Pacific Northwest
NATIONAL LABORATORY

Model inversion for complex physical systems using low-dimensional surrogates

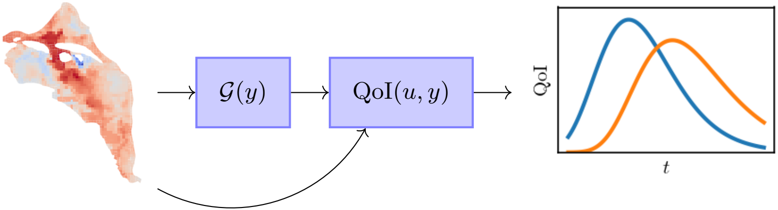
SIAM 2023 Mathematical Problems in Industry Workshop

David A. Barajas-Solano (he/him)

U.S. DEPARTMENT OF ENERGY BATTELLE
PNNL is operated by Battelle for the U.S. Department of Energy

Model inversion is critical for predictive modeling

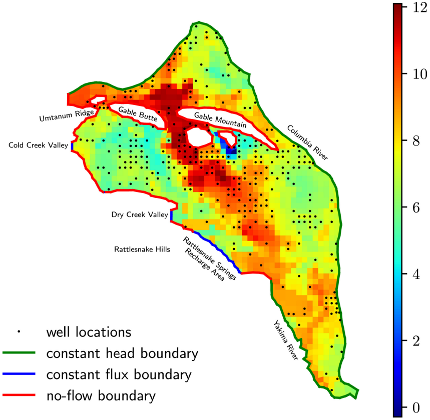
BVP models of physical systems often involve heterogeneous parameter fields $y(x)$



It is necessary to accurately estimate $y(x)$ in order to maintain confidence in the predictive capacity of our models

Hanford Site

- DOE is responsible for the Hanford Site, “one of the largest cleanup efforts in the world, managing the legacy of five decades of nuclear weapons production”
- Calibrating groundwater flow and contaminant transport models of the Hanford Site is crucial for evaluating remediation strategies and performing exposure assessments
- A network of sparsely-distributed observation wells collect measurements of hydraulic pressure and tracer breakthrough curves from tracer experiments

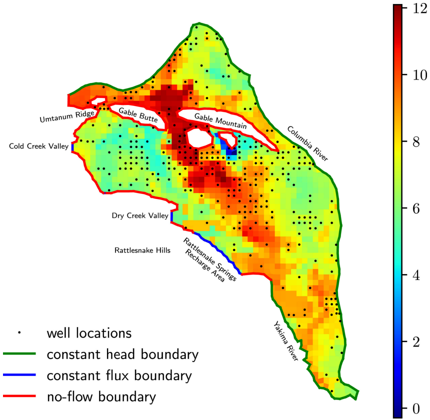


Simplified Hanford Site flow model

The geochemistry and groundwater flow and transport dynamics at the Hanford Site are very complex. We consider the following simplified flow model:

- Two-dimensional stationary flow
- Confined aquifer model
- Cell-centered finite volume discretization with $\sim 1,400$ cells
- For this model, flow is fully determined by the BCs and the transmissivity field $T(x)$
- Hydraulic pressure measurements at obs. wells

Objective: Assuming BCs are known, estimate the log-transmissivity field $y := \log T(x)$ from the pressure measurements



PDE-constrained formulation of the inverse problem

The true parameter field $y_{\text{ref}}(x)$ is usually not observable directly. Rather, we can observe a vector function $\mathbf{h}(u, y)$:

$$\hat{\mathbf{u}} = \mathbf{h}(u_{\text{ref}}, y_{\text{ref}}) + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_u^2 \mathbf{I})$$

which induces the *observable response function* $\mathbf{g}(y) := \mathbf{h}(\mathcal{G}(y), y)$

$$\min_y \frac{1}{2\sigma_u^2} \|\hat{\mathbf{u}} - \mathbf{g}(y)\|_2^2 + \gamma \rho(y)$$

To make it amenable to numerical treatment, we introduce the *finite-dimensional representation*

$$y(x) \approx \tilde{y}(x; \boldsymbol{\xi}), \quad \boldsymbol{\xi} \in \mathbb{R}^{N_\xi}$$

