Notes

SCK team

1 Notation

- Average emission per source: $x_1, \ldots, x_m$ where $m = 200$.
- Measured concentrations $y_1, \ldots, y_n$ where $n = 4636$. These are consecutive measurements for several measurement stations.
- $M_{i,(j,\Delta)}$ is a block in the sensitivity matrix M that represents the contribution of x_j from $\Delta - 1$ days ago to the observed concentration y_i.
- Scaling factors $s_1, \ldots, s_m$ (in `scalings.csv`). These are put in a diagonal matrix S.
- $\mathbf{1}_q = [1 \ \ldots \ 1]^T \in \mathbb{R}^q$ is the constant vector consisting of only ones.
- $\otimes$ is the Kronecker product. In particular,

$$\underbrace{\begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}}_{\in \mathbb{R}^p} \otimes \underbrace{\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}}_{\in \mathbb{R}^q} = \underbrace{\begin{bmatrix} a_1 \\ \vdots \\ a_1 \\ a_2 \\ \vdots \\ a_2 \\ \vdots \\ a_p \end{bmatrix}}_{\in \mathbb{R}^{pq}}.$$

A standard rule of thumb in numerical linear algebra is that the Kronecker product serves a mostly theoretical purpose and is rarely calculated explicitly. See, e.g. [1, Section 12.3].

2 Model

2.1 Initial model

With $n = 1, \ldots, 4636$ and $m = 1, \ldots, 200$, the initial model is

$$y_i \approx \hat{y}_i = \sum_{j=1}^{m} \sum_{\Delta=1}^{15} M_{i,(j,\Delta)} \frac{x_j}{s_j}.$$

That is, for each emission source j, we simulate the contribution of 15 days, and add this together for all j. A matrix formulation of this reads as follows:

$$\begin{aligned} \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{bmatrix} &= \begin{bmatrix} M_{1,(1,1)} & M_{1,(1,2)} & \cdots & M_{1,(m,15)} \\ \vdots & \vdots & & \vdots \\ M_{n,(1,1)} & M_{n,(1,2)} & \cdots & M_{n,(m,15)} \end{bmatrix} \begin{bmatrix} s_1^{-1} x_1 \\ \vdots \\ s_1^{-1} x_1 \\ s_2^{-1} x_2 \\ \vdots \\ s_n^{-1} x_m \end{bmatrix} \\ &= \begin{bmatrix} M_{1,(1,1)} & M_{1,(1,2)} & \cdots & M_{1,(n,15)} \\ \vdots & \vdots & & \vdots \\ M_{m,(1,1)} & M_{m,(1,2)} & \cdots & M_{m,(n,15)} \end{bmatrix} \left(\begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \otimes \mathbf{1}_{15} \right). \end{aligned} \tag{1}$$

Now we present a different interpretation. Instead of $S^{-1}x \otimes \mathbf{1}_{15}$, we make a vector $\tilde{x}_{365} := S^{-1}x \otimes \mathbf{1}_{365}$. This is a vector of length 365×200 that gives the emission from day 1 to day 365 for each source. This is assumed to be constant for all days. Therefore, $\tilde{x}_{365}$ consists of $m = 200$ blocks, each of which is a constant vector of length 365.

We can write a system of equations that is logically the same as eq. (1) but uses $\tilde{x}_{365}$. This gives a larger linear system. Because the emission from most dates has a negligible contribution to the observed concentrations, many coefficients in the larger matrix will be zero. For every measurement i, we have a linear relationship that looks something like this:

$$\hat{y}_i = \begin{bmatrix} \cdots & M_{i,(1,1)} & \dots & M_{i,(1,15)} & 0 & \dots & M_{i,(2,1)} & \dots & M_{i,(2,15)} & \dots \end{bmatrix} (S^{-1}x \otimes \mathbf{1}_{365}).$$

The presence of zeros in the matrix indicates that emissions from the distant past do not contribute to the observed concentrations.

