DAKOTA 101

Sensitivity Analysis

http://dakota.sandia.gov/

Learning goals:

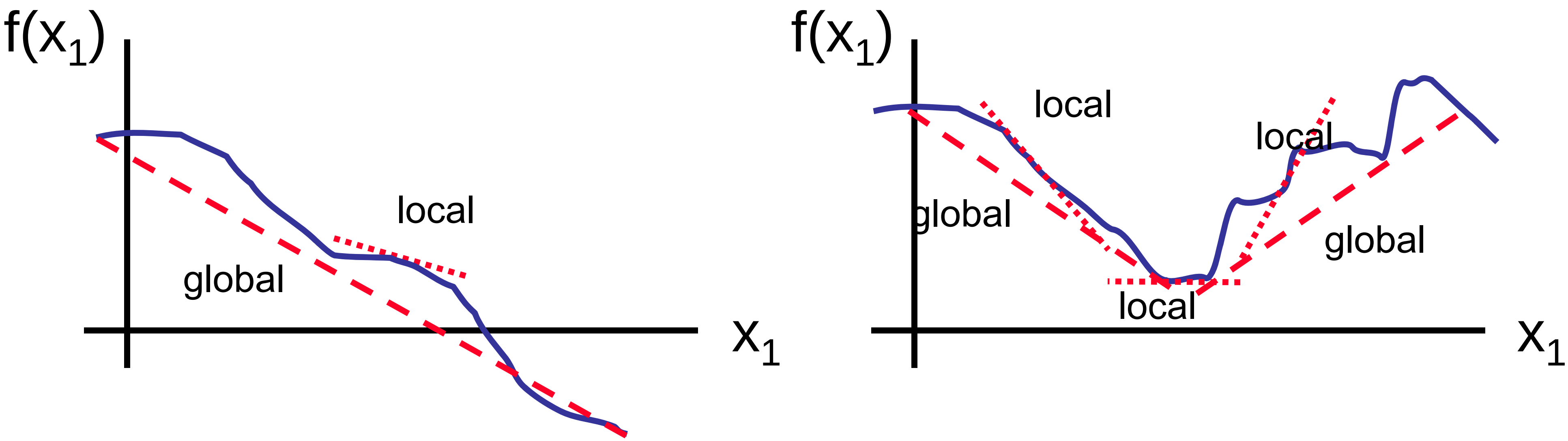
- **Define sensitivity analysis and know when and why to apply it**
- **Use DAKOTA to perform sensitivity analyses; understand potential pitfalls**

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

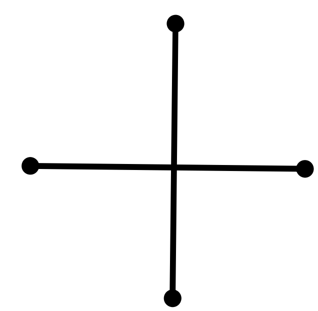
Sensitivity Analysis and UQ Terminology

- **Sensitivity Analysis (SA):**
  - How do code outputs vary due to changes in code inputs?
  - Local sensitivity: code output gradient (derivative) data for a specific set (or sets) of code input parameter values
  - Global sensitivity: the general trends of the code outputs over the full range of code input parameter values (linear, quadratic, etc.)
- **Uncertainty Quantification (UQ):**
  - Aleatoric UQ: What are the statistics or probability distributions on code outputs, given probability distributions on code inputs?
    - Estimate Probability [f(x) > f*], e.g., probability of system failure
  - Epistemic UQ: What are the possible/plausible code outputs?
- **Validation:**
  - How “close” are my code output predictions to experimental data, including UQ for both?
- **Quantification of margins and uncertainties (QMU):**
  - How “close” are my code output predictions (including UQ) to the system’s required performance level?

