Chapter 5
Construction of the likelihood

Now if we define the the posterior and prior measures as

$$d\mu\left(m\right) \stackrel{\text{def}}{=} \pi\left(m|y\right)d\lambda\left(m\right), \quad d\nu\left(m\right) \stackrel{\text{def}}{=} \pi\left(m\right)d\lambda\left(m\right),$$

Again, from our interpretation in Figure 4.1, $d\mu\left(m\right)$ and $d\nu\left(m\right)$ can be considered as the differential areas around m under the curves $\pi\left(m|y\right)$ and $\pi\left(m\right)$, respectively.

we can rewrite the Bayes formula as

$$\boxed{\frac{d\mu}{d\nu}\left(m\right) = \frac{\pi\left(y|m\right)}{\pi\left(y\right)} \propto \pi\left(y|m\right).} \tag{5.1}$$

Note that we can ignore $\pi\left(y\right)$ in the last expression since $\pi\left(y\right)$ is independent of m, and hence can be considered as a proportional constant.

It is important to point out that, unlike the standard form (4.6) *that is only valid for finite dimensional cases, the Bayes formula in the form* (5.1) *is also valid for infinite dimensional problems. Of course, when both prior and posterior measures admit a density with respect to a reference measure, Lebesgue measure for example, then* (5.1) *reduces to* (4.6).

Definition 5.1. The *conditional mean* is defined as

$$\mathbb{E}\left[m|y\right] = \int_S m\pi\left(m|y\right)\,dm.$$

Exercise 5.1. Show that

$$\mathbb{E}\left[m\right] = \int_T \mathbb{E}\left[m|y\right]\pi\left(y\right)\,dy.$$

•

5.1 Construction of likelihood

In this section, we present a popular approach to construct the likelihood. We begin with the additive noise case. The ideal deterministic model is given by

$$y = h(m),$$

where $y \in \mathbb{R}^r$. But due to random additive noise e, we have the following statistical model instead

$$y^{obs} = h(m) + e, \tag{5.2}$$

where y^{obs} is the actual observation rather than $y = f(m)$. Since the noise comes from external sources, in this note, it is assumed to be independent of m. In the likelihood modeling, we pretend to have realization(s) of m and the task is to construct the distribution of y^{obs}. From (5.2), one can see that the randomness in y^{obs} is the randomness in e shifted by an amount $h(m)$, see Figure 5.1, and hence

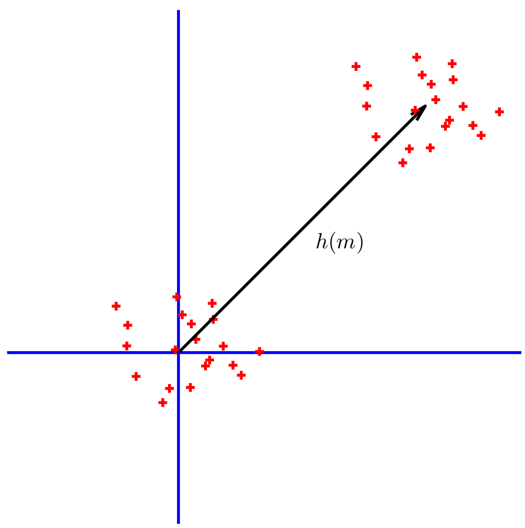

Fig. 5.1 Likelihood model with additive noise.

