

1- Gaussian process regression (GPR)

función: $f(x) \sim GP(m(x), K(x, \tilde{x}^*))$, donde

$m(x) = 0$ y $K(x, \tilde{x}^*)$ covarianza.

Optimización: $\log p(y | X, \theta) = -\frac{1}{2} y^T K^{-1} y \dots$

$$\dots - \frac{1}{2} \log |K| - \frac{n}{2} \log 2\pi$$

donde K = matriz de covarianza

y = observaciones.

2- Gaussian process classification (GPC)

función: $p(y=1|x) = \sigma(f(x))$ donde

$\sigma(\cdot)$ es una función sigmoide.

Optimización: se utiliza aproximación mediante inferencia variacional o técnicas, como laplace approximation o expectation propagation (EP) para aproximar el posterior.

3- Variational gaussian process (VGP)

función: Distribución Variacional $q(f)$ se aproxima a la posterior $p(f|X, y)$

Optimización: Se optimiza el lower bound (ELBO)

$$\text{ELBO} = E_{q(f)} (\log p(y|f)) - \text{KL}(q(f) || p(f)),$$

donde $\text{KL}(\cdot)$ es la divergencia KL entre la distribución variacional $q(f)$ y la verdadera $p(f)$.

4- Sparse Gaussian Process Regression (SGPR)

función: Igual que en GPR, pero se utilizan un conjunto de m puntos de apoyo Z , donde $m \ll n$.

Optimización: Se maximiza un bound variacional o la log-verosimilitud marginal, pero usando los inducing points para reducir la complejidad a $O(nm^2)$ en lugar de $O(n^3)$.

5 - Stochastic Variational gaussian process (SVGP)

función: Similar a VGP, pero utiliza inducing point para aproximar la posterior.

Optimización: Optimiza un ELBO similar al de VGP, pero aplicando técnicas de gradiente estocásticas sobre mini batches de datos.