A Multivariate Interpolation and Regression Enhanced Kriging Surrogate Model

Komahan Boopathy
Markus P. Rumpfkeil

21^{st} **AIAA Computational Fluid Dynamics Conference, San Diego, California**

University of Dayton
Department of Mechanical and Aerospace Engineering

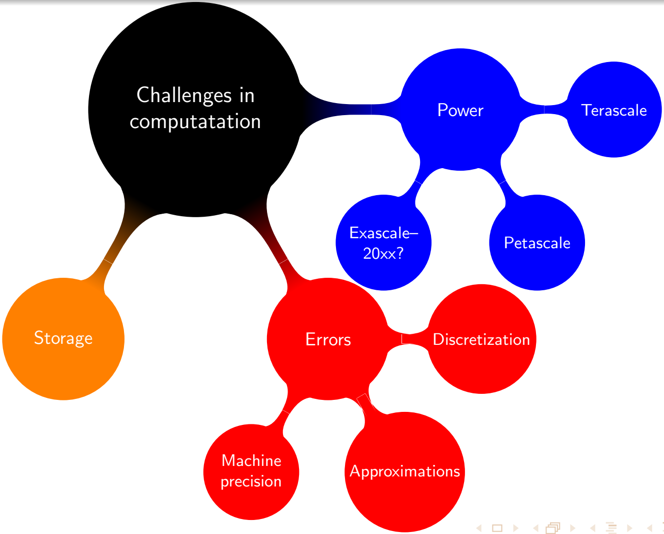
Outline

Table of Contents

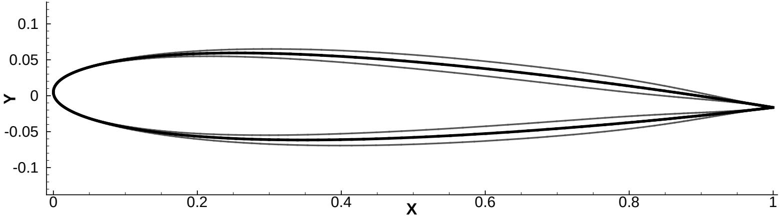
Introduction and Motivation I

- **Analysis:**
  - Theoretical
  - Experimental
  - Computational
- **Advancements:**
  - Hardware (processor speed, multi-core systems)
  - Software (parallel programming)
  - Algorithms and other tools (sophisticated methods)

Introduction and Motivation II



Introduction and Motivation III

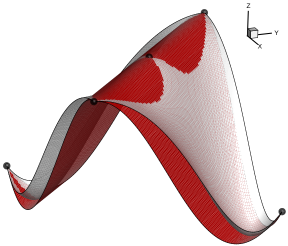


Optimization:

- Many design iterations – can be very expensive
- Highly coupled with several disciplines
- Time consuming to do physical testing and infeasibility

Introduction and Motivation IV

- **How to alleviate computational burden?**
  - Surrogate models / Meta models/ Response surfaces



Surrogate Model

Approximation of the exact function using interpolation and/or extrapolation

Introduction and Motivation V

Active Research

- **Enhance the existing surrogates**
  - Training point selection
  - Higher order information (gradients & Hessian)
  - Variable-fidelity (multi-fidelity)

Introduction and Motivation V

Active Research

- **Enhance the existing surrogates**
  - Training point selection
  - Higher order information (gradients & Hessian)
  - Variable-fidelity (multi-fidelity)
- **Develop new surrogates**
  - Robust & versatile

Table of Contents

Training Point Selection

Domain based sampling

- Monte-Carlo
- Latin Hypercube
- Delaunay Triangulation

Training Point Selection

Domain based sampling

- Monte-Carlo
- Latin Hypercube
- Delaunay Triangulation

Response based (adaptive)

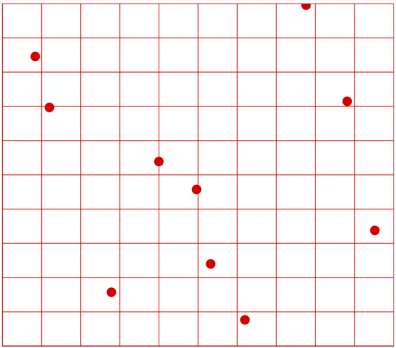
- Distance / Function values / Gradients / Physics

