$$\underline{w}_{LS} = \arg\min_{\underline{w}} E(\underline{w}) \rightarrow \frac{\partial E(\underline{w})}{\partial w_k} = \phi \quad \forall k = 0, 1, \dots, M$$

(*)

$$\sum_{n=1}^{N} \left(\underline{\phi}^T(x_n)\, \underline{w} - t_n \right) \phi_k(x_n) = \phi$$

$$k=0 \quad \sum_{n=1}^{N} \underline{\phi}^T(x_n)\, \phi_0(x_n) \cdot \underline{w} = \sum_{n=1}^{N} \phi_0(x_n)\, t_n$$

$$k=1 \quad \vdots \qquad \vdots$$

$$k=M \quad \sum_{n=1}^{N} \underline{\phi}^T(x_n)\, \phi_M(x_n)\, \underline{w} = \sum_{n=1}^{N} \phi_M(x_n)\, t_n$$

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

CHECK IT!

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

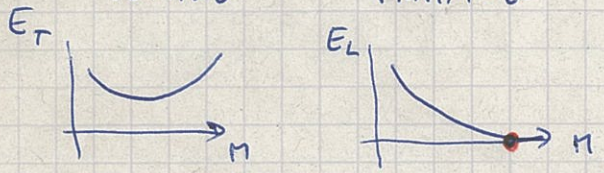
(BISHOP 1.1)

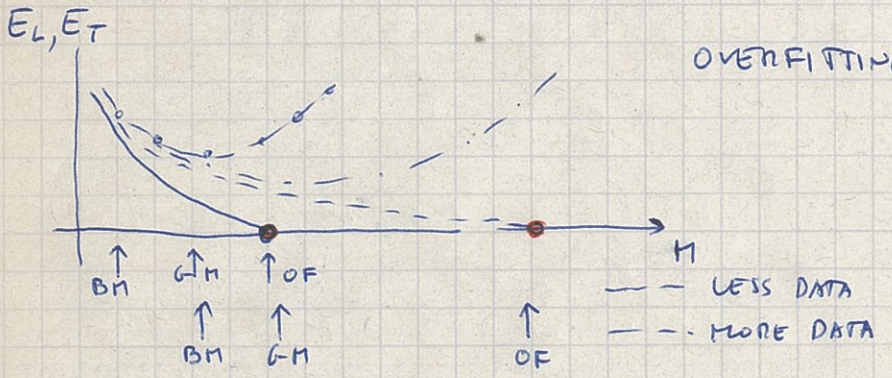
OVERFITTING

OF DATA $\longleftrightarrow$ # OF PARAMETERS IN THE MODEL (COMPLEXITY OF THE MODEL)

TESTING — TRAINING

$$E_{RMS} = \sqrt{2E(\underline{w}^*)/N}$$





OVERFITTING
- EXACT ON NOISY DATA — MODELS THE NOISE
- BAD ON TEST — CANNOT GENERALIZE

(LINEAR REGRESSION)

LEARNABLES — LINEAR-IN-PARAMETERS — ANALYTIC SOLUTION

— OVERFITTING — SELECTING MIDDLE COMPLEXITY MODELS, HOW?