2. Extrapolation with model outside calibration domain

\[ Y_i = f(x_i, q) + \delta(x_i, q_{\text{dis}}) + \epsilon_i \]

\( \delta \) ensures that \( q \) is estimated in a statistically consistent manner in the calibration domain, may be inaccurate outside domain without restrictions on \( \delta \) or prior information.

Computed model \( f(x_i, q) + \delta(x_i, q_{\text{dis}}) \)

is accurate inside calibration domain but

has minimal predictive capability outside this domain ("overfitting")

Sufficient prior structure on \( \delta \)

Surrogate models (§1.3)

Construct representations quantifying primary features of the high-fidelity model while being computationally efficient for:

- Bayesian model calibration, uncertainty quantification, design, optimization...

Regression or interpolation-based models

Inter-fit models, response surface models, emulators, meta-models, approximation models

Model \( y = f(q), q \in \mathbb{R}^p \)
draw samples to construct input-output relations (based on interpretation or regression theory)

...non-invasive method, can use software for large scale applications to generate data...

\[ M \text{ realizations: } y_m = f(q^m), \quad m = 1 \ldots M \]

Choose \( q^m \) e.g. by Monte Carlo, critical approx \( f(q) \) is treated as black box

Construct emulator \( f(q) \) approximating \( f(q) \)

Approximate fine scale behavior by statistical model

\[ y_m = f(q^m) + \varepsilon_m, \quad m = 1 \ldots M \]

random variable

\( y_m \) realizations

\( \varepsilon_m \) i.i.d., \( \sim N(0, \sigma^2) \)

**Quadratic response surface model**

\[ f(q, \beta) = \beta_0 + \sum_{i=1}^{P} \beta_i q_i + \sum_{i=1}^{P} \sum_{j=i}^{P} \beta_{ij} q_i q_j \]

\( P = \frac{(P+1)(P+2)}{2} \) coefficients \( \beta_0, \beta_i, \beta_{ij} \)

Need \( M > P \) samples

\[ X = \begin{bmatrix} q_1 & q_1 \cdot q_1 & (q_1)^2 & q_1 \cdot q_2 & \cdots & q_1 \cdot q_P \\ q_1^2 & q_1 \cdot (q_1)^2 & (q_1)^2 & q_1 q_2 & \cdots & q_1 q_P \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ q_P & q_P \cdot q_1 & (q_P)^2 & q_P q_2 & \cdots & q_P q_P \end{bmatrix} \in \mathbb{R}^{M \times P} \]
Least squares approximation:

\[ \beta = (X^TX)^{-1}X^T y_s \]

Suitable for optimization problems: analytic solutions of optimum

Kriging model:

\[ \hat{f}(q, \beta) = g^T(q) \beta + Z(q) \]

Trend function \rightarrow Gaussian process \rightarrow kriging model

Ordinary kriging: \( g^T(q) \beta = \beta_0 \)

At sample points: \( y_m = f(q^m, \beta_0), \ y_s = (y_1^s, \ldots, y_n^s) \)

\( Z \) stationary random process: \( E(Z) = 0 \)

\( E(Z^2) = \sigma^2 \)

\( \text{cov}(Z(q^i), Z(q^j)) = \sigma^2 R(q^i, q^j) + \sigma_0^2 \delta(q^i - q^j) \)

where

\( \text{minimum function} \)

\( \delta(q^i - q^j) = \begin{cases} 1 & q^i = q^j \\ 0 & \text{otherwise} \end{cases} \)

\( R(q^i, q^j) = \exp \left( -\sum_{k=1}^{p} \Theta_k |q^{i_k} - q^{j_k}|^{\phi_k} \right) \)

\( 0 < \phi_k < 2, \ \Theta_k > 0 \)

\[ \hat{f}(q, \beta_0) = \beta_0 + r^T(q) R^{-1}(y_s - \beta_0 \mathbf{1}) \]