Monte-Carlo Sampling

Monte-Carlo

- Random number generator
- Very simple to program
- No control over locations

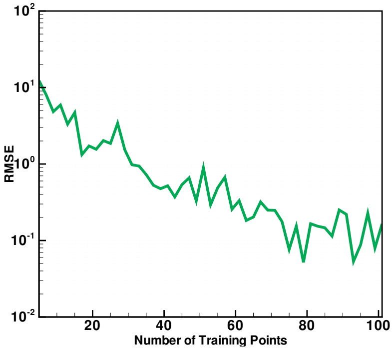
Latin Hypercube Sampling



Latin Hypercube

- McKay - while designing computer experiments
- Equal probability
- N^M bins in the design space
- No two points lie in the same bin

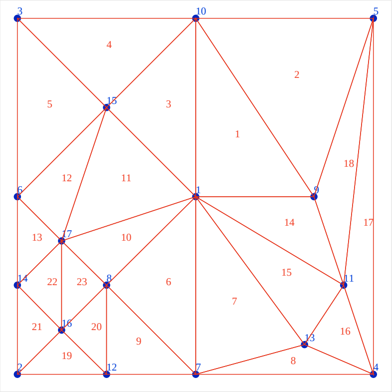
Latin Hypercube Sampling



Typical convergence history

- Random fluctuations
- Each data point is expensive to obtain
- Waste of computational time
- Need for monotonicity

Delaunay Triangulation



Delaunay Triangulation

- Geometrical method
- Split into hyper triangles
- Poor scaling to higher dimensions

Kriging Surrogate

- Originated in geological statistics
- Predicts the function by stochastic processes
- Highly non-linear and multi-modal functions
- The basic formulation of Kriging is given as,

$$\tilde{f}(x) = \mu + Z(x)$$

  $\rightarrow$ μ models the mean behavior

  $\rightarrow$ $Z(x)$ models the local variations using a Gaussian process
- Variants:
  - Direct: Gradient/Hessian terms are included in the formulation (correlation between func-grad, func-Hess, grad-grad, etc.)
  - Indirect: Same formulation as original Kriging but additional samples are created by using gradient/Hessian information

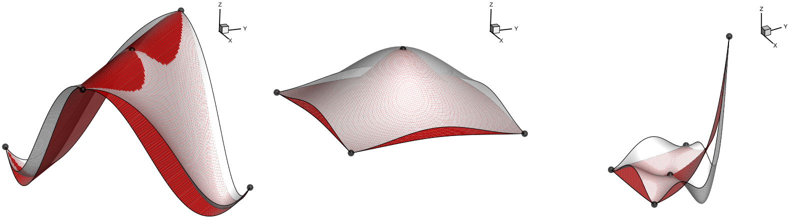
Multivariate Interpolation and Regression

- Based on Taylor series expansion
- Mathematically,

$$\tilde{f} = \sum_{i=1}^{N_v} a_{vi}(x) f(x_{vi}) + \sum_{i=1}^{N_g} a_{gi}(x) \nabla f(x_{gi})$$

  - N_v, N_g is the number of function and func-grad data points
  - a_{vi} and a_{gi} are the basis functions
  - f and ∇f are the function f and gradient values
- **Tunable parameters:** Taylor order n and others

Analytic Test Functions

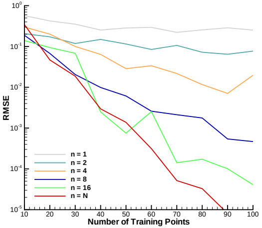


Cos   Runge   Exponential

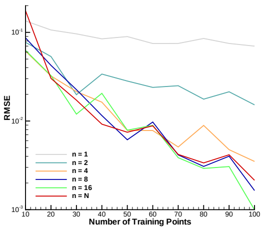
Analytic test functions on hypercube $[-2, 2]^M$

1. Cosine: $f_1(x_1, \ldots, x_M) = \cos(x_1 + \ldots + x_M)$
2. Runge: $f_2(x_1, \ldots, x_M) = \frac{1}{1+x_1^2+\ldots+x_M^2}$
3. Exponential: $f_3(x_1, \ldots, x_M) = e^{(x_1+\ldots+x_M)}$

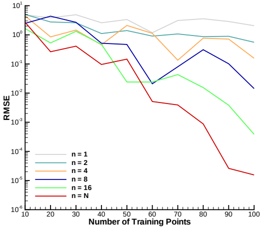
Effect of Taylor order (2D)



Cosine



Runge

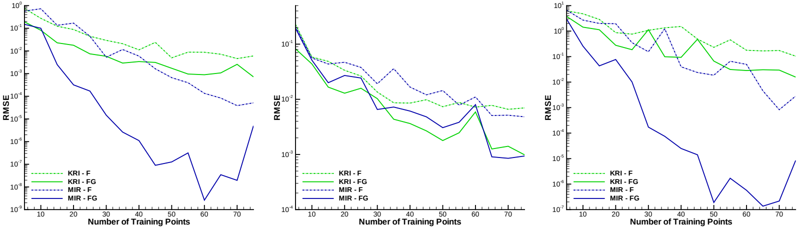


Exponential

Remarks:

- Higher *n* can corrupt the solution as well
- Higher *n* mandates more computational time
- Choice of an optimum Taylor order: tedious task

Original Kriging vs. MIR in two dimensions



Cosine | Runge | Exponential

Remarks:

- **Advantage:** Accuracy, convergence rate
- **Disadvantage:** Computationally intensive, tunable params.

