Q. Batch Techniques, such as maximum likelihood solution (3.15) which involve processing the entire training set in one go, can be computationally costly for large data sets.

If the data sets are sufficiently large we can use sequential algorithms called online algorithms in which data points are considered one at a time and the model parameters updated after each such presentation.

Sequential Learning is also appropriate for realtime applications in which the data observations are arriving in a continuous stream, and predictions must be made before all the data points are seen.

We can obtain sequential learning algorithm by applying the technique of stochastic gradient descent, as follows.

If the error function comprises the sum over data points $E = \sum_n E_n$ then after presentation of pattern n, the stochastic gradient descent algorithm updates parameter vector w using

$$w^{\tau+1} = w^{\tau} - \eta \nabla E_n$$

Scanned with CamScanner