\( R_{ij} = R(q^i, q^j), \ r(q) = R(q, q) \in \mathbb{R}^M \)

\( \mathbf{1} = (1, \ldots, 1)^T \in \mathbb{R}^M \)

\[ \beta_0(\theta, y) = (\mathbf{1}^T R^{-1} \mathbf{1})^{-1} \mathbf{1}^T R^{-1} y_s \]

Least squares estimate for \( \beta_0 \)
Minimize $Y = c^TY_S$, $c = (c_1, c_2, ..., c_m)$

$c = \text{min arg } E(c^TY_S - Y)^2$  \text{(*)}

Constraint $E(c^T Y_S) = E(Y)$ \text{(**) $\sigma^2$}

\[(*) \Rightarrow E \left( \sum_{i=1}^{M} \sum_{j=1}^{M} c_i y_{S_i} y_{S_j} c_j - 2 \sum_{i=1}^{M} c_i y_{S_i} y_{Y} y_{Y} \right) = \sigma^2 (1 + c^T R c - 2c^T r)
\text{(**) $\Rightarrow$ E(\sum_{i=1}^{M} c_i y_{S_i}) = E(Y) \Rightarrow \sum_{i=1}^{M} c_i = 1$}

\[\text{min } c^T R c - 2c^T r \quad c^T c = 1\]

With Lagrange multiplier $\lambda$

\[\begin{pmatrix} 0 & 1^T \\ 1 & R \end{pmatrix} \begin{pmatrix} \lambda \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\]

With solution

\[\hat{y}(q, \beta_0) = \hat{y} = \beta_0 + r^T R^{-1}(y_S - \beta_0 1)\]

\[\beta_0 = (1^T R^{-1} 1)^{-1} 1^T R^{-1} y_S\]

At $q_j^i: v_i(q_j^i) = R(q_i^j, q_j^i)$, $r^T R^{-1} = (0 ... 1 ... 0) = e_j^T$

\[\Rightarrow \hat{y}_j = \beta_0 + c_j^T (y_S - \beta_0 1) = y_{e_j} = y_j\]
Radial basis functions

\[ f(q; \beta_0) = \sum_{m=1}^{M} f_m \psi_m(q) + \beta_0 = \Psi^T q + \beta_0 \]

\[ \psi_m(q) = \psi(11q^m - q_11) = \psi(r^m), \quad m = 1 \ldots M \]

\[ \psi(r^m) = \begin{cases} 
  e^{-r^m / 2\sigma^2}, & n = 1, 2, 3 \\
  r^m, & n = 1, 2, 3 \\
  r^m \ln r_0 & 
\end{cases} \]

Interpolation condition: \( f(q^m; \beta_0) = y_m, \quad m = 1 \ldots M \)

\[ \sum_{m=1}^{M} f_m = 0 \quad \beta_0 \) takes the constant part \]

\[ \sum_{m=1}^{M} \sum_{k=1}^{M} f_k \psi_k(q^m) = y_s - \beta_0 \quad \phi_{mk} \sum_{m=1}^{M} \sum_{k=1}^{M} f_k \psi_k(q^m) = \phi y_s - \beta_0 \]

\[ \phi f = y_s - \beta_0 \]

\[ f(q; \beta_0) = \beta_0 + \psi^T(q) \phi^{-1}(y_s - \beta_0) \]

Compare with kriging model

\[ r_i(q) = R(q^i, q^i) \leftrightarrow \psi_i(q) = \psi(11q^i - q_11) \]

\[ R_{ij} = R(q^i, q^j) \leftrightarrow \phi_{ij} = \psi_j(q^i) = \psi(11q^i - q_11) \]

\[ = \exp(-\sum \theta_k |q^i - q^j|_{k}) \]

\[ \sum_{m=1}^{M} f_m = \phi^{-1}(y_s - \beta_0) = 0 \Rightarrow \beta_0 = (\phi^T \phi)^{-1} \phi^T y_s \]
Evolutionary PDE
\[ \frac{\partial u}{\partial t} = N(u, q) + F(q), \quad x \in D, \quad t > 0 \]

