Mathematics — "from Eq. (4.14) to Eq. (4.20)"

03 Feb, 2021 (Alexandra Sempuski)

Or, "Can I do anything right this semester? Perhaps I can!"

Beginning at the expression for the least-squares likelihood function describing our experimental measurements with some added Gaussian noise:

$$\text{prob}(\{D_k\}|\{A_j, x_j\}, M, I) \propto \exp\left(-\frac{\chi^2}{2}\right), \quad (4.14)$$

where χ^2 can be expressed as: $\chi^2 = \sum_{k=1}^{N} \left(\frac{F_k - D_k}{\sigma_k}\right)^2$, and

$\{\sigma_k\}$ is the Gaussian noise, while $F_k = \int dx \cdot G(x) \cdot R(x_k - x) + B(x)$.

Model equation

Background term.

After defining these quantities, we know we need to get the posterior pdf for M in the end, so we define that next, using Bayes' Theorem:

$$G(x) = \sum_{j=1}^{M} A_j f(x, x_j), \text{ and}$$

$$f(x, x_j) = \exp\left[-\frac{(x-x_j)^2}{2W^2}\right].$$

→ Width.

$$\text{prob}(M|\{D_k\}, I) = \frac{\text{prob}(\{D_k\}|M, I)\,\text{prob}(M|I)}{\text{prob}(\{D_k\}|I)}.$$

A uniform prior for each M will drop it such that

$$\text{prob}(M|\{D_k\}, I) \propto \underbrace{\text{prob}(\{D_k\}|M, I)}. \quad (4.15)$$

This is the marginal likelihood, because $\{A_j, x_j\}$ is not explicitly stated → implies we've marginalized over it.

Marginal integral can be written as: (4.16)

$$\text{prob}(\{D_k\}|M, I) = \iiint \cdots \int \text{prob}(\{D_k\}, \{A_j, x_j\}|M, I) \cdot d^M A_j \cdot d^M x_j.$$

→ Why here, you ask? →

Well, we need to have this variable on the lhs of the conditional sign to marginalise over it, so we need this expression! However, how do we get that?

↳ By finding a likelihood & a prior:

$$\mathrm{prob}(\{D_k\}, \{A_j, x_j\} | M, I) = \mathrm{prob}(\{D_k\} | \{A_j, x_j\}, M, I) \times$$
$$\times \underbrace{\mathrm{prob}(\{A_j, x_j\} | M, I)}_{\text{Need a prior on these parameters.}}.$$

Already defined in (4.14)!

For the prior, we choose reasonably ranged uniform pdfs on the x_j & A_j:

$$x_{min} \le x_j \le x_{max} \quad \& \quad 0 \le A_j \le A_{max}. \qquad (4.17).$$

Amplitude cannot be negative here.

Otherwise, the prior is zero.

$$\therefore \mathrm{prob}(\{A_j, x_j\} | M, I) = [(x_{max} - x_{min}) . A_{max}]^{-M}. \qquad (4.18).$$

(M lines)

∴ Our marginal likelihood becomes:

$$\mathrm{prob}(\{D_k\} | M, I) = \int\int\int \cdots \int \exp\left(-\frac{\chi^2}{2}\right) \cdot [(x_{max} - x_{min}) \cdot A_{max}]^{-M} \cdot d^M A_j \cdot d^M x_j.$$

∴ We get a posterior of:

$$\mathrm{prob}(M | \{D_k\}, I) \propto [(x_{max} - x_{min}) A_{max}]^{-M} \cdot \int\int\int \cdots \int \exp\left(-\frac{\chi^2}{2}\right) \cdot d^M A_j \cdot d^M x_j. \qquad (4.19).$$

$\pi)^M$

Now we make some approximations, namely we assume a set of 2M parameters $\vec{X}_0 = \{A_{0j}, x_{0j}\}$ which have the optimal fit to the data (the χ^2 is minimised with these parameters).
Employing an old method of Taylor expanding about this point:

$$\chi^2 \simeq \chi^2_{min} + \cancelto{0}{\vec{\nabla}\chi^2(\vec{X}_0)} \cdot (\vec{X}-\vec{X}_0) + \frac{1}{2}(\vec{X}-\vec{X}_0)^T \cdot \vec{\nabla}\vec{\nabla}\chi^2(\vec{X}_0) \cdot (\vec{X}-\vec{X}_0) + \ldots$$

Putting this expansion back in the exponential, we see that: it equals:

$$\exp\left(-\frac{\chi^2_{min}}{2}\right) \cdot \exp\left[-\frac{1}{4}(\vec{X}-\vec{X}_0)^T \cdot \vec{\nabla}\vec{\nabla}\chi^2(\vec{X}_0) \cdot (\vec{X}-\vec{X}_0)\right].$$

In the posterior we now have:

$$\text{prob}(M|\{D_k\}, I) \propto \exp\left(-\frac{\chi^2_{min}}{2}\right) \cdot \underbrace{\iiint \cdots \int \exp\left(-\frac{1}{4}(\vec{X}-\vec{X}_0)^T \cdot \vec{\nabla}\vec{\nabla}\chi^2(\vec{X}_0) \cdot (\vec{X}-\vec{X}_0)\right) \cdot d^{2M}X_j}.$$

In the Appendix of Silvia, (A.19) gives a general formula for this type of Gaussian integral: $\dfrac{(2\pi)^{N/2}}{\sqrt{\det(\underline{\underline{H}})}}$, where $N =$ number of dimensions.

real symmetric matrix

For us, the top factor will be $(4\pi)^{N/2}$ instead of $(2\pi)^{N/2}$, due to the $1/4$ factor in the exponential, instead of the generic $1/2$.
Here $N = 2M$, so we get:

$$\text{prob}(M|\{D_k\}, I) \propto \exp\left(-\frac{\chi^2_{min}}{2}\right) \cdot \frac{(4\pi)^{\cancel{2}M/\cancel{2}}}{\sqrt{\det(\underbrace{\vec{\nabla}\vec{\nabla}\chi^2}_{\text{Our matrix.}})}} \Rightarrow \exp\left(-\frac{\chi^2_{min}}{2}\right) \cdot \frac{(4\pi)^M}{\sqrt{\det(\vec{\nabla}\vec{\nabla}\chi^2)}}.$$

Lastly, to account for any possible location of each M, we multiply by M! and get:

$$\boxed{\text{prob}(M|\{D_k\}, I) \propto \frac{M! \cdot (4\pi)^M}{[(x_{max}-x_{min})A_{max}]^M \cdot \sqrt{\det(\vec{\nabla}\vec{\nabla}\chi^2)}} \cdot \exp\left(-\frac{\chi^2_{min}}{2}\right). \quad (4.20).}$$