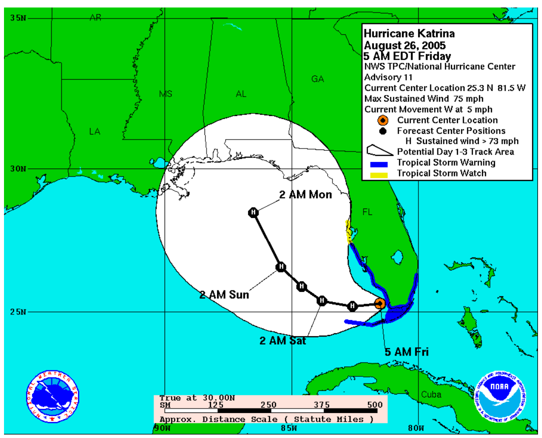
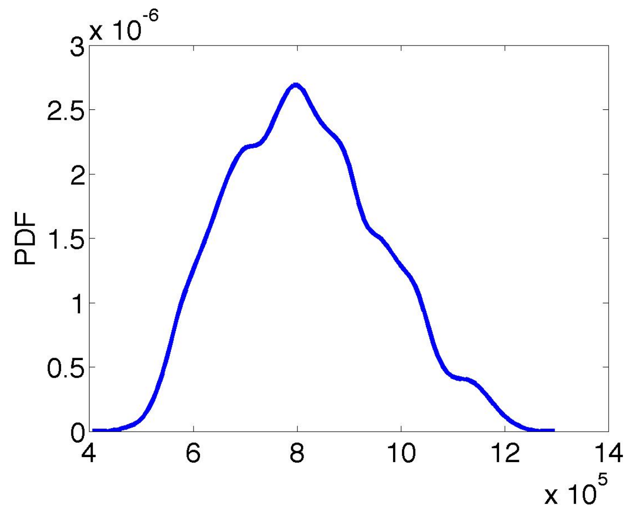
Global Sensitivity Analysis, Active Subspaces, and Response Surface Construction

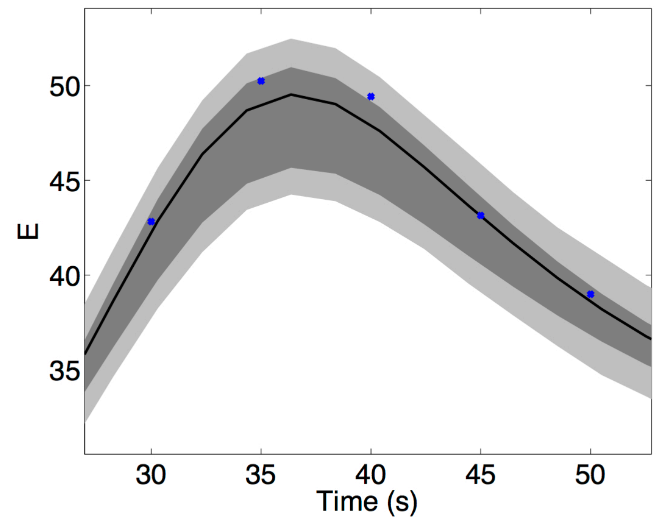
Ralph C. Smith

Department of Mathematics

North Carolina State University







Support: Air Force grant AFOSR FA9550-15-0299

NSF CMMI-1306290

DOE Consortium for Advanced Simulation of LWR (CASL)

NNSA Consortium for Nonproliferation Enabling Capabilities (CNEC)

Parameter Selection Techniques

Motivation: Consider spring model

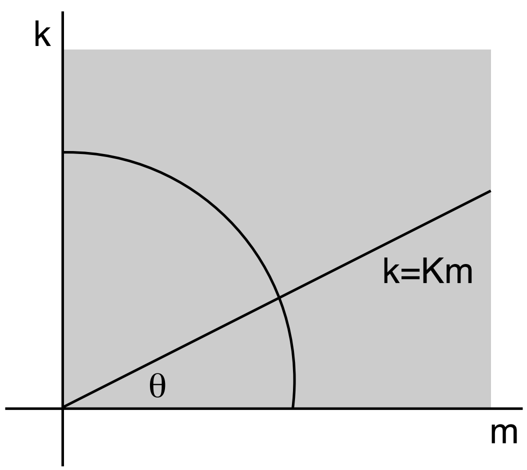
$$m\frac{d^2z}{dt^2} + kz = 0$$

$$z(0) = z_0\ ,\ \ \frac{dz}{dt}(0) = 0$$

with solution $z(t) = z_0 \cos(\sqrt{k/m} \cdot t)$.

Observation: Parameters q = [k,m] not uniquely determined by displacement data.

