Surrogate-based modeling of electrical systems

Ivo Couckuyt
Professor Ir. Tom Dhaene

Outline

- Introduction
  - Surrogate modeling
- Surrogate-based Optimization
- Integrating surrogate models in optimization
- Statistical infill criteria
- Conclusion

A New Science Paradigm

- thousand years ago: experimental science
  - description of natural phenomena
- last few hundred years: theoretical science
  - Newton's laws, Maxwell's equations...
- last few decades: computational science
  - simulation of complex phenomena
- today: e-Science or data-centric science
  - massive computing
  - large data exploration and mining
  - unify: theory, experiment, and simulation

(With thanks to Jim Gray)

Surrogate Modeling Lab – www.surmo.intec.ugent.be
Department of Information Technology (INTEC) - IBCN

Surrogate model?

- system modeling
  - real world
    - I/O system
    - stimulus / response
  - simulation model
    - approximation
    - discretization
  - surrogate model
    - metamodel, RSM, emulator
    - scalable analytical model
    - "model of model"
Adaptive sampling

- traditional approach
  - uniform sampling
  - oversampling
  - undersampling

- adaptive sampling
  - optimal sample distribution
  - Sequential design

Adaptive modeling

- traditional approach
  - local approximation
  - overmodeling
  - undermodeling

- adaptive modeling
  - global approximation
  - optimal model complexity

Methodology

1. Initial data points
2. Run simulation code
3. Identify new data points
4. Create surrogate models
5. Tune parameters
6. Estimate accuracy
7. Improvement?
   - Yes
   - No
   - Accuracy reached?
     - No
     - Yes
     - Terminate

Outline

- Introduction
- Surrogate-based Optimization
  - The basic idea
- Integrating surrogate models in optimization
- Statistical infill criteria
- Use cases
- Conclusion
Surrogate-based Optimization

- Expensive simulation code
  - Reduce the number of simulations

=> Surrogate-Based Optimization (SBO)
  - Surrogate used to expedite search of global optimum
  - Global accuracy of surrogate not a priority
  - Surrogate model is a tool to an end, i.e., optimization

Outline

- Surrogate-based Optimization
- Integrating surrogate models in optimization
- Statistical infill criteria
- Use cases
- Conclusion

Multi-fidelity methods

- Multiple simulation codes for the same problem
  - Varying accuracy (fidelity)

- Exploit!
  - "Fuse" such multi-fidelity models

- Introduce scaling factors
  - Correct low-fidelity model to agree with the high-fidelity model on a set of points
  - Additive or multiplicative factors
  - Zero-order, first-order, etc. scaling
Multi-fidelity methods (cont’d)

- Space mapping
  - Maps the input spaces of the simulation models so that the optima align in the design space
  - Variants: input, output, manifold, ... space mapping
- Efficient
- Not completely black-box

Cokriging
- Inherently a multi-fidelity surrogate model
- Combines low- and high-fidelity data to increase prediction accuracy

Outline

- Surrogate-based Optimization
- Integrating surrogate models in optimization
- Statistical infill criteria
- Use cases
- Conclusion

Methodology

1. Initial data points
2. Run simulation code
3. Identify new data points
4. Create surrogate models
5. Tune parameters
6. Estimate accuracy
7. Improve?
   - Yes: Accuracy reached
   - No: Terminate
8. Yes: Terminate
   - No: Accuracy reached
Methodology (cont'd)

- Surrogate model is cheap
- Sequential design = optimization
  - Select new points based on information that the surrogate model provides
- Surrogate model is a global approximation
  - But not necessarily globally accurate!
- Simple approach
  - Select the optimum of the surrogate as the next sample iterate

Simple Gaussian Process

- Gaussian Process
  - flexible!

Simple Polynomial

- Polynomial
  - Fixed structure (quadratic)

Simple Gaussian Process (cont'd)

- But...
**Gaussian Process provides uncertainty!**

**Kriging**
- Gaussian Process-based
  \[ Y(x) = f(x) + Z(x) \]
- Hyperparameters theta
  - Equal to number of dimensions
  - Classical optimization problem
- Provides uncertainty about the predicted function values!

=> Statistical infill criteria
- Balances exploration vs exploitation
- Popularized in the Efficient Global Optimization (EGO) method
**Illustration Gaussian Process**

- Unknown model
- Data points \( f(x) \)
- Minimum over all data points: \( f_{\text{min}} \)
- Surrogate model

\[
Y(x) \sim N(\mu(x), \sigma^2(x))
\]

\[
y = \mu(x), \sigma^2 = \sigma^2(x)
\]

**Probability of Improvement**

\[
\text{Pol}(x) = P(Y(x) \leq f_{\text{min}}) = \int_{-\infty}^{f_{\text{min}}} \phi(Y(x)) \, dY
\]

**Illustration Gaussian Process**

- Unknown model
- Data points \( f(x) \)
- Minimum over all data points: \( f_{\text{min}} \)
- Surrogate model
- Gaussian PDF at \( x=0.5 \)
- Prediction mean at \( x=0.5 \)

\[
Y(x) \sim N(\mu(x), \sigma^2(x))
\]

\[
y = \mu(x), \sigma^2 = \sigma^2(x)
\]

**Expected Improvement**

- Pol is already useful
  - But doesn't quantify the amount of improvement!