Weak form
\[ \int_D \frac{\partial u}{\partial t} v \, dx + \int_D N(u, q) S(v) \, dx = \int_D F(q) v \, dx \quad \forall v \in V \]

Surrogate model
\[ u(t, x, q) = \sum_{k=0}^\infty \sum_{j=1}^J U_{jk}(t) \phi_j(x) \psi_k(q) \]

\[ \phi_j(x) \quad \text{FEM basis functions} \]
\[ \psi_k(q) = L_k(q) \quad \text{Lagrange polynomials} \]
\[ \psi_k(q^m) = \delta_{km} \]

Insert (2) into (1) at \( q = q^m \) \( \Rightarrow \) equations for \( u_{jk} \)

Parameter space is discretized

Projection-based methods

\[ V^J = \text{span} \{ \phi_j \} \subset V \quad \text{in (1)} \]

\[ u^J(t, x, q) = \sum_{j=1}^J U_{jk}(t) \phi_j(x) \]

\[ \int_D \frac{\partial u^J}{\partial t} \phi_k \, dx + \int_D N(u^J, q) S(\phi_k) \, dx = \int_D F(q) \phi_k \, dx \quad l = 1, \ldots, J \]

Reduce approximation space

\[ V^{J_r} = \text{span} \{ \phi_{j_r} \} \subset V, \quad J_r \ll J \]
with approximation
\[ \hat{u}(t, x, q) = \hat{u}^j(t, x, q) = \sum_{j=1}^{J} u_j(t) \phi_j^r(x) \]

(1) \[ \frac{\partial u^j}{\partial t} + \frac{\partial}{\partial x} \left( \frac{1}{D} F(q) \phi_j^r \right) = \int_{D} N(u^j, q) S(\phi_j^r) dx = \frac{1}{D} F(q) \phi_j^r dx \]

How are \( \phi_j^r \) chosen?

**Eigenfunctions or modal expansions**

analytic or numerical eigenfunctions in space

**Separation of variables**

\[
\begin{align*}
\frac{\partial T}{\partial x} &= -\alpha \frac{\partial^2 T}{\partial x^2} \quad 0 < x < L, \quad t > 0 \\
T(t, 0) &= T(t, L) = 0, \quad T(0, x) = T_0(x) \\
T(t, x) &= \sum_{j=1}^{J} u_j(t) \phi_j(x) \\
\phi_j(x) &= \sin \frac{j \pi x}{L}, \quad \text{few basis functions needed} \\
\end{align*}
\]

**Snapshot-based methods**

**Snapshot set:** solutions generated at different independent variables (e.g.) or parameter values

\[ u(t_m, x, q), \quad u(t, x, q^m) \]

sampling techniques must generate basis that is sufficiently rich to incorporate all expected system dynamics

**Monte Carlo sampling for** \( q^m \)}
Proper orthogonal decomposition (POD)
numerical solutions \( \{ u_m(x) \}_{m=1}^M, \) all \( x \in \mathcal{D} \)
e.g. solutions at \( t_m \)

Consider deviations from the mean \( \bar{u} \)

\[
\bar{u} = \frac{1}{M} \sum_{m=1}^{M} u_m(x), \quad v_m = u_m - \bar{u}
\]

POD is algorithm to compress information in \( v_m \)

Construct basis functions \( \phi(x) = \sum_{m=1}^{M} a_m v_m(x) \)

\( a_m \) maximize \( \frac{1}{M} \sum_{m=1}^{M} |\langle v_m, \phi \rangle|^2 \) \( \quad (3) \)
such that \( \langle \phi, \phi \rangle = \| \phi \|^2 = 1 \)

\[
\langle f, g \rangle = \int_{\mathcal{D}} f(x)g(x)dx
\]

\[
C(x, y) = \frac{1}{M} \sum_{m=1}^{M} v_m(x)v_m(y)
\]

\[
R\phi = \int_{\mathcal{D}} C(x, y)\phi(y)dy
\]

\[
\Rightarrow \langle R\phi, \phi \rangle = \frac{1}{M} \sum_{m=1}^{M} |\langle v_m, \phi \rangle|^2 \quad (4)
\]

\[
\langle R\phi, \psi \rangle = \langle \phi, R\psi \rangle, \text{ } R \text{ is symmetric}
\]