Admissible Parameter Space: $\mathbb{Q} = (0,\infty) \times (0,\infty)$



Note: Determination of slope equivalent to specifying θ

$$I(q) = \{\theta = \arctan(k/m) \,|\, 0 < \theta < \pi/2\}$$

$$NI(q) = \left\{r = \sqrt{k^2+m^2} \,|\, r > 0\right\}$$

Note: $\mathbb{Q} = I(q) \oplus NI(q)$

Global Sensitivity Analysis

Example: Portfolio model

$$Y = c_1 Q_1 + c_2 Q_2$$

Note:

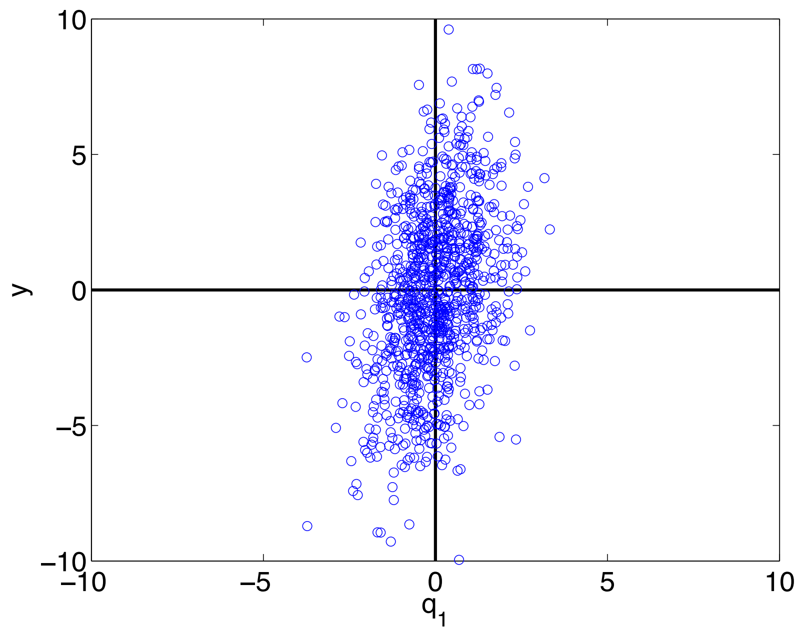
- Q_1 and Q_2 represent hedged porfolios
- c_1 and c_2 amounts invested in each portfolio

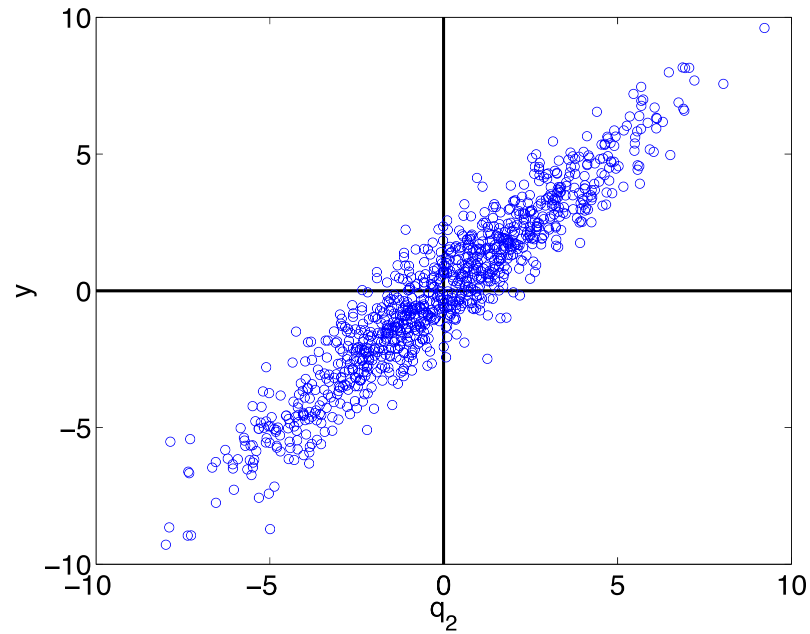
Take

$$c_1 = 2 \; , \;\; c_2 = 1$$

$Q_1 \sim N(0, \sigma_1^2)$ with $\sigma_1 = 1$

$Q_2 \sim N(0, \sigma_2^2)$ with $\sigma_2 = 3$





Local Sensitivities: Alun Lloyd

$$s_i \equiv \frac{\partial Y}{\partial Q_i} \Rightarrow s_1 = 2 > s_2 = 1$$

Solutions:

- Response correlation
- Variance methods
- Random sampling of local sensitivities

Variance-Based Methods

Variances:

$$D_i = \int_0^1 f_i^2(q_i)dq_i$$

$$D_{ij} = \int_0^1 \int_0^1 f_{ij}^2(q_i, q_j)dq_i dq_j.$$

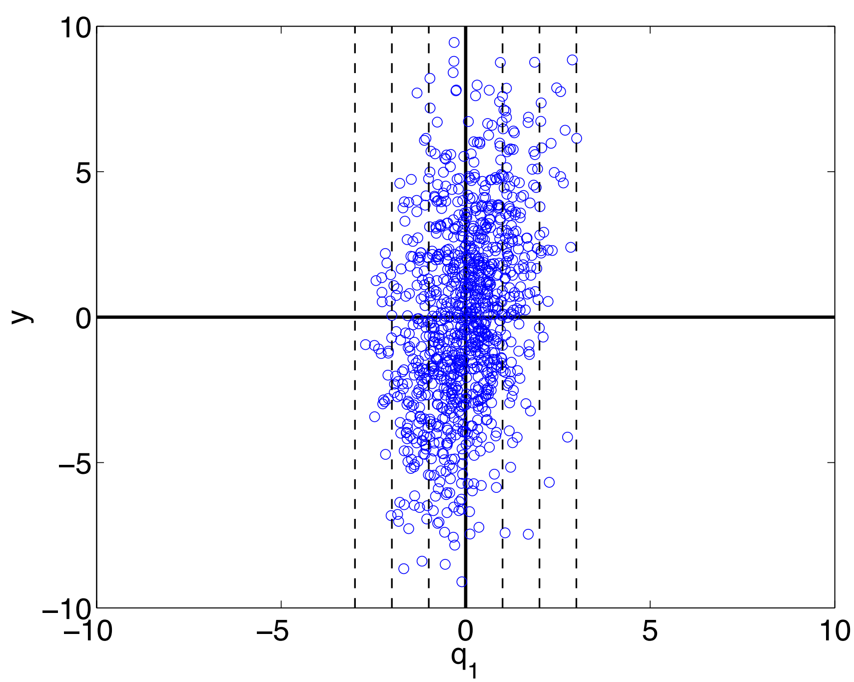
$$D = \text{var}(Y) = \int_\Gamma f^2(q)dq - f_0^2$$

