Solutions Sheet

Nina Fischer and Yannick Zelle

January 25, 2022

Exercise 1

We have :

$$\begin{array}{cc} p(\lambda) & \text{Prior} \\ p(r|\lambda) & \text{Likelihood} \end{array}$$

Thus we can derive our Posterior:

$$\begin{aligned} p(\lambda|r) &= \frac{p(\lambda)p(r|\lambda)}{\int_0^\infty p(\lambda)p(r|\lambda)d\lambda} \\ &\propto p(\lambda)p(r|\lambda) \\ &= e^{-\lambda}\frac{\lambda^{r-1}}{r!} \end{aligned}$$

This funcions corresponds to P^* in the source. We will therefore assign:

$$P^*(\lambda) := e^{-\lambda}\frac{\lambda^{r-1}}{r!}$$

We will consider:

$$\begin{aligned} L(\lambda) &= -ln(P^*(\lambda)) \\ &= \lambda - (r-1)ln(\lambda) + ln(r) \\ &\propto \lambda - (r-1)ln(\lambda) \end{aligned}$$

We will now search for the extrema of L and therefore determine first and second derivative:

$$\frac{\delta L}{\delta \lambda} = 1 - \frac{(r-1)}{\lambda}$$

$$\frac{\delta^2 L}{\delta \lambda^2} = \frac{(r-1)}{\lambda^2}$$

Setting the first derivative to 0 gives us $\lambda_0 = r-1$. We plug this into our second derivative:

$$\frac{\delta^2 L(\lambda_0)}{\delta \lambda^2} = \frac{1}{r-1} > 0$$

Note that this implies that λ_0 minimises L and because the logarithm is a monotonic increasing function maximizes P^*. We therefore set:

$$\lambda_{MAP} = \lambda_0$$

and

$$A = \frac{\delta^2 L(\lambda_0)}{\delta \lambda^2} = \frac{1}{r-1}$$

Therefore we can determine the Gaussian which approximates P^*:

$$g(\lambda) = \frac{1}{\sqrt{2\pi(r-1)}} e^{-\frac{1}{2(r-1)}(\lambda-(r-1))^2}$$

We can read of the hyperparameters and therefore state :

$$p(\lambda|r) \approx N(\lambda|r-1, r-1)$$

Exercise 2

We want to show the equivalence of

$$p(\tilde{f}_i|f_i) = \mathcal{N}(\tilde{f}_i|f_i, \sigma^2) \text{ with } p(y_i = 1|\tilde{f}_i) = \begin{cases} 1, & \text{if } \tilde{f}_i \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

and

$$p(y_i = 1|f_i) = \sigma_{probit}(f_i) = \tfrac{1}{\sqrt{2\pi}} \textstyle\int_{-\infty}^{f_i} exp(-\frac{t^2}{2})dt.$$

Therefore we integrate:

$$\int_{-\infty}^{\infty} p(\tilde{f}_i|f_i)p(y_i = 1|\tilde{f}_i)d\tilde{f}_i = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(\tilde{f}_i - f_i)^2}{2\sigma^2})p(y_i|\tilde{f}_i)d\tilde{f}_i$$

Because of the definition of $p(y_i|\tilde{f}_i)$ we can rewrite:

$$\begin{aligned}\int_{-\infty}^{\infty} p(\tilde{f}_i|f_i)p(y_i = 1|\tilde{f}_i)d\tilde{f}_i &= \int_{0}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(\tilde{f}_i - f_i)^2}{2\sigma^2})d\tilde{f}_i \\ &= \int_{0}^{f_i} \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{t^2}{2\sigma^2})dt\end{aligned}$$

Thus we have

$$\int_{-\infty}^{\infty} p(\tilde{f}_i|f_i)p(y_i = 1|\tilde{f}_i)d\tilde{f}_i = \frac{1}{\sigma\sqrt{2\pi}} \int_{0}^{f_i} exp(-\frac{t^2}{2\sigma^2})dt$$

and to match the function exactly we choose $\sigma^2 = 1$:

$$\frac{1}{1\sqrt{2\pi}} \int_0^{f_i} exp(-\frac{t^2}{2 \cdot 1^2})dt = \frac{1}{\sqrt{2\pi}} \int_0^{f_i} exp(-\frac{t^2}{2})dt$$

