On the Consistency of Calibration Parameter Estimation in Deterministic Computer Experiments

Rui Tuo and C. F. Jeff Wu

Chinese Academy of Sciences & Georgia Institute of Technology

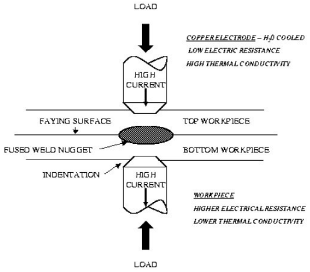
Oct 19, 2012

Calibration Problems

- Consider a computer experiment problem with both computer output and physical response.
  - Physical experiment has control variables.
  - Computer code is deterministic.
  - Computer input involves control variables and calibration parameters.
- Calibration parameters represent inherent attributes of the physical system, which cannot be controlled in physical experiment.
- In many cases, the true value of the calibration parameters cannot be measured physically.
- Kennedy and O'Hagan (2001) describe the calibration problems as:
  - "Calibration is the activity of adjusting the unknown (calibration) parameters until the outputs of the (computer) model fit the *observed* data."

A Spot Welding Example

- Consider a spot welding example from Bayarri et al. (2007). Two sheets of metal are compressed by water-cooled copper electrodes under an applied load.
- Control variables
  - Applied load L
  - Direct current of magnitude C
- Calibration parameter
  - Contact resistance at the faying surface



Notation

- Denote the control variable by x, and the calibration parameter by θ.
- For simplicity, assume that the physical response y^p has *no* random error. Denote the computer output by y^s.
    - y^p is a deterministic function of x and y^s is a deterministic function of (x, θ).
- Calibration problems can be formulated as

$$y^p(x) = y^s(x, \theta_0) + \delta(x), \tag{1}$$

where θ_0 is the true value of the calibration parameter and δ is the *discrepancy* between the physical response and the computer model.

Identifiability

- The true calibration parameter θ_0 is unidentifiable because both θ_0 and δ are unknown.
    - For any given θ, $\epsilon(x,\theta) = y^p(x) - y^s(x,\theta)$ solves equation (1).
    - “A lack of (likelihood) identifiability, ..., persists independently of the prior assumptions and will typically lead to inconsistent estimation in the asymptotic sense.” (Wynn, 2001).
- The identifiability issue is also observed by Bayarri et al. (2007) and Han et al. (2009) etc.

- Let $\epsilon(x,\theta) = y^p(x) - y^s(x,\theta)$.

Definition

Define the L^2 distance projection of θ by

$$\theta^* = \operatorname{argmin}_{\theta\in\Theta} \|\epsilon(\cdot,\theta)\|_{L^2(\Omega)},$$

where Θ is the domain for θ, and Ω is the experimental region.

- We treat θ^* as the "true" calibration parameter and the problem becomes well defined.

Kennedy-O'Hagan Method

- Kennedy and O'Hagan (2001) proposed a modeling framework for calibration problems.
- Main idea:
  - Consider the model

$$y^p(x) = y^s(x, \theta_0) + \delta(x).$$

  - Choose a prior distribution for θ_0.
  - Assume that $y^s(\cdot, \cdot)$ and $\delta(\cdot)$ are independent realizations of Gaussian processes. Then the posterior distribution of θ_0 can be obtained.
- By imposing such a stochastic structure, there is no identifiability problem.

Frequentist Version of Kennedy-O'Hagan Method

- The log-likelihood function is given by

$$l(\theta, \sigma^2) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2}\log|\Phi_{\mathbf{x}}| - \frac{1}{2\sigma^2}\epsilon(\mathbf{x}, \theta)^{\mathrm{T}}\Phi_{\mathbf{x}}^{-1}\epsilon(\mathbf{x}, \theta),$$

  where $\Phi_{\mathbf{x}}$ is the covariance matrix.
- Given $\Phi_{\mathbf{x}}$, the MLE for θ is

$$\hat{\theta} = \operatorname*{argmin}_{\theta\in\Theta} \epsilon(\mathbf{x}, \theta)^{\mathrm{T}}\Phi_{\mathbf{x}}^{-1}\epsilon(\mathbf{x}, \theta).$$

- We refer to this method as the *KO calibration*. This frequentist version is asymptotically equivalent to the Bayesian estimation, provided that the support of the prior distribution for θ is sufficiently wide.