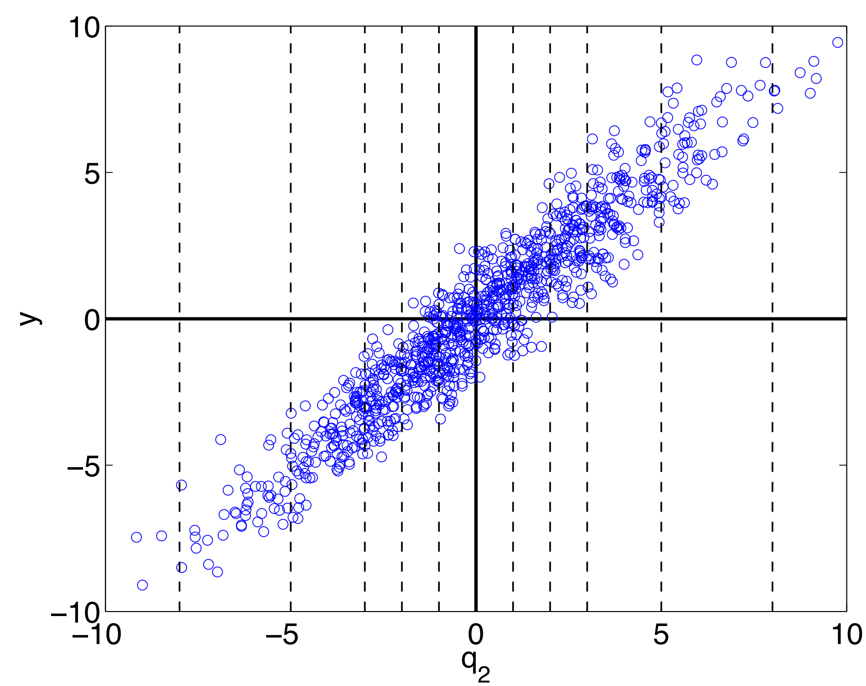
Sobol Indices:

$$S_i = \frac{D_i}{D} \quad , \; S_{ij} = \frac{D_{ij}}{D} \quad , \; i,j = 1, \cdots, p$$

$$S_{T_i} = S_i + \sum_{j=1}^{p} S_{ij}$$

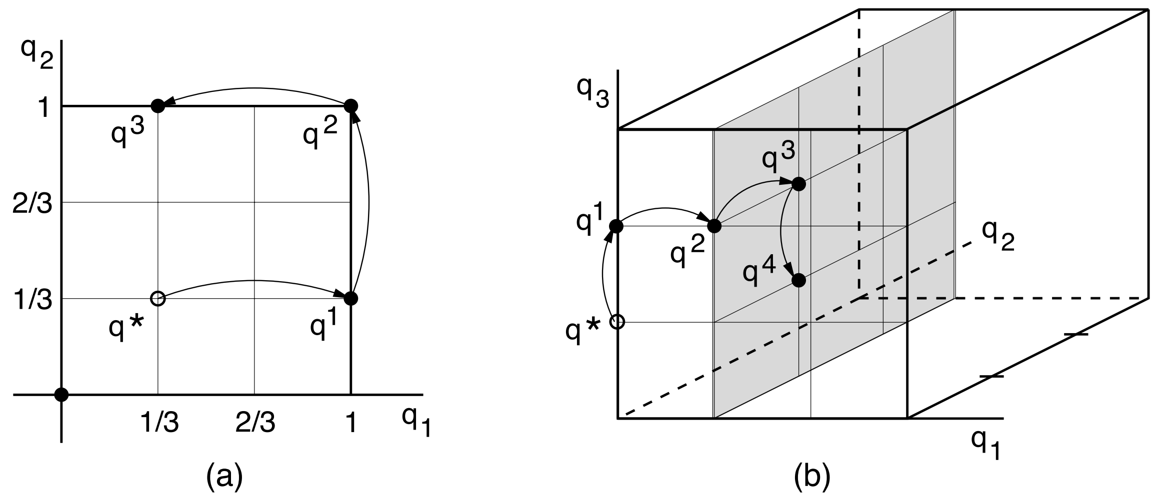
Statistical Interpretation: $D_i = \text{var}[\mathbb{E}(Y|q_i)] \Rightarrow S_i = \dfrac{\text{var}[\mathbb{E}(Y|q_i)]}{\text{var}(Y)}$





Morris Screening

Example: Consider uniformly distributed parameters on $\Gamma = [0,1]^p$



Elementary Effect:

$$d_i^j = \frac{f(q^j + \Delta e_i) - f(q^j)}{\Delta} \qquad i^{th} \text{ parameter, } j^{th} \text{ sample}$$

$$\Delta \in \left\{ \frac{1}{\ell - 1}, \cdots, 1 - \frac{1}{\ell - 1} \right\} \qquad \ell \text{ is level}$$

