Design of Experiments and Sensitivity analysis

Course and practical application

Author

Institut

OpenMOLE

June 14, 2019

Outline

1. Introduction
2. Basic experiments
3. DOE Samplings
4. Sensitivity analysis
5. Application in OpenMOLE

1. Introduction
2. Basic experiments
3. DOE Samplings
4. Sensitivity analysis
5. Application in OpenMOLE

Design of Experiments

- Interactive model exploration by hand and the need for preliminary experiments - The Design of Experiments as the definition of tasks to extract information from the simulation model - Example: NetLogo behavior space: basic grid DOE - Sensitivity analysis as an advanced DOE

Remark 1: terminology strongly depends on disciplines and practices
*Remark 2: these are generally **preliminary experiments** to prepare more elaborated, question-related, experiments*

1. Introduction
2. Basic experiments
3. DOE Samplings
4. Sensitivity analysis
5. Application in OpenMOLE

Basic experiments

*Provide explicitly sampling points on which the model (or its replication task) will be run: notion of **direct sampling** in OpenMOLE (corresponds to DOE in the literature)*
- full samplings - elaborated sampling for high dimensions given a low computational budget (**the curse of dimensionality**)

One factor at a time

All factors have nominal values and a discrete variation set, in which each is varied while others remaining fixed
- *when model is slow - or computational budget highly limited* - *does not capture interaction between parameters, and highly dependent on nominal values* - *seen as a bad practice* BUT *useful for models taking significant time, and prone to thematic interpretation*
Example where One-At-a-Time fails

Grid sampling

Ensemble product of discrete variation ranges for factors (usually a regular grid but not necessarily)
quickly limited by the curse of dimensionality - in practice still powerful with a quick model and a low number of parameters
naive approach, i.e. done by many "simulation-newcomers" such as economics or some parts of physics

1. Introduction
2. Basic experiments
3. DOE Samplings
4. Sensitivity analysis
5. Application in OpenMOLE

DOE Samplings

Computational limitations => need specific methods to efficiently sample the parameter space
The field of Design of Experiments has proposed different methods for numerical experiments given limited computational resources
Examples: Sobol sequence (quicker convergence of integral estimation), Latin Hypercube Sampling, Orthogonal sampling

Latin Hypercube Sampling

Minimizing discrepency: intuitively being spread evenly accross the parameter space (def of discrepancy)

|x||||| |:–:|:–:|:–:|:–:|:–:| ||x|||| |||||x| ||||x|| |||x|||

Latin cube: one point in each row and column; hypercube generalization in any dimension

Sobol sequence

Quasi-random sequences with low discrepancy (also Halton sequences e.g.)
Estimate integral in $1/N$ instead of $1/\sqrt{N}$ with random sampling
Constructed recursively (using bit representations).
TODO illustration in 2D

1 Introduction

2 Basic experiments

3 DOE Samplings

4 Sensitivity analysis

5 Application in OpenMOLE

Sensitivity analysis

How to summarize model sensitivity and isolate principal factors ?
Examples: Morris and Saltelli methods

Morris method

Idea : Sample trajectories in the parameter space in a One-At-a-Time manner. Screening method isolating *elementary effects*
[Saltelli et al., 2004]
- isolate local effects of factors - more efficient than point sampling to get individual effects - useful as a first experiment to understand the relative influence of factors
[Campolongo et al., 2011] propose to extend the method with Sobol sequences

Saltelli method

Estimation of relative and conditional variances

$$ST_i = \frac{E_{\mathbf{X}\sim i}\left[Var(Y|\mathbf{X}\sim i)\right]}{Var(Y)}$$

1. Introduction
2. Basic experiments
3. DOE Samplings
4. Sensitivity analysis
5. Application in OpenMOLE

OpenMOLE syntax: Direct sampling

val explo = DirectSampling(evaluation = model, sampling = ...)

Example of samplings

- One-factor sampling “‘ sampling = OneFactorSampling((x1 in (0.0 to 1.0 by 0.2)) nominal 0.5, (x2 in (0.0 to 1.0 by 0.2)) nominal 0.5) “‘
- Grid sampling “‘ sampling = (x1 in (0.0 to 1.0 by 0.5)) x (x2 in (0.0 to 1.0 by 0.5)) “‘
- LHS Sampling

“‘ sampling = LHS(100,x1 in (0.0,1.0),x2 in (0.0,1.0)) “‘

- Sobol sampling

“‘ sampling = SobolSampling(100,x1 in (0.0,1.0),x2 in (0.0,1.0)) “‘

—-
Saltelli
(method in itself)
“‘ val sen = SensitivitySaltelli(//evaluation = (model on env), evaluation = (model on env by 1000), samples = 100000, inputs = Seq(humanFollowProbability in (0.0,1.0), humanInformedRatio in (0.0,1.0),humanInformProbability in (0.0,1.0)), outputs = Seq(peakTime, peakSize, totalZombified,halfZombified, spatialMoranZombified,spatialDistanceMeanZombified,spatialEntropyZombified,spatialSlopeZom) “‘
—
Morris
(example from market) “‘ SensitivityMorris(evaluation = modelExec on envLocal hook storeSimuCSV, inputs = Seq(inputNumberOfCars in (1.0, 41.0), inputAcceleration in (0.0, 0.0099), inputDeceleration in (0.0, 0.099)), outputs = Seq(outputSpeedMin, outputSpeedMax), repetitions = 100, levels = 5) “‘
—- Practical application

Your turn to run some direct samplings and/or sensitivity analysis
- given the described zombie model, what first experiment beyond stochasticity would be relevant ? - write a script - explore results (using e.g. the OpenMOLE GUI plots)
resources: - one script running directsampling - example of grid explo results - example of Saltelli
—- **Reserve**
-> results of direct sampling
— References
Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M. (2004). Sensitivity analysis in practice: a guide to assessing scientific models. Chichester, England.

References I


Campolongo, F., Saltelli, A., and Cariboni, J. (2011).
From screening to quantitative sensitivity analysis. a unified approach.
Computer Physics Communications, 182(4):978–988.


Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (2004).
Sensitivity analysis in practice: a guide to assessing scientific models.

Chichester, England.