Examples of Sensitivity Analysis



- **Sensitivity analysis examines variations in f(x_1) due to perturbations in x_1.**
  - **Local sensitivities are typically partial derivatives.**
    - **Given a specific x1, what is the slope at that point?**
  - **Global sensitivities are typically found via regression.**
    - **What is the trend of the function over all values of x1?**



Goals of Sensitivity Analysis

- Identify most important variables and their interactions
  - Provide a focus for resources
    - Model development
    - Code development
    - Uncertainty characterization
  - Provide a basis for constructing surrogate models
  - Select variables for use in optimization
  - Screening: Identity the most important variables, downselect for further analysis

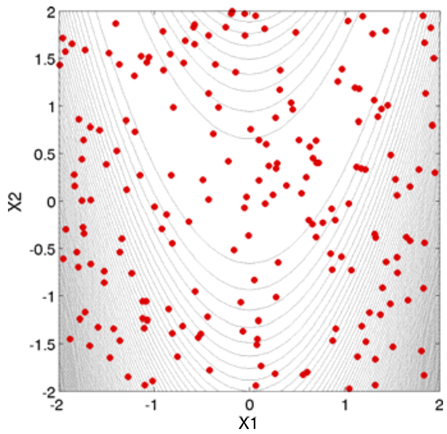
- Can have the side effect of identifying code and model issues

Exercise: Determine trends relative to parameters for cantilever problem

- **Use Jaguar to construct and run a sampling study to determine most influential parameters for cantilever**
  - **6 uniform variables:**

Variable	R	E	X	Y	beam_width	beam_thickness
Upper bound	48000	4.50E+07	700	1200	2.2	2.2
Lower Bound	32000	1.50E+07	300	800	2	2

  - **3 responses**
  - **100 samples**



- **Review correlations**
  - **Simple correlation:** measures the strength and direction of a linear relationship between variables
  - **Partial correlation:** like simple correlation but adjusts for the effects of the other variables
  - **Rank correlations:** simple and partial correlations performed on "rank" of data

Potential Solution:
Sensitivity Analysis for Cantilever

```
# DAKOTA INPUT FILE - dakota_sa_cantilever.in
strategy,
    single_method
    tabular_graphics_data
method,
    sampling
    sample_type lhs
    seed =52983
    samples = 100
variables,
    uniform_uncertain =6
      upper_bounds   48000   45.E+6  700.  1200.  2.2    2.2
      lower_bounds    32000.   15.E+6  300.    800.   2.0     2.0
      descriptors  'R' 'E' 'X' 'Y' 'beam_width' 'beam_thickness'
interface,
    direct
    analysis_driver = 'mod_cantilever‘
responses,
    num_response_functions = 3
    response_descriptors = 'weight' 'stress' 'displ'
    no_gradients
    no_hessians
```


Sandia National Laboratories

SA Results for Cantilever

Simple Correlation Matrix

	R	*E*	*X*	*Y*	beam width	beam thickness	*weight*	*stress*	*displ*
R	1.000								
E	-0.022	1.000							
X	0.012	-0.007	1.000						
Y	0.020	0.017	-0.027	1.000					
beam width	0.009	-0.009	-0.017	-0.014	1.000				
beam thickness	0.003	-0.013	0.038	-0.025	-0.012	1.000			
weight	0.011	-0.016	0.014	-0.027	0.703	0.703	1.000		
stress	-0.345	0.022	0.557	0.579	-0.303	-0.339	-0.457	1.000	
displ	0.009	-0.879	0.085	0.293	-0.125	-0.164	-0.207	0.313	1.000

Partial Correlation Matrix between Input and Output:

	weight	stress	displ
R	0.137	-0.990	-0.058
E	-0.026	0.015	-0.954
X	-0.009	0.996	0.313
Y	0.052	0.996	0.735
beam_width	1.000	-0.984	-0.421
beam_thickness	1.000	-0.989	-0.525