Global Sensitivity Measures: r samples

$$\mu_i^* = \frac{1}{r} \sum_{j=1}^{r} |d_i^j(q)|$$

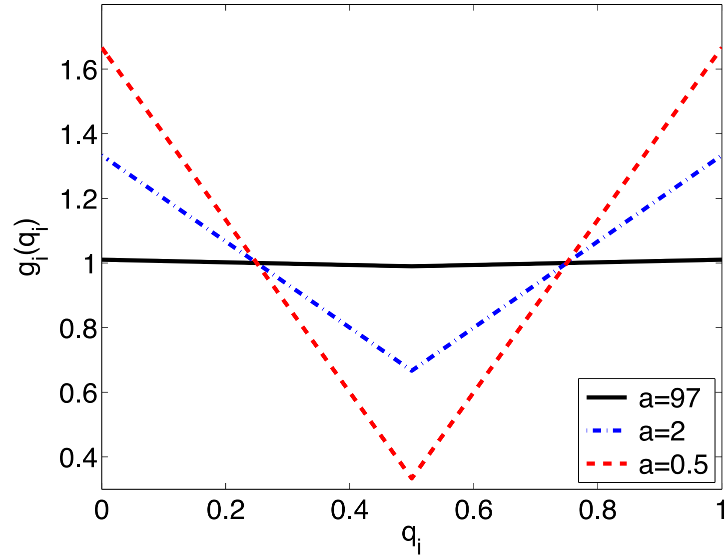
$$\sigma_i^2 = \frac{1}{r-1} \sum_{j=1}^{r} \left(d_i^j(q) - \mu_i \right)^2 \ , \ \mu_i = \frac{1}{r} \sum_{j=1}^{r} d_i^j(q)$$

Global Sensitivity Analysis

Example: Sobol function

$$Y = \prod_{i=1}^{p} g_i(Q_i) \quad , \; g_i(Q_i) = \frac{|4Q_i - 2| + a_i}{1 + a_i}$$

where $a_i \geq 0$ are fixed, deterministic coefficients that determine relative importance of parameters.



Note: For $Q_i \sim \mathcal{U}(0,1), i = 1, \cdots, p$.

$$D_i = \text{var}[\mathbb{E}(Y|q_i)] = \frac{1}{3(1+a_i)^2}$$

$$D_{ij} = \text{var}[\mathbb{E}(Y|q_i, q_j)] - D_i - D_j = D_i D_j$$

$$D = \text{var}(Y) = -1 + \prod_{i=1}^{p}(1 + D_i)$$

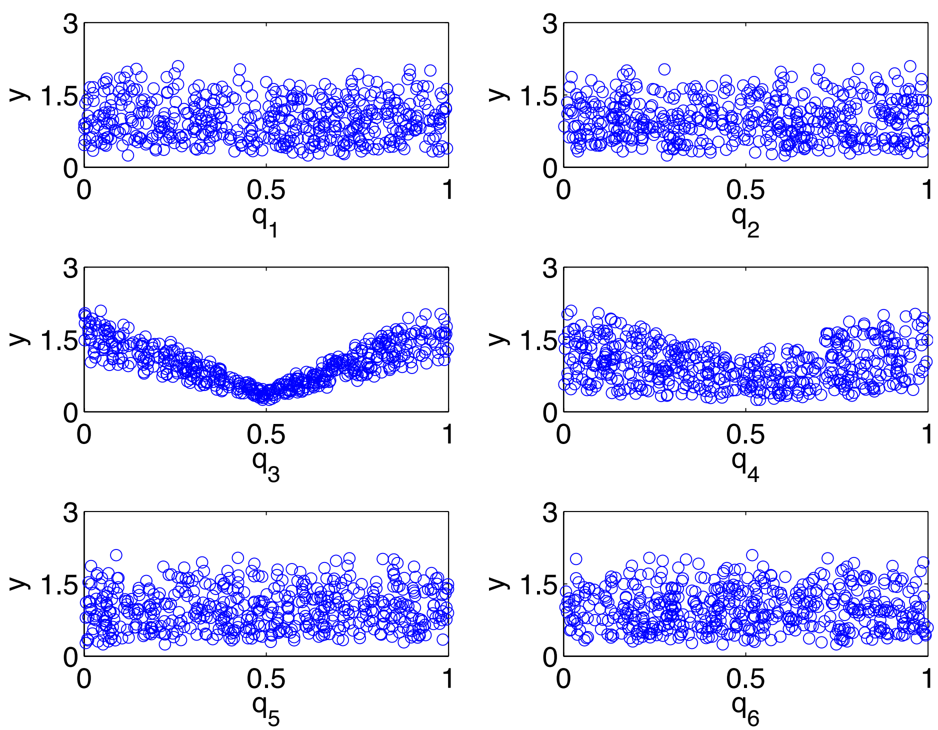
Global Sensitivity Analysis

Sobol Indices:

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6
a_i	78	12	0.5	2	97	33
D_i	5.3×10^{-5}	2.0×10^{-3}	1.5×10^{-1}	3.7×10^{-2}	3.5×10^{-5}	2.9×10^{-4}
S_i	2.8×10^{-4}	1.0×10^{-2}	7.7×10^{-1}	1.9×10^{-1}	1.8×10^{-4}	1.5×10^{-3}
S_{T_i}	3.3×10^{-4}	1.2×10^{-2}	8.0×10^{-1}	2.2×10^{-1}	2.1×10^{-4}	1.8×10^{-3}

