we have that

$$p(w|x,t,\alpha,\beta) = \frac{p(t|x,w,\beta)p(w|\alpha)}{p(x,t,\alpha,\beta)}$$

We are trying to maximize the posterior to find w, which means we have to maximize $p(t|x,w,\beta)p(w|\alpha)$. Suppose $p(w|\alpha)$ is a normal distribution, we have:

$$p(w|\alpha) = N(w|0,\alpha^(-1)I) = ((\frac{\alpha}{2pi})^(M+1)/2)exp(-\frac{\alpha}{1}w^(T)w)$$

So

$$p(w|x,t,\alpha,\beta)\alpha p(t|x,w,\beta)p(w|a)\alpha exp(-\frac{\beta}{2}\sum_{n=1}^{N}(y(x_n,w)-t_n)^2 - \frac{\alpha}{2}w^(T)w)$$

we find that the maximum of the posterior by the minimum of

$$\frac{\beta}{2}\sum_{n=1}^{N}(y(x_n,w)-t_n)^2 + \frac{\alpha}{2}w^(T)w$$

or minimize

$$Q = (\| Xw - t \|)^2 + \lambda w^(T)w$$

we have

$\frac{dQ}{dw} = \frac{d}{dw}(Xw-t)^(T)(Xw-t)+2\lambda(Iw) = \frac{d}{dw}(w^(T)X^(T)-t^(T))(Xw-t)+2\lambda Iw = X^(T)Xw+X^(T)Xw-2X^(T)t+2\lambda(Iw) = 0$
$<->X^(T)t = w(X^(T)X+\lambda(I))$
$<->w = (X^(T)X+\lambda(I))^(-1)X^(T)t$