Optimal design of experiment for model discrimination

Oliwer Sliczniuk

February 6, 2024

Notes about the optimal model discrimination

1 What information do we have?

- Process model: $\dot{x} = f(t, x, p)$

- Measurement equation: $\dot{y} = g(x)$

- Likelihood function with respect to the dataset

- Parameter estimation

- Sensitivity equations: $\dot{S} = J_x \dot{x} + J_p$

2 Model discrimination

Let's consider two probability distributions $p(y_1)$ and $p(y_2)$ where each represents the Gaussian probability density function for a model.

$$p(Y|y_1) = \prod_{i=1}^{n_t} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(\frac{\sum (Y - y_1(t, x, p))^2}{2\sigma_1^2} \right) \tag{1}$$

If the ratio of two probability distributions is considered to indicate the measure of similarity, then $\ln\left(\frac{p(Y|y_1)}{p(Y|y_2)}\right)$ becomes a measure of the odds in favour of choosing hypothesis H_1 ($p(Y|y_1)$ is a true model) over hypothesis H_2 ($p(Y|y_2)$ is a true model). Alternatively, the ratio can be interpreted as the information in favour of hypothesis H_1 as opposed to the hypothesis H_2. The so-called 'weight of evidence' or expected information in favour of choosing H_1 over H_2 can be defined through the Kullback–Leibler divergence and is represented by:

$$I(1:2) = \int_{-\infty}^{\infty} p(Y|y_1) \ln\left(\frac{p(Y|y_1)}{p(Y|y_2)} \right) dY \tag{2}$$

The above equation can be written more explicitly as

$$I(1:2) = \int_{-\infty}^{\infty} p(Y|y_1) \left[\sum_{i=1}^{n_t} \left(\ln \left(\frac{\sigma_2}{\sigma_1} \right) - \frac{(Y_i - y_{1i})^2}{2\sigma_1^2} + \frac{(Y_i - y_{2i})^2}{2\sigma_2^2} \right) \right] dY$$

$$= \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \ln \left(\frac{\sigma_2}{\sigma_1} \right) \right) dY - \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \frac{(Y_i - y_{1i})^2}{2\sigma_1^2} \right) dY$$

$$+ \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \frac{(Y_i - y_{2i})^2}{2\sigma_2^2} \right) dY \tag{3}$$

As the expected error is constant for all the measurements: $\mathbb{E}[(Y_i - y_{1i})^2] = \mathbb{E}[\sigma_1^2]$, which simplifies the equation:

$$I(1:2) = \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \ln \left(\frac{\sigma_2}{\sigma_1} \right) \right) dY - \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(\frac{1}{2} p(Y|y_1) \right) dY$$

$$+ \sum_{i=1}^{n_t} \int_{-\infty}^{\infty} \left(p(Y|y_1) \frac{(Y_i - y_{2i})^2}{2\sigma_2^2} \right) dY \tag{4}$$

The first two terms can be simplified by taking a constant in front of integrals and by noticing that $\int p(x)dx = 1$.

$$I(1:2) = n_t \ln \left(\frac{\sigma_2}{\sigma_1} \right) - \frac{n_t}{2} + \sum_{i=1}^{n_t} \frac{1}{2\sigma_2^2} \int_{-\infty}^{\infty} \left(p(Y|y_1) \left(Y_i^2 - 2Y_i y_{2i} + y_{2i}^2 \right) \right) dY \tag{5}$$

$$I(1:2) = n_t \ln \left(\frac{\sigma_2}{\sigma_1} \right) - \frac{n_t}{2} + \sum_{i=1}^{n_t} \frac{1}{2\sigma_2^2} \int_{-\infty}^{\infty} \left(p(Y|y_1) \times Y_i^2 \right) dY$$

$$- \sum_{i=1}^{n_t} \frac{2y_{2i}}{2\sigma_2^2} \underbrace{\int_{-\infty}^{\infty} \left(p(Y|y_1) \times Y_i \right) dY}_{\text{expected value} = y_{1i}} + \sum_{i=1}^{n_t} \frac{y_{2i}^2}{2\sigma_2^2} \underbrace{\int_{-\infty}^{\infty} p(Y|y_1) dY}_{=1} \tag{6}$$

The remaining integral can be solved by recognizing that $\sigma^2 = \int_{-\infty}^{\infty} X^2 p(X)dX - \mathbb{E}[(X)]^2$, which leads to $\int_{-\infty}^{\infty} Y_i^2 p(Y|y_{1i})dY = y_{1i}^2 + \sigma_1^2$.

