CS 3220: Prelim 2 example questions

Instructor: Anil Damle

Whats in this document

These questions are intended to be somewhat representative of the type of questions that could arise on the second prelim. They are focused on topics that have not been covered by HW problems.

Question 1

Say you are given independent and identically distributed samples $X_1, X_2, \ldots, X_n$ from a uniform distribution on $[0, \theta]$ and θ is unknown. What is the MLE for θ?

Solution

Here we have that the likelihood function can be written as

$$L_n(\theta) = \frac{1}{\theta^n}$$

where $\theta \geq \max_i X_i$. The constraint stems from the fact that for all i we must have $\theta \geq X_i$ since we observed X_i. The pdf of a uniform is constant, hence L_n does not depend on the X_i directly. The MLE $\widehat{\theta}$ is therefore the value of θ that maximizes $(1/\theta)^n$ over $\theta \geq \max_i X_i$. Since $(1/\theta)^n$ monotonically decays as $\theta \to \infty$ we conclude that

$$\widehat{\theta} = \max_{i=1,2,\ldots,n} X_i.$$

Question 2

Consider a sequence of independent and identically distributed random vectors $X_1, X_2, \ldots, X_n$ where $X_i \in \mathbb{R}^d$ and $n \geq d$. Let $\bar{\mu}_\ell$ be the sample mean of $(X_i)_\ell$ and $\bar{\sigma}_\ell^2$ be the sample variance of $(X_i)_\ell$. Show that if any two entries of the vectors X_i are perfectly correlated, *i.e.*

$$\frac{\sum_{i=1}^{n}[(X_i)_\ell - \bar{\mu}_\ell][(X_i)_k - \bar{\mu}_k]}{\bar{\sigma}_k \bar{\sigma}_\ell} = 1$$

for some $k \neq \ell$, then $\widehat{\Sigma} = \widehat{X}\widehat{X}^T$ has rank $< d$.

Solution

First, observe that equation given is equivalent to saying (recall that $\widehat{X}(k,:)$ is a row vector)

$$\frac{\widehat{X}(\ell,:)(\widehat{X}(k,:))^T}{\|\widehat{X}(\ell,:)\|_2 \|\widehat{X}(\ell,:)\|_2} = 1. \quad (1)$$

Note that for any two vectors u and v, if $u^T v = \|u\|_2 \|v\|_2$ then

$$\begin{aligned}
\left\| \left(I - \frac{vv^T}{\|v\|_2^2} \right) u \right\|_2 &= u^T \left(I - \frac{vv^T}{\|v\|_2^2} \right) u \\
&= \|u\|_2^2 - (u^T v)^2 / \|v\|_2^2 \\
&= \|u\|_2^2 \|v\|_2^2 - (u^T v)^2 \\
&= 0.
\end{aligned}$$

In other words, we may conclude that if rows ℓ and k of $\widehat{X}$ satisfy (1) then they are colinear. Therefore, there exists a non-zero vector $w \in \mathbb{R}^d$ such that $w^T \widehat{X} = 0$. From this is follows that $\widehat{\Sigma} w = 0$, it follows that the null space of $\widehat{\Sigma}$ is non-trivial and therefore $\widehat{\Sigma}$ has rank $< d$.