

Given spike and slab prior $p(w|b, \sigma_{sp}^2, \sigma_{sl}^2) = \begin{cases} \mathcal{N}(w|0, \sigma_{sp}^2) ; b=0 \\ \mathcal{N}(w|0, \sigma_{sl}^2) ; b=1 \end{cases}$

This is modelled by bernoulli b , s.t. $b=1$ with probability $\pi = 1/2$
and $b=0$ with probability $(1-\pi) = 1/2$

Then, $p(w|b, \sigma_{sp}^2, \sigma_{sl}^2) = \mathcal{N}(w|0, (\sigma_{sp}^2)^{1-b} (\sigma_{sl}^2)^b)$

Marginal prior $p(w|\sigma_{sl}^2, \sigma_{sp}^2) = \sum_{b \in \{0,1\}} p(w|b, \sigma_{sl}^2, \sigma_{sp}^2) p(b)$

$$= \frac{1}{2} \mathcal{N}(w|0, \sigma_{sp}^2) + \frac{1}{2} \mathcal{N}(w|0, \sigma_{sl}^2)$$

$$= \frac{1}{2} \left[\frac{1}{\sqrt{2\pi\sigma_{sp}^2}} \exp\left\{-\frac{w^2}{2\sigma_{sp}^2}\right\} + \frac{1}{\sqrt{2\pi\sigma_{sl}^2}} \exp\left\{-\frac{w^2}{2\sigma_{sl}^2}\right\} \right]$$

A mixture of 2 gaussians

plot for $p(w|\sigma_{sl}^2=100, \sigma_{sp}^2=1)$

Now, considering a noisy version of w , modelled by $x = w + \epsilon$, $\epsilon \sim \mathcal{N}(\epsilon|0, \sigma^2)$

\Rightarrow since w and ϵ are independent variables,

and $w \sim \mathcal{N}(w|0, (\sigma_{sp}^2)^{1-b} (\sigma_{sl}^2)^b)$, $\epsilon \sim \mathcal{N}(\epsilon|0, \sigma^2)$

we can write x as $x \sim \mathcal{N}(x|0, \sigma^2 + (\sigma_{sp}^2)^{1-b} (\sigma_{sl}^2)^b)$

Using the results that $x = w + \epsilon \Rightarrow \mu_x = \mu_w + \mu_\epsilon$ and $\Sigma_x = \Sigma_w + \Sigma_\epsilon$

Thus,

$$p(b, x | \sigma_{sp}^2, \sigma_{sl}^2) = \mathcal{N}(x|0, \sigma^2 + (\sigma_{sp}^2)^{1-b} (\sigma_{sl}^2)^b) p(b)$$