Bayesian inference and Data assimilation

Exercise 5

Jin W. Kim (jin.won.kim@uni-potsdam.de)

Problem 1a There are many ways to understand this problem, but let us consider an intuitive mass transfer problem.

Suppose we have three buckets filled with 1/3 unit of sand. These buckets are on a certain height—a_1, a_2 and a_3—from the baseline. The goal is to re-distribute the sand into another set of buckets at height b_1, b_2 and b_3. The energy requirement to move a unit mass from one bucket to another is proportional to the square of height difference. The goal is to find a strategy that minimizes energy consumption.

Let D_{ij} be the energy consumption per unit mass from a_i to b_j. Then D_{ij} is expressed by the following matrix:

$$[D_{ij}] = \begin{pmatrix} (a_1 - b_1)^2 & (a_1 - b_2)^2 & (a_1 - b_3)^2 \\ (a_2 - b_1)^2 & (a_2 - b_2)^2 & (a_2 - b_3)^2 \\ (a_3 - b_1)^2 & (a_3 - b_2)^2 & (a_3 - b_3)^2 \end{pmatrix} = \begin{pmatrix} 0.25 & 1 & 4 \\ 0.25 & 0 & 9 \\ 2.25 & 1 & 16 \end{pmatrix}$$

Note that the third column, the cost from anywhere to b_3 dominates every other scenarios. Hence we start from here (re-arrange trick). We want to fill b_3 as much from a_1, and if it is not possible, then use some amount from a_2 and so on.

In the problem setting, we have exactly 1/3 unit in a_1 and exactly 1/3 will be transferred into b_3. Hence we assign the mass as follows:

$$[T_{ij}] = \begin{pmatrix} 0 & 0 & 1/3 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 0 \end{pmatrix}$$

We repeat for the rest of the buckets. Now cost to fill b_1 dominates one required to fill b_2, we optimize b_1 first by assign 1/3 from a_2 to b_1, namely, $t_{21} = 1/3$. The resulting optimal strategy is therefore

$$[T^*_{ij}] = \begin{pmatrix} 0 & 0 & 1/3 \\ 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \end{pmatrix}$$

In Example 2.29 in the lecture notes, they also change the order of indices to visualize the strategy. By rearranging, we have

$[D_{ij}]$	b_3	b_1	b_2
a_1	4	0.25	1
a_2	9	0.25	0
a_3	16	2.25	1

$[T_{ij}]$	b_3	b_1	b_2
a_1	1/3	0	0
a_2	0	1/3	0
a_3	0	0	1/3

The strategy is simply put as much mass as allowed on the diagonal elements, and 0 mass on off-diagonals. This strategy naturally incurs the sparse structure of the optimal coupling.

Problem 2a We want to find $\{(b_i, c_i) : i = 1, \dots, M\}$ such that

$$\int_0^1 f(x)\,\mathrm{d}x = \sum_{i=1}^{M} b_i f(c_i)$$

for all $f(x) = a_0 + a_1 x + \ldots + a_{p-1}x^{p-1}$.

- For $M = 1$ and $p = 2$, the left-hand side is

$$\int_0^1 a_0 + a_1 x\,\mathrm{d}x = a_0 + \frac{1}{2}a_1 x^2\Big|_0^1 = a_0 + \frac{1}{2}a_1$$

  Meanwhile the right-hand side is

$$b_1 f(c_1) = a_0 b_1 + a_1 b_1 c_1$$

  Since a_k is arbitrary, we conclude

$$\begin{cases} b_1 = 1 \\ b_1 c_1 = \frac{1}{2} \end{cases} \qquad \Longrightarrow \qquad b_1 = 1,\ c_1 = \frac{1}{2}$$

- For $M = 2$ and $p = 3$, the left-hand side is

$$\int_0^1 a_0 + a_1 x + a_2 x^2\,\mathrm{d}x = a_0 + \frac{1}{2}a_1 x^2 + \frac{1}{3}a_2 x^3\Big|_0^1 = a_0 + \frac{1}{2}a_1 + \frac{1}{3}a_2$$

  Meanwhile the right-hand side is

$$\begin{aligned} b_1 f(c_1) + b_2 f(c_2) &= a_0 b_1 + a_1 b_1 c_1 + a_2 b_1 c_1^2 + a_0 b_2 + a_1 b_2 c_2 + a_2 b_2 c_2^2 \\ &= a_0(b_1 + b_2) + a_1(b_1 c_1 + b_2 c_2) + a_2(b_1 c_1^2 + b_2 c_2^2) \end{aligned}$$

