< What is a Gaussian Process >

: A Gaussian process is a non parametric model used for regression. It doesn't
(e.g. a,b in $y = ax + b$)
a fixed form. Instead it defines a distribution over functions.

"I don't know the exact function that maps input to output, but I'll assume that any finite collection is jointly Gaussian distributed.

* Key intuition

Suppose your trying to model an unknown function $y = f(x)$
You only have a few observed input–output pairs.

$x_1, x_2 .. x_n \quad y_1, y_2 \cdots y_n$

A Gaussian process gives a distribution over all possible functions that
Fit the known data points. & Follow a certain smoothness on structure controlled by a kernal func.

* Core Components.

1. Mean functions.

$m(x) = \mathbb{E}(f(x)) = 0$ usually assumed zero.

2. Covariance Function (Kernel)

$K(x, x') = Cov(f(x), f(x'))$

If x, x' are close $K(x, x')$ is large → outputs are highly correlated (smoothness)

* How GP regression Works.

1. Given training data X_{train} Y_{train}.
2. Want to predict at test points X_{test}
3. $\begin{bmatrix} y_{train} \\ y_{test} \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X_{train}, X_{train}) & K(X_{train}, X_{test}) \\ K(X_{test}, X_{train}) & K(X_{test}, X_{test}) \end{bmatrix}\right)$

μ : predicted value / σ^2 : uncertainty variance at that point.

* Advantages

Uncertainty Quantification : not only predictions but also their confidence intervals.
Non parametric : No need to choose a specific model shape.
Small dataset : Powerful with few data points.