Reproducing Kernel Hilbert Space

- To study the asymptotic behavior of the KO calibration, we need the reproducing kernel Hilbert spaces, also known as the native spaces, as the mathematical tool.
- Suppose Φ is a symmetric positive definite function on Ω. Define the linear space

$$F_\Phi(\Omega) = \{\sum_{i=1}^{N} \alpha_i \Phi(\cdot, x_i) : N \in \mathbb{N}, \alpha_i \in \mathbf{R}, x_i \in \Omega \text{ for } 1 \leq i \leq n\}$$

  and equip this space with the bilinear form

$$\left\langle \sum_{i=1}^{N} \alpha_i \Phi(\cdot, x_i), \sum_{j=1}^{M} \beta_j \Phi(\cdot, y_j) \right\rangle_\Phi := \sum_{i=1}^{N} \sum_{j=1}^{M} \alpha_i \beta_j \Phi(x_i, y_j).$$

- The native Hilbert function space is defined as the closure of $F_\Phi(\Omega)$, denoted as $\mathcal{N}_\Phi(\Omega)$.

Filling distance

- As a property of design, we define the filling distance as

$$h_{\mathbf{x},\Omega} := \sup_{x\in\Omega} \min_{x_i\in\mathbf{x}} \|x - x_i\|.$$

- The design minimizing the filling distance is known as the *minimax distance design*.
- To develop an asymptotic theory, we assume that we have a sequence of designs, denoted by $\mathcal{D}_n$. Let the filling distance of $\mathcal{D}_n$ be h_n.

Limiting Value of KO Calibration

- By the definition of the native norm, we have

$$\epsilon(\mathbf{x},\theta)^{\mathsf{T}}\Phi_{\mathbf{x}}^{-1}\epsilon(\mathbf{x},\theta)=\|\hat{\epsilon}(\cdot,\theta)\|^2_{\mathcal{N}_\Phi(\Omega)},$$

  where $\hat{\epsilon}$ is the interpolate of ϵ given by the Gaussian process model.
- Therefore, under mild conditions we have

Theorem

If there exists a unique θ' such that

$$\|\epsilon(\cdot,\theta')\|_{\mathcal{N}_\Phi(\Omega)}=\inf_{\theta\in\Theta}\|\epsilon(\cdot,\theta)\|_{\mathcal{N}_\Phi(\Omega)}.$$

Then $\hat{\theta}_n\to\theta'$ as $h_n\to 0$.

Comparison between two norms (1)

- In general, the limiting value θ' of the KO calibration differs from the L^2 distance projection θ^* of θ.
- In order to study the difference $\theta' - \theta^*$, let us consider the difference between $\|\cdot\|_{L^2(\Omega)}$ and $\|\cdot\|_{\mathcal{N}_\Phi(\Omega)}$.

Comparison between two norms (2)

- Define the integral operator $\kappa(f) = \int_\Omega \Phi(\cdot, x) f(x) dx$ for $f \in L^2(\Omega)$. Denote the eigenvalues of κ by $\lambda_1 \geq \lambda_2 \geq \cdots$.
- Let f_i be the eigenfunction associated with λ_i and $\|f_i\|_{L^2(\Omega)} = 1$. Then

$$\|f_i\|^2_{\mathcal{N}_\Phi(\Omega)} = \langle f_i, \lambda^{-1} f_i \rangle_{L^2(\Omega)} = \lambda_i^{-1},$$

  where the first equality follows from the identity $\|\kappa(f)\|^2_{\mathcal{N}_\Phi(\Omega)} = \langle f, \kappa(f) \rangle_{L^2(\Omega)}$ for any $f \in L^2(\Omega)$ (Wendland, 2005).
- Since $\lim_{k\to\infty} \lambda_k = 0$, we obtain $\sup_{f \in \mathcal{N}_\Phi(\Omega)} \frac{\|f\|_{\mathcal{N}_\Phi(\Omega)}}{\|f\|_{L^2(\Omega)}} = \infty$.
- There are some functions with very small L^2 norm but their native norm is bounded away from zero. Therefore, the KO calibration can give results that are far from the L^2 projection.

An Illustrative Example (1)

- Let $\Omega = [-1, 1]$ and $\Phi(x_1, x_2) = \exp\{-(x_1 - x_2)^2\}$.
- Consider a calibration problem with a three-level calibration parameter, corresponding to computer codes with discrepancy $\epsilon_1, \epsilon_2, \epsilon_3$.
- Suppose that ϵ_1 and ϵ_2 are the first and second eigenfunctions of κ with L^2 norms $\sqrt{20}$ and $\epsilon_3 = \sin \pi x$.

