BAYESIAN ESTIMATION

SUPERVISED LEARNING.

www.ictp.it

WARNING: GOAL IS NOT ESTIMATE.

$\hat{\theta}$: RANDOM VARIABLE

WHAT WE KNOW / OBSERVE

- TRAINING DATA = $\{\vec{x}_n, t_n\}_{n=1}^{N}$ & TEST DATA = $\{\vec{x}'_m\}_{m=1}^{M}$

WHAT WE DON'T KNOW

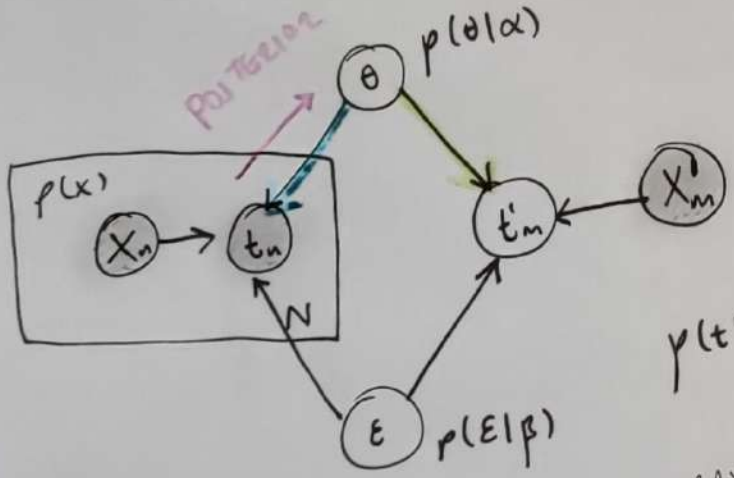
- DISTRIBUTION WITH UNKNOWN PARAMETER θ.
- TARGET FOR TEST DATA $\{t'_m\}_{m=1}^{M}$

ASSUMPTIONS

- GAUSSIAN NOISE WITH $\frac{1}{\beta}$ NOISE VARIANCE
- PRIOR → GAUSSIAN WITH $\frac{1}{\alpha}$ NOISE VARIANCE.

OBJECTIVE: $p(t'|X', t, X)$ → dependence on the data.

NOT TO BE CONFUSED WITH ESTIMATE!



FROM UNSUPERVISED:

$$p(t'|X', t, X) = \int d\theta \cdot p(t'|X', \theta)\, p(\theta|X_n, t_n)$$

BAYES: $p(\theta|X_n, t_n) = \dfrac{p(\theta)\, p(t_n|\theta, X_n)}{p(t_n)}$

BAYES ON target.

$$p(t'|X', t, X) = \int d\theta \cdot p(t'|X', \theta) \cdot \frac{p(\theta)\, p(t_n|\theta, X_n)}{p(t_n)}$$

$$p(t'|X', t, X) \propto \int d\theta \cdot p(t'|X', \theta) \prod_n p(\theta)\, p(t_n|\theta, X_n)$$

$p(t'|X', \theta)$ ← ESTIMATE. $p(\theta)$ ← prior. $p(t_n|\theta, X_n)$ ← likelihood

Likelihood
$p(t_n|X_n, \theta)$

posterior
$p(\theta|X_n, t_n)$

ESTIMATE
$p(t'|X', \theta)$