Max of (3) or (4) is given by eigenvector \( \phi \) with
max eigenvalue \( \lambda \)

\[
R\phi = \lambda \phi \text{ (scaled) such that } \| \phi \| = 1
\]
Insert \( \phi(y) = \sum_{k=1}^{M} a_k \psi_k(y) \) into

\[
\int_{D} C(x,y) \phi(y) \, dy = \lambda \phi(x)
\]

\[
\Rightarrow \frac{1}{M} \sum_{m=1}^{M} \left( \sum_{k=1}^{D} \psi_k(y) \psi_m(y) \, dy \right) a_k \psi_m(x) = \sum_{m=1}^{M} \lambda a_m \psi_m(x)
\]

Let \( K_{mk} = \frac{1}{M} \int_{D} \psi_m(y) \psi_k(y) \, dy \)

\[
a = (a_1, a_2, \ldots, a_M)^T
\]

\[
\Rightarrow Ka = \lambda a \quad K = K^T \quad \text{pos. def}
\]

Orthogonal eigenvectors \( a_1, a_2, \ldots, a_M \)

eigenvalues \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M \geq 0 \)

Maximum of (3) is \( \phi_1(x) = \sum_{m=1}^{M} a_m^1 \psi_m(x) \)

\[
\phi_j(x) = \sum_{m=1}^{M} a_m^j \psi_m(x), \quad j = 2, \ldots, M
\]

Orthonormal basis functions

\[
\phi_j(x) = \sum_{m=1}^{M} \frac{a_m^j}{\sqrt{\lambda_m}} \psi_m(x), \quad j = 1, \ldots, M
\]

Compare with Karhunen-Loève expansion for a random field \( \alpha(x, w) \)
POD with discrete observations

Snapshots $u_m \in \mathbb{R}^n, M \leq n$

$$v_m = u_m - \bar{u}$$

Snapshot matrix $A = (v_1, \ldots, v_M) \in \mathbb{R}^{n \times M}$

$$K = \frac{1}{M} A^T A \in \mathbb{R}^{M \times M}, \text{ rank } A = r = M$$

$$K_{mj} = \frac{1}{M} v_m^T v_j$$

POD basis $\{\phi_j\}_{j=1}^r$ by eigenvalue problem

$$K \phi_j = \lambda_j \phi_j \Rightarrow \phi_j = \frac{1}{\sqrt{\lambda_j}} A v_j$$

Relation to SVD:

$$A = U \Sigma V^T, U \in \mathbb{R}^{n \times n}, V \in \mathbb{R}^{M \times M}$$

$U, V$ orthogonal

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, D = \text{diag}(\sigma_1, \ldots, \sigma_r) \in \mathbb{R}^{n \times n}$$

here $r = M$

$$K = \frac{1}{M} A^T A = \frac{1}{M} V \Sigma^2 V^T, V = (\tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_M)$$

$$\Rightarrow K \tilde{v}_j = \frac{1}{M} \sigma_j^2 \tilde{v}_j \Rightarrow \lambda_j = \frac{1}{M} \sigma_j^2$$

POD basis in $\tilde{u}$ $\phi_j = \frac{1}{\sqrt{\lambda_j}} A \tilde{v}_j = \frac{\sigma_j}{\sqrt{\lambda_j}} \tilde{v}_j$

Reduced representation $\tilde{u} = \sum_{j=1}^r u_j \phi_j$
High-dimensional model representation (HDMR)

