Q4 (a) given details $\longrightarrow$

$$\left\{ \begin{array}{l} P(\sigma^2) = IG(a_0, b_0) \\ P(w \mid \sigma^2, \Phi) = \mathcal{N}(0, \gamma \sigma^2 I) \end{array} \right\}$$

$a_0 = 0.1$
$b_0 = 0.00001$
$\gamma = 0.001/N$ ↳ no. of training samples.

$\Phi \equiv$ design matrix $\longrightarrow$ dimensions $(N, K+1)$

K = degree of the polynomial.

$$P(w, \sigma^2 \mid \Phi) = P(w \mid \sigma^2, \Phi)\, P(\sigma^2)$$

$$= \mathcal{N}(0, \gamma I \sigma^2) \cdot IG(a_0\ b_0)$$

$$= NIG(0, \gamma I, a_0, b_0)$$

↳ normal inverse gamma.

$$= \frac{b_0^{a_0}}{(2\pi)^{K+1/2} |\gamma I|^{1/2}\, \Gamma(a_0)} \left(\frac{1}{\sigma^2}\right)^{a_0 + \frac{K+1}{2} + 1} \times \exp\left(\frac{-1}{\sigma^2}\left(b_0 + \frac{1}{2}\frac{W^T W}{\gamma}\right)\right)$$

(note $W^T (\gamma I)^{-1} W = \frac{W^T W}{\gamma}$)

$$\propto \left(\frac{1}{\sigma^2}\right)^{a_0 + \frac{K+1}{2} + 1} \cdot \exp\left(\frac{-1}{\sigma^2}\left(b_0 + \frac{W^T W}{\gamma}\right)\right)$$

↳ this is the normal inverse gamma prior.

* to find the marginal posterior, posterior

$$P(\beta, \sigma^2 \mid y) \propto \left(\frac{1}{\sigma^2}\right)^{a_0 + \frac{n+K+1}{2} + 1} \exp\left(-\frac{1}{\sigma^2}\left(b_0 + \frac{1}{2}\left\{\frac{w^T w}{\gamma} + (y - \phi w)^T (y - \phi w)\right)\right)\right)$$

$$= \left(\frac{1}{\sigma^2}\right)^{a + \frac{n+K+1}{2} + 1} \exp\left(-\frac{1}{\sigma^2}\left(b^* + \frac{1}{2}(w - \mu^*)^T V^{*-1} (w - \mu^*)\right)\right)$$

$$\mu^* = \left(\tfrac{1}{\gamma} I + \phi^T \phi\right)^{-1} \left(\cancel{\tfrac{1}{\gamma} I \times 0} + \phi^T y\right)$$

$$= \left(I/\gamma + \phi^T \phi\right)^{-1} \phi^T y$$

$$V^* = \left(I/\gamma + \phi^T \phi\right)^{-1}$$

$$a^* = a_0 + n/2$$

$$b^* = b_0 + \frac{1}{2}\left[y^T y - \mu^{*T} V^{*-1} \mu^*\right]$$

↳ note that the marginal posterior distribution of σ^2 is $IG(a^* b^*)$

↳ to find the marginal posterior on the weights we have to integrat σ^2 out.

$$P(w|y) = \int P(w,\sigma^2|y)\, d\sigma^2$$

$$= \int NIG(a^*, b^*, \mu^*, V^*)\, d\sigma^2$$

$$\propto \int \left(\frac{1}{\sigma^2}\right)^{a^*+1} \exp\left(\frac{-1}{\sigma^2}\left[b^* + \frac{1}{2}(w-\mu^*)^T V^{*-1}(w-\mu^*)\right]\right) d\sigma^2$$

$$\propto \left[1 + \frac{(w-\mu^*)^T V^{*-1}(w-\mu^*)}{2b^*}\right]^{-a^* + \frac{K+1}{2}}$$

proportional to a multivariate t-distrb^n.

with parameters

$$\nu^* = 2a^*$$

$$\Sigma^* = \frac{b^*}{a^*} V^*$$

our next goal is to compute the marginal distrb^n of y.

$$P(y) = \int P(y|w, \sigma^2)\, P(w,\sigma^2)\, dw\, d(\sigma^2)$$

$$= \int N(\phi \bar{w}, \sigma^2 I_n) \times NIG\left(0, \cancel{\tfrac{1}{\gamma} I}^{\,\gamma I}, a_0, b_0\right) d\bar{w}\, d\sigma^2$$

$$= \frac{b_0^{a_0}}{(2\pi)^{\frac{k+1}{2}} |\gamma I|^{1/2} \Gamma(a)} \int \left(\frac{1}{\sigma^2}\right)^{a^* + \frac{k+1}{2} + 1} \times \exp\left(\frac{-1}{\sigma^2}\left[b^* + \frac{1}{2}(w - \mu^*)^T V^{*-1} (w - \mu^*)\right]\right)$$

$$= \frac{b_0^{a_0}}{\Gamma(a)\, (2\pi)^{\frac{n+k+1}{2}} \sqrt{|\gamma I|}} \times \frac{\Gamma(a^*)\, (2\pi)^{k+1/2} \sqrt{|V^*|}}{(b^*)^{a^*}}$$

$$= \frac{b_0^{a_0}\, \Gamma(a_0 + n/2) \sqrt{|V^*|}}{(2\pi)^{n/2}\, \Gamma(a_0) \sqrt{|\gamma I|}} \times \left[b + \frac{1}{2}\left(y^T y - \mu^* V^{*-1} \mu^*\right)\right]^{-k'}$$

where $k' = \left(a_0 + \frac{n}{2}\right)$

to find the predictive distrb^n ⟶

$$P(\tilde{y}|y) = \int P(\tilde{y}|W\sigma^2)\, P(W, \sigma^2|y)\, dW\, d\sigma^2$$

$$= \int N(\tilde{\Phi} W, \sigma^2 I)\, NIG(\mu^* V^* a^* b^*)$$

$$= t_{2a^*}\left(\tilde{\Phi}\mu^*, \frac{b^*}{a^*}(I + \tilde{\Phi} V^* \tilde{\Phi}^T)\right)$$