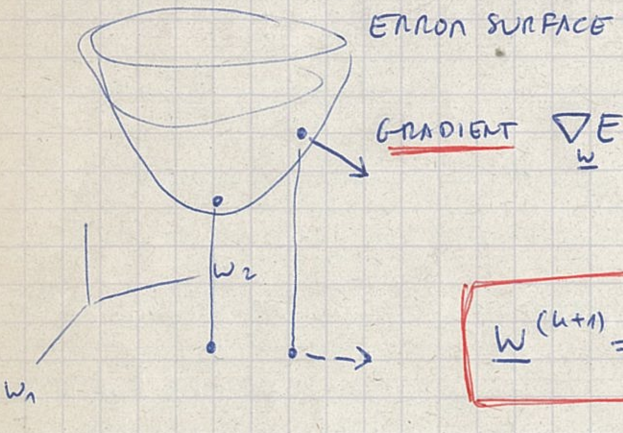
$$\left(\underline{t} - y(x, \underline{w}^*) \right) \perp y(x, \underline{w}^*)$$

OR

$$(\underline{t} - \underline{y})^T \underline{y} = \phi$$

SEQUENTIAL LEARNING — NOT LINEAR-IN-PARAMETERS

(BISHOP 3.1.3)



GRADIENT $\nabla_{\underline{w}} E = \left(\frac{\partial E}{\partial w_0}, \frac{\partial E}{\partial w_1}, \cdots \frac{\partial E}{\partial w_M} \right)$

IN OPTIMUM $\nabla_{\underline{w}} E \equiv \phi$

$$\boxed{\underline{w}^{(k+1)} = \underline{w}^{(k)} - \eta \nabla_{\underline{w}} E(\underline{w})}$$

LMS ALGORITHM: $\underline{w}(k+1) = \underline{w}(k) + 2\mu\, \varepsilon(k)\, \underline{\Phi}(x_k)$ $\quad \varepsilon(k) = t_k - y(x_k, \underline{w}^{(k)})$

(LIN-IN-PARAM)

$0 < \mu < \frac{1}{\lambda_{MAX}}$ — MAX. EIGENVALUE OF $\mathcal{E}\{ \underline{\Phi}(x) \underline{\Phi}^T(x) \}$