BAYES LINEAR REGRESSION

- REGULARIZED LS & ML SOLUTION: HOW TO DETERMINE λ?
- GENERAL SOLUTION: CROSS VALIDATION — TRIAL-AND-ERROR
- GOAL: AVOID OVERFITTING: TO FIND THE PROPER MODEL COMPLEXITY

BAYES LINEAR REGRESSION — FROM TRAINING DATA

BASIC PRINCIPLE: MODEL PARAMETER VECTOR $\underline{w}$ IS A RANDOM VARIABLE
ITS STARTING (PRIOR) PROBABILITY DISTR. IS KNOWN
(MOST OFTEN GAUSSIAN $N(\underline{m}_0, \underline{\underline{\Sigma}}_{w_0})$)

PRIOR: $N(\underline{m}_0, \underline{\underline{\Sigma}}_{w_0})$

$$p(\underline{w}) = \frac{1}{(2\pi)^{M/2} |\underline{\underline{\Sigma}}_{w_0}|^{1/2}} \exp\left[-\frac{1}{2}(\underline{w}-\underline{m}_0)^T \underline{\underline{\Sigma}}_{w_0}^{-1} (\underline{w}-\underline{m}_0)\right]$$

BAYES-RULE: DETERMINING THE POSTERIOR

$$\boxed{p(\underline{w}|D)} = \frac{\overbrace{p(D|\underline{w})}^{\text{LIKELIHOOD}}\, p(\underline{w})}{\underbrace{P(D)}_{\text{FEASIBLE CHOICES}}} = \frac{p(D|\underline{w})\, p(\underline{w})}{\int p(D|\underline{w})\, p(\underline{w})\, d\underline{w}} \boxed{\propto p(D|\underline{w})\, p(\underline{w})}$$

(CONJUGATED PRIOR!)
(NONINFORMATIVE PRIOR!)

$$\boxed{\underline{w}^*_{MAP} = \arg\max_{\underline{w}} p(\underline{w}|D)}$$

MAXIMUM A POSTERIORI

$$\beta = \frac{1}{\sigma^2}$$

GENERAL SOLUTION:
(INDEPENDENT DATA)

$$p(t|\underline{x}, \underline{w}, \beta) = N(t|y(\underline{x}, \underline{w}), \beta^{-1})$$

$G \cdot G = G$
$\downarrow$
posterior

$$\begin{cases} p(\underline{t}|\underline{X}, \underline{w}, \beta) = \prod_{n=1}^{N} N(t_n|\underline{w}^T \underline{\phi}(x_n), \beta^{-1}) \\ p(\underline{w}) = N(\underline{w}|\underline{m}_0, \underline{\underline{\Sigma}}_{w_0}) \end{cases}$$

$$p(\underline{w}|\underline{t}) = N(\underline{w}|\underline{m}_N, \underline{\underline{\Sigma}}_N) \qquad \text{(CHECK IT)}$$