Maximum a posteriori estimator

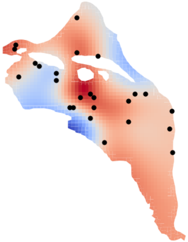
$$\boldsymbol{\xi}_{\text{MAP}} := \arg\min_y \frac{1}{2\sigma_u^2} \|\hat{\mathbf{u}} - \tilde{\mathbf{g}}(\boldsymbol{\xi})\|_2^2 + \gamma R(\boldsymbol{\xi}), \quad \tilde{\mathbf{g}}(\boldsymbol{\xi}) := \mathbf{g}(\tilde{y}(\cdot; \boldsymbol{\xi}))$$

Finite-dimensional representations of the parameter field

The naïve ("direct") representation: For a cell-centered finite volumes discretization

$$\tilde{y}(x;\mathbf{y}) := \langle \mathbf{1}_{\mathsf{FV}}(x), \mathbf{y} \rangle \,, \quad [\mathbf{1}_{\mathsf{FV}}(x)]_1 = 1_{\Omega_i}(x) = \begin{cases} 1 & \text{if } x \in \ i\text{th cell } \Omega_i \\ 0 & \text{otherwise} \end{cases}$$

Pilot points/Inducing points: $\tilde{y}(x)$ given by interpolating between values at "pilot points" $\mathbf{y} = y(X_{\mathsf{PP}})$ via Kriging/Gaussian process regression (GPR)



$$\tilde{y}(x;\mathbf{y}) := m(x) + c(x, X_{\mathsf{PP}}) C^{-1}(X_{\mathsf{PP}}, X_{\mathsf{PP}}) \left[\mathbf{y} - m(X_{\mathsf{PP}})\right]$$

where $m(x)$ and $C(x, x')$ are the Kriging/GPR mean and covariance kernel

Kosambi-Karhunen-Loève expansion (KKLEs)

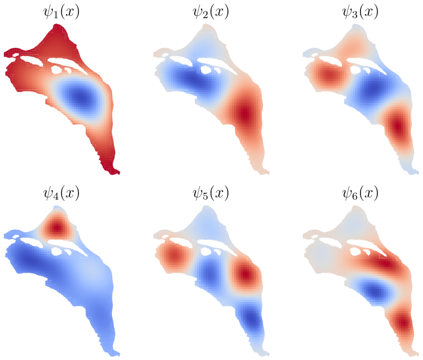
KKLEs are expansions of the form

$$\tilde{y}(x;\boldsymbol{\xi}) := m(x) + \langle \boldsymbol{\psi}(x), \boldsymbol{\xi} \rangle$$

where $[\boldsymbol{\psi}(x)]_i = \lambda_i \varphi_i(x)$, and $\{\lambda_i, \varphi_i(x)\}_{i=1}^{N_\xi}$ are the solutions to the eigenproblem

$$\int C(x,x')\,\varphi(x')\,\mathrm{d}x' = \lambda \varphi(x)$$

For the simplified Hanford Site model, we model the log-transmissivity field using a KKLE with $N_\xi = 10^3$



Log-transmissivity KKL eigenfunctions

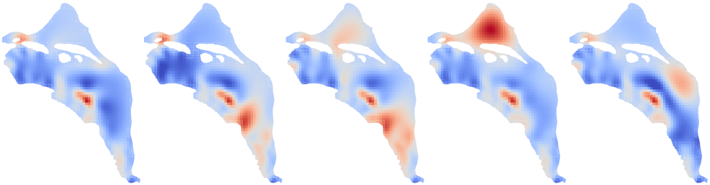
Challenges in model inversion

1. **High dimensionality**: High heterogeneity requires large N_ξ
2. **Scaling of cost with N_ξ**: Cost of inversion algorithms in general increases with N_ξ
3. **Expensive evaluation of $\tilde{g}(\boldsymbol{\xi})$**: Evaluating the observable response requires a query $\mathcal{G}(\tilde{y}(x;\boldsymbol{\xi}))$, which in turn requires a usually expensive forward solve
4. **Ill-posedness of the inverse problem**

Ill-posedness of the inverse problem

It is often the case in inverse problems that many $\boldsymbol{\xi}$ result in the same response $\tilde{\mathrm{g}}(\boldsymbol{\xi})$ up to observation error, because either:

- The observations are affected by loss of memory (e.g., ICs are forgotten) or destruction of information (due to e.g., diffusion), or
- $\boldsymbol{\xi} \mapsto \tilde{\mathrm{g}}(\boldsymbol{\xi})$ is surjective



All these log-transmissivity fields have the same response at 50 observation wells...

