CS6190: Probabilistic Modeling Homework 1
Exponential Families, Conjugate Priors

Gopal Menon

29 January, 2018

Written Part

1. **Expectation of sufficient statistics:** Consider a random variable X from a continuous exponential family with natural parameter $\eta = (\eta_1, \ldots, \eta_n)$. Recall this means the pdf is of the form:

$$p(x) = h(x) \exp\left(\eta \cdot T(x) - A(\eta)\right)$$

(a) Show that $E\left[T(x)|\eta\right] = \nabla A(\eta) = \left(\frac{\partial A}{\partial \eta_1}, \ldots, \frac{\partial A}{\partial \eta_d}\right)$.

Hint: Start with the identity $\int p(x)dx = 1$, and take the derivative with respect to η.

$$\begin{aligned}
\int p(x)dx &= 1 \\
\Rightarrow \int h(x) \exp\left(\eta \cdot T(x) - A(\eta)\right) dx &= 1 \\
\nabla \left(\int h(x) \exp\left(\eta \cdot T(x) - A(\eta)\right) dx \right) &= 0 \\
\int h(x) \exp\left(\eta \cdot T(x) - A(\eta)\right) . \left(T(x) - \nabla A(\eta)\right) dx &= 0 \\
\int T(x) h(x) \exp\left(\eta \cdot T(x) - A(\eta)\right) dx &= \nabla A(\eta) \int h(x) \exp\left(\eta \cdot T(x) - A(\eta)\right) dx \\
\int T(x) p(x) dx &= \nabla A(\eta) \int p(x) dx \\
E\left[T(x)\right] &= \nabla A(\eta)
\end{aligned}$$

(b) Verify this formula works for the Gaussian distribution with unknown mean, μ, and known variance, σ^2.

Hint: Start by thinking about what the natural parameter η and the function $A(\eta)$ are, then verify that the expectation of the Gaussian is the same as $\nabla A(\eta)$.

The pdf for a Gaussian distribution with unknown mean, μ, and known variance, σ^2 is given by:

$$
\begin{aligned}
p(x) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \\
&= \frac{1}{\sqrt{2\pi}}\exp\left[\ln\left(\frac{1}{\sigma}\right)\right]\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \\
&= \frac{1}{\sqrt{2\pi}}\exp\left[-\ln(\sigma)-\frac{(x-\mu)^2}{2\sigma^2}\right] \\
&= \frac{1}{\sqrt{2\pi}}\exp\left[-\ln(\sigma)-\frac{x^2}{2\sigma^2}+\frac{2x\mu}{2\sigma^2}-\frac{\mu^2}{2\sigma^2}\right] \\
&= \frac{1}{\sqrt{2\pi}}\exp\left[\frac{x\mu}{\sigma^2}-\frac{x^2}{2\sigma^2}-\left(\ln(\sigma)+\frac{\mu^2}{2\sigma^2}\right)\right] \\
&= \frac{1}{\sqrt{2\pi}}\exp\left[\begin{bmatrix}\frac{\mu}{\sigma^2}\\ \frac{-1}{2\sigma^2}\end{bmatrix}\cdot\begin{bmatrix}x\\ x^2\end{bmatrix}-\left(\ln(\sigma)+\frac{\mu^2}{2\sigma^2}\right)\right]
\end{aligned}
$$

Based on the form of the exponential family pdf, we can see that:

$$
\begin{aligned}
A(\eta) &= \ln(\sigma)+\frac{\mu^2}{2\sigma^2}, \\
\eta &= \begin{bmatrix}\frac{\mu}{\sigma^2}\\ \frac{-1}{2\sigma^2}\end{bmatrix}, \text{ and} \\
T(x) &= \begin{bmatrix}x\\ x^2\end{bmatrix}
\end{aligned}
$$

Based on the equation for $E\left[T(x)\right]$ above:

$$
\begin{aligned}
E\left[T(x)\right] &= \nabla A(\eta) \\
\Rightarrow E\left[\begin{bmatrix}x\\ x^2\end{bmatrix}\right] &= \begin{bmatrix}\frac{\partial A}{\partial \eta_1}\\ \frac{\partial A}{\partial \eta_2}\end{bmatrix}, \text{ where } \eta_1=\frac{\mu}{\sigma^2}, \text{ and } \eta_2=\frac{-1}{2\sigma^2} \\
\begin{bmatrix}E\left[x\right]\\ E\left[x^2\right]\end{bmatrix} &= \begin{bmatrix}\frac{\partial\left(\ln(\sigma)+\frac{\mu^2}{2\sigma^2}\right)}{\partial\eta_1}\\ \frac{\partial\left(\ln(\sigma)+\frac{\mu^2}{2\sigma^2}\right)}{\partial\eta_2}\end{bmatrix} \\
&= \begin{bmatrix}\frac{\partial\left(\ln(\sigma)+\frac{\sigma^2}{2}\frac{\mu^2}{\sigma^4}\right)}{\partial\eta_1}\\ \frac{\partial\left(\ln\left(\sqrt{\frac{-1}{2\eta_2}}\right)-\mu^2\eta_2\right)}{\partial\eta_2}\end{bmatrix} \\
&= \begin{bmatrix}\frac{\partial\left(\ln(\sigma)+\frac{\sigma^2}{2}\eta_1^2\right)}{\partial\eta_1}\\ \frac{\partial\left(\ln\left(\frac{i}{\sqrt{2}\sqrt{\eta_2}}\right)-\mu^2\eta_2\right)}{\partial\eta_2}\end{bmatrix} \\
&= \begin{bmatrix}\frac{\sigma^2}{2}.2\eta_1\\ \frac{\sqrt{2}\sqrt{\eta_2}}{i}.\frac{i}{\sqrt{2}}.\frac{-1}{2}\eta_2^{\frac{-3}{2}}-\mu^2\end{bmatrix} \\
&= \begin{bmatrix}\frac{\sigma^2}{2}.2\frac{\mu}{\sigma^2}\\ \frac{-1}{2\eta_2}-\mu^2\end{bmatrix} \\
&= \begin{bmatrix}\mu\\ \sigma^2-\mu^2\end{bmatrix} \\
\Rightarrow E\left[x\right] &= \mu, \text{ and} \\
E\left[x^2\right] &= \sigma^2-\mu^2, \text{ which is true since } \sigma^2=E\left[x^2\right]-\mu^2
\end{aligned}
$$