Statistical Machine Learning

Lecture 02
Non-Parametric Bayesian Fundamentals

Spring 2021

Sharif University of Technology

$D: \{x_1, x_2, \ldots x_n\}$ ← Data

Model: $P(X|\theta)$ ← likelihood function

$\underline{\theta}$: parameter Random / unknown

$P(\theta)$ ← prior on θ

Bayes Model: $P(\theta|X) = \dfrac{P(X|\theta)P(\theta)}{P(x)}$

posterior

θ finite size ← parametric

θ Not finite ← non parametric Bayesian (NPB) non-par Baysian = NPB

D ← data set observation

↳ pdf. $P(D|\theta)$ → prior knowledge

likelihood → parameter

$P(\theta)$ → prior

posterior

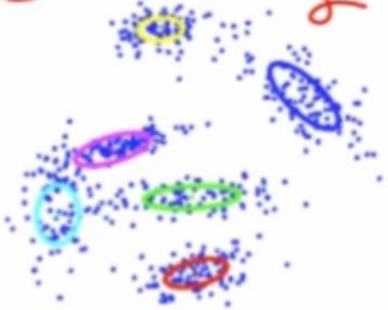
model

$$P(\theta|D) = \frac{P(D|\theta)\,P(\theta)}{P(D)}$$

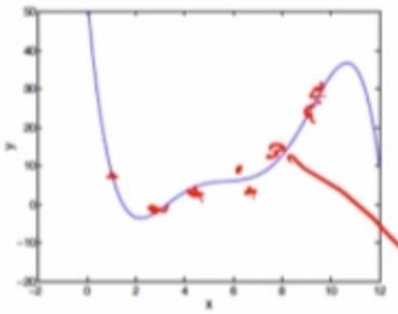
Bayes

Clustering

D



Regression



Polynomial with order m

m ← # of Gaussians in Mixture model

$$\alpha G_1 + \beta G_2 + \gamma G_3 = G \quad \text{GMM}$$

$$p(m|D) = \frac{p(D|m)\,p(m)}{p(D)} \qquad p(D|m) = \int p(D|\theta, m)\, p(\theta|m)\, d\theta$$

① $p(m)$ ← select

② given D use Bayes Rule $p(m|D)$ ← infer

NPB models

- Bayesian models are powerful when your prior captures your belief.

- fix parameters for complex Dataset (multimodel) → is inflexible ⇒ unreasonable inference

