2. Extrapolation with model outside calibration domain

$$Y_i = f(x_i, q) + \delta(x_i, q_{dis}) + \varepsilon_i$$

δ ensures that q is estimated in a statistically consistent manner in the calibration domain, may be inaccurate outside domain without restrictions on δ or prior information.
Combined model $f(x_i, q) + \delta(x_i, q_{dis})$ is accurate inside calibration domain but has minimal predictive capability outside this domain ("overfitting"), insufficient prior structure on δ

Surrogate models (§13)

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

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

Model $\quad y = f(q), \quad q \in \Gamma \subset \mathbb{R}^p$

draw samples to construct input-output relations based on interpolation or regression theory

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

M realizations : $y_m = f(q^m)$, $m = 1 \cdots M$
Choose q^m e.g. by Monte Carlo, critical for approx.
$f(q)$ is treated as black box

Construct emulator $\tilde{f}(q)$ approximating $f(q)$

Approximate fine scale behavior by statistical model

$$Y_m = \tilde{f}(q^m) + \varepsilon_m, \quad m = 1 \cdots M$$

random variable

y_m realizations
ε_m iid, $\sim N(0, \sigma_0^2)$

Quadratic response surface model

$$\tilde{f}(q, \beta) = \beta_0 + \sum_{i=1}^{p} \beta_i q_i + \sum_{i=1}^{p} \beta_{ii} q_i^2 + \sum_{i=1}^{p} \sum_{j>i}^{p} \beta_{ij} q_i q_j$$

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

Need $M > P$ samples

$$X = \begin{bmatrix} 1 & q_1^1 & \cdots & q_p^1 & (q_1^1)^2 & \cdots & (q_p^1)^2 & q_1^1 q_2^1 & \cdots & q_{p-1}^1 q_p^1 \\ 1 & q_1^2 & & & & & & & & \\ \vdots & & & \vdots & & & & & & \\ 1 & q_1^M & & q_p^M & (q_1^M)^2 & \cdots & (q_p^M)^2 & q_1^M q_2^M & & q_{p-1}^M q_p^M \end{bmatrix}$$

$\in \mathbb{R}^{M \times P}$

Least squares approximation

$$\beta = (X^T X)^{-1} X^T y_s$$

Suitable for optimization problems:
analytic solutions of optimum

Kriging model

$$\tilde{f}(q, \beta) = g^T(q)\beta + Z(q)$$

trend function — Gaussian process error model

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

at sample points: $y_m = \tilde{f}(q^m, \beta_0)$, $y_s = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix}$

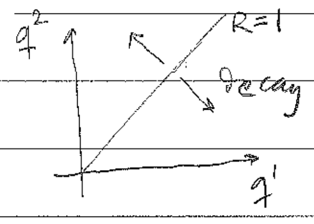
Z stationary random process, $E(Z) = 0$, $E(Z^2) = \sigma^2$

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

correlation function — measurement noise

where

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



$$R(q^i, q^j) = \exp\left(-\sum_{k=1}^{p} \theta_k |q_k^i - q_k^j|^{\gamma_k}\right)$$

$$0 < \gamma_k \leq 2, \quad \theta_k \geq 0$$

$$\tilde{f}(q, \beta_0) = \beta_0 + r^T(q) R^{-1}(y_s - \beta_0 \mathbb{1})$$

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

$\mathbb{1} = (1, 1 \cdots 1)^T \in \mathbb{R}^M$

$$\beta_0(\theta, \gamma) = [\mathbb{1}^T R^{-1} \mathbb{1}]^{-1} \mathbb{1}^T R^{-1} y_s$$

least squares estimate for β_0

6.5

Minimize $Y = c^T Y_s$, $c = (c_1, c_2 \cdots c_m)$

$$c = \text{minarg}\, E\left(c^T Y_s - Y\right)^2 \qquad (*)$$

constraint $E(c^T Y_s) = E(Y)$ $\qquad (\star)$

$$(*) \Rightarrow E\left(\sum_{i=1}^{M}\sum_{j=1}^{M} c_i \overbrace{Y_{s_i} Y_{s_j}}^{\sigma^2 R(q^i, q^j)} c_j - 2\sum_{i=1}^{M} c_i \underbrace{Y_{s_i} Y}_{\sigma^2 R(q, q^i)} + \overset{\sigma^2}{\overset{\downarrow}{Y^2}}\right)$$