$\pi_{y^{obs}|m}\left(y^{obs}|m\right) = \pi_e\left(y^{obs} - h(m)\right)$. More rigorously, assume that both y^{obs} and e are random variables on a same probability space, we have

$$\int_A \pi_{y^{obs}|m}\left(y^{obs}|m\right) dy^{obs} \overset{\text{def}}{=} \mu_{y^{obs}|m}(A) \overset{(4.1)}{=} \mu_e(A - h(m))$$
$$= \int_{A-h(m)} \pi_e(e)\, de \overset{\text{change of variable}}{=} \int_A \pi_e\left(y^{obs} - h(m)\right) dy^{obs}, \quad \forall A \subset S,$$

which implies

$$\pi_{y^{obs}|m}\left(y^{obs}|m\right) = \pi_e\left(y^{obs} - h(m)\right).$$

Exercise 5.2. We consider the following multiplicative noise case

$$y^{obs} = eh(m). \tag{5.3}$$

Show that the likelihood for multiplicative noise model (5.3) has the following form

$$\pi_{y^{obs}|m}\left(y^{obs}|m\right) = \frac{\pi_e\left(y^{obs}/h(m)\right)}{h(m)}, \quad h(m) \neq 0. \tag{5.4}$$

Now look at the following multiplicative noise model

$$y_i^{obs} = e_i \times m_i, \quad i = 1, \dots, n,$$

where each e_i is independent, identically distributed by the following log-normal distribution

$$W_i = \log(e_i) \sim \mathcal{N}\left(w_0, \sigma^2\right), \quad w_0 = \log(\alpha_0).$$

Determine the likelihood

$$\pi_{\mathbf{y}^{obs}|\mathbf{m}}\left(\mathbf{y}^{obs}|\mathbf{m}\right),$$

where $\mathbf{y}^{obs} \stackrel{\text{def}}{=} \left[y_1^{obs}, \dots, y_n^{obs}\right]$ and $\mathbf{m} \stackrel{\text{def}}{=} [m_1, \dots, m_n]$. •

Exercise 5.3. Can you generalize the result for the noise model $e = g\left(y^{obs}, h(x)\right)$? •

For concreteness, let us consider the following one dimensional deblurring (deconvolution) problem

$$g(s_j) = \int_0^1 a(s_j, t) f(t)\, dt + e(s_j), \quad 0 \le j \le n,$$

where $a(s,t) = \frac{1}{\sqrt{2\pi\beta^2}} \exp(-\frac{1}{2\beta^2}(t-s)^2)$ is a given kernel, and $s_j = j/n, j = 0, \dots, n$ the mesh points. Our task is to reconstruct $f(t) : [0,1] \to \mathbb{R}$ from the noisy observations $g(s_j)$, $j = 0, \dots, n$. To cast the function reconstruction problem, which is in infinite dimensional space, into a reconstruction problem in $\mathbb{R}^n$, we discretize $f(t)$ on the same mesh and use simple rectangle method for the integral. Let us define $y^{obs} = [g(s_0), \dots, g(s_n)]^T$, $m = (f(s_0), \dots, f(s_n))^T$, and $\mathscr{A}_{i,j} = a(s_i, s_j)/n$, then the discrete deconvolution problem reads

$$y^{obs} = \mathscr{A} m + e.$$

Here, we assume $e \sim \mathcal{N}\left(0, \sigma^2 I\right)$, where I is the identity matrix in $\mathbb{R}^{(n+1)\times(n+1)}$. Since Section 5.1 suggests the likelihood of the form

$$\pi\left(y^{obs}|m\right) = \mathcal{N}\left(\mathscr{A}m, \sigma^2 I\right) \propto \exp\left(-\frac{1}{2\sigma^2}\left(y^{obs} - \mathscr{A}m\right)^T\left(y^{obs} - \mathscr{A}m\right)\right),$$

the Bayesian solution to our inverse problem is, by virtue of the Bayes formula (4.6), given by

$$\boxed{\pi_{\text{post}}\left(m|y^{obs}\right) \propto \exp\left(-\frac{1}{2\sigma^2}\left(y^{obs} - \mathscr{A}m\right)^T\left(y^{obs} - \mathscr{A}m\right)\right) \times \pi_{\text{prior}}(m)}, \quad (5.5)$$

where we have ignored the denominator $\pi\left(y^{obs}\right)$ since it does not depend on the parameter of interest m. Thus, the posterior is "completely determined" once the prior is given and this is the subject of the next section.