Outer-Product Emulator (OPE) and Calibration for Multivariate Output

February 12, 2021

General notations:

- m – number of simulation/computer model runs; moderate.
- M – output dimensionality, e.g. the number of grid points where the displacement is calculated; large.
- n – number of data samples observed, e.g. the number of beams. For now we have $n = 1$.
- N – number of data points observed at each sample; large.
- $\boldsymbol{x}$ – the vector of controllable inputs (locations/grid coordinates); $\dim(\boldsymbol{x}) = p$.
- $\boldsymbol{\theta}$ – uncontrollable inputs/calibration parameters; $\dim(\boldsymbol{\theta}) = q$.

Calibration statistical model – relationship between the field measurement y and the computer model at location $\boldsymbol{x}_i$, with the "true" value of the unobserved parameters $\boldsymbol{\theta}$.

$$y(\boldsymbol{x}_i) = \eta(\boldsymbol{x}_i, \boldsymbol{\theta}) + \delta(\boldsymbol{x}_i) + \varepsilon_i \tag{1}$$

The joint vector of observed $y_{\tilde{i}} = y(\boldsymbol{x}_{\tilde{i}})$, $\tilde{i} = 1, \dots, N$
and computational outputs $\eta_{ij} = \eta(\boldsymbol{xc}_i, \boldsymbol{tc}_j)$, $i = 1, \dots, M$, $j = 1, \dots, m$:

$$D = [y_{\tilde{1}}, \dots, y_{\tilde{N}}, \eta_{11}, \dots, \eta_{1m}, \eta_{21}, \dots, \eta_{2m}, \eta_{M1}, \dots, \eta_{Mm}]^T; \dim D = N + mM.$$

Both the simulator and the discrepancy terms are modelled as (independent) Gaussian processes with constant means and covariance functions k_η and k_δ respectively:

$$\begin{aligned} \eta &\sim \mathrm{GP}(\mu, k_\eta(\boldsymbol{x}, \boldsymbol{t}; \boldsymbol{x}', \boldsymbol{t}')), \\ \delta &\sim \mathrm{GP}(0, k_\delta(\boldsymbol{x}; \boldsymbol{x}')). \end{aligned}$$

We assume the separability of k_η in $\boldsymbol{x}$ and $\boldsymbol{t}$, such that

$$k_\eta(\boldsymbol{x}, \boldsymbol{t}; \boldsymbol{x}', \boldsymbol{t}') = k_x(\boldsymbol{x}; \boldsymbol{x}') \times k_t(\boldsymbol{t}; \boldsymbol{t}').$$

The covariance matrix of the joint vector D is:

$$\Sigma_D = \Sigma_\eta + \begin{pmatrix} \Sigma_\delta + \Sigma_\varepsilon & 0 \\ 0 & 0 \end{pmatrix}, \tag{2}$$

where Σ_η is an $(N+Mm)\times(N+Mm)$ matrix, with each element being the k_η function evaluated for the pairs of $(\boldsymbol{x},\boldsymbol{t})$ across both data points $(\boldsymbol{x}_{\tilde{i}},\boldsymbol{tf})$ (first N components) and computational model inputs $(\boldsymbol{x}_i,\boldsymbol{t}_j)$ (last Mm components). By $\boldsymbol{tf}$ here we denote the "current" value of $\boldsymbol{\theta}$ at a particular computational/evaluation step.
Σ_δ and Σ_ε are $N\times N$ covariance matrices for the discrepancy and error terms, evaluated at data controllable inputs $\boldsymbol{x}_{\tilde{i}}$.

Let's write out the total covariance matrix, one block at a time.

I. The $[1:N]\times[1:N]$ block of $\boldsymbol{\Sigma_D}$:

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{D}}^{(11)} = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{\mathbf{(11)}} + \boldsymbol{\Sigma_\delta} + \boldsymbol{\Sigma_\varepsilon} = &\Big\{k_x(\boldsymbol{x}_{\tilde{i}};\boldsymbol{x}'_{\tilde{j}})\Big\}_{\tilde{i},\tilde{j}=1}^{N} + \Big\{k_\delta(\boldsymbol{x}_{\tilde{i}};\boldsymbol{x}'_{\tilde{j}})\Big\}_{\tilde{i},\tilde{j}=1}^{N} + \sigma^2\boldsymbol{I}_N\\ =&\boldsymbol{K}^{\boldsymbol{x}}_{\tilde{\boldsymbol{N}}\tilde{\boldsymbol{N}}} + \boldsymbol{K}^{\boldsymbol{\delta}}_{\tilde{\boldsymbol{N}}\tilde{\boldsymbol{N}}} + \boldsymbol{K}^{\boldsymbol{\varepsilon}}_{\tilde{\boldsymbol{N}}\tilde{\boldsymbol{N}}}\end{aligned}$$