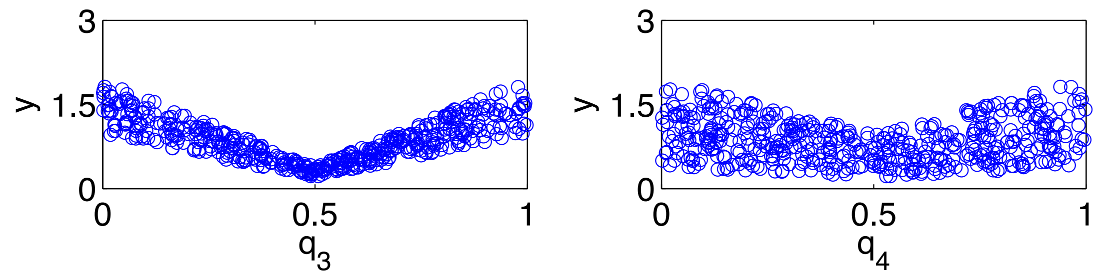
Morris Indices: With $\ell = 4, \Delta = \frac{2}{3}, r = 4$

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6
μ_i	-0.006	-0.078	-0.130	-0.004	0.012	-0.004
μ_i^*	0.056	0.277	1.760	1.185	0.035	0.099
σ_i	0.064	0.321	2.049	1.370	0.041	0.122



Insensitive Parameters: Take

$$q_1 = q_2 = q_5 = q_6 = \frac{1}{2}$$



SIR Disease Example

SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \gamma k I S \quad , \; S(0) = S_0 \qquad \text{Susceptible}$$

$$\frac{dI}{dt} = \gamma k I S - (r + \delta) I \quad , \; I(0) = I_0 \qquad \text{Infectious}$$

$$\frac{dR}{dt} = rI - \delta R \quad , \; R(0) = R_0 \qquad \text{Recovered}$$

Note: Parameter set $q = [\gamma, k, r, \delta]$ is not identifiable

Assumed Parameter Distribution:

$$\gamma \sim \mathcal{U}(0,1) \; , \; k \sim \mathrm{Beta}(\alpha, \beta) \; , \; r \sim \mathcal{U}(0,1) \; , \; \delta \sim \mathcal{U}(0,1)$$

Infection Coefficient	Interaction Coefficient	Recovery Rate	Birth/death Rate

Response:

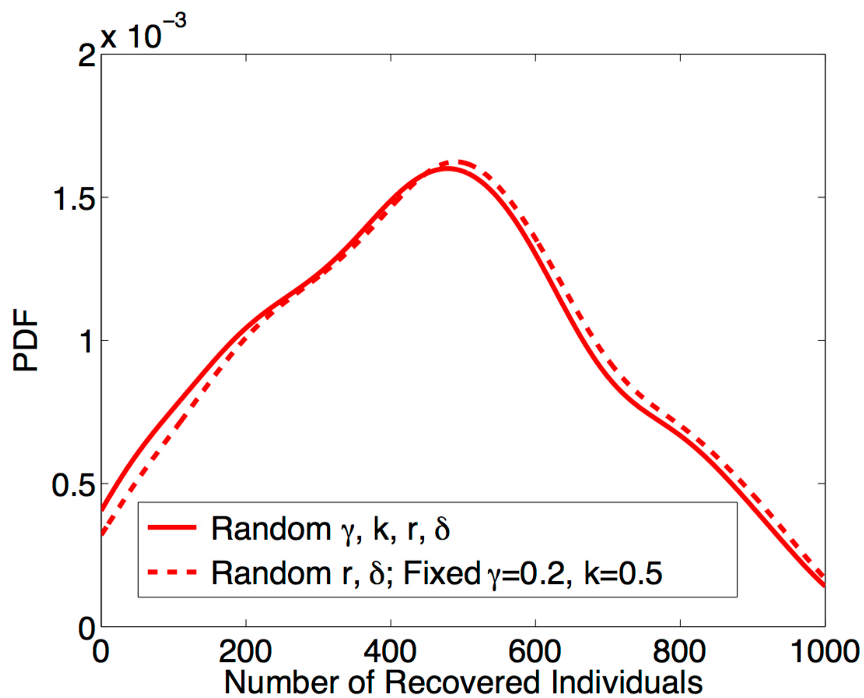
$$y = \int_0^5 R(t,q)\,dt$$

SIR Disease Example

Global Sensitivity Measures:

		γ	k	r	δ
Sobol	S_i	0.0997	0.0312	0.7901	0.1750
	S_{T_i}	-0.0637	-0.0541	0.5634	0.2029
	μ_i^* $(\times 10^3)$	0.2532	0.2812	2.0184	1.2328
Morris	σ_i $(\times 10^3)$	0.9539	1.6245	6.6748	3.9886

