Central ML Problem

23 September 2024 18:21

Data

N rows of data of dimension D.

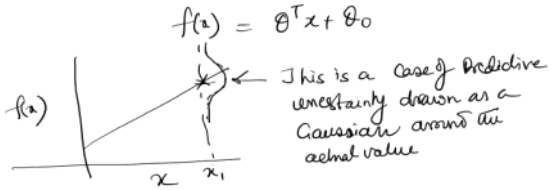
$X \in \mathbb{R}^{N \times D}$

Model

We're trying to find a f^n approximation of a predictor f^n

$f: \mathbb{R}^D \to \mathbb{R}$ (regression)

Predictor f^n = "probability distrib" that models uncertainity"



Learning.

Find a model $f(x)$ + its corresponding parameters.

① Prediction / Inference Phase

② Training / Param. Estimation Phase

③ Hyperparameter tuning / model selection phase

Phase 1 + 2 : Predictor f^n + parameters.

FUNCTION

- principle of empirical risk minimization
- statistical model (MLE)
- Hill climbing AND gradient descent.

PROBABILISTIC MODEL

- Bayesian inference.

In training, simulate behavior of predictor on unseen data

- Cross validation (non probabilistic)
- Regularization ("simple explan^n fans)

Phase 3 : Model Selection

Structure of predictor
- hyperparam optimiz^n
- cross val.

Class of prob. dist.
- probabilistic ML.

Statistical vs ML Model.

Samples drawn from popl^n	
Model fitted on sample	Models can work w/ less data (low accuracy tho.)
Parametric in nature	Non parametric.
MULTICOLLINEARITY >>>	If MC, weight adjusts automatically.

Namita Achyuthan