→ Non parametric models are better

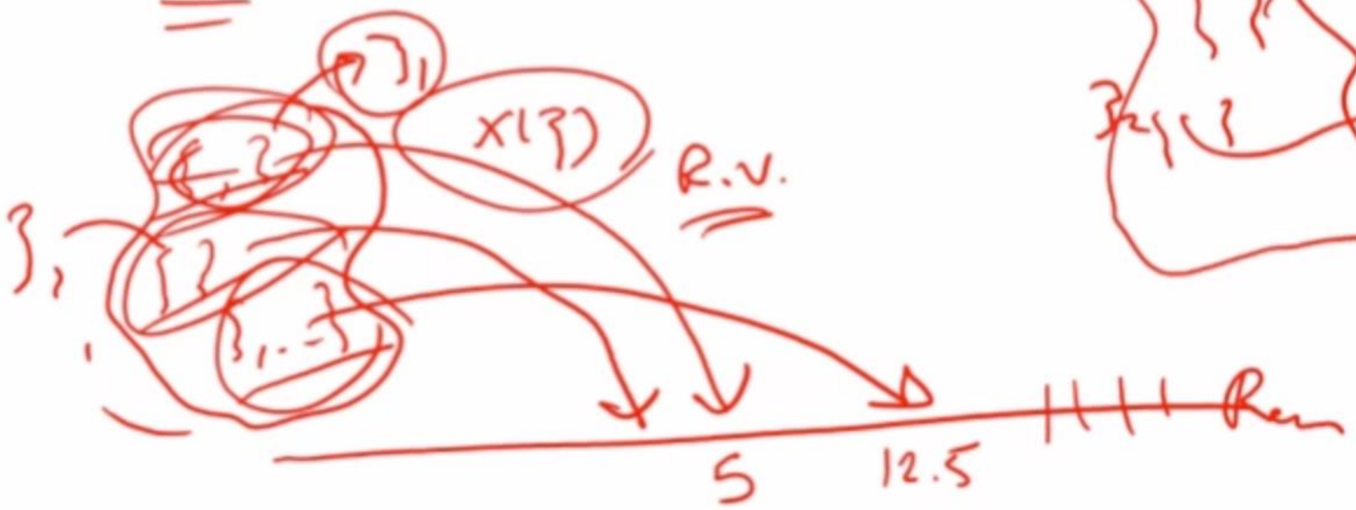
⇒ NPB is more flexi

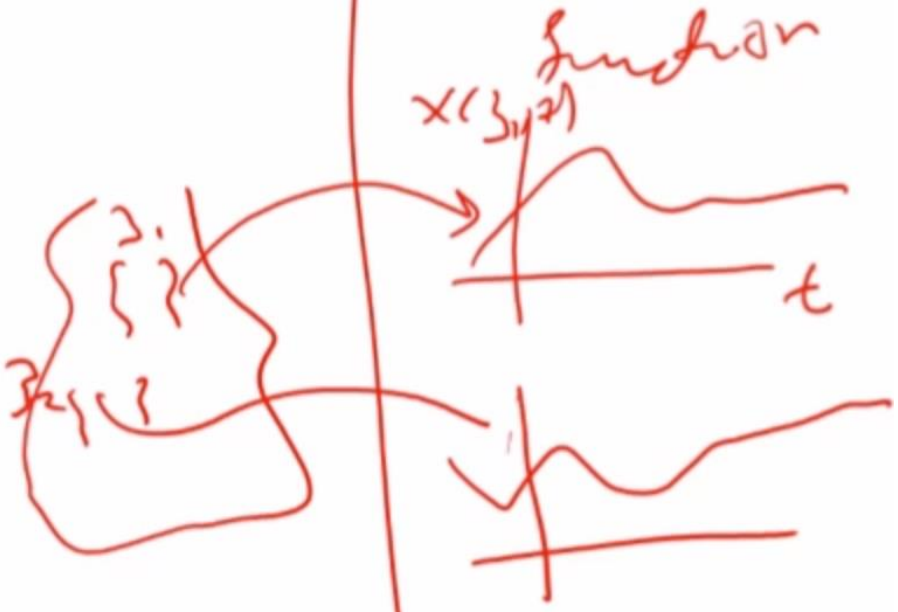
* Gaussian process (GP)

