Bayesian logistic regression

Lecture 10

(K=2) here

Assume $p(\vec{w}) = \mathcal{N}(\vec{w} \mid \vec{0}, \underbrace{\alpha^{-1}\mathbb{1}}_{"S_0"})$

$$p(\vec{w}\mid\vec{t}) \sim p(\vec{t}\mid\vec{w})\, p(\vec{w})$$

Then $\log p(\vec{w}\mid\vec{t}) = -\frac{\alpha}{2}\vec{w}^T\vec{w} +$

$$+ \sum_{n=1}^{N}\left[t_n \log y_n + (1-t_n)\log(1-y_n)\right] + \text{const}(\vec{w}), \quad (*)$$

where $y_n = \sigma(\vec{w}^T\cdot\vec{\varphi}_n)$ $\quad t_n = \{0,1\}$

$(*)$ is at max when $\vec{w} = \vec{w}_{MAP}$; then

$$S_N \equiv -\vec{\nabla}_{\vec{w}}\vec{\nabla}_{\vec{w}} \log p(\vec{w}\mid\vec{t}) = \alpha\mathbb{1} + \sum_{n=1}^{N} y_n(1-y_n)\vec{\varphi}_n\vec{\varphi}_n^T.$$

Thus $\underset{\text{posterior}}{p(\vec{w}\mid\vec{t})} \Rightarrow q(\vec{w}\mid\vec{t}) = \mathcal{N}(\vec{w}\mid\vec{w}_{MAP}, S_N)$.

Next, predictive distribution:

$$p(C_1\mid\vec{\varphi},\vec{t}) = \int d\vec{w}\, p(C_1\mid\vec{\varphi},\vec{w})\, p(\vec{w}\mid\vec{t}) \simeq$$

$$\simeq \int d\vec{w}\, \sigma(\underbrace{\vec{w}^T\vec{\varphi}}_{"a})\, q(\vec{w}\mid\vec{t})$$

($\vec{\varphi}$: here input vector of features)

$$p(C_2\mid\vec{\varphi},\vec{t}) = 1 - p(C_1\mid\vec{\varphi},\vec{t}).$$

Use $\sigma(\vec{w}^T\vec{\varphi}) = \int da\, \sigma(a)\, \delta(a - \vec{w}^T\vec{\varphi})$ to

obtain: $p(C_1\mid\vec{\varphi},\vec{t}) \simeq \int da\, \sigma(a)\, p(a)$, where

(since $p(\vec{w}\mid\vec{t})$ replaced by $q(\vec{w}\mid\vec{t})$)

$$p(a) = \int d\vec{w}\, q(\vec{w}\mid\vec{t})\, \delta(a - \vec{w}^T\vec{\varphi})$$

So, compute

$$p(a) = \frac{1}{(2\pi)^{M/2}} \frac{1}{|S_N|^{1/2}} \int d\vec{w} \, \frac{1}{2\pi} \int dk \, e^{ik(a - \vec{w}^T \vec{\varphi})} \times$$

$$\times \, e^{-\frac{1}{2}(\vec{w} - \vec{w}_{MAP})^T S_N^{-1} (\vec{w} - \vec{w}_{MAP})} \underset{\vec{w}' = \vec{w} - \vec{w}_{MAP}}{=}$$

$$= \frac{1}{(2\pi)^{M/2}} \frac{1}{|S_N|^{1/2}} \frac{1}{2\pi} \int dk \, e^{ik(a - \vec{\varphi}^T \vec{w}_{MAP})} \times$$

$$\times \int d\vec{w}' e^{-ik\vec{\varphi}^T \vec{w}'} e^{-\frac{1}{2} \vec{w}'^T S_N^{-1} \vec{w}'} \boxed{=}$$

If M is symm. & invertible,

transpose

$$\vec{x}^T M \vec{x} - 2\vec{b}^T \vec{x} = (\vec{x} - M^{-1}\vec{b})^T M (\vec{x} - M^{-1}\vec{b}) - \vec{b}^T M^{-1} \vec{b}$$

Here,
$$\begin{cases} \vec{x} \to \vec{w}', \\ M \to \frac{S_N^{-1}}{2} \Rightarrow M^{-1} = 2S_N, \\ 2\vec{b}^T \to ik\vec{\varphi}^T \Rightarrow \vec{b} \to \frac{ik}{2}\vec{\varphi}. \end{cases}$$

$$\boxed{=} \frac{1}{(2\pi)^{M/2}} \frac{1}{|S_N|^{1/2}} \frac{1}{2\pi} \int dk \, e^{ik(a - \vec{\varphi}^T \vec{w}_{MAP})} e^{-\frac{k^2}{4} \vec{\varphi}^T (2S_N) \vec{\varphi}} \times$$

$$\times \underbrace{\int d\vec{w}' e^{-(\vec{w}' - ikS_N\vec{\varphi})^T \frac{S_N^{-1}}{2} (\vec{w}' - ikS_N\vec{\varphi})}}_{(2\pi)^{\frac{M}{2}} |S_N|^{1/2}} =$$

$$= \frac{1}{2\pi} \int dk \, e^{ik(a - \vec{\varphi}^T \vec{w}_{MAP})} e^{-\frac{k^2}{2} \vec{\varphi}^T S_N \vec{\varphi}}$$