Choice of local and global surrogate

	Kriging	MIR

Choice of local and global surrogate

	Kriging	MIR
Computational time	Less	Very high

Choice of local and global surrogate

	Kriging	MIR
Computational time	Less	Very high
Hessian capability	Yes	No (research area)

Choice of local and global surrogate

	Kriging	MIR
Computational time	Less	Very high
Hessian capability	Yes	No (research area)
Variable fidelity support	Yes	No (research area)

Choice of local and global surrogate

	Kriging	MIR
Computational time	Less	Very high
Hessian capability	Yes	No (research area)
Variable fidelity support	Yes	No (research area)
Convergence rate	Low	High

Choice of local and global surrogate

	Kriging	MIR
Computational time	Less	Very high
Hessian capability	Yes	No (research area)
Variable fidelity support	Yes	No (research area)
Convergence rate	Low	High
Tunable parameters	Absent	Present

Choice of local and global surrogate

	Kriging	MIR
Computational time	Less	Very high
Hessian capability	Yes	No (research area)
Variable fidelity support	Yes	No (research area)
Convergence rate	Low	High
Tunable parameters	Absent	Present

Our theme: Use MIR to guide global Kriging

Adaptive Training Point Selection

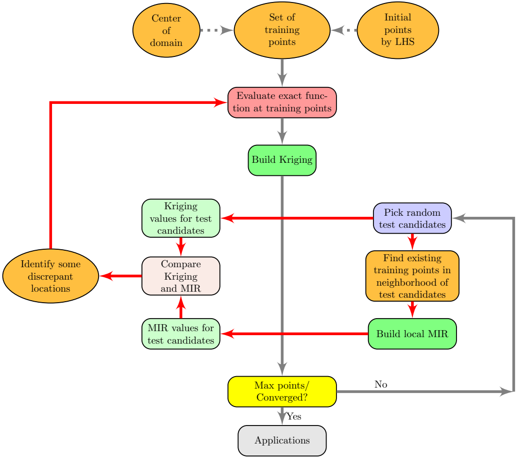


Table of Contents

Conclusion

- **Summary:**
  - Made use of local surrogate for training point selection
  - Applied to multi-dimensional test functions
  - Showed improvement for monotonic convergence behavior
  - Showed Variable-fidelity results
- **Future Work:**
  - Where to use gradient information?
  - How to use variable-fidelity data efficiently?
- **Potential Applications:**
  - Aerospace design & optimization
  - Uncertainty quantification
  - Aerodynamic databases

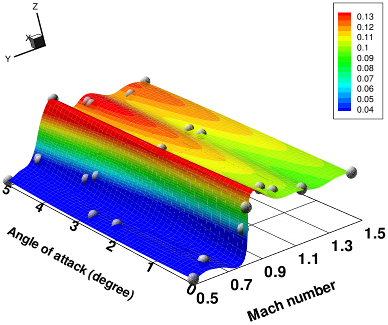
Acknowledgments

1. Wataru Yamazaki – Kriging surrogate
2. Qiqi Wang – MIR source code

Selected Bibliography

Any Questions?

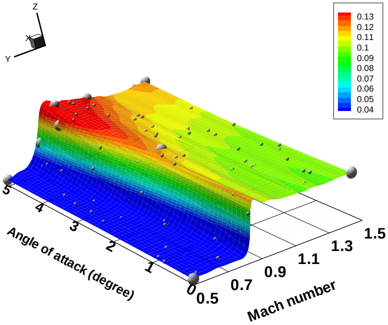
Kriging Drag Database - High Fidelity Model



Kriging Drag Database

- 25 Euler evaluations
- Fine mesh $19,548$ elements
- Adaptive sampling strategy
- Not computationally expensive
- Nicely captures transonic behavior