In many cases we find that if we move in parameter space along certain $r \ll N_{\xi}$ "effective coordinates", the observable response changes, but if we move in orthogonal directions, it doesn't

Pacific Northwest NATIONAL LABORATORY

Surrogate models of the observable response eliminate the need to query forward solvers

- The main source of computational cost is forward solver queries to evaluate $\tilde{\mathrm{g}}(\boldsymbol{\xi})$, but what if we could (approximately) evaluate $\tilde{\mathrm{g}}(\boldsymbol{\xi})$ without a forward solve?
- A *surrogate model* for $\tilde{\mathrm{g}}(\boldsymbol{\xi})$ is a regression function $f(\boldsymbol{\xi}) \approx \tilde{\mathrm{g}}(\boldsymbol{\xi})$ that can be evaluated faster than a forward solve, usually constructed in a data-driven fashion, i.e., from training data

$$\mathbf{U}_{\text{train}} = \begin{bmatrix} \tilde{\mathrm{g}}(\boldsymbol{\xi}^{(1)}) \\ \vdots \\ \tilde{\mathrm{g}}(\boldsymbol{\xi}^{(N_{\text{train}})}) \end{bmatrix}, \quad \boldsymbol{\Xi}_{\text{train}} = \begin{bmatrix} — & \boldsymbol{\xi}^{(1)} & — \\ & \vdots & \\ — & \boldsymbol{\xi}^{N_{\text{train}}} & — \end{bmatrix}, \boldsymbol{\xi}^{(i)} \sim p(\boldsymbol{\xi})$$

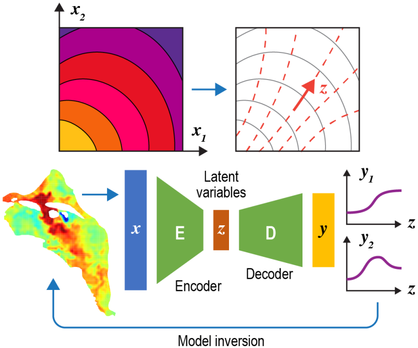
- Great! Let's train a surrogate model for $\tilde{\mathrm{g}}(\boldsymbol{\xi})$ and call it a day... right?
- **Unfortunately** the cost of regression in terms of the amount of training data required also scales with N_{ξ}, and the training data comes from forward solver queries, so we're back to square one

Low-dimensional surrogate modeling can help address the challenges in surrogate modeling

We say that a function has a *low-dimensional* structure if it can be approximated in the form

$$f(\boldsymbol{\xi}) = (f \circ \mathbf{z})(\boldsymbol{\xi}), \text{ where}$$

- The **encoder transformation** $\boldsymbol{\xi} \mapsto \mathbf{z}(\boldsymbol{\xi})$ transforms $\boldsymbol{\xi}$ into a set of $r \ll N_{\xi}$ "effective coordinates" or "latent variables" $\mathbf{z}$
- The **decoder transformation** $\mathbf{z} \mapsto f(\mathbf{z})$ transforms the effective coordinates to the original output



If (big if) the observable response has a low-dimensional structure, we can try to construct surrogate models for model inversion with a encoder-decoder structure, hopefully reducing the cost of regression because the regression function $f(\mathbf{z})$ is r-dimensional

Ridge function models

For the linear scalar observable response $\tilde{g}(\boldsymbol{\xi}) = b + \langle \mathbf{a}, \boldsymbol{\xi} \rangle$, its variation is restricted to the one-dimensional linear subspace $\mathrm{span}\{\mathbf{a}\}$, so that

$$\tilde{g}(\boldsymbol{\xi}) = b + \|\mathbf{a}\|\eta, \quad \eta = \langle \hat{\mathbf{a}}, \boldsymbol{\xi} \rangle, \quad \hat{\mathbf{a}} := \mathbf{a}/\|\mathbf{a}\|$$

Based on this observation, we assume that the variation of $\tilde{g}$ is mostly over a r-dimensional subspace V with $r \ll N_\xi$ (and that its variation over $V^\perp$ can be disregarded). We consider the *ridge function* model

$$\tilde{g}(\boldsymbol{\xi}) \approx f(\mathbf{A}\boldsymbol{\xi})$$

where the rows of the *rotation matrix* $\mathbf{A}$, $\hat{\mathbf{a}}_1, \ldots, \hat{\mathbf{a}}_r$ form an orthonormal basis for V (so that $\mathbf{A}\mathbf{A}^\top = \mathbf{I}$), and $f(\cdot)$ is a r-dimensional nonlinear regression function

