EECS 189 Introduction to Machine Learning
Fall 2020

DIS9

This discussion was released **Friday, October 30**.

1 k-means demo

Work through the Jupyter notebook in DataHub.

2 MLE of Multivariate Gaussian

In lecture, we discussed uses of the multivariate Gaussian distribution. We just assumed that we knew the parameters of the distribution (the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$). In practice, though, we will often want to estimate μ and Σ from data. (This will come up even beyond regression-type problems: for example, when we want to use Gaussian models for classification problems.) This problem asks you to derive the Maximum Likelihood Estimate for the mean and variance of a multivariate Gaussian distribution, under some conditions.

In homework, you will look at a more general version of this same problem.

(a) Let $\mathbf{X}$ have a multivariate Gaussian distribution with mean $\mu \in \mathbb{R}^d$ and covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$. **Write the log likelihood of drawing the n i.i.d. samples $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ from X given Σ and μ.**

Recall that the probability density function of the aforementioned multivariate Gaussian distribution is

$$P(\mathbf{x}|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\{-\frac{1}{2}(\mathbf{x}-\mu)^\top \Sigma^{-1}(\mathbf{x}-\mu)\}.$$

(b) **Prove that the maximum likelihood estimates of μ and Σ are the sample mean and covariance.** For simplicity, compute the MLE estimate only over diagonal PSD matrices as candidates for Σ, and show that each diagonal element of Σ is the sample covariance over the values of the samples in that dimension.

Contributors:

- Anant Sahai
- Josh Sanz
- Katia Patkin
- Philipp Moritz
- Rahul Arya
- Soroush Nasiriany