RELEVANCE VECTOR MACHINES

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

$$y(\underline{x}) = \sum_{i=1}^{M} w_i \phi_i(\underline{x}) = \underline{w}^T \phi(\underline{x})$$

LIKELIHOOD

$$p(\underline{t} \mid \underline{X}, \underline{w}, \beta) = \prod_{n=1}^{N} p(t_n \mid \underline{x}_n, \underline{w}, \beta^{-1})$$

$$y(\underline{x}) = \sum_{n=1}^{N} w_n k(\underline{x}, \underline{x}_n) + b$$

$$p(\underline{w} \mid \underline{\alpha}) = \prod_{i=1}^{M} \mathcal{N}(w_i \mid 0, \alpha_i^{-1})$$

SEPARATE PRIORS !

POSTERIOR

$$p(\underline{w} \mid \underline{t}, \underline{X}, \alpha, \beta) = \mathcal{N}(\underline{w} \mid \underline{m}, \underline{\underline{\Sigma}})$$

$\underline{m} = \ldots$

$\underline{\underline{\Sigma}} = \ldots$

OPTIMAL ESTIMATE OF $\underline{\alpha}, \beta$ FROM DATA

EVIDENCE APPROXIMATION

$$p(\underline{t} \mid \underline{X}, \underline{\alpha}, \beta) = \int \underbrace{p(\underline{t} \mid \underline{X}, \underline{w}, \beta)}_{\text{LIKELIHOOD}} \underbrace{p(\underline{w} \mid \underline{\alpha})}_{\text{PRIOR}} d\underline{w}$$

$$\frac{\partial \ln \square}{\partial \underline{\alpha}} \qquad \partial \beta$$

$\underline{\alpha}$: $\alpha_i \uparrow \infty$ $\quad p(w_i \mid \ldots) \sim \mathcal{N}(0, 0)$

$\alpha_i < \infty$ RELEVANCE VECTOR