

Incomplete data problem

① \Rightarrow Here ML estimation is made difficult by the absence of some part of data in more familiar and simple data structure

② Solⁿ = Expectation Maximization
1 iteration of EM algorithm

① Expectation Step (E-Step)

② Maximization Step (M-Step)

③ closely related to ad hoc approach to estimation with missing data

\rightarrow where the parameters are estimated by filling in initial values for the missing data.

\rightarrow the latter are then updated by their predicted values using these initial parameters

\rightarrow the parameters are then re-estimated and so on, proceeding iteratively until convergence.

④ Main idea of EM Algorithm = Associate with given incomplete-data, a complete data problem for which ML estimation is computationally more tractable
(reformulating the problem in terms of easily solvable data problem)

Algorithm \rightarrow

EM Algorithm)

① E-step = manufacturing data for the complete data problem

using observed data set of incomplete data problem & current value of parameter, so that

② the simpler M-step computation can be applied to "completed" data set

③ \Rightarrow Log likelihood of the complete-data problem that is "manufactured" in E-step

④ as it is based on unobservable data, it is replaced by its conditional expectation given the observed data, E-step is affected using the current fit for the unknown parameter.

$$\left[\frac{\partial \log L(\Psi)}{\partial \Psi} = 0 \right]$$

★ EM Algorithm \Rightarrow problem of solving the incomplete data likelihood equation

indirectly by proceeding in terms of complete data log likelihood function $\log L_c(\Psi)$

x