Exercise 3

(a) *Proof.* Let $k(x_1, x_2) = C$ with $C \in \mathbb{R}_{>0}$. Then for $x \in \mathbb{R}^n$ we have:

$$x^T k_{\mathbf{xx}} x = C(\sum_{i=1}^{n} x_i)(\sum_{j=1}^{n} x_j)$$

We will show that this sum is greater or equal to 0. To show that let I be the set of indices from 1 to n. Let further be :

$$P \subseteq I := \{i \in I : x_i \geq 0\}$$

$$N \subseteq I := \{i \in I : x_i < 0\}$$

Then we can write :

$$C(\sum_{i=1} x_i)(\sum_{j=1} x_j) = C(\sum_{i\in P} x_i + \sum_{j \in N} x_j)(\sum_{l \in P} x_l + \sum_{k \in N} x_k)$$

We can now distinguish two cases:

Case 1: $\sum_{i\in P} x_i \geq \sum_{j\in N} |x_j|$ Then we have

$$C(\underbrace{\sum_{i\in P} x_i + \sum_{j \in N} x_j}_{\geq 0})(\underbrace{\sum_{l \in P} x_l + \sum_{k \in N} x_k}_{\geq 0}) \geq 0$$

Case 2: $\sum_{i\in P} x_i < \sum_{j\in N} |x_j|$ Then we have

$$C\,\underbrace{(\underbrace{\sum_{i\in P} x_i + \sum_{j \in N} x_j}_{<0})(\underbrace{\sum_{l \in P} x_l + \sum_{k \in N} x_k}_{<0})}_{>0} > 0$$

So we have

$$x^T k_{\mathbf{xx}} x = C(\sum_{i=1}^{n} x_i)(\sum_{j=1}^{n} x_j) \geq 0$$

And k is thus positive semidefinite. $\square$

(b) *Proof.* Let $k(x_1, x_2) = x_1 \cdot x_2$ with $X = \mathbb{R}$.

It follows:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i \cdot c_j \cdot k(x_i, x_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i \cdot c_j \cdot x_i \cdot x_j$$
$$= \sum_{i=1}^{n} c_i \cdot x_i \sum_{i=1}^{n} c_i \cdot x_i \qquad = (\sum_{i=1}^{n} c_i \cdot x_i)^2 \geq 0$$

Thus k is positive semidefinite. □

(c) *Proof.* Let $k(x_1, x_2) = x_1 + x_2$ with $X = \mathbb{R}$. We have x $\in \mathbb{R}$. So with x=-1 it is:

$$(-1) \cdot k(-1, -1) \cdot (-1) = (-1) \cdot (-2) \cdot (-1) = -2 < 0$$

Therefore $k(x_1, x_2) = x_1 + x_2$ is not positive semidefinite and thus is not a kernel. □

(d) *Proof.* Let $k(x_1, x_2) = 5 \cdot x_1^T \cdot x_2$ with $X = \mathbb{R}^D$.

It follows:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i \cdot c_j \cdot k(x_i, x_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i \cdot c_j \cdot 5 \cdot x_i^T \cdot x_j$$
$$= 5 \cdot (\sum_{i=1}^{n} c_i \cdot x_i)^T \cdot (\sum_{i=1}^{n} c_i \cdot x_i) \quad = 5 \cdot \| \sum_{i=1}^{n} c_i \cdot x_i\|_2^2 \geq 0$$

Thus k is positive semidefinite. □

(e) *Proof.* Let $k(x_1, x_2) = (x_1^T \cdot x_2 + 1)^2$ with $X = \mathbb{R}^N$.

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i \cdot c_j \cdot k(x_i, x_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i \cdot c_j \cdot (x_i \cdot x_j + 1)$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i \cdot c_j \cdot x_i \cdot x_j + c_i \cdot c_j$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \cdot x_j \cdot c_i \cdot c_j + \sum_{i=1}^{n}\sum_{j=1}^{n} c_i \cdot c_j$$
$$= (\sum_{i=1}^{n} c_i \cdot x_i)^T \sum_{j=1}^{n} c_j \cdot x_j + (\sum_{i=1}^{n} c_i)^T \sum_{j=1}^{n} c_j$$
$$= || \sum_{i=1}^{n} c_i \cdot x_i||_2^2 + || \sum_{i=1}^{n} c_i||_2^2 \geq 0$$

Thus k is positive semidefinite.

□