The final form of the Kullback–Leibler divergence becomes:

$$I(1:2) = n_t \ln\left(\frac{\sigma_2}{\sigma_1}\right) - \frac{n_t}{2} + \sum_{i=1}^{n_t}\frac{y_{1i}^2 + \sigma_1^2}{2\sigma_2^2}\sum_{i=1}^{n_t}\frac{2y_{2i}y_{1i}}{2\sigma_2^2} + \sum_{i=1}^{n_t}\frac{y_{2i}^2}{2\sigma_2^2}$$

$$= n_t \ln\left(\frac{\sigma_2}{\sigma_1}\right) - \frac{n_t}{2} + \frac{n_t}{2\sigma_1^2} + \sum_{i=1}^{n_t}\frac{1}{\sigma_2^2}(y_{1i} - y_{2i})^2 \tag{7}$$

$$I(2:1) = n_t \ln\left(\frac{\sigma_1}{\sigma_2}\right) - \frac{n_t}{2} + \frac{n_t}{2\sigma_2^2} + \sum_{i=1}^{n_t}\frac{1}{2\sigma_1^2}(y_{1i} - y_{2i})^2 \tag{8}$$

While Kulback-Liebler divergence is a statistical distance, it is not a metric on the space of probability distributions. While metrics are symmetric and generalize linear distance, satisfying the triangle inequality, divergences are asymmetric in general and generalize squared distance. In general, $I(1:2) \neq I(2:1)$. By taking into account that the Kullbacl-Liebler divergence is additive for independent distribution, the function J for model discrimination can be defined as

$$J(1,2) = I(1:2) + I(2:1) = \int_{-\infty}^{\infty}[p(Y|y_2) - p(Y|y_1)]\ln\frac{p(Y|y_1)}{p(Y|y_2)}dy$$

$$= \frac{n_t(\sigma_1^2 - \sigma_2^2)}{2\sigma_1^2\sigma_2^2} + \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2} \times \sum_{i=1}^{n_t}(y_{1i} - y_{2i})^2 \tag{9}$$

The first term of J is independent of changes in y_1 and y_2, while the second term is equivalent to the sum of squared differences between two model outputs. By maximizing J, the y_1 and y_2 are spread apart. Although both models were fitted with the same dataset, they employ structurally different extraction kinetic terms. These structural differences lead to different outputs, particularly in regions not covered by the dataset. In this work, both models can be manipulated by flow rate, inlet temperature and pressure to find the dynamic operating conditions such that model outputs that differ from each other the most. The optimization problem can be formulated as follow

$$\hat{\theta}_{MLE} = \arg\max_{F,T_{in},P} J(y_1(t, x_1, u), y_2(t, x_2, u))$$

$$\text{s.t.} \quad \dot{x}_1 = G_1(x_1(t); u)$$

$$\dot{x}_2 = G_2(x_2(t); u)$$

$$y_1 = g_1(t, x_1)$$

$$y_2 = g_2(t, x_2) \tag{10}$$

$$1.67 \times 10^{-5} \leq F \leq 6.67 \times 10^{-5}$$

$$30°C \leq T_{in} \leq 40°C$$

$$100 \text{ bar} \leq T_{in} \leq 200 \text{ bar}$$

By employing the dynamic operating conditions determined by the optimization to the real system and comparing the result of a new experiment with responses from both models, the model discrimination can be concluded.