II. The $[1:N]\times[N+1:N+Mm]$ block of $\boldsymbol{\Sigma_D}$: $\boldsymbol{\Sigma}_{\boldsymbol{D}}^{(12)} = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{(12)}$

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{D}}^{(12)} = &\begin{bmatrix} k_x(\boldsymbol{x}_{\tilde{1}},\boldsymbol{xc}_1) & k_x(\boldsymbol{x}_{\tilde{1}},\boldsymbol{xc}_2) & \dots & k(\boldsymbol{x}_{\tilde{1}},\boldsymbol{xc}_M)\\ k_x(\boldsymbol{x}_{\tilde{2}},\boldsymbol{xc}_1) & k_x(\boldsymbol{x}_{\tilde{2}},\boldsymbol{xc}_2) & \dots & k(\boldsymbol{x}_{\tilde{2}},\boldsymbol{xc}_M)\\ \vdots & \vdots & \vdots & \vdots\\ k_x(\boldsymbol{x}_{\tilde{N}},\boldsymbol{xc}_1) & k_x(\boldsymbol{x}_{\tilde{N}},\boldsymbol{xc}_2) & \dots & k(\boldsymbol{x}_{\tilde{N}},\boldsymbol{xc}_M)\end{bmatrix} \otimes \begin{bmatrix}k_t(\boldsymbol{tc}_1,\boldsymbol{tf}), & k_t(\boldsymbol{tc}_2,\boldsymbol{tf}), & \dots, & k_t(\boldsymbol{tc}_m,\boldsymbol{tf})\end{bmatrix}\\ =&\boldsymbol{K}^{\boldsymbol{x}}_{\tilde{\boldsymbol{N}}\boldsymbol{M}} \otimes \boldsymbol{K}^{\boldsymbol{t}}_{\boldsymbol{1m}}\end{aligned}$$

III. The $[N+1:N+Mm]\times[1:N]$ block of $\boldsymbol{\Sigma_D}$: $\boldsymbol{\Sigma}_{\boldsymbol{D}}^{(21)} = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{(21)} = [\boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{(12)}]^T$

$$\boldsymbol{\Sigma}_{\boldsymbol{D}}^{(21)} = [K^x_{\tilde{N}M}\otimes K^t_{1m}]^T = [K^x_{\tilde{N}M}]^T \otimes [K^t_{1m}]^T = \boldsymbol{K}^{\boldsymbol{x}}_{\boldsymbol{M}\tilde{\boldsymbol{N}}} \otimes \boldsymbol{K}^{\boldsymbol{t}}_{\boldsymbol{m1}}$$

IV. The $[N+1:N+Mm]\times[N+1:N+Mm]$ block of $\boldsymbol{\Sigma_D}$:

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{D}}^{(22)} = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{(22)} = &\begin{bmatrix} k_x(\boldsymbol{xc}_1,\boldsymbol{xc}_1) & k_x(\boldsymbol{xc}_1,\boldsymbol{xc}_2) & \dots & k_x(\boldsymbol{xc}_1,\boldsymbol{xc}_M)\\ k_x(\boldsymbol{xc}_2,\boldsymbol{xc}_1) & k_x(\boldsymbol{xc}_2,\boldsymbol{xc}_2) & \dots & k_x(\boldsymbol{xc}_2,\boldsymbol{xc}_M)\\ \vdots & \vdots & \vdots & \vdots\\ k_x(\boldsymbol{xc}_M,\boldsymbol{xc}_1) & k_x(\boldsymbol{xc}_M,\boldsymbol{xc}_2) & \dots & k_x(\boldsymbol{xc}_M,\boldsymbol{xc}_M)\end{bmatrix} \otimes\\ &\begin{bmatrix} k_x(\boldsymbol{tc}_1,\boldsymbol{tc}_1) & k_x(\boldsymbol{tc}_1,\boldsymbol{tc}_2) & \dots & k_x(\boldsymbol{tc}_1,\boldsymbol{tc}_m)\\ k_x(\boldsymbol{tc}_2,\boldsymbol{tc}_1) & k_x(\boldsymbol{tc}_2,\boldsymbol{tc}_2) & \dots & k_x(\boldsymbol{tc}_2,\boldsymbol{tc}_m)\\ \vdots & \vdots & \vdots & \vdots\\ k_x(\boldsymbol{tc}_m,\boldsymbol{tc}_1) & k_x(\boldsymbol{tc}_m,\boldsymbol{tc}_2) & \dots & k_x(\boldsymbol{tc}_m,\boldsymbol{tc}_m)\end{bmatrix}\\ =&\boldsymbol{K}^{\boldsymbol{x}}_{\boldsymbol{MM}} \otimes \boldsymbol{K}^{\boldsymbol{t}}_{\boldsymbol{mm}}\end{aligned}$$