Result: Large number of interactions: $k \sim \text{Beta}(2, 7)$



Exercise 1: Reproduce these results using the following codes

- SIR_Saltelli
- SIR_morris.m

Exercise 2: To simulate with few interactions, run with

$$k \sim \text{Beta}(0.2, 15)$$

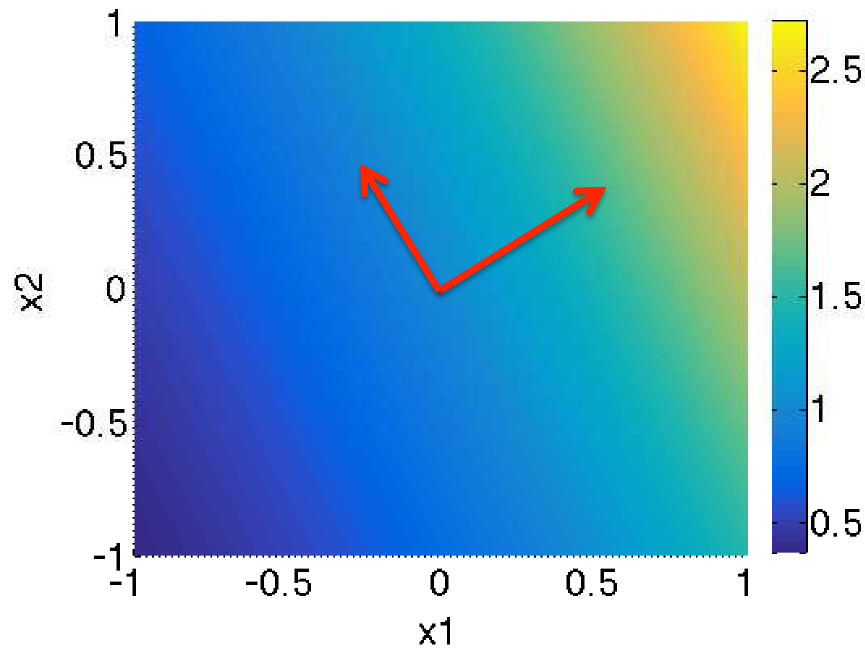
Active Subspace Construction

Note:

- Functions may vary significantly in only a few directions
- "Active" directions may be linear combination of inputs

Example: $y = \exp(0.7x_1 + 0.3x_2)$

- Varies most in [0.7, 0.3] direction
- No variation in orthogonal direction



Strategy:

- *Linearly parameterized problems:* Employ SVD or QR decomposition;
- *Nonlinear problems:* Construct approximate gradient matrix and employ SVD or QR.

Parameter Space Reduction Techniques: Linear Problems

Second Issue: Models depends on very large number of parameters – e.g., millions – but only a few are “significant”.

Linear Algebra Techniques: Linearly parameterized problems

$$y = Aq \ , \ q \in \mathbb{R}^p \ , \ y \in \mathbb{R}^m$$

Singular Value Decomposition (SVD):

$$A = U\Sigma V^T \ , \ \Sigma = [S \quad 0]$$

$$S = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix} \ , \ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq \varepsilon$$

Rank Revealing QR Decomposition: $A^T P = QR$

Problem: Neither is directly applicable when m or p are very large; e.g., millions.

Solution: Random range finding algorithms.

Random Range Finding Algorithms: Linear Problems

Algorithm: Halko, Martinsson and Tropp, SIAM Review, 2011

1. Choose ℓ random inputs q^i and compute outputs $y^i = Aq^i$ which are compiled in the $m \times \ell$ matrix Y.

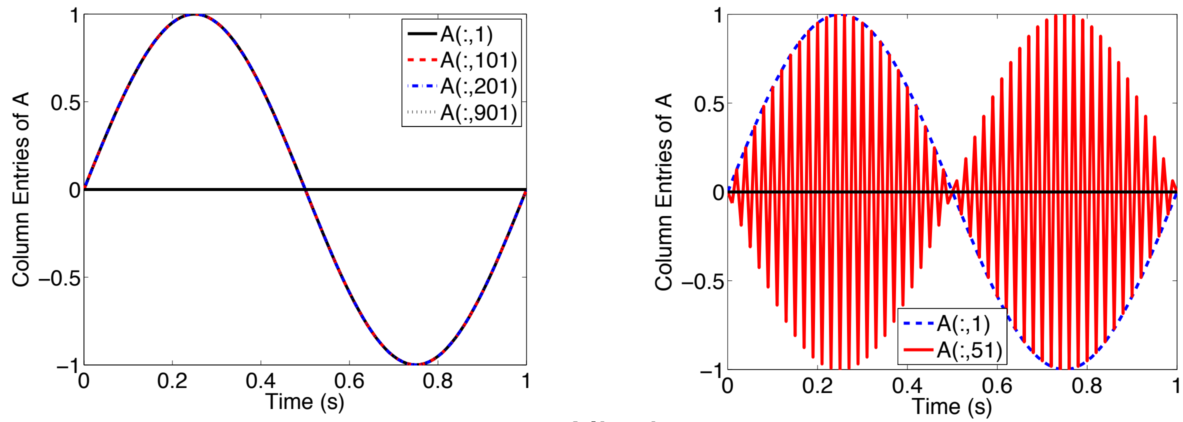
2. Take a pivoted QR factorization $Y = QR$ to construct a matrix Q whose columns form an orthonormal basis for the range of Y.