  Since a_k is arbitrary, we conclude

$$b_1 + b_2 = 1 \tag{1}$$

$$b_1 c_1 + b_2 c_2 = \frac{1}{2} \tag{2}$$

$$b_1 c_1^2 + b_2 c_2^2 = \frac{1}{3} \tag{3}$$

  Multiply c_2 on (1) and subtract from (2) to obtain

$$b_1(c_1 - c_2) = \frac{1}{2} - c_2 \tag{4}$$

  Multiply c_2 on (2) and subtract from (3) to obtain

$$b_1(c_1^2 - c_1 c_2) = \frac{1}{3} - \frac{1}{2}c_2 \tag{5}$$

Since the left-hand side of (5) equals to $c_1 \cdot b_1(c_1 - c_2)$, substitute (4) to get

$$c_1\Big(\frac{1}{2} - c_2\Big) = \frac{1}{12} + \frac{1}{2}\Big(\frac{1}{2} - c_2\Big)$$

By rearranging terms, one obtains

$$\Big(c_1 - \frac{1}{2}\Big)\Big(\frac{1}{2} - c_2\Big) = \frac{1}{12} \tag{6}$$

Observe that (4) is equivalent to

$$b_1\Big[\Big(c_1 - \frac{1}{2}\Big) + \Big(\frac{1}{2} - c_2\Big)\Big] = \frac{1}{2} - c_2 \tag{7}$$

Therefore we can solve for $\tilde{c}_2 := \dfrac{1}{2} - c_2$ in terms of b_1 from (6) and (7):

$$b_1\Big(\frac{1}{12\tilde{c}_2} + \tilde{c}_2\Big) = \tilde{c}_2 \quad \Longrightarrow \quad \tilde{c}_2 = \frac{1}{2\sqrt{3}}\sqrt{\frac{b_1}{b_2}}$$

Upon substituting this to (6),

$$c_1 - \frac{1}{2} = \frac{1}{2\sqrt{3}}\sqrt{\frac{b_2}{b_1}}$$

In summary, we have

$$c_1 = \frac{1}{2} + \frac{1}{2\sqrt{3}}\sqrt{\frac{b_2}{b_1}}, \quad c_2 = \frac{1}{2} - \frac{1}{2\sqrt{3}}\sqrt{\frac{b_1}{b_2}}$$

We want these values to be in $[0, 1]$, and therefore from $c_1 \leq 1$,

$$\frac{1}{2} + \frac{1}{2\sqrt{3}}\sqrt{\frac{b_2}{b_1}} \leq 1 \quad \Longrightarrow \quad \frac{b_2}{b_1} \leq 3$$

and from $c_2 \geq 0$,

$$\frac{1}{2} - \frac{1}{2\sqrt{3}}\sqrt{\frac{b_1}{b_2}} \geq 0 \quad \Longrightarrow \quad \frac{b_1}{b_2} \leq 3$$

Since $b_2 = 1 - b_1$, we conclude the admissible solution is given by

$$\frac{1}{4} \leq b_1 \leq \frac{3}{4}, \quad b_2 = 1 - b_1, \quad c_1 = \frac{1}{2} + \frac{1}{2\sqrt{3}}\sqrt{\frac{b_2}{b_1}}, \quad c_2 = \frac{1}{2} - \frac{1}{2\sqrt{3}}\sqrt{\frac{b_1}{b_2}}$$

Problem 2b We want to find the decomposition of the form

$$f(x_1, x_2) = f_0 + f_1(x_1) + f_2(x_2) + f_{12}(x_1, x_2)$$

where

$$\begin{aligned} f_0 &= \iint_{[0,1]^2} f(x_1, x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\ f_1(x_1) &= \int_{[0,1]} f(x_1, x_2)\,\mathrm{d}x_2 - f_0 \\ f_2(x_2) &= \int_{[0,1]} f(x_1, x_2)\,\mathrm{d}x_1 - f_0 \\ f_{12}(x_1, x_2) &= f(x_1, x_2) - f_1(x_1) - f_2(x_2) - f_0 \end{aligned}$$

Given the function $f(x_1, x_2) = 12x_1 + 6x_2 - 6x_1x_2$, we can compute

$$\begin{aligned} \int_{[0,1]} f(x_1, x_2)\,\mathrm{d}x_2 &= \int_{[0,1]} 12x_1 + 6x_2 - 6x_1x_2\,\mathrm{d}x_2 \\ &= 12x_1x + 3x^2 - 3x_1x^2\Big|_0^1 \\ &= 9x_1 + 3 \end{aligned}$$