$$= \sigma^2\left(1 + c^T R c - 2 c^T r\right)$$

$$(\star) \Rightarrow \quad E\left(\sum_{i=1}^{M} c_i Y_{s_i}\right) = E(Y) \Rightarrow \sum_{i=1}^{M} c_i = 1$$

$$\min_c \; c^T R c - 2 c^T r$$
$$\mathbb{1}^T c = 1$$

With Lagrange multiplier λ

$$\begin{pmatrix} 0 & \mathbb{1}^T \\ \mathbb{1} & R \end{pmatrix}\begin{pmatrix} \lambda \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ r \end{pmatrix}$$

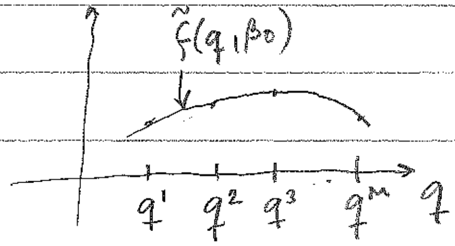
with solution

$$\tilde{f}(q, \beta_0) = \hat{y} = \beta_0 + r^T R^{-1}(y_s - \beta_0 \mathbb{1})$$

$$\beta_0 = (\mathbb{1}^T R^{-1} \mathbb{1})^{-1} \mathbb{1}^T R^{-1} y_s$$

At q^j : $r_i(q^j) = R(q^i, q^j)$, $r^T R^{-1} = (0 \cdots \overset{j}{1} \cdots 0) = e_j^T$

$$\Rightarrow \hat{y}_j = \beta_0 + e_j^T (y_s - \beta_0 \mathbb{1}) = y_{s_j} = y_j$$



Radial basis functions

$$\tilde{f}(q,\beta_0) = \sum_{m=1}^{M} f_m \psi_m(q) + \beta_0 = \psi^T(q) f + \beta_0$$

$$\psi_m(q) = \psi(\|q^m - q\|) = \psi(r^m), \quad m = 1 \cdots M$$

$$\psi(r^m) = \begin{cases} e^{-r_m/2\sigma^2} \\ r_m^n, \; n = 1,2,3 \\ r_m^2 \ln r_m \end{cases}$$

Interpolation condition: $\tilde{f}(q^m, \beta_0) = y_m, \; m = 1 \cdots M$

$$\sum_{m=1}^{M} f_m = 0, \quad \beta_0 \text{ takes the constant part}$$

vector of observations

$$\Rightarrow \sum_{k=1}^{M} f_k \underbrace{\psi_k(q^m)}_{\phi_{mk}} = y_s - \beta_0 \mathbb{1}$$

const. scaling of ψ_k

$$\sum_{k=1}^{M} f_k \overbrace{\sum_{m=1}^{M} \psi_k(q^m)} = \sum_{i=1}^{M} y_i - M\beta_0 = 0$$

$$\sum_{m=1}^{M} \sum_{k=1}^{M} f_k \psi_k(q^m) = \mathbb{1}^T y_s - \beta_0 \underbrace{\mathbb{1}^T \mathbb{1}}_{M} = 0$$

$$\phi f = y_s - \beta_0 \mathbb{1}$$

$$\Rightarrow \tilde{f}(q,\beta_0) = \beta_0 + \psi^T(q) \underbrace{\phi^{-1}(y_s - \beta_0 \mathbb{1})}_{f}$$

Compare with kriging model

$$r_i(q) = R(q^i, q) \longleftrightarrow \psi_i(q) = \psi(\|q^i - q\|)$$

$$R_{ij} = R(q^i, q^j) \longleftrightarrow \phi_{ij} = \psi_j(q^i) = \psi(\|q^i - q^j\|)$$

$$= \exp\left(-\sum \theta_k |q_k^i - q_k^j|^{\gamma_k}\right)$$

$$\sum_{m=1}^{M} f_m = \mathbb{1}^T \phi^{-1}(y_s - \beta_0 \mathbb{1}) = 0 \Rightarrow \beta_0 = (\mathbb{1}^T \phi^{-1} \mathbb{1})^{-1} \mathbb{1}^T \phi^{-1} y_s$$