Example: $y_i = \sum_{k=1}^{p} q_k \sin(2\pi k t_i) \; , \; i = 1, \cdots, m$

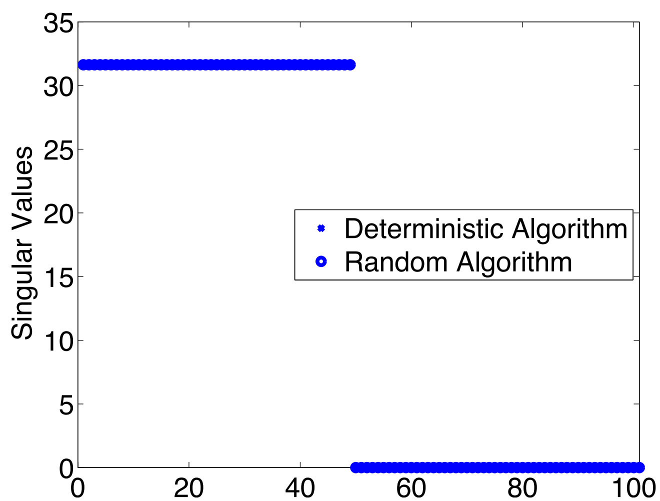
$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \sin(2\pi t_1) & \cdots & \sin(2\pi p t_1) \\ \vdots & & \vdots \\ \sin(2\pi t_m) & \cdots & \sin(2\pi p t_m) \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \\ q_p \end{bmatrix}$$

Random Range Finding Algorithms: Linear Problems

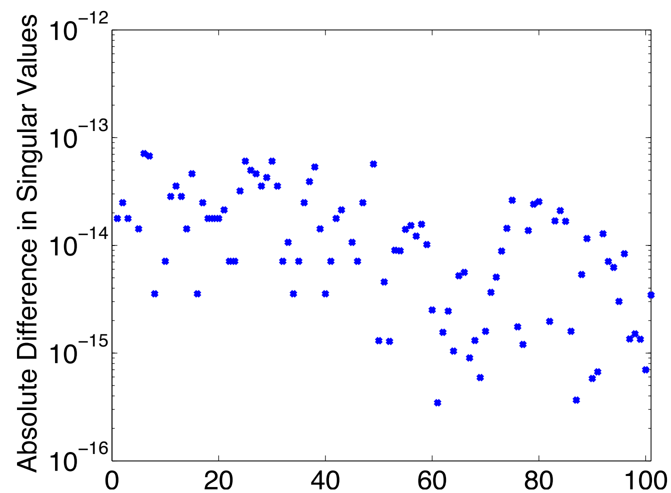
Example: m = 101, p = 1000: Analytic value for rank is 49



Aliasing



Singular Values

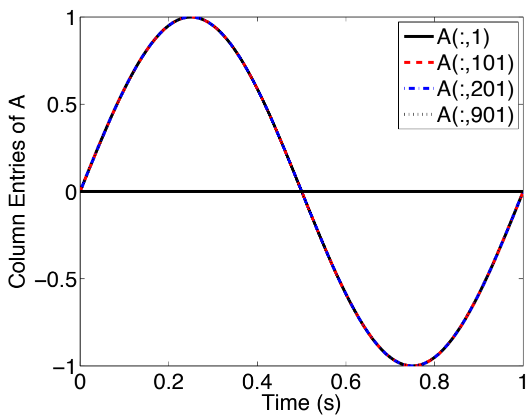


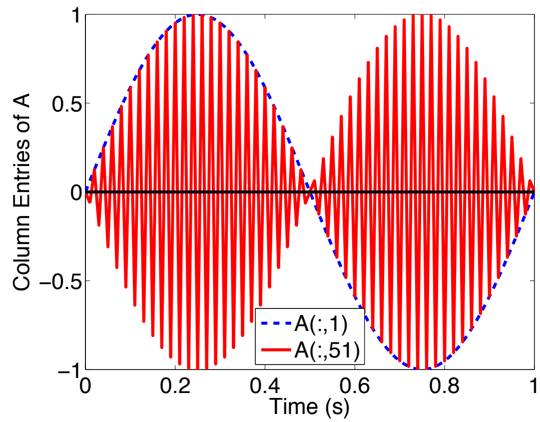
Absolute Difference Between Singular Values

Example: m = 101, p = 1,000,000: Random algorithm still viable

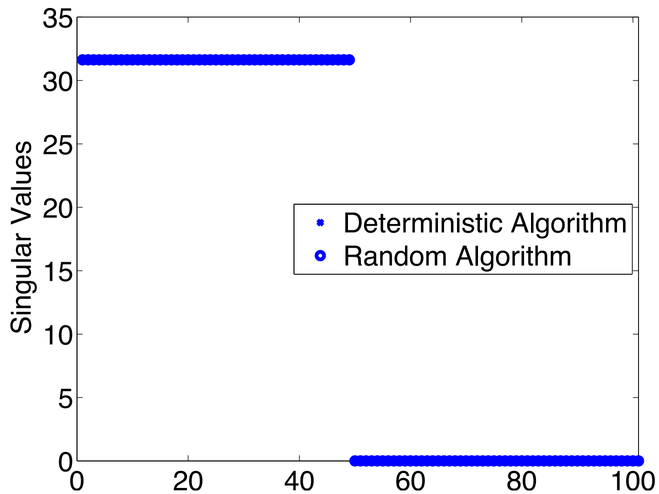
Random Range Finding Algorithms: Linear Problems