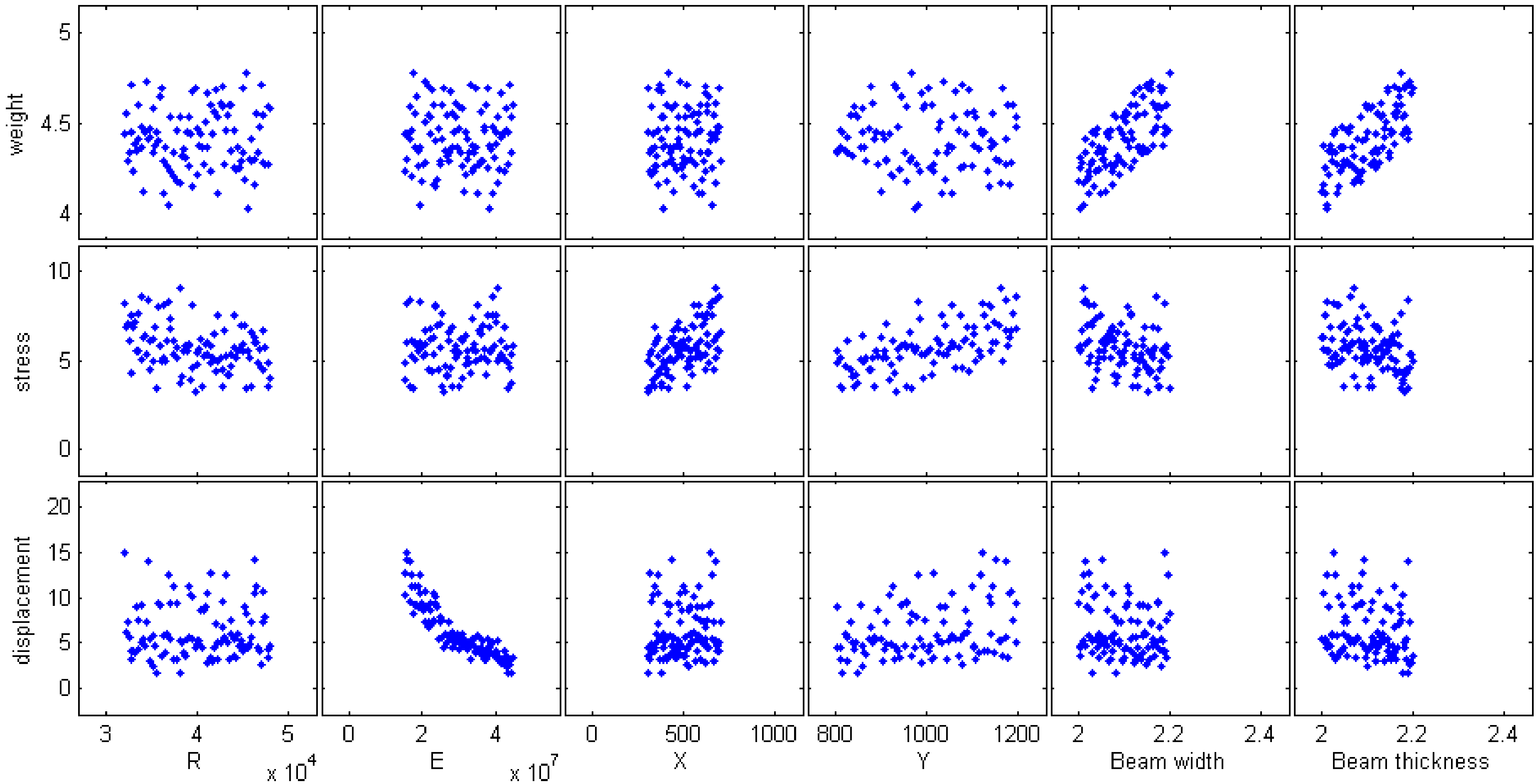
Partial Rank Correlation Matrix between Input and Output:

	weight	stress	displ
R	-0.071	-0.837	-0.056
E	0.082	-0.085	-0.981
X	0.179	0.924	0.531
Y	-0.055	0.934	0.824
beam_width	0.981	-0.800	-0.559
beam_thickness	0.980	-0.838	-0.753