The final block-representation of the total covariance matrix:

$$\boldsymbol{\Sigma_D} = \begin{bmatrix} \boldsymbol{K}^{\boldsymbol{x}}_{\tilde{\boldsymbol{N}}\tilde{\boldsymbol{N}}} + \boldsymbol{K}^{\boldsymbol{\delta}}_{\tilde{\boldsymbol{N}}\tilde{\boldsymbol{N}}} + \boldsymbol{K}^{\boldsymbol{\varepsilon}}_{\tilde{\boldsymbol{N}}\tilde{\boldsymbol{N}}} & \boldsymbol{K}^{\boldsymbol{x}}_{\tilde{\boldsymbol{N}}\boldsymbol{M}} \otimes \boldsymbol{K}^{\boldsymbol{t}}_{\boldsymbol{1m}}\\ \boldsymbol{K}^{\boldsymbol{x}}_{\boldsymbol{M}\tilde{\boldsymbol{N}}} \otimes \boldsymbol{K}^{\boldsymbol{t}}_{\boldsymbol{m1}} & \boldsymbol{K}^{\boldsymbol{x}}_{\boldsymbol{MM}} \otimes \boldsymbol{K}^{\boldsymbol{t}}_{\boldsymbol{mm}}\end{bmatrix} \tag{3}$$

Blockwise inversion:

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \begin{bmatrix} (\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} & -(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \\ -\boldsymbol{D}^{-1}\boldsymbol{B}(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} & (\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1} \end{bmatrix}$$

Note: $\boldsymbol{\Sigma_D}$ is symmetrical, $\boldsymbol{B} = \boldsymbol{C}^T$ and the corresponding blocks in $\boldsymbol{\Sigma_D^{-1}}$ are also symmetrical.
1. The $(1,1)$, $N \times N$ block of the inverse matrix is

$$\begin{aligned} \boldsymbol{\Sigma_D^{-1}}[1,1] &= \left(\boldsymbol{\Sigma_D^{(11)}} - \boldsymbol{\Sigma_D^{(21)T}} \left[\boldsymbol{\Sigma_D^{(22)}}\right]^{-1} \boldsymbol{\Sigma_D^{(21)}}\right)^{-1} \\ &= (\boldsymbol{Q_1^T Q_1})^{-1} = \boldsymbol{Q_1^{-1} Q_1^{-T}}, \end{aligned}$$

where $\boldsymbol{Q_1}$ is the Cholesky decomposition of the matrix to be inverted here – in a similar way as matrix D in eq. (3.6a) in Rougier, (2008).
2. The $(1,2)$, $N \times Mm$ block of the inverse matrix (and the transposed $(2,1)$-block) is

$$\begin{aligned} \boldsymbol{\Sigma_D^{-1}}[1,2] &= -\boldsymbol{Q_1^{-1} Q_1^{-T}} \left[\boldsymbol{K^{x}_{\tilde{N}M}} \otimes \boldsymbol{K^{t}_{1m}}\right] \left[\boldsymbol{K^{x}_{MM}} \otimes \boldsymbol{K^{t}_{mm}}\right]^{-1} \\ &= -\boldsymbol{Q_1^{-1} Q_1^{-T}} \left[\boldsymbol{K^{x}_{\tilde{N}M} K^{-x}_{MM}} \otimes \boldsymbol{K^{t}_{1m} K^{-t}_{mm}}\right] \\ &= -\boldsymbol{Q_1^{-1} Q_1^{-T} R}. \end{aligned}$$

3. The $(2,2)$, $Mm \times Mm$ block

$$\begin{aligned} \boldsymbol{\Sigma_D^{-1}}[2,2] &= \left(\boldsymbol{\Sigma_D^{(22)}} - \boldsymbol{\Sigma_D^{(12)T}} \left[\boldsymbol{\Sigma_D^{(11)}}\right]^{-1} \boldsymbol{\Sigma_D^{(12)}}\right)^{-1} \\ &= (\boldsymbol{Q_2^T Q_2})^{-1} = \boldsymbol{Q_2^{-1} Q_2^{-T}} \end{aligned}$$

The whole matrix:

$$\boldsymbol{\Sigma_D^{-1}} = \begin{bmatrix} \boldsymbol{Q_1^{-1} Q_1^{-T}} & -\boldsymbol{Q_1^{-1} Q_1^{-T} R} \\ -\boldsymbol{R^T Q_1^{-1} Q_1^{-T}} & \boldsymbol{Q_2^{-1} Q_2^{-T}} \end{bmatrix}$$