Evolutionary PDE

$$\frac{\partial u}{\partial t} = \mathcal{N}(u,q) + F(q), \quad x \in D, \; t \geq 0$$

Weak form

$$(1) \quad \int_D \frac{\partial u}{\partial t} v \, dx + \int_D N(u,q) S(v) \, dx = \int_D F(q) v \, dx \qquad v \in V$$

Surrogate model

$$(2) \quad \tilde{u}(t,x,q) = \sum_{k=0}^{K} \sum_{j=1}^{J} u_{jk}(t) \phi_j(x) \psi_k(q)$$

$\phi_j(x)$ FEM basis functions

$\psi_k(q) = L_k(q)$ Lagrange polynomials

$$\psi_k(q^m) = \delta_{km}$$

Insert (2) into (1) at $q = q^r \Rightarrow$ equations for u_{jr}

Parameter space is discretized

Projection based methods

$V^J = \text{span}\{\phi_j\} \subset V$ in (1)

$$u^J(t,x,q) = \sum_{j=1}^{J} u_{jk}(t) \phi_j(x)$$

$$(1) \Rightarrow \int_D \frac{\partial u^J}{\partial t} \phi_\ell \, dx + \int_D N(u^J, q) S(\phi_\ell) \, dx = \int_D F(q) \phi_\ell \, dx$$

$\ell = 1 \ldots J$

Reduce approximation space

$$V^{J_r} = \text{span}\{\phi_j^r\} \subset V, \quad J_r \ll J$$

with approximation

$$\tilde{u}(t,x,q) = u^{J_r}(t,x,q) = \sum_{j=1}^{J_r} u_j(t)\,\phi_j^r(x)$$

$$(1) \Rightarrow \int_D \frac{\partial u^{J_r}}{\partial t}\phi_\ell^r\,dx + \int_D N(u^{J_r},q)\,S(\phi_\ell^r)\,dx = \int_D F(q)\,\phi_\ell^r\,dx$$

How are ϕ_ℓ^r chosen?

<u>Eigenfunctions or modal expansions</u>

analytic or numerical eigenfunctions in space

separation of variables

$$\begin{cases} \dfrac{\partial T}{\partial x} = \alpha \dfrac{\partial^2 T}{\partial x^2} & 0 < x < L,\ t > 0 \\ T(t,0) = T(t,L) = 0, & T(0,x) = T_0(x) \end{cases}$$

$$T(t,x) = \sum u_j(t)\,\phi_j^r(x)$$

$$\phi_j^r(x) = \sin\frac{j\pi x}{L}\,,$$ few basis functions needed

<u>hit</u> 150424

<u>Snapshot-based methods</u>

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

$$u(t_m, x, q),\ u(t, x, q^m)$$

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 $x \in D$

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 $\phi(x) = \sum_{m=1}^{M} a_m v_m(x)$

$$a_m \text{ maximize } \frac{1}{M}\sum_{m=1}^{M} |\langle v_m, \phi\rangle|^2 \qquad (3)$$

such that $\langle \phi, \phi \rangle = \|\phi\|^2 = 1$

$$\langle f, g\rangle = \int_D f g \, dx$$

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

$$R\phi = \int_D C(x,y)\phi(y)\,dy$$

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

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

Max of (3) or (4) is given by eigenvector ϕ with max eigenvalue λ

$R\phi = \lambda\phi$, scaled such that $\|\phi\| = 1$

Insert $\phi(y) = \sum_{k=1}^{M} a_k v_k(y)$ into

$$\int_D C(x,y)\,\phi(y)\,dy = \lambda\,\phi(x)$$

$$\Rightarrow \frac{1}{M}\sum_{m=1}^{M}\left(\sum_{k=1}^{M}\int_D v_k(y)\,v_m(y)\,dy\; a_k\right) v_m(x) = \sum_{m=1}^{M} \lambda\, a_m v_m(x)$$