Therefore

$$f_0 = \int_0^1 9x_1 + 3\,\mathrm{d}x_1 = \frac{15}{2}$$

and

$$f_1(x_1) = 9x_1 + 3 - \frac{15}{2} = 9x_1 - \frac{9}{2}$$

Meanwhile,

$$\begin{aligned} f_2(x_2) &= \int_{[0,1]} 12x_1 + 6x_2 - 6x_1x_2\,\mathrm{d}x_1 - f_0 \\ &= 6x^2 + 6x_2x - 3x_2x^2\Big|_0^1 - \frac{15}{2} \\ &= 3x_2 - \frac{3}{2} \end{aligned}$$

In consequence,

$$\begin{aligned} f_{12}(x_1, x_2) &= f(x_1, x_2) - f_1(x_1) - f_2(x_2) - f_0 \\ &= 12x_1 + 6x_2 - 6x_1x_2 - 9x_1 + \frac{9}{2} - 3x_2 + \frac{3}{2} - \frac{15}{2} \\ &= 3x_1 + 3x_2 - 6x_1x_2 - \frac{3}{2} \end{aligned}$$

In order to compute the variance contribution, we note that

$$\mathbb{E}[X_i] = \frac{1}{2}, \quad \mathrm{Var}[X_i] = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

for $i = 1, 2$. Also X_1 and X_2 are independent, hence $\operatorname{Cov}[X_1, X_2] = 0$. Now it is straightforward that

$$\sigma_1^2 = \operatorname{Var}[f_1(X_1)] = \operatorname{Var}\left[9X_1 - \frac{9}{2}\right] = 81\operatorname{Var}[X_1] = \frac{27}{4}$$

and

$$\sigma_2^2 = \operatorname{Var}[f_2(X_2)] = \operatorname{Var}\left[3X_2 - \frac{3}{2}\right] = 9\operatorname{Var}[X_2] = \frac{3}{4}$$

Now for σ_{12}^2,

$$\begin{aligned}
\sigma_{12}^2 &= \operatorname{Var}[f_{12}(X_1, X_2)] \\
&= \operatorname{Var}\left[3X_1 + 3X_2 - 6X_1X_2 - \frac{3}{2}\right] \\
&= \operatorname{Var}[3X_1] + \operatorname{Var}[3X_2] + \operatorname{Var}[6X_1X_2] \\
&\quad + 2\operatorname{Cov}[3X_1, 3X_2] - 2\operatorname{Cov}[3X_1, 6X_1X_2] - 2\operatorname{Cov}[3X_2, 6X_1X_2] \\
&= 9\operatorname{Var}[X_1] + 9\operatorname{Var}[X_2] + 36\operatorname{Var}[X_1X_2] - 36\operatorname{Cov}[X_1, X_1X_2] - 36\operatorname{Cov}[X_2, X_1X_2]
\end{aligned}$$

Note that $\operatorname{Cov}[X_1, X_1X_2] = \operatorname{Cov}[X_2, X_1X_2]$ due to the symmetry. Now compute

$$\begin{aligned}
\operatorname{Var}[X_1X_2] &= \mathbb{E}[(X_1X_2)^2] - \mathbb{E}[X_1X_2]^2 \\
&= \mathbb{E}[X_1^2]\mathbb{E}[X_2]^2 - (\mathbb{E}[X_1]\mathbb{E}[X_2])^2 \\
&= \left(\frac{1}{12} + \frac{1}{4}\right)\left(\frac{1}{12} + \frac{1}{4}\right) - \left(\frac{1}{4}\right)^2 \\
&= \frac{7}{144}
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{Cov}[X_1, X_1X_2] &= \mathbb{E}[X_1X_1X_2] - \mathbb{E}[X_1]\mathbb{E}[X_1X_2] \\
&= \mathbb{E}[X_1^2]\mathbb{E}[X_2] - \mathbb{E}[X_1]^2\mathbb{E}[X_2] \\
&= \left(\frac{1}{12} + \frac{1}{4}\right)\frac{1}{2} - \left(\frac{1}{2}\right)^2\frac{1}{2} \\
&= \frac{1}{24}
\end{aligned}$$

Collecting the results,

$$\begin{aligned}
\sigma_{12}^2 &= 9\frac{1}{12} + 9\frac{1}{12} + 36\frac{7}{144} - 72\frac{1}{24} \\
&= \frac{1}{4}
\end{aligned}$$

Note that

$$\sigma_1^2 = \frac{27}{4} > \sigma_2^2 = \frac{3}{4} > \sigma_{12}^2 = \frac{1}{4}$$

and thus we note that f_{12} contributes least on the variance of $f(X_1, X_2)$, as desired.