

Lecture Three

March 8, 2016

1 EM-Algorithm

Bias-Variance Dilemma

1.1 Gaussian

Since gaussians are simple we like to use them.

- Supervise Learning: training with labels
- Unsupervised Learning: training without labels

The idea of unsupervised learning involves the idea of clustering. The main assumption being that the data models fit into local clusters that can be registered.

Each class is described by a gaussian: $\vec{\mu}, \Sigma$.

EM:

1. Pick $\mu_1, \mu_2, \dots, \mu_3$ randomly. The covariance matrix is round at initialization. This implies: $\Sigma_1 = \Sigma_2 = \dots = \Sigma_n = I$.
2. **Expectation Step:** Assign data points to the classes according to the randomly defined Gaussians.
3. **Maximization Step:** Compute the center of the Assignment. Then you move center of each gaussian μ to the center of the assignments. Update $\mu_1, \mu_2, \dots, \mu_n$ as the centroids of the preliminary class assignment. Update $\Sigma_1, \Sigma_2, \dots, \Sigma_n$

2 Homework

Perform the EM algorithm on the entire training set of digits to compute the the different classes not knowing the original labels.

Perform the calculations on both the training set and test set.