

Class notes on ML.

Consider the standard Normal $N(\mu, \sigma^2)$. The log likelihood is $-\frac{1}{2} \log \sigma^2 - \frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}$. So the score is

$$dl/d(\mu, \sigma^2) = \begin{pmatrix} \frac{(x-\mu)}{\sigma^2} \\ -\frac{1}{2} \frac{1}{\sigma^2} + \frac{1}{2} \frac{(x-\mu)^2}{\sigma^4} \end{pmatrix}.$$

We get the Hessian by differentiating again

$$H = \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{(x-\mu)}{\sigma^4} \\ \cdot & \frac{1}{2} \frac{1}{\sigma^4} - \frac{(x-\mu)^2}{\sigma^6} \end{pmatrix}.$$

Take the expectation (which will be the limit when the number of observations in the likelihood function goes to infinity)

$$H = \begin{pmatrix} -\frac{1}{\sigma^2} & 0 \\ 0 & -\frac{1}{2} \frac{1}{\sigma^4} \end{pmatrix}.$$

We see that the variance of the estimates of μ and σ^2 are σ^2 and $2\sigma^2$, resp.

What if we take the outer product of the scores, i.e., $dl dl'$? We have

$$\begin{pmatrix} \frac{(x-\mu)}{\sigma^2} \\ -\frac{1}{2} \frac{1}{\sigma^2} + \frac{1}{2} \frac{(x-\mu)^2}{\sigma^4} \end{pmatrix} \begin{pmatrix} \frac{(x-\mu)}{\sigma^2}, & -\frac{1}{2} \frac{1}{\sigma^2} + \frac{1}{2} \frac{(x-\mu)^2}{\sigma^4} \end{pmatrix}.$$

which is

$$\begin{pmatrix} \frac{(x-\mu)^2}{\sigma^4} & -\frac{1}{2} \frac{(x-\mu)}{\sigma^4} + \frac{1}{2} \frac{(x-\mu)^3}{\sigma^6} \\ \cdot & \frac{1}{4} \frac{1}{\sigma^4} + \frac{1}{4} \frac{(x-\mu)^4}{\sigma^8} - \frac{1}{2} \frac{(x-\mu)^2}{\sigma^6} \end{pmatrix}.$$

Take the expectation and get

$$E dl dl = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{4} \frac{1}{\sigma^4} + \frac{3}{4} \frac{1}{\sigma^4} - \frac{1}{2} \frac{1}{\sigma^4} \end{pmatrix}.$$

Collecting terms, we see that we get $E dl dl' = -EH$. Pretty cool, no?