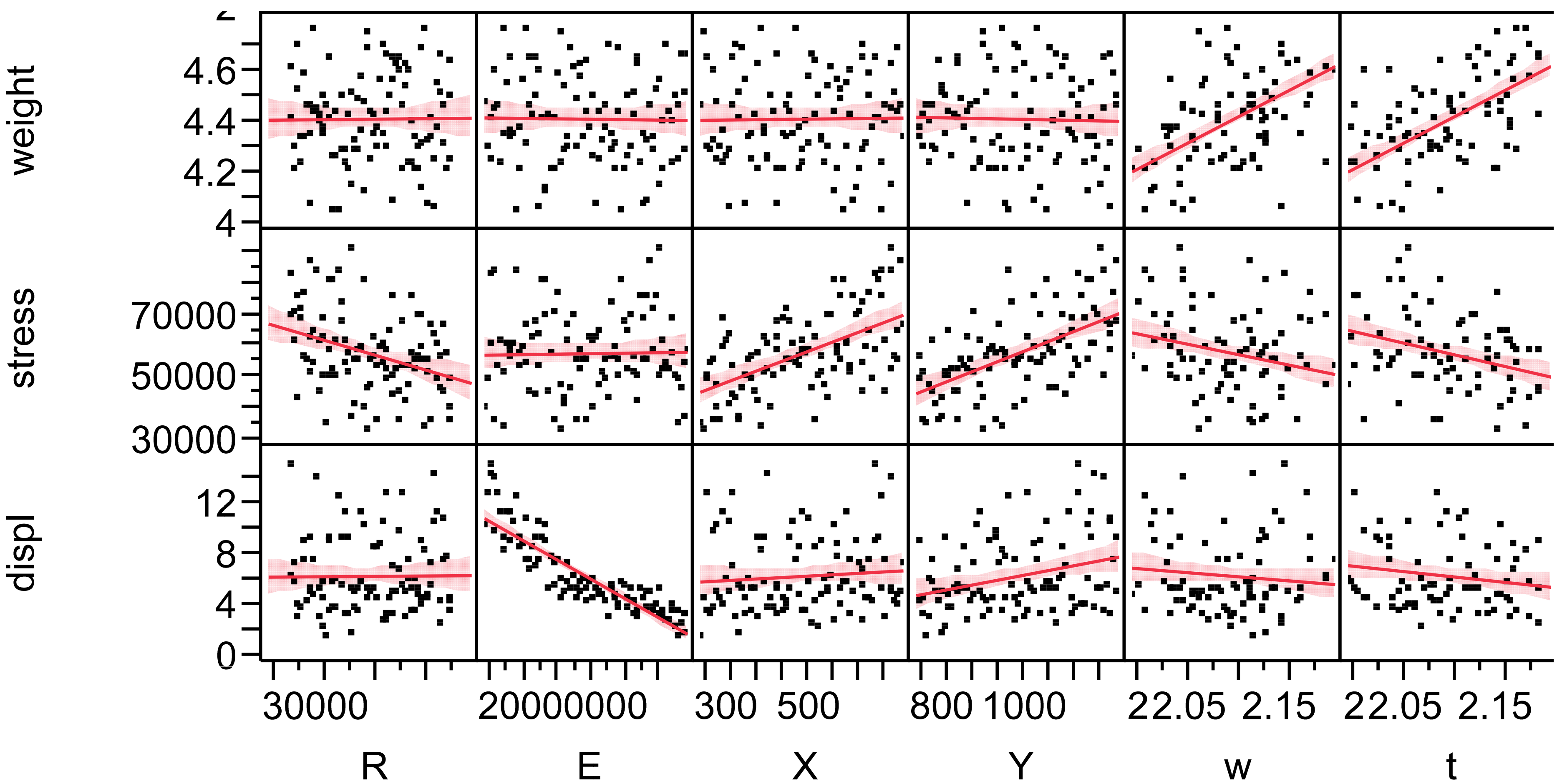
Beam width and thickness are important contributors to all outputs, several other variables also rate highly on partial correlations.

