Bayesian Linear Regression (BLR)

$\longrightarrow$ Gaussian Process

$$f(x) = w^T \phi(x)$$

- $\phi(x)$: fixed feature map, or can be understood as a set of basis functions
- $w \sim N(0, \Sigma_p)$: prior on weights

Posterior over function values

Given training inputs X, targets y and Gaussian noise, we get a posterior predictive distribution at a test input x_*:

$$p(f_* \mid x_*, X, y) = N(\mu_*, \sigma_*^2)$$

where • $\mu_* = \phi(x_*)^T \mathbb{E}[w]$ • $\sigma_*^2 = \phi(x_*)^T \mathbb{V}[w] \phi(x_*)$

Kernel Trick

$$k(x, x') = \phi(x) \Sigma_p \phi(x')$$

$\Rightarrow$ Infinite features $\rightarrow$ Gaussian Process, how?

Define a distribution over function values directly, using their covariance $f(x) \sim \mathcal{GP}(0, k(x, x'))$

A GP is

A collection of random variables, any finite number of which have a joint Gaussian Distribution