Using $\int dx\, e^{ax^2+bx} = \sqrt{-\frac{\pi}{a}}\, e^{-b^2/4a}$,

we obtain:

$$\begin{cases} x \to k, \\ a \to -\frac{1}{2}\vec{\varphi}^T S_N \vec{\varphi}, \\ b \to i(a - \vec{\varphi}^T \vec{w}_{MAP}) \end{cases}$$, so that

$$p(a) = \underbrace{\frac{1}{2\pi}\sqrt{\frac{2\pi}{\vec{\varphi}^T S_N \vec{\varphi}}}}_{\frac{1}{\sqrt{2\pi\sigma_a^2}}}\, e^{-\frac{(a-\vec{\varphi}^T\vec{w}_{MAP})^2}{2\underbrace{\vec{\varphi}^T S_N \vec{\varphi}}_{\sigma_a^2}}} = \mathcal{N}(a|\underbrace{\mu_a}_{\vec{w}_{MAP}^T\vec{\varphi}}, \sigma_a^2).$$

Thus

$$p(C_1|\vec{\varphi},\vec{t}) = \int da\, \underbrace{\sigma(a)}_{\text{use } \simeq \Phi(\lambda a)} \mathcal{N}(a|\mu_a, \sigma_a^2) \simeq$$

$$\simeq \Phi\left(\frac{\mu_a}{(\lambda^{-2}+\sigma_a^2)^{1/2}}\right) \simeq \sigma\left(\frac{\mu_a}{(1+\underbrace{\lambda^2}_{=\frac{\pi}{8}}\sigma_a^2)^{1/2}}\right).$$

shown before

Note that the DB ($p(C_1|\vec{\varphi},\vec{t}) = p(C_2|\vec{\varphi},\vec{t}) = 0.5$)

is given by $\underbrace{\mu_a}_{=\vec{w}_{MAP}^T\vec{\varphi}} = 0$, linear in feature space

Now, let's show that

$$\int da\, \varphi(\lambda a)\, \mathcal{N}(a \mid \mu_{\bullet}, \sigma_{\bullet}^2) = \varphi\left(\frac{\mu}{\sqrt{\frac{1}{\lambda^2}+\sigma^2}}\right)$$

RHS:

$$\varphi\Bigg(\underbrace{\frac{\mu}{\sqrt{\frac{1}{\lambda^2}+\sigma^2}}}_{=z}\Bigg) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} d\theta\, e^{-\theta^2/2}, \quad \text{s.t.}$$

$$\frac{d\varphi}{d\mu} = \frac{d\varphi}{dz}\frac{dz}{d\mu} = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{\mu^2}{2\left(\frac{1}{\lambda^2}+\sigma^2\right)}}\, \frac{1}{\sqrt{\frac{1}{\lambda^2}+\sigma^2}}$$

LHS:

use $\frac{a-\mu}{\sigma} = z \Rightarrow a = \mu + \sigma z$

$da = \sigma dz$

$$\frac{1}{\sqrt{2\pi\sigma^2}}\, \sigma \int_{-\infty}^{\infty} dz\, \underbrace{\varphi(\lambda(\mu+\sigma z))}_{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\lambda(\mu+\sigma z)} d\theta\, e^{-\theta^2/2}}\, e^{-z^2/2}$$

Then

$$\frac{d}{d\mu}(\text{LHS}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dz\, e^{-\frac{\lambda^2(\mu+\sigma z)^2}{2}}\, \lambda\, e^{-z^2/2} =$$

$$= \frac{\lambda}{2\pi}\, e^{-\frac{\lambda^2\mu^2}{2}} \int dz\, e^{-\lambda^2\mu\sigma z}\, e^{-\frac{z^2}{2}(1+\lambda^2\sigma^2)} =$$

Here, $\begin{cases} a = -\frac{1}{2}(1+\lambda^2\sigma^2) \\ b = -\lambda^2\mu\sigma \end{cases}$

$$= \frac{\lambda}{2\pi}\, e^{-\frac{\lambda^2\mu^2}{2}} \sqrt{\frac{2\pi}{1+\lambda^2\sigma^2}}\, e^{\frac{\lambda^4\mu^2\sigma^2}{2(1+\lambda^2\sigma^2)}} =$$

$$\Leftrightarrow \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\lambda^{-2}+\sigma^2}} e^{-\underbrace{\frac{\lambda^2\mu^2}{2}\left[1-\frac{\lambda^2\sigma^2}{1+\lambda^2\sigma^2}\right]}_{\frac{\lambda^2\mu^2}{2}\frac{1}{1+\lambda^2\sigma^2} = \frac{\mu^2}{2}\frac{1}{\lambda^{-2}+\sigma^2}}}$$

So, $\frac{d}{d\mu}(\text{LHS}) = \frac{d}{d\mu}(\text{RHS}) \Rightarrow \text{LHS} = \text{RHS} + \text{const}(\mu)$

Consider the $\mu \to +\infty$ limit:

$$\lim_{\mu\to\infty} \text{RHS} = 1$$

$$\lim_{\mu\to\infty} \text{LHS} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dz \, e^{-\frac{z^2}{2}} = 1$$

$\phi(\lambda(\mu+\sigma z)) \to 1$ as $\mu \to +\infty$

Thus $\text{const}(\mu) = 0 \Rightarrow \text{LHS} = \text{RHS}$