$$\hat{\theta}_{MAP} = \arg\max_{\theta} p(\theta|D_N) = \arg\min_{\theta}\left[-\ln p(\theta|D_N)\right]$$

$$= \arg\min_{\theta}\Big[\underbrace{-\ln p(D_N|\theta)}_{V_{ML}} \underbrace{-\ln p(\theta)}_{\frac{1}{\sigma_\theta^2}\|\theta\|^2}\Big]$$

FOR GAUSSIAN PRIOR

TYPICALLY:

$\cancel{p(\theta)\ as}$ $\theta \sim N(\emptyset, \sigma_\theta^2)$

$\sim N(\theta_0, \sigma_\theta^2)$

ETC.

$\boxed{\text{LIKE REGULARIZED ML, BUT WITH WELL CHOOSEN REGULARIZATION CONSTANT}}$

AS $N \cdot N \sim N$, FOR GAUSSIAN LIKELIHOOD AND GAUSSIAN PRIOR AND I.I.D. NOISES & PARAMETER COMPONENTS

$$\hat{\theta}_{MAP} = \left(\underline{\underline{\Phi}}^T \underline{\underline{\Phi}} + \frac{1}{\sigma_\theta^2}\underline{\underline{I}}\right)^{-1} \underline{\underline{\Phi}}^T \underline{y}$$

REGULARIZATION

→ INTRODUCES BIAS

→ LESS VARIANCE

BETTER MSE

- TO BE EXPECTED:

$$\left[MSE(\hat{\underline{\theta}}_{MAP}) < MSE(\hat{\underline{\theta}}_{ML})\right]$$

- IF PRIOR = UNIFORM (CONSTANT)

$$\left[\hat{\underline{\theta}}_{MAP} \to \hat{\underline{\theta}}_{ML}\right]$$

GENERAL CASE (DEPENDENT NOISES & PARAMETERS)

IF $p(\underline{x}) = N(\underline{x}|\underline{\mu}, \underline{\underline{\Lambda}}^{-1}) = N(\underline{x}|\underline{\mu}, \underline{\underline{\Sigma}}_x)$ COVARIANCE ↔ PRECISION

AND

$\cancel{p(\underline{y}|\underline{x}) = N(\underline{y}|\underline{\underline{A}}\underline{\mu}+\underline{b}, \underline{\underline{L}}^{-1}+\underline{\underline{A}}\,\underline{\underline{\Lambda}}^{-1}\underline{\underline{A}}^T)}$

~~THEN~~ $p(\underline{y}|\underline{x}) = N(\underline{y}|\underline{\underline{A}}\,\underline{x}+\underline{b}, \underline{\underline{L}}^{-1})$

THEN

$$p(\underline{y}) = N(\underline{y}|\underline{\underline{A}}\underline{\mu}+\underline{b}, \underline{\underline{L}}^{-1}+\underline{\underline{A}}\,\underline{\underline{\Lambda}}^{-1}\underline{\underline{A}}^T)$$