統計學習初論（105-2）

Self-Assessment Homework 3C

國立台灣大學資管系

Question 1 (Exercise 4.25)

4.25 ($\star\star$) Suppose we wish to approximate the logistic sigmoid $\sigma(a)$ defined by (4.59) by a scaled probit function $\Phi(\lambda a)$, where $\Phi(a)$ is defined by (4.114). Show that if λ is chosen so that the derivatives of the two functions are equal at $a = 0$, then $\lambda^2 = \pi/8$.

Solution:

4.25 From (4.88) we have that

$$\begin{aligned}\left.\frac{\mathrm{d}\sigma}{\mathrm{d}a}\right|_{a=0} &= \sigma(0)(1-\sigma(0)) \\ &= \frac{1}{2}\left(1-\frac{1}{2}\right)=\frac{1}{4}.\end{aligned} \tag{170}$$

Since the derivative of a cumulative distribution function is simply the corresponding density function, (4.114) gives

$$\begin{aligned}\left.\frac{\mathrm{d}\Phi(\lambda a)}{\mathrm{d}a}\right|_{a=0} &= \lambda\mathcal{N}(0|0,1) \\ &= \lambda\frac{1}{\sqrt{2\pi}}.\end{aligned}$$

Setting this equal to (170), we see that

$$\lambda=\frac{\sqrt{2\pi}}{4} \quad \text{or equivalently} \quad \lambda^2=\frac{\pi}{8}.$$

This is illustrated in Figure 4.9.

Question 2 (Exercise 4.26)

4.26 ($\star\star$) In this exercise, we prove the relation (4.152) for the convolution of a probit function with a Gaussian distribution. To do this, show that the derivative of the left-hand side with respect to μ is equal to the derivative of the right-hand side, and then integrate both sides with respect to μ and then show that the constant of integration vanishes. Note that before differentiating the left-hand side, it is convenient first to introduce a change of variable given by $a = \mu + \sigma z$ so that the integral over a is replaced by an integral over z. When we differentiate the left-hand side of the relation (4.152), we will then obtain a Gaussian integral over z that can be evaluated analytically.

Solution:

4.26 First of all consider the derivative of the right hand side with respect to μ, making use of the definition of the probit function, giving

$$\left(\frac{1}{2\pi}\right)^{1/2} \exp\left\{-\frac{\mu^2}{2(\lambda^{-2}+\sigma^2)}\right\} \frac{1}{(\lambda^{-2}+\sigma^2)^{1/2}}.$$

Now make the change of variable $a = \mu + \sigma z$, so that the left hand side of (4.152) becomes

$$\int_{-\infty}^{\infty} \Phi(\lambda\mu + \lambda\sigma z) \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2}z^2\right\} \sigma \,\mathrm{d}z$$

where we have substituted for the Gaussian distribution. Now differentiate with respect to μ, making use of the definition of the probit function, giving

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}z^2 - \frac{\lambda^2}{2}(\mu+\sigma z)^2\right\} \sigma \,\mathrm{d}z.$$

The integral over z takes the standard Gaussian form and can be evaluated analytically by making use of the standard result for the normalization coefficient of a Gaussian distribution. To do this we first complete the square in the exponent

$$\begin{aligned}
&-\frac{1}{2}z^2 - \frac{\lambda^2}{2}(\mu+\sigma z)^2 \\
&= -\frac{1}{2}z^2(1+\lambda^2\sigma^2) - z\lambda^2\mu\sigma - \frac{1}{2}\lambda^2\mu^2 \\
&= -\frac{1}{2}\left[z + \lambda^2\mu\sigma(1+\lambda^2\sigma^2)^{-1}\right]^2 (1+\lambda^2\sigma^2) + \frac{1}{2}\frac{\lambda^4\mu^2\sigma^2}{(1+\lambda^2\sigma^2)} - \frac{1}{2}\lambda^2\mu^2.
\end{aligned}$$

Integrating over z then gives the following result for the derivative of the left hand side

$$
\begin{aligned}
& \frac{1}{(2\pi)^{1/2}} \frac{1}{(1+\lambda^2\sigma^2)^{1/2}} \exp\left\{ -\frac{1}{2}\lambda^2\mu^2 + \frac{1}{2} \frac{\lambda^4\mu^2\sigma^2}{(1+\lambda^2\sigma^2)} \right\} \\
= \quad & \frac{1}{(2\pi)^{1/2}} \frac{1}{(1+\lambda^2\sigma^2)^{1/2}} \exp\left\{ -\frac{1}{2} \frac{\lambda^2\mu^2}{(1+\lambda^2\sigma^2)} \right\}.
\end{aligned}
$$

Thus the derivatives of the left and right hand sides of (4.152) with respect to μ are equal. It follows that the left and right hand sides are equal up to a function of σ^2 and λ. Taking the limit $\mu \to -\infty$ the left and right hand sides both go to zero, showing that the constant of integration must also be zero.