HDMR or Sobol representation of \( f \) in \( y = f(q) \)

\[
f(q) = f_0 + \sum_{i=1}^{p} f_i(q_i) + \sum_{1 \leq i < j \leq p} f_{ij}(q_i, q_j) + \ldots
\]

\[
+ \sum_{1 \leq i_1 \leq i_2 \leq \ldots \leq i_p \leq n} f_{i_1 \ldots i_p}(q_{i_1}, \ldots, q_{i_p}) + \ldots + f_{123\ldots p}(q_{i_1}, \ldots, q_{i_p})
\]

For mean response, \( q \in [0,1]^p \)

\( \Rightarrow \) exact representation

In practice, approximate expansion

\[
f(q) \approx f_0 + \sum_{i=1}^{p} f_i(q_i) + \sum_{1 \leq i < j \leq p} f_{ij}(q_i, q_j)
\]

**ANOVA-HDMR**

\[
f_0 = \frac{1}{n} \int_{\Gamma} f(q) \, dq , \quad \Gamma = [0,1]^p
\]

\[
f_i(q_i) = \frac{1}{n} \int_{\Gamma^{p-1}} f(q) \, dq_{\bar{i}} - f_0
\]

\[
\Gamma^{p-1} = [0,1]^{p-1} , \quad dq_{\bar{i}} = dq_1 dq_2 \ldots dq_{i-1} dq_{i+1} \ldots dq_p
\]

\[
f_{ij}(q_i, q_j) = \frac{1}{n} \int_{\Gamma^{(p-2)}} f(q) \, dq_{\bar{ij}} - f_i(q_i) - f_j(q_j) - f_0
\]

\( \text{En so: integration over } \Gamma \text{ numerically or MC} \)

**Random sampling** **RS-HDMR**
\[
fo = \frac{1}{R} \sum_{r=1}^{R} f(q^r) \quad \text{for } q^r \sim U(0,1)^p
\]

Exponential growth in sample points to determine \(f_i(q_i), f_{ij}(q_i,q_j), \ldots\)

Assume \(f_i(q_i) = \sum_{k=1}^{K} a_k \phi_k(q_i)\)

\[
f_{ij}(q_i,q_j) = \sum_{k=1}^{K} \sum_{l=1}^{K} B_{ij}^{kl} \phi_k(q_i) \phi_l(q_j)
\]

Tensor product basis functions: \(\phi_{kl}(q_i,q_j) = \phi_k(q_i) \phi_l(q_j)\)

\(\phi_k\) are Legendre polynomials shifted to \([0,1]\)

Orthogonality of \(\phi_k \Rightarrow a_k^2 \approx \frac{1}{R} \sum_{r=1}^{R} f(q^r) \phi_k(q^r)\)

\[
a_k \approx \frac{1}{2k+1}
\]

Surrogate-based Bayesian model calibration

Likelihood function

\[
\Pi(y|q) = \frac{1}{(2\pi \sigma^2)^{n/2}} e^{-\frac{1}{2} \sum_{i=1}^{n} (y_i - \hat{f}_i(q))^2}
\]

Errors \(\varepsilon_i\) iid, \(\varepsilon_i \sim N(0,\sigma^2)\)

\[
SSQ = \sum_{i=1}^{n} (y_i - \hat{f}_i(q))^2
\]

Replace \(\hat{f}_i(q)\) with surrogate evaluation \(\hat{f}_i(q)\)
\[ \tilde{\pi}(y|q) = \frac{1}{(2\pi \sigma^2)^{n/2}} e^{-\tilde{SS}_q / 2\sigma^2} \]

\[ \tilde{SS}_q = \sum_{i=1}^{n} (y_i - \tilde{f}_i(q))^2 \]

Exercises: 13.1, 13.2

Local sensitivity analysis (§14)

Quantify contributions of parameters and determine effect of variations in parameters.

* Is model robust or sensitive to various parameters?
* Can model be simplified by fixing insensitive parameters?
* Find regimes in parameter space with optimal impact on output and its uncertainties.
* Guide experimental design to measurement regimes that have the greatest impact on parameter sensitivity.