ensem

Recall

D: $X_1, X_2, \dots\dots X_n$

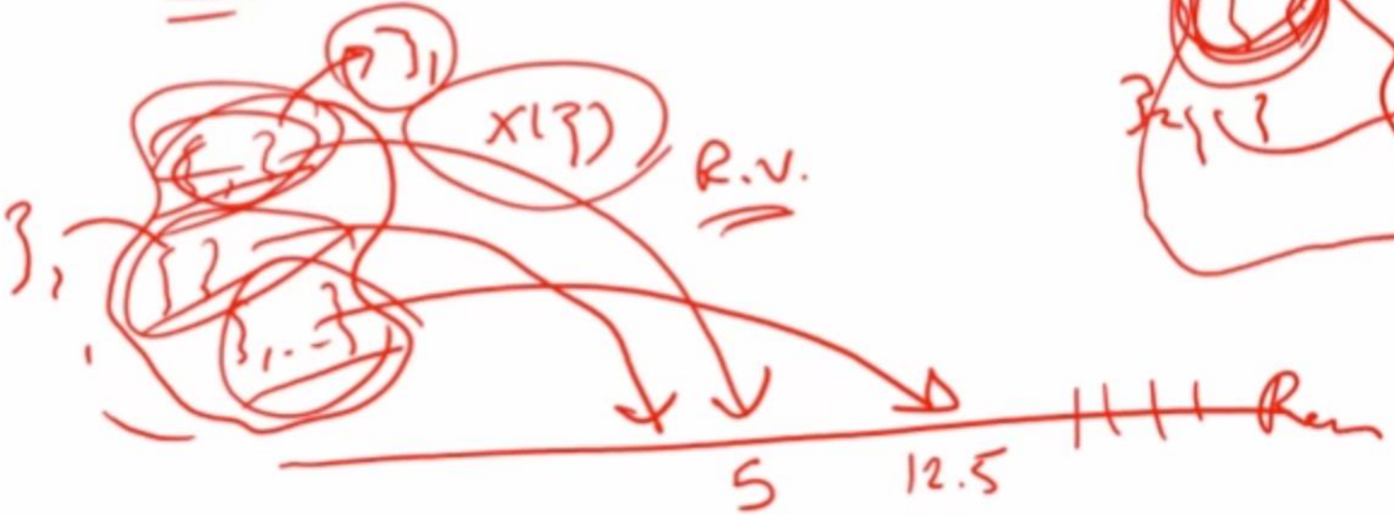


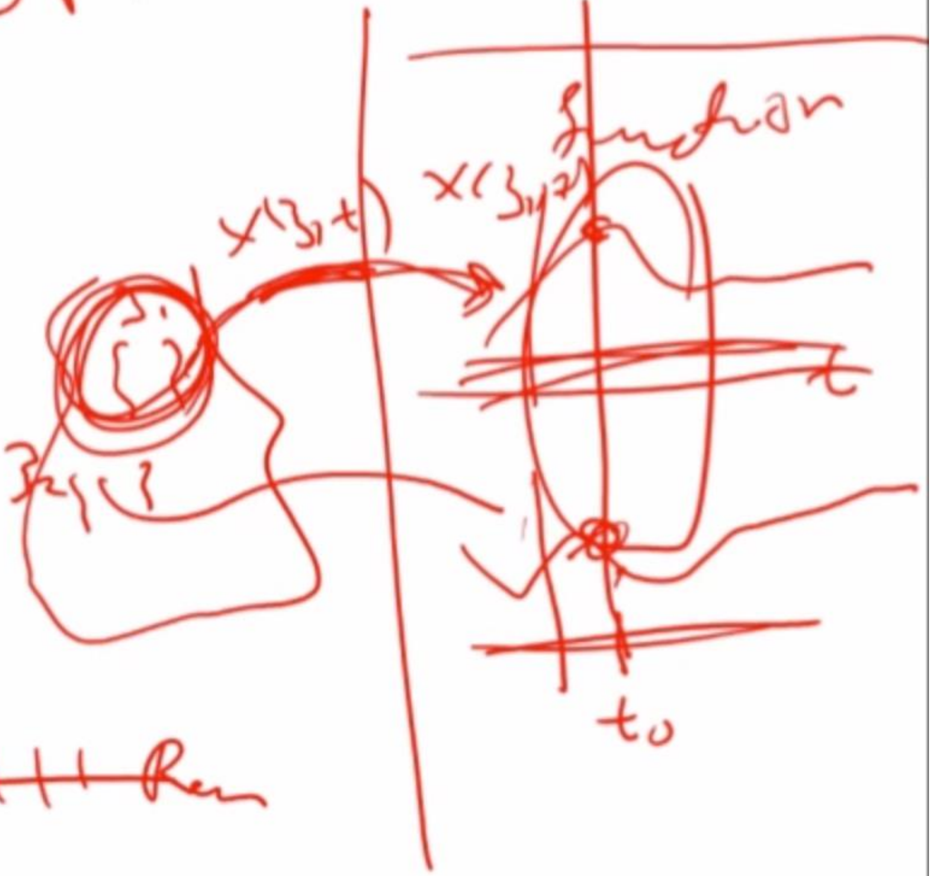


* Gaussian process (GP)

Recall

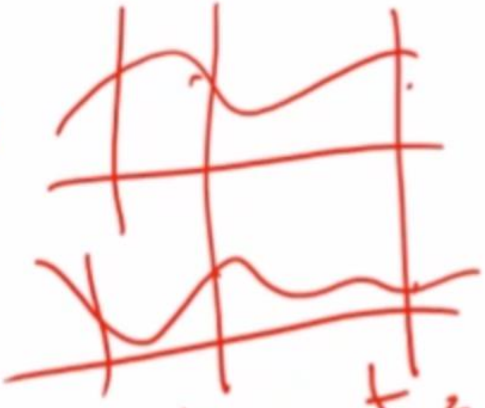
D: $X_1, X_2, \ldots\ldots X_n$





Stationary process

Strict ✓ hard to achive



WSS → $X(t)$ ← R.P.

① $E[X(t)] \triangleq$ not a fn of t
$= \mu$

② $R(t_1, t_2) = R(\underbrace{t_1 - t_2}_{\tau}) = R(\tau)$