- Expected Improvement (EI)
  - The first moment of Pol
  \[
  I(x) = \max(Y(x) - f_{\text{min}}, 0)
  \]

\[
E[I(x)] = \int_{-\infty}^{f_{\text{min}}} I(x) \phi(Y(x)) \, dY
\]

\[
E[I(x)] = \begin{cases} 
(f_{\text{min}} - \bar{y}) \cdot \phi \left( \frac{f_{\text{min}} - \bar{y}}{\hat{s}} \right) + \hat{s} \cdot \phi \left( \frac{f_{\text{min}} - \bar{y}}{\hat{s}} \right) & \text{if } \hat{s} > 0 \\
0 & \text{if } \hat{s} = 0
\end{cases}
\]

- Good trade-off between exploitation and exploration
Use case 1

- Branin function
  - 2D benchmark function for optimization

- Experimental setup
  - 21 initial point (latin hypercube)
  - Kriging
  - Expected improvement
  - Stopping criterion: 1% relative error w.r.t. optimum

Use case 1 (cont'd)

- Mathematical function:
  - 3 global optimums
  - \( x^* = (-\pi, 12.275), (\pi, 2.275), (9.42478, 2.475) \)
  - \( f(x^*) = 0.397887 \)
Use case 1 (cont'd)

- Prediction
- Expected Improvement
- Prediction variance

Use case 2

- Textile antenna
  - Inverse problem: identify material properties
- Software: ADS Momentum
- Inputs (2D)
  - Material properties
- Output
  - Error between measurements and simulation

Diagram:
- Input to Simulator
- Output to Error function
- Error function to Cost function

Use case 2 (cont'd)

- Initial design of 14 samples

Chart:
- Minimum cost function value (NFE)
  - kriging (MLE)
  - Final cost function value
  - Number of samples: 15 to 45
Use case 2 (cont'd)

Pareto optimization

- MultiObjective Surrogate-Based Optimization (MOSBO)
  - Perhaps even more promising than SBO!
- ParEGO
  - Traditional EGO approach but...
  - Cost functions are aggregated into a weighted sum
  - Weights randomized every iteration!
- Or extending the concept of Pol and El...

Probability of Improvement

- Improving on one point
  - In f1, f2, both objectives
- Shaded area (P_augm)
- Hatched area (P_dom)

Uniformity of front = the search strategy

Probability of Improvement (cont'd)

- k levels of improvement
  - k=0: augmenting the Pareto front
  - 0<k=N_par: dominating k points
Pareto optimization (cont’d)

- Probability of improvement
  - \( P_{augm} \Rightarrow \text{chance of augmenting current pareto front} \)
  - \( P_{dom} \Rightarrow \text{chance of improving on the current pareto front} \)
  - More general: levels of improvement
    - Chance of improving on at least \( k \) points of the current pareto front
    - 'k levels of improvement': \( k \) needs to be chosen

Irrespective of scale of objectives!

Use case 3

- Veldhuizen and Lamont's MOP2 function (VLMOP2)
  - Multiobjective optimization benchmark problem
  - 2 outputs

- Initial design of 16 samples

- Pareto Expected Improvement

Expected improvement

- Amount of improvement over centroid (moment arm) w.r.t. closest point on pareto front
- Also 2 flavours: dominating or augmenting

Encourages uniformity

Though, scaling still important (unavoidable)
Current work

- Generalization of statistical criterions
  - 'Directly' solve inverse problems
  - Include manufacturing uncertainty
  - Identify quasi-optimal regions
  - ...

- Examples:
  - Generalized Pol
    \[ gPol(x) = P(a \leq Y(x) \leq b) = \int_a^b \phi(Y(x)) \, dY \]
  - Expected Distance
    \[ D(x) = |Y(x) - \bar{f}_{min}| \]
    \[ E[D(x)] = \int_{-\infty}^{\infty} f(x) \, \phi(Y(x)) \, dY \]

Use case 4

- Branin function
- Output range of interest: [20, 35]
  - \( \phi \) = inverse problem
- Initial design of 8 samples (LHD)
- Use gPol to select new samples
  - No optimization possible!
  - \( n=10000 \) and \( k=1 \)

\[ \text{Generate n candidates} \rightarrow \text{Select m best candidates} \]
Use case 5

- Differential pair on-board microstrip (above PEC ground)
  - 10kHz -> 10 Ghz

- Inputs
  - Spacing between conductors
  - Width of conductors

- Output
  - Relative error between output and 100 Ohm

- Goal: 100 Ohm differential impedance at HF
  - Relative error < 0.02

- Extra: Manufacturing uncertainty on inputs

Use case 5 (cont'd)

- Add input uncertainty
  - Account for a deviation of epsilon ε over the inputs
  - Due to manufacturing limitations

- Expected gPol
  - \( f(x_0) = \int_A gPoI(x) \cdot \beta(x) dY \)

- A denotes an area in the input space
  - \( A = x_0 - \epsilon \to x_0 + \epsilon \)

Outline

- Surrogate-based Optimization
- Integrating surrogate models
- Statistical infill criteria
- Conclusion

Conclusion

- SBO methods are quite powerful
- Solves real-life problems in an efficient way
  - Minimizes number of evaluations
  - Reduces/Eliminates noise
  - Provides sensitivity, robustness analysis, ...

- Careful! Many traps to fall into...
  - Needs decent understanding of the methodology (in case of failure)
  - Curse of dimensionality
  - ...
International Workshop on

Advances in Modeling and Optimization of High Frequency Structures

August 21-22, 2010

Reykjavik University, Iceland

HÁSKÓLINN Í REYKJAVÍK
REYKJAVIK UNIVERSITY

Engineering Optimization & Modeling Center

EOMC