Abstract

In this thesis we describe improved methods of estimating mean and scatter from multivariate data. As we know, the sample mean and the sample variance matrix are non-robust estimators, which means that even a small amount of measurement errors can seriously affect the resulting estimate. We can deal with that problem using MCD estimator (minimum covariance determinant), that finds a sample variance matrix only from a selection of data, specifically those with the smallest determinant of this matrix. This estimator can be also very helpful in outlier detection, which is used in many applications. Moreover, we will introduce the MVE estimator (minimum volume ellipsoid). We will discuss some of the properties and compare these two estimators.