Exercise: Input/Output Scatter Plots for Cantilever

The `dakota_tabular.dat` file can be used in Mintab, JMP, Excel, etc., to generate scatter plots



Exercise: Input/Output Scatter Plots for Cantilever



Additional Sensitivity Analysis Capabilities

- **Variance-based decomposition (via sampling or PCE)**
  - **Goal: Apportion uncertainty in responses to uncertainty in inputs**
  - **Expensive: K*(N+2) simulations required, K = # samples, N = # variables, recommended K ≥ 100**
  - **Exercise: Modify the sensitivity analysis method to perform variance-based decomposition on the cantilever problem**
- **Main Effects/Analysis of Variance (ANOVA)**
  - **Goal: Determine effect of a variable on mean behavior**
  - **Uses design of experiments: Coverage of space (e.g., space filling, interior, boundaries/extremes, etc.) varies by design**
  - **Exercise: Modify the sensitivity analysis method to perform a main effects analysis using an orthogonal array on the cantilever problem**

Exercise: Explore Other SA Methods for Cantilever

SAME INPUT DECK,
JUST CHANGE METHOD SPEC

```
method,
sampling
   sample_type lhs
   seed =52983
   samples = 100
```

LHS Sampling

```
method,
   dace oas
   main_effects
   seed =52983
   samples = 500
```

Main Effects Analysis using Orthogonal Arrays

```
method,
sampling
   sample_type lhs
   seed =52983
   samples = 500
   variance_based_decomp
```

Variance-based Decomposition using LHS Sampling

```
method,
   psuade_moat
   partitions = 3
   seed =52983
   samples = 100
```

Morris One-At-a-Time

Additional Sensitivity Analysis Capabilities

Global sensitivity indices for each response function:			
weight	Sobol	indices:	
	Main	Total	
	0.00	0.00	**R**
	0.00	0.00	**E**
	0.00	0.00	**X**
	0.00	0.00	**Y**
	0.49	0.51	**beam_width**
	0.51	0.52	**beam_thickness**
stress	Sobol	indices:	
	Main	Total	
	0.16	0.13	**R**
	0.00	0.00	**E**
	0.37	0.36	**X**
	0.39	0.36	**Y**
	0.08	0.08	**beam_width**
	0.11	0.12	**beam_thickness**
displ	Sobol	indices:	
	Main	Total	
	0.00	0.00	**R**
	0.90	0.92	**E**
	0.02	0.02	**X**
	0.07	0.08	**Y**
	0.02	0.01	**beam_width**
	0.04	0.05	**beam_thickness**

Variance-based decomposition

Response Function 1

ANOVA	Table	for	Factor	(Variable)	4	
Source	of	Sum	of	Mean	Sum	
Variation	DoF	Squares	of	Squares	Fdata	
Between	Groups	22	2.18E-03	9.89E-05	**3.22E-03**	**Y**
Within	Groups	506	1.55E+01	3.07E-02		
Total	528	1.55E+01				
ANOVA	Table	for	Factor	(Variable)	5	
Source	of	Sum	of	Mean	Sum	
Variation	DoF	Squares	of	Squares	Fdata	
Between	Groups	22	7.80E+00	3.55E-01	**2.32E+01**	**Beam Width**
Within	Groups	506	7.73E+00	1.53E-02		
Total	528	1.55E+01				
ANOVA	Table	for	Factor	(Variable)	6	
Source	of	Sum	of	Mean	Sum	
Variation	DoF	Squares	of	Squares	Fdata	
Between	Groups	22	7.70E+00	3.50E-01	**2.26E+01**	**Beam Thickness**
Within	Groups	506	7.84E+00	1.55E-02		
Total	528	1.55E+01				

Main Effects Analysis

Same relative ranking across methods

DAKOTA Sensitivity Analysis

- Parameter study, design and analysis of computer experiments, and general sampling methods:
  - Single and multi-parameter studies (grid, vector, centered)
  - DDACE (grid, sampling, orthogonal arrays, Box-Behnken, CCD)
  - FSUDACE (Quasi-MC, CVT)
  - PSUADE (Morris designs)
  - Monte Carlo, Latin hypercube sampling (with correlation or variance analysis, including variance-based decomposition)
  - Mean-value with importance factors

- DAKOTA outputs basic statistics on responses, including mean, standard deviation, and correlations; tabular output can be analyzed with any third-party statistics package
- Determine main effects and key parameter *interactions*
- In SA, one typically does not make a distribution assumption