STATS 240 HW1

Statistical Methods

October 25, 2021

1 Problem 1

Let $\mathbf{X} = (X_1, X_2 ... X_n)$ be n observations for which the joint density function $f_\Theta(\mathbf{x})$ depends on unknown parameter Θ. We have to show that:

$$E(\nabla log f_\Theta(\mathbf{x})) = 0$$

The gradient is with respect to the parameter vector Θ, i.e. component wise derivatives. We will use Θ to denote vector gradient and avoid writing component wise derivatives in every step.

First, look at the term inside the expectation. Simplify it as follows:

$$\nabla log f_\Theta(\mathbf{x})$$

$$= \frac{1}{f_\Theta(\mathbf{x})} f'_\Theta(\mathbf{x})$$

where we have used the simple chain rule. Now, we have to take the expected value of this. Since the expectation is with respect to the random variable $\mathbf{x}$, use the joint density function $f_\Theta(\mathbf{x})$. Therefore,

$$E(\nabla log f_\Theta(\mathbf{x}))$$

$$= E(\frac{1}{f_\Theta(\mathbf{x})} f'_\Theta(\mathbf{x}))$$

$$= \int \frac{1}{f_\Theta(\mathbf{x})} f'_\Theta(\mathbf{x}) f_\Theta(\mathbf{x}) \, d\mathbf{x}$$

$$= \int f'_\Theta(\mathbf{x}) \, d\mathbf{x}$$

Since f is given to be smooth, we can apply the Leibniz integral rule on the above result to get

$$\int f'_\Theta(\mathbf{x}) \, d\mathbf{x}$$

$$= \frac{d}{d\Theta} \int f_\Theta(\mathbf{x}) \, d\mathbf{x}$$

Now, note that $f_\Theta(\mathbf{x})$ is a joint density function, and hence

$$= \int f_\Theta(\mathbf{x}) \, d\mathbf{x} = 1$$

which gives that

$$= \frac{d}{d\Theta} \int f_\Theta(\mathbf{x}) \, d\mathbf{x} = 0$$

This proves the desired result

$$E(\nabla log f_\Theta(\mathbf{x})) = 0$$

For the second part, we have to show that

$$E(-\nabla^2 log f_\Theta(\mathbf{x})) = cov(\nabla log f_\Theta(\mathbf{x}))$$

We use the result from above and apply the chain rule again

$$\nabla^2 log f_\Theta(\mathbf{x})$$

$$= \nabla \frac{1}{f_\Theta(\mathbf{x})} f'_\Theta(\mathbf{x})$$

$$= \frac{1}{f_\Theta(\mathbf{x})} f''_\Theta(\mathbf{x}) - (\frac{1}{f_\Theta(\mathbf{x})} f'_\Theta(\mathbf{x}))^2$$

Rearrange the terms from the above equation and simplify to get that

$$f''_\Theta(\mathbf{x}) = \nabla^2 log f_\Theta(\mathbf{x}).f_\Theta(\mathbf{x}) + \frac{1}{f_\Theta(\mathbf{x})}(f'_\Theta(\mathbf{x}))^2$$

From above, we have that $\nabla log f_\Theta(\mathbf{x}) = \frac{1}{f_\Theta(\mathbf{x})} f'_\Theta(\mathbf{x})$ and hence

$$f''_\Theta(\mathbf{x}) = \nabla^2 log f_\Theta(\mathbf{x}).f_\Theta(\mathbf{x}) + (\nabla log f_\Theta(\mathbf{x}))^2.f_\Theta(\mathbf{x})$$

Now we integrate both sides:

$$\int f''_\Theta(\mathbf{x})\, d\mathbf{x} = \int \nabla^2 log f_\Theta(\mathbf{x}).f_\Theta(\mathbf{x})\, d\mathbf{x} + \int (\nabla log f_\Theta(\mathbf{x}))^2.f_\Theta(\mathbf{x})\, d\mathbf{x}$$

The LHS value is zero by the same procedure as above (since f is smooth, we can differentiate under the integral again using the Leibniz rule). For the RHS, the first term is exactly $E(-\nabla^2 log f_\Theta(\mathbf{x}))$ and the second term is $cov(\nabla log f_\Theta(\mathbf{x}))$. Therefore,

$$E(\nabla^2 log f_\Theta(\mathbf{x})) + cov(\nabla log f_\Theta(\mathbf{x})) = 0$$

and by linearity of expectation (allowing us to take the negative sign inside)

$$E(-\nabla^2 log f_\Theta(\mathbf{x})) = cov(\nabla log f_\Theta(\mathbf{x}))$$

By
Aajeya Jajoo
Student ID: 06446042