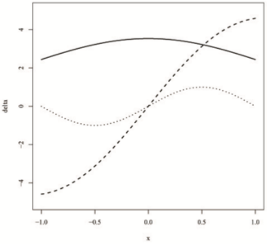


Figure : The solid and dashed lines are the first and second eigenfunction of κ. The dotted line shows the function $\sin \pi x$.

An Illustrative Example (2)

- By Definition 1, the third computer code is the L^2 distance projection.
- The KO calibration with the correlation function Φ gives a different ranking:

$$\begin{aligned}
\epsilon_1^T \Phi_{\mathbf{x}} \epsilon_1 &= 12.594 \\
\epsilon_2^T \Phi_{\mathbf{x}} \epsilon_2 &= 57.908 \\
\epsilon_3^T \Phi_{\mathbf{x}} \epsilon_3 &= 436.268
\end{aligned}$$

- $|\epsilon_3(x)|$ is smaller than $|\epsilon_1(x)|$ and $|\epsilon_2(x)|$ for every x, i.e., the point-wise predictive error for the third computer code is uniformly smaller than the first two. This gives a good justification for choosing the L^2 norm rather than the native norm.

Consistent Calibration

Definition

Suppose θ^* is the unique solution of

$$\|\epsilon(\cdot,\theta^*)\|_{L^2(\Omega)} = \min_{\theta\in\Theta} \|\epsilon(\cdot,\theta)\|_{L^2(\Omega)}.$$

For the design $\mathcal{D}_n$, let $\hat{\theta}_n$ be an estimator of the calibration parameter. The estimator $\hat{\theta}_n$ is *asymptotically consistent* if $\hat{\theta}_n \to \theta^*$ as $h_n \to 0$.

- The aim is to find a consistent estimator of the calibration parameter.

Modified KO Calibration

- We only consider the *stationary* Gaussian process models, i.e., $\Phi(x_1, x_2) = R(x_1 - x_2)$.
- Consistency for calibration cannot be achieved if R is fixed.
- We assume that R has a correlation parameter ϕ satisfying $R(x; \phi) = R(\phi x; 1)$ for any $\phi > 0$ and x. Most correlation families like Gaussian or Matérn family satisfy the conditions.
- Now we assume that the correlation parameter ϕ_n is a *fixed sequence of constants*, not unknown parameter to be estimated from data.
- Define the *modified KO calibration* $\hat{\theta}_n$ by the MLE under the correlation function $\Phi(x_1, x_2) = R(x_1 - x_2, \phi_n)$.

Consistency Conjecture

- We expect that $\phi_n \to +\infty$ is a necessary condition for consistency.
- If ϕ_n increases too fast, the interpolator does not converge. Haaland and Qian (2011) gives an error bounds on the interpolate

$$\|\epsilon - \hat{\epsilon}_n\|_{\mathcal{N}_{R_\phi}(\Omega)} \leq C_R(\phi h_{X,\Omega})^{k/2}\|\epsilon\|_{\mathcal{N}_{R_\phi * R_\phi}(\Omega)}.$$

  This result reveals that another necessary condition for the consistency is $\phi_n h_n \to 0$.
- We state the following conjecture:
  Under regularity conditions, the modified KO calibration is asymptotically consistent if $\phi_n \to +\infty$ and $\phi_n h_n \to 0$.

- Consider the three-level calibration example again.
- Define $\Phi(x_1, x_2; \phi) = \exp\{-\phi(x_1 - x_2)^2\}$ for $\phi = 1, 2, 3, 4$.
- The following table shows the values of $\epsilon_i^T \Phi_{X,\phi}^{-1} \epsilon_i$.

	$\phi = 1$	$\phi = 2$	$\phi = 3$	$\phi = 4$
$\epsilon_1^T \Phi_{X,\phi}^{-1} \epsilon_1$	12.59418	14.96617	17.46962	19.64678
$\epsilon_2^T \Phi_{X,\phi}^{-1} \epsilon_2$	57.90827	44.70162	46.05707	47.69874
$\epsilon_3^T \Phi_{X,\phi}^{-1} \epsilon_3$	436.2677	26.35112	8.998971	6.259807

Ongoing Work

- Proof of the conjecture.
- Convergence rate for the modified KO calibration.
- Extensions to physical experiments with measurement error.