Ex 2.5

Conjugate Prior $\quad P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)}$ with $P(\theta)$ + $P(\theta|x)$ being the same form

the likelihood $P(x|\theta)$ has products of $\mu^x(1-\mu)^{1-x}$ so if we use the form $\mu^a(1-\mu)^b$ for the prior then the posterior will be of the same form

$$\Rightarrow \quad P(\mu|a,b) = C\,\mu^{a-1}(1-\mu)^{b-1}$$

not sure where -1 comes from

we need to find the normalisation C

$\mu \in [0,1]$ so

$$\frac{1}{C} = \int_0^1 P(\mu|a,b) = \int_0^1 \mu^{a-1}(1-\mu)^{b-1}\,d\mu$$

W.t.S. $\quad \frac{1}{C} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

have $\quad \Gamma(a) = \int_0^\infty u^{(a-1)}e^{-u}\,du$

$$\Gamma(a)\Gamma(b) = \int_0^\infty e^{-x}x^{a-1}\,dx\int_0^\infty e^{-y}y^{b-1}\,dy$$

$$= \int_0^\infty\int_0^\infty e^{-(x+y)}x^{a-1}y^{b-1}\,dx\,dy$$

sub $t = y+x$

$y = t$

$$= \int_0^\infty\int_t^\infty e^{-t}x^{a-1}(t-x)^{b-1}\,dx\,dt$$

$$= \int_0^\infty \int_t^\infty e^{-t} (\mu t)^{a-1} (t - \mu t)^{b-1} \, t \, d\mu \, dt$$

sub $x = \mu t$

$\frac{dx}{d\mu} = t$

$$= \int_0^\infty \int_t^\infty e^{-t} \, t^a \mu^{a-1} \, t^{b-1} (1-\mu)^{b-1} \, d\mu \, dt$$

$x = t \quad \mu = 0$

$\infty \quad x = \infty$

2.19

Show $\underline{\underline{\Sigma}} = \sum_{i=1}^{D} \lambda_i \underline{u}_i \underline{u}_i^T$ with $\underline{\Sigma} \underline{u}_i = \lambda_i \underline{u}_i$

When Σ is real and symmetric i.e $\Sigma = \Sigma^T$ the eigenvectors from an orthonormal basis for $\mathbb{R}^D$

so $\langle \underline{u}_i | \underline{u}_j \rangle = \delta_{ij}$

$$\Sigma = I \Sigma \qquad (I\Sigma)_{ij} = \sum_k I_{ik} \Sigma_{kj}$$

$$= \sum_k \underline{u}_i^T \underline{u}_k \Sigma_{kj}$$

$$= \sum_{k,l} U_{li} U_{kl} \Sigma_{kj}$$

$\Sigma = \Sigma^T$

$$= \sum_{kl} U_{li} (U^T)_{lk} \Sigma_{jk}$$

Also $(\Sigma \underline{u}_a)_b = \sum_c \Sigma_{bc} U_{ac}$

$$= \lambda_a U_{ab}$$

$$\Rightarrow (\Sigma)_{ij} = \sum_l U_{li} \lambda_l U^T_{lj}$$

$$\Rightarrow \Sigma = \sum_l \lambda_l \underline{u}_l \underline{u}_l^T$$

2.36)

$$\sigma_{ML}^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n-\mu)^2 = \frac{1}{N}\sum_{n=1}^{N}x_n^2 - \frac{2\mu}{N}\sum_{n=1}^{N}x_n + \frac{1}{N}\sum_{n=1}^{N}\mu^2$$

$$= \frac{1}{N}\sum_{n=1}^{N-1}x_n^2 + \frac{x_N^2}{N} - \frac{2\mu}{N}\sum_{n=1}^{N-1}x_n - \frac{2\mu x_n}{N} + \mu^2$$

$$B\,(N-1)\mu^2 + \frac{\mu^2}{N} = 1$$

$$= \frac{1}{N}\left(\sum_{n=1}^{N-1}(x_n^2 - 2\mu x_n + \mu)\right) + \frac{x_N^2}{N} - \frac{4\mu x_n}{N} + \frac{\mu^2}{N}$$

$$B(N-1) + \frac{1}{N} = 1$$

$$B = \frac{1-\frac{1}{N}}{N-1} = \frac{N}{(N-1)^2}$$

$$= \frac{N-1}{N}\sigma_{ML}^{(N-1)2} + \frac{(x_N-\mu)^2}{N}$$

get it in robbins monroe form

$$= \sigma_{ML}^{(N-1)2} - \frac{1}{N}\sigma_{ML}^{(N-1)2} + \frac{(x_N-\mu)^2}{N}$$

$$\sigma_{ML}^{(N)2} = \text{"} \quad - \frac{1}{N}\left(\sigma_{ML}^{(N-1)2} - \frac{(x_N-\mu)^2}{N}\right)$$

robbins monroe

$$\theta^{(N)} = \theta^{(N-1)} + a_{N-1}\frac{\partial}{\partial\theta^{(N-1)}}\ln(p(x_n|\theta^{(N-1)}))$$

$$\ln(p) = -\ln\sigma^{(N-1)} - \frac{1}{2}\ln 2\pi - \frac{1}{2}\left(\frac{x_n-\mu^{(N-1)}}{\sigma^{(N-1)}}\right)^2$$

$$\frac{\partial \ln p}{\partial\sigma^{N-1}} = -\frac{1}{\sigma^{N-1}} + \frac{(x_n-\mu^{(N-1)})^2}{\sigma^{(N-1)3}}$$

satisfies conditions

So set $a_{N-1} = N\sigma^{(N-1)3}$ and the forms are equivalent

wrong do w.r.t. σ^2 not σ

2.46

$$W(\Lambda) = B|\Lambda|^{\nu-D-1} \exp\left(-\frac{1}{2} Tr(W^{-1}\Lambda)\right)$$

post $\propto$ like $\times$ prior

like given by $C|\Lambda|^{1/2} e^{-\frac{1}{2}(x-\mu)^T\Lambda(x-\mu)}$

post $\propto B|\Lambda|^{\nu-D-1} e^{-\frac{1}{2}Tr(W^{-1}\Lambda)} \; C|\Lambda|^{-1/2} e^{-\frac{1}{2}(x-\mu)^T\Lambda(x-\mu)}$

$$= B|\Lambda|^{\nu-D-3/2} e^{-\frac{1}{2}\left(Tr(W^{-1}\Lambda)+(x-\mu)^T\Lambda(x-\mu)\right)}$$

Trace is linear function so we have the form of the prior $|\Lambda|^c e^{-L(\Lambda)}$ with L linear