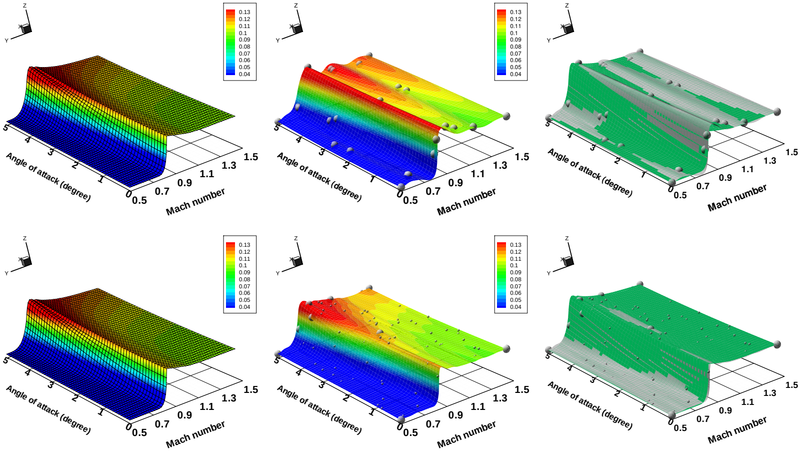
Kriging Drag Database - Variable Fidelity Model



Variable Fidelity

- 9 High fid. training points adaptively
- Fine mesh 19,548 elements
- 64 Low fid. training points via LHS
- Coarse mesh 4,433 elements

Drag Database



Euler | Our Kriging | Both combined

Direct Kriging

- Gradient/Hessian terms are included in the formulation
  - Function value estimated using a linear combination of function, gradient and Hessian values
  - Minimize mean-squared-error (MSE) between exact and estimated function value
  - Final form of the gradient/Hessian enhanced direct Cokriging:

$$\hat{\mathcal{J}}(D) = \mu + r^T(D)R^{-1}(Y - \mu I)$$

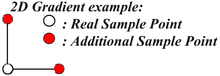
where

$\mu = (I^T R^{-1} I)^{-1}(I^T R^{-1} Y)$	constant mean term
R	correlation matrix between samples
$Y = \left(\mathcal{J}(D_1), \ldots, \left.\frac{d\mathcal{J}}{dD}\right\|_{D_1}, \ldots, \left.\frac{d^2\mathcal{J}}{dD^2}\right\|_{D_1}, \ldots \right)$	vector of sample point information
$r(D)$	correlation between D and samples

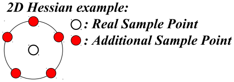
  - Determine required derivatives of correlation function (up to fourth order) with automatic differentiation

Indirect Kriging

- Additional samples are created by using gradient and Hessian information



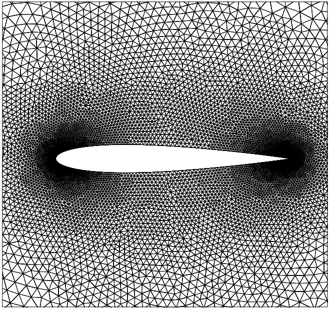
$$\mathcal{J}_{add} = \mathcal{J}(D_i) + \left.\frac{d\mathcal{J}}{dD}\right|_{D_i} \Delta D$$



$$\mathcal{J}_{add} = \mathcal{J}(D_i) + \left.\frac{d\mathcal{J}}{dD}\right|_{D_i} \Delta D + \frac{1}{2}\Delta D^T \left.\frac{d^2\mathcal{J}}{dD^2}\right|_{D_i} \Delta D$$

  - Major parameters: distance between real and additional points ΔD and number of additional points per real sample point
  - Worse R matrix conditioning with smaller distances and larger number of additional points
    $\rightarrow$ Severe trade-offs for these parameters

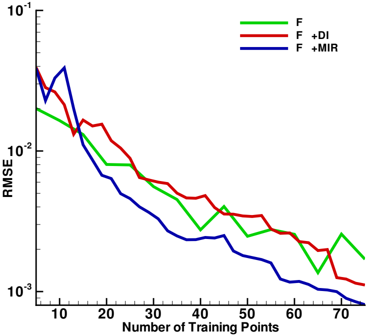
Test Case



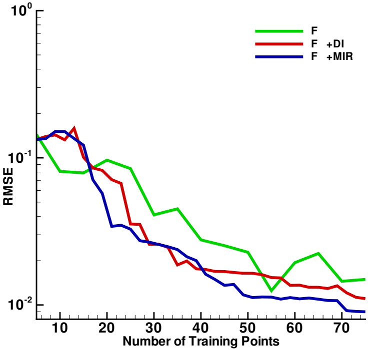
Problem Setup

- NACA0012 airfoil
- Eulerian flow solver
- Cell-centered second-order accurate finite-volume approach
- $0.5 < M < 1.5$ and $0^\circ < \alpha < 5^\circ$
- Fine mesh $19,548$ elements
- Coarse mesh $4,433$ elements

