

Part 2: Regression

$$Y \in \mathbb{R}, \quad X \in \mathbb{R}^D$$

"true" model (unknown "data generating process"):

- deterministic: $Y = f^*(X)$ f^* : true generative process
- non-deterministic: $Y = f^*(X, \eta)$ η : random numbers

for fixed X , $f^*(X, \eta)$ will not always produce the same result, because η will be different in every run

\Rightarrow determine the posterior prob of all possible outcomes Y :

$$p(Y|X) \Leftrightarrow Y = \underset{\substack{\uparrow \\ \text{deterministic fct. with random argum}}}{f^*(X, \eta)} \text{ with } X \perp \eta$$

"reparameterization trick" of $p(Y|X)$

- simplest case: additive noise: $Y = f^*(X, \eta) = f(X) + \eta$

\Rightarrow task: find a model $\hat{f}(X) \approx f(X)$

- general case: outcome depends on the noise in a more complicated way

example: post-additive noise: $Y = g(f(X) + \eta)$

- nature of the function $f(X)$:

- simplest case: linear

$$f(X) = X \cdot \underset{\substack{\leftarrow \text{column vector of weights} \\ \leftarrow \text{row vector}}}{w}$$

"linear regression"

- otherwise: "non-linear regression"

- noise η can always be uniform, because $\eta' = Q(\eta)$ can have arbitrary distr. with $Q(\eta)$ the "quantile function" as sorted into $f^*(X, \eta)$

Zoo of regression methods

model class $f(x)?$

linear

non-linear

noise?

noise?

additive gaussian

other

additive gaussian

other

only y noisy?

no (x also noisy)

yes

is the linear system underdetermined? $D > N?$

$D > N?$

yes ($D > N$)

no ($D < N$)

is the noise variance equal for all y_i ?

no

yes

ordinary least squares

is the variance σ_i^2 known for all i ?

no

yes

weighted least squares

heteroscedastic linear regression

total least squares

regularized / constrained least squares

non-linear least squares

regression tree / forest

linear regression in an "augmented" feature space
 $\tilde{x} = \phi(x)$

kernel methods, Gaussian processes

non-linear $f(x)$ (hand-crafted, neural network)
 \Rightarrow can use linear regression on \tilde{x}

Poisson noise

transform to Gaussian noise
 $\tilde{y} = 2\sqrt{y + 3/8}$ "Anscombe transform"
 \Rightarrow ordinary least squares

known parametric noise distribution

maximum-likelihood estimator for that noise

outliers

robust regression

outlier removal