Local sensitivity analysis focuses on variability of response when parameters are perturbed about a nominal value.

Construction of local sensitivities:
1. finite difference approximation
2. solution of sensitivity equations
3. automatic differentiation (AD)
\[
\frac{dy}{dt} = f(y, t, q)
\]

What is \( \frac{dy}{dq} \)?

1. Finite difference approximation
\[
\frac{dy}{dq_k} \approx \frac{y(t, q+\Delta q) - y(t, q)}{\Delta q}
\]
where \( \Delta q_k = \frac{h_i}{h} \), \( h_k = h(\frac{i}{2}) \).

2. \( \frac{\partial \frac{dy}{dq}}{\partial q} = \frac{\partial (y(q))}{\partial q} = \frac{\partial f(y(q))}{\partial q} + \frac{\partial f(y(q))}{\partial y(q)} \frac{dy}{dq} \)

\[
\Rightarrow \frac{\partial \frac{dy}{dq}}{\partial q} = \frac{\partial f(y(q))}{\partial q} + \frac{\partial f(y(q))}{\partial y(q)} \frac{dy}{dq}
\]

Solve for \( \frac{dy}{dq} \).

3. AD can compute \( \frac{\partial \bar{G}}{\partial q} \) where \( \bar{G}(q) \) is computed in subroutine

Forward sensitivity analysis (FSAP)

Matrix system, solution \( \phi \), observation \( y \), mean parameter values \( q \)

\[(\forall) \quad A(q) \phi = s(q), \quad y = C(q) \phi = C(q) A^{-1}(q) s(q) \in \mathbb{R}^p
\]

nominal values \( A = A(q), \ C = C(q), \ s = s(q) \)

If \( y \in \mathbb{R}^p \) then \( C \in \mathbb{R}^{(N-1) \times p} \)

work to solve (\( \forall \)) the expensive part

for \( p \) changes of \( q \): (\( \forall \)) is solved \( p \) times
\[ \delta y = C^T \delta \phi + \delta C^T \delta \bar{\phi} \quad (1st \ variation) \]

\[(*) \quad A \delta \phi + S A \delta \phi = \delta s \]

To compute \( \delta y \):

\[ V \phi = \delta \phi \Rightarrow \delta \phi : A \delta \phi = \delta s - S A \delta \phi \Rightarrow \delta \phi = A^{-1} (\delta s - S A \delta \phi) \]

\[ \delta y = C^T A^{-1} (\delta s - S A \delta \phi) + \delta C^T A^{-1} \delta \phi \]

Perturb \( \psi \) by \( \delta \psi \Rightarrow S A, S C, \delta s \Rightarrow \delta y \)

Repeated solution of \((*)\) for large \( A \) and different \( \delta q \) may be expensive.

**Adjoint sensitivity analysis procedure (ASAP)**

Using perturbations

\[(*) \quad A^T \psi = C \quad (adjoint \ sensitivity \ eq.) \]

Multiply \((*)\) by \( \psi^T \)

\[ \psi^T A \delta \phi = \psi^T (\delta s - S A \delta \phi) \Leftrightarrow \langle \psi, \delta s - S A \delta \phi \rangle = \langle \psi, A \delta \phi \rangle \]

\[ \langle \psi, A \delta \phi \rangle = \langle A^T \psi, \delta \phi \rangle = \langle C, \delta \phi \rangle \]

\[ \Rightarrow \delta y = \delta C^T \phi + \psi^T (\delta s - S A \delta \phi) \quad (\&) \]

Solve \((*)\) once for \( \bar{\psi} \), then compute \( \delta y \) from \((\&)\)
With $C \in \mathbb{R}^{(N-1)\times 1}$, $\Rightarrow \nu$ right hand sides
Solve ($\#) \nu$ times
ASAP more efficient than FSAP when number of parameters $p$ exceeds the number of responses $\nu$