Ridge function models

For the linear scalar observable response $\tilde{g}(\boldsymbol{\xi}) = b + \langle \mathbf{a}, \boldsymbol{\xi} \rangle$, its variation is restricted to the one-dimensional linear subspace $\mathrm{span}\{\mathbf{a}\}$, so that

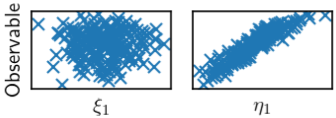
$$\tilde{g}(\boldsymbol{\xi}) = b + \|\mathbf{a}\|\eta, \quad \eta = \langle \hat{\mathbf{a}}, \boldsymbol{\xi} \rangle \,, \quad \hat{\mathbf{a}} := \mathbf{a}/\|\mathbf{a}\|$$

Based on this observation, we assume that the variation of $\tilde{g}$ is mostly over a r-dimensional subspace V with $r \ll N_\xi$ (and that its variation over $V^\perp$ can be disregarded). We consider the *ridge function* model

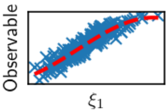
$$\tilde{g}(\boldsymbol{\xi}) \approx f(\mathbf{A}\boldsymbol{\xi})$$

where the rows of the *rotation matrix* $\mathbf{A}$, $\hat{\mathbf{a}}_1, \ldots, \hat{\mathbf{a}}_r$ form an orthonormal basis for V (so that $\mathbf{A}\mathbf{A}^\top = \mathbf{I}$), and $f(\cdot)$ is a r-dimensional nonlinear regression function

1. Find the rotation matrix $\mathbf{A}$



2. Find the regression function $f(\cdot)$



Pacific Northwest
NATIONAL LABORATORY

Algorithms for estimating rotation matrices

- **Basis adaptation (BA)**: For a scalar observable $\tilde{g}(\boldsymbol{\xi})$ and training data

$$\mathbf{u}_{\text{train}} = \begin{bmatrix} \tilde{g}(\boldsymbol{\xi}^{(1)}) \\ \vdots \\ \tilde{g}(\boldsymbol{\xi}^{(N_{\text{train}})}) \end{bmatrix}, \quad \boldsymbol{\Xi}_{\text{train}} = \begin{bmatrix} \text{---} & \boldsymbol{\xi}^{(1)} & \text{---} \\ & \vdots & \\ \text{---} & \boldsymbol{\xi}^{N_{\text{train}}} & \text{---} \end{bmatrix},$$

we pose the linear model approximation $\mathbf{u} \approx b + \boldsymbol{\Xi}\mathbf{a}$, and find $\mathbf{a}$ via linear regression (ordinary least squares, LASSO, etc.). We set $\hat{\mathbf{a}} = \mathbf{a}/||\mathbf{a}||$

Algorithms for estimating rotation matrices

- **Basis adaptation (BA)**: For a scalar observable $\tilde{g}(\boldsymbol{\xi})$ and training data

$$\mathbf{u}_{\text{train}} = \begin{bmatrix} \tilde{g}(\boldsymbol{\xi}^{(1)}) \\ \vdots \\ \tilde{g}(\boldsymbol{\xi}^{(N_{\text{train}})}) \end{bmatrix}, \quad \boldsymbol{\Xi}_{\text{train}} = \begin{bmatrix} \text{---} & \boldsymbol{\xi}^{(1)} & \text{---} \\ & \vdots & \\ \text{---} & \boldsymbol{\xi}^{N_{\text{train}}} & \text{---} \end{bmatrix},$$

  we pose the linear model approximation $\mathbf{u} \approx b + \boldsymbol{\Xi}\mathbf{a}$, and find $\mathbf{a}$ via linear regression (ordinary least squares, LASSO, etc.). We set $\hat{\mathbf{a}} = \mathbf{a}/||\mathbf{a}||$
- Other algorithms include **Sliced inverse regression (SIR)** and **Active Subspaces (AS)**
- SIR and AS are formulated for $r > 1$ but in its original form BA is formulated for $r = 1$
- SIR and BA only require zeroth-order information, but AS requires gradient information

In summary

What we want

- Train **low-dimensional surrogate models** of the observable stationary pressure response of a simplified model of the Hanford Site
- Use these surrogate models to estimate the reference log-transmissivity field from observations of the pressure response

What we bring

- The reference log-transmissivity field
- $20,000$ sets of KKLE coefficients corresponding to log-transmissivity fields, and the corresponding observable response at 323 observation wells
- *Some* Python code (among other stuff, the reference FV solver) and Jupyter notebooks, and a `conda-forge` environment file