Now assume that in each station, there is one observation per day, named y_{kt}, where $k = 1, \dots, K$ where K is the number of stations and $t = 1, \dots, 365$. The index $t = 1$ corresponds to the most recent measurement. Then the full system reads as follows:

$$\begin{bmatrix} \hat{y}_{1,1} \\ \hat{y}_{1,2} \\ \vdots \\ \hat{y}_{2,1} \\ \vdots \\ \hat{y}_{K,365} \end{bmatrix} = \underbrace{\begin{bmatrix} M_{(1,1),(1,1)} & M_{(1,1),(1,2)} & \cdots & M_{(1,1),(1,15)} & 0 & \cdots & \cdots & M_{(1,1),(2,1)} & \cdots & 0 \\ 0 & M_{(1,2),(1,1)} & \cdots & M_{(1,2),(1,14)} & M_{(1,2),(1,15)} & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\ M_{(2,1),(1,1)} & M_{(2,1),(1,2)} & \cdots & M_{(2,1),(1,15)} & 0 & \cdots & \cdots & M_{(2,1),(2,1)} & \cdots & 0 \\ \vdots & & & & & & & & & \\ \vdots & & & & & & & & & \end{bmatrix}}_{:=M_{shift}} \tilde{x}_{365} \tag{2}$$

where

$$\tilde{x}_{365} := S^{-1}x \otimes \mathbf{1}_{365} = \begin{bmatrix} s_1^{-1}x_1 \\ \vdots \\ s_1^{-1}x_1 \\ s_2^{-1}x_2 \\ \vdots \\ s_m^{-1}x_m \end{bmatrix}.$$

We could reorganise the vector $\tilde{x}_{365}$ to the transposed expression $\mathbf{1}_{365} \otimes (S^{-1}x)$, which is a vector containing all emissions on day 1, followed by all emissions on day 2, etc. A similar permutation of y could be chosen. This has the effect of permuting the rows and columns of M_{shift}. A permutation of the variables gives a model that is essentially the same as eq. (2) but may have advantages when it comes to implementation or computational efficiency. We will not consider this.

2.2 Model with time-dependent correction to the emissions

If we weigh each day and each emission source by some factor, we are applying a transformation

$$\begin{bmatrix} s_1^{-1}x_1 \\ \vdots \\ s_1^{-1}x_1 \\ s_2^{-1}x_2 \\ \vdots \\ s_m^{-1}x_m \end{bmatrix} \mapsto \begin{bmatrix} w_{1,1}s_1^{-1}x_1 \\ \vdots \\ w_{1,365}s_1^{-1}x_1 \\ w_{2,1}s_2^{-1}x_2 \\ \vdots \\ w_{m,365}s_m^{-1}x_m \end{bmatrix}$$

which can be written concisely as

$$W(S^{-1}x \otimes \mathbf{1}_{365}) \quad \text{where} \quad W = \begin{bmatrix} w_{1,1} & 0 & \cdots & 0 \\ 0 & w_{1,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & w_{m,365} \end{bmatrix}.$$

Using this weighted vector of emissions instead of $\tilde{x}_{365}$ in eq. (2) gives

$$\hat{y} = M_{shift}W(S^{-1}x \otimes \mathbf{1}_{365}). \tag{3}$$

In this expression, the unknown is W and should be fit to the data. There are a plethora of numerical techniques for solving linear equations where the unknown is a matrix. In the following, we provide a theoretical description of a system that can be solved for W.

We can put constraints on W of the form $W \in \mathcal{W}$, where $\mathcal{W}$ is some set of constrained diagonal matrices. Suppose we have a parametrisation $\Gamma : \mathbb{R}^k \to \mathcal{W}, u \mapsto \Gamma(u)$. In this case, eq. (3) can be reformulated as a linear operation applied to $\Gamma(u)$ as follows:

$$\hat{y} = ((S^{-1}x \otimes \mathbf{1}_{365})^T \otimes M_{shift})\mathrm{vec}\,\Gamma(u).$$

If Γ is a linear map, the above is a linear expression in u. Then, the minimisation of $\|y - \hat{y}\|^2$ can be solved using numerical linear algebra.

References

[1] Gene H Golub and Charles F Van Loan. *Matrix computations*. Vol. 3. Baltimore: JHU press, 2013.