Let $K_{mk} = \frac{1}{M}\int_D v_m(y)\,v_k(y)\,dy$

$$a = (a_1, a_2 \cdots a_M)^T$$

$$\Rightarrow K a = \lambda a \quad , K = K^T, \text{ pos. def.}$$

Orthogonal eigenvectors $a^1, a^2, \cdots a^M$
eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_M \geq 0$

Maximum of (3) is $\phi_1(x) = \sum_{m=1}^{M} a^1_m v_m(x)$

$$\phi_j(x) = \sum_{m=1}^{M} a^j_m v_m(x), \quad j = 2, \cdots M$$

Orthonormal basis functions

$$\phi_j(x) = \sum_{m=1}^{M} \frac{1}{\sqrt{M\lambda_j}}\, a^j_m v_m(x), \quad j = 1, \cdots M$$

Compare with Karhunen-Loève expansion for a random field $\alpha(x,\omega)$

POD with Discrete observations

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

$v_m = u_m - \bar{u}$

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

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

$K_{mk} = \frac{1}{M} v_m^T v_k$

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

$$K w_j = \lambda_j w_j \Rightarrow \phi_j^r = \frac{1}{\sqrt{M\lambda_j}} A w_j$$

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

$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$, $D = \mathrm{diag}(\sigma_1 \cdots \sigma_r) \in \mathbb{R}^{r \times r}$

here $r = M$

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

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

POD basis is $\phi_j^r = \frac{1}{\sqrt{M\lambda_j}} A \tilde{v}_j = \frac{\sigma_j}{\sqrt{\sigma_j^2}} \tilde{v}_j$

Reduced representation $\tilde{u} = \sum_{j=1}^{J_r} u_j \phi_j^r$

High-dimensional model representation (HDMR)

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

$$f(q)=f_0+\sum_{i=1}^{p} f_i(q_i)+\sum_{1\le i\le j\le p} f_{ij}(q_i,q_j)+\cdots$$

$$+\sum_{1\le i_1\le\cdots\le i_s\le p} f_{i_1\cdots i_s}(q_{i_1},\cdots q_{i_s})+\cdots+\underbrace{f_{123\cdots p}(q_1,\cdots q_p)}_{\text{higher order effects}}$$

f_0 mean response, $q\in[0,1]^p$

$\Rightarrow$ exact representation

In practice, approximate expansion

$$f(q)\approx f_0+\sum_{i=1}^{p} f_i(q_i)+\sum_{1\le i\le j\le p} f_{ij}(q_i,q_j)$$

ANOVA-HDMR

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

$$f_i(q_i)=\int_{\Gamma^{p-1}} f(q)\,dq_{\sim i}-f_0$$

$\Gamma^{p-1}=[0,1]^{p-1}$, $dq_{\sim i}=dq_1dq_2\cdots dq_{i-1}dq_{i+1}\cdots dq_p$

$$f_{ij}(q_i,q_j)=\int_{\Gamma^{p-2}} f(q)\,dq_{\sim\{ij\}}-f_i(q_i)-f_j(q_j)-f_0$$

For f_0: integration over Γ numerically or MC

Random sampling RS-HDMR

$$f_0 = \frac{1}{R}\sum_{r=1}^{R} f(q^r), \quad q^r \sim U(0,1)^p$$

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

Assume
$$f_i(q_i) = \sum_{k=1}^{K} \alpha_k^i \psi_k(q^i)$$

$$f_{ij}(q_i, q_j) = \sum_{k=1}^{K}\sum_{\ell=1}^{L} \beta_{k\ell}^{ij} \psi_{k\ell}(q_i, q_j)$$

tensor product basis functions: $\psi_{k\ell}(q_i, q_j) = \psi_k(q_i)\psi_\ell(q_j)$

ψ_k are Legendre polynomials shifted to $[0,1]$

Orthogonality of $\psi_k \Rightarrow \alpha_k^i \approx \frac{1}{\gamma_k}\frac{1}{R}\sum_{r=1}^{R} f(q^r)\psi_k(q_i^r)$

$$\gamma_k = \frac{1}{2k+1}$$

Surrogate-based Bayesian model calibration

Likelihood function
$$\pi(y|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}$$

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

$$SS_q = \sum_{i=1}^{n}(y_i - f_i(q))^2$$

Replace $f_i(q)$ 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)

$$\frac{dy}{dt} = f(y, q)$$

What is $\frac{\partial y}{\partial q}$?