Convergence History

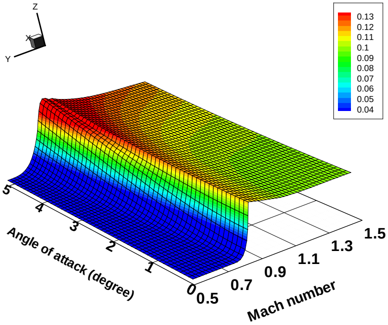


Drag



Lift

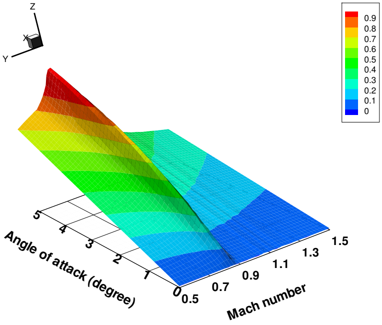
Exact Drag Database



Exact Drag Database

- Solves Euler Equations (Inviscid)
- Cartesian mesh - α vs. M
- 2601 nodes
- Computationally expensive

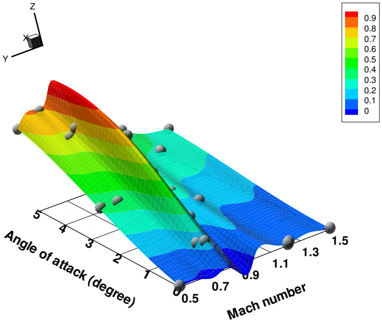
Exact Lift Database



Exact Lift Database

- Solves Euler Equations (Inviscid)
- Cartesian mesh - α vs. M
- 2601 nodes
- Computationally expensive

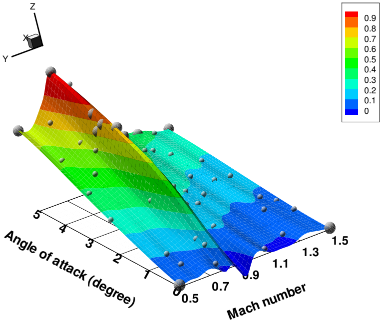
Kriging Lift Database - High Fidelity Model



Kriging Lift Database

- 25 Euler evaluations
- Fine mesh 19, 548 elements
- Adaptive sampling strategy
- Not computationally expensive
- Nicely captures transonic behavior

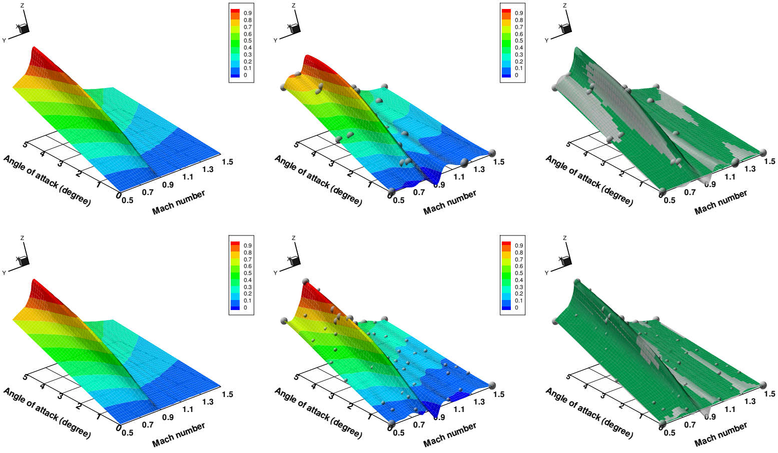
Kriging Lift Database - Variable Fidelity Model



Variable Fidelity

- 15 High fid. training points adaptively
- Fine mesh 19, 548 elements
- 40 Low fid. training points via LHS
- Coarse mesh 4, 433 elements

Lift Database



Euler | Our Kriging | Both combined

Observations

RMSE comparisons for Kriging models

RMSE	High-fidelity	Variable-fidelity

Observations

RMSE comparisons for Kriging models

RMSE	High-fidelity	Variable-fidelity
Drag Coefficient	0.45868×10^{-2}	0.38118×10^{-2}

Observations

RMSE comparisons for Kriging models

RMSE	High-fidelity	Variable-fidelity
Drag Coefficient	0.45868×10^{-2}	0.38118×10^{-2}
Lift Coefficient	0.32746×10^{-1}	0.27735×10^{-1}