Multivariate Normal Maximum Likelihood

1 Why

What of the generalization to a multivariate normal.

2 Result

Proposition 1. *Let* $(x^1, \dots, x^n)$ *be a dataset in* R^d. *Let* f *be a multivariate normal density with mean*

$$\frac{1}{n}\sum_{k=1}^{d} x^k$$

and covariance

$$\frac{1}{n}\sum_{k=1}^{n}\left(x^k - \frac{1}{n}\sum_{k=1}^{n} x^k\right)\left(x^k - \frac{1}{n}\sum_{k=1}^{n} x^k\right)^\top.$$

Then f *is a maximum likelihood multivariate normal density.*

Proof. We express the log likelihood

$$\sum_{k=1}^{n} -\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu) - \frac{1}{2}\log(2\pi)^d - \frac{1}{2}\log \mathbf{det}\,\Sigma$$

Let $P = \Sigma^{-1}$. The $\log \mathbf{det}\,\Sigma$ is $\log \mathbf{det}\,P^{-1}$ is $\log\left(\mathbf{det}\,P\right)^{-1}$ is $-\log \mathbf{det}\,P$. Use matrix calculus to get

$$\frac{\partial \ell}{\partial P} = \sum_{k=1}^{n}(x^k - \mu)(x^k - \mu)^\top - P^{-1}.$$

□

We call these two objects the *maximum likelihood mean* and *maximum likelihood covariance* of the dataset.