Matthew Scielman
995367342

CSC2506 Assignment 1

(1) 1) let $f = w^T x$, $p(w) \sim N(w \mid 0, \Sigma)$

show $p(f \mid \underline{x})$ Gaussian.

pf: Using moment generating functions, $m_f(t) = m_{w^T x}(t) = E(e^{(w^T x)t}) = E(e^{(x^T w)t})$
$= E(e^{w^T(xt)}) = m_w(xt) = \exp\{\frac{1}{2}(xt)^T \Sigma (xt)\} = \exp\{\frac{1}{2} t^2 x^T \Sigma x\}$
$\Rightarrow f \sim N(0, x^T \Sigma x)$ by Correspondence of mgf to probability distribution functions.

2) $E(f) = E(w^T x) = E(x^T w) = x^T E(w) = 0$, $Cov(f) = Cov(x^T w) = x^T cov(w) x = x^T \Sigma x$.

3) Using mgf again and that $\varepsilon \perp f$, $m_{f+\varepsilon}(t) = m_f(t) m_\varepsilon(t) = \exp\{\frac{1}{2} t^2 x^T \Sigma x\} \exp\{\frac{1}{2} t^2 0.001\}$
$= \exp\{\frac{1}{2} t^2 (x^T \Sigma x + 0.001)\} \Rightarrow f \sim N(0, x^T \Sigma x + 0.001)$

(2) To find the Conjugate prior of $X \sim N(\mu, \lambda^{-1})$ we seek to find a prior such that multiplication with the likelihood retains the same functional form.

we take $\ell(\mu, \lambda \mid x) = P(x \mid \mu, \lambda) \propto \lambda^{1/2} \exp\{-\frac{\lambda}{2}(x-\mu)^2\}$,

and we note that our prior $P(\mu, \lambda) = P(\mu \mid \lambda) P(\lambda)$ must maintain the same functional form as $\ell(\mu, \lambda \mid x)$.

we can see that for any α, β, $\ell(\mu, \lambda \mid x)$ has the same functional form as $\lambda^\alpha \exp\{-\frac{\lambda}{\beta}\}$, so this would be a prior for the marginal distribution of λ

$\Rightarrow P(\lambda \mid \alpha, \beta) = \Gamma(\alpha, \beta)$

Also notice that, for a fixed λ, to retain Conjugacy, $P(\mu \mid \lambda) \propto \lambda^{1/2} \exp\{-\frac{\lambda}{2}(x^2 - 2x\mu - \mu^2)\}$, a quadratic exponential $\Rightarrow$ $\mu \mid \lambda \propto N(\gamma, c)$, where $c \propto \lambda^{-1}$, (since λ^{-1} appears everywhere c would, and so, λ^{-1} must be proportional to c to maintain Conjugacy with the likelihood).

finally, we have that our prior $P(\mu, \lambda) = P(\mu \mid \lambda) P(\lambda) = N(\mu \mid \gamma, c) \Gamma(\lambda \mid \alpha, \beta)$
and if we take $c = (J\lambda)^{-1} \Rightarrow P(\mu, \lambda) = N(\mu \mid \gamma, (J\lambda)^{-1}) \Gamma(\lambda \mid \alpha, \beta)$, as needed ■