How the SVM work?

Kernel function

Taylor expansion

Why PCA work? Why SVD give the correct u vectors?

Multivariate variate gaussian distribution in anomaly detection algorithm.

Why there is a sigma matrix?