NEML CW

Ben James, Johns Noble

March 2025

Contents

1 Introduction

It is often very natural to attempt to model data using gaussian distribution. Gaussian distributions are well defined for data points which lie in euclidean space for any number of dimensions. We however know that most data should be modelled as to be taken from a manifold. There are methods to do manifold learning but once we have a manifold the notion of probability distributions need to be altered such that we are defining a probability measure across a manifold. It does not make sense for a region outside the manifold to possess non-zero measure. The Locally Adaptive Normal Distribution (LAND) aims to attempt to do this. Once we have a notion of normal distribution on a manifold given some mean and covariance, We can try and fit distribution to the data using MLE approaches. Finally, we will be able to use the fact that we can fit gaussians to data to be able to perform EM algorithms over data to fit a mixture of LANDS over data. TODO: Write something about how manifold learning is handled i.e. Do we know the manifold which data will be distributed on or is this something that needs to be learnt using graphs?

2 Constructing Metrics

3 Parametric Inference

4 Extending to Mixture Models