1. Finite Difference approximation

$$\frac{\partial y}{\partial q_k} \approx \frac{y(t, q+h_k) - y(t,q)}{|h_k|} \qquad h_k = h\begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k$$

2. $\frac{\partial}{\partial q_k}\frac{\partial y}{\partial t} = \frac{\partial}{\partial q_k} f(y,q) = \frac{\partial f}{\partial q_k}(y,q) + \frac{\partial f}{\partial y}(y,q)\frac{\partial y}{\partial q_k}$

$$\Rightarrow \frac{\partial}{\partial t}\left(\frac{\partial y}{\partial q_k}\right) = \frac{\partial f}{\partial q_k}(y,q) + \frac{\partial f}{\partial y}(y,q)\left(\frac{\partial y}{\partial q_k}\right)$$

Solve for $\frac{\partial y}{\partial q_k}$

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

<u>Forward sensitivity analysis (procedure) (FSAP)</u>

Matrix system, solution ϕ, observation y
mean parameter values $\bar{q}$

$$(\nabla) \qquad A(q)\,\phi = s(q)$$
$$y = C^T(q)\,\phi = C^T(q)\,A^{-1}(q)\,s(q) \in \mathbb{R}^{\nu}$$

nominal values $\bar{A} = A(\bar{q})$, $\bar{C} = C(\bar{q})$, $\bar{s} = s(\bar{q})$

If $y \in \mathbb{R}^{\nu}$ then $C \in \mathbb{R}^{(N-1)\times\nu}$
work to solve (∇) the expensive part
for p changes of q : (∇) is solved p times

$$\delta y = \bar{C}^T \delta\phi + \delta C^T \bar{\phi} \qquad \text{(1st variation)}$$

(*) $\quad \bar{A}\delta\phi + \delta A\bar{\phi} = \delta s$

To compute δy:

$\bar{A}\bar{\phi} = \bar{s} \Rightarrow \bar{\phi}$

$\delta\phi$: $\bar{A}\delta\phi = \delta s - \delta A\bar{\phi} \Rightarrow \delta\phi = \bar{A}^{-1}(\delta s - \delta A\bar{\phi})$

$$\Rightarrow \delta y = \underbrace{\bar{C}^T \bar{A}^{-1}}_{\psi^T}(\delta s - \delta A \bar{A}^{-1}\bar{s}) + \delta C^T \underbrace{\bar{A}^{-1}\bar{s}}_{\bar{\phi}}$$

Perturb q by $\delta q \Rightarrow \delta A, \delta C, \delta s \Rightarrow \delta y$

Repeated solution of (*) for large $\bar{A}$ and different δq may be expensive

Adjoint sensitivity analysis procedure (ASAP)

Using perturbations

($\star$) $\quad \bar{A}^T\psi = \bar{C} \quad$ (adjoint sensitivity eq.)

Multiply (*) by ψ^T

$$\psi^T\bar{A}\delta\phi = \psi^T(\delta s - \delta A\bar{\phi}) \Leftrightarrow \langle\psi, \delta s - \delta A\bar{\phi}\rangle = \langle\psi, \bar{A}\delta\phi\rangle$$

$$\langle\psi, \bar{A}\delta\phi\rangle = \langle\bar{A}^T\psi, \delta\phi\rangle = \langle\bar{C}, \delta\phi\rangle$$

$$\Rightarrow \delta y = \delta C^T\bar{\phi} + \psi^T(\delta s - \delta A\bar{\phi}) \qquad (\&)$$

Solve ($\star$) once for $\bar{q}$, then compute δy from (&)

With $\bar{C} \in \mathbb{R}^{(N-1)\times\nu}$ $\Rightarrow$ ν right hand sides

Solve ($\star$) ν times

ASAP more efficient than FSAP when number of parameters p exceeds the number of responses ν