Example: m = 101, p = 1000

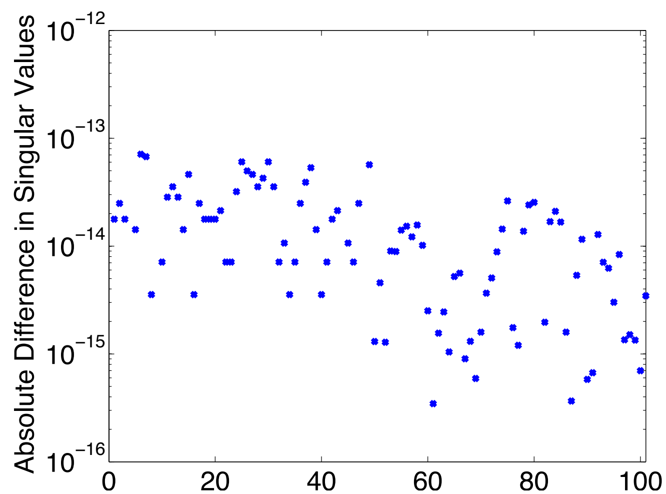




Aliasing



Singular Values



Absolute Difference Between Singular Values

Exercise 3:

- Reproduce these results using the code aliasing.m. You can experiment with the size of the random sample.
- Modify the code to use R to compute the rank.
- See how far you can increase the dimension.

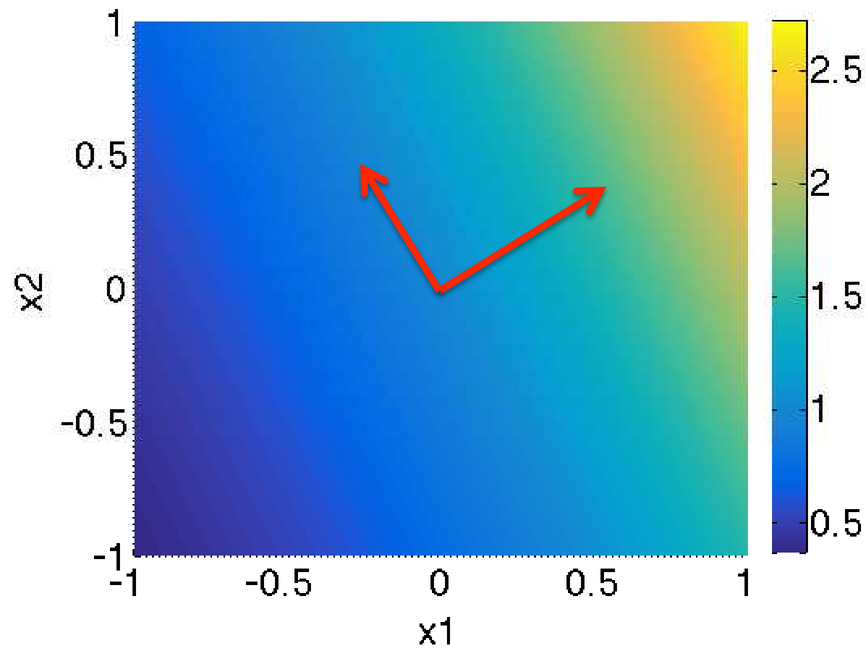
Active Subspace Construction

Note:

- Functions may vary significantly in only a few directions
- “Active” directions may be linear combination of inputs

Example: $y = \exp(0.7x_1 + 0.3x_2)$

- Varies most in [0.7, 0.3] direction
- No variation in orthogonal direction



Strategy:

- *Linearly parameterized problems:* Employ SVD or QR decomposition;
- *Nonlinear problems:* Construct approximate gradient matrix and employ SVD or QR.

Gradient-Based Active Subspace Construction

Active Subspace: See [Constantine, SIAM 2015]. Consider

$$f = f(\mathbf{x}),\ \mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^m$$

and

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = \left[\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_m}\right]^T$$

Construct outer product

$$\mathbf{C} = \int (\nabla_x f)(\nabla_x f)^T \rho dx$$

Partition eigenvalues: $\mathbf{C} = \mathbf{W} \Lambda \mathbf{W}^T$

$$\Lambda = \begin{bmatrix} \Lambda_1 & \\ & \Lambda_2 \end{bmatrix},\ \mathbf{W} = [\mathbf{W}_1\ \mathbf{W}_2]$$

Rotated Coordinates:

$$\mathbf{y} = \mathbf{W}_1^T \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{z} = \mathbf{W}_2^T \mathbf{x} \in \mathbb{R}^{m-n}$$

Active Variables

Active Subspace: Range of eigenvectors in $\mathbf{W}_1$

Gradient-Based Active Subspace Construction

Active Subspace: Construction based on random sampling

1. Draw M independent samples $\{\mathbf{x}^j\}$ from ρ
2. Evaluate $\nabla_{\mathbf{x}} f_j = \nabla_{\mathbf{x}} f(\mathbf{x}^j)$
3. Approximate outer product

$$C \approx \widetilde{C} = \frac{1}{M} \sum_{j=1}^{M} (\nabla_{\mathbf{x}} f_j)(\nabla_{\mathbf{x}} f_j)^T$$

Note: $\widetilde{C} = GG^T$ where $G = \frac{1}{\sqrt{M}}[\nabla_{\mathbf{x}} f_1, \cdots, \nabla_{\mathbf{x}} f_M]$

4. Take SVD of $G = \mathbf{W}\sqrt{\mathbf{\Lambda}}\mathbf{V}^T$
   - Active subspace of dimension n is first n columns of W

Research: Develop efficient algorithm for codes that do not have adjoint capabilities

Note: Finite difference approximations tempting but not very effective

Strategy: Algorithm based on adaptive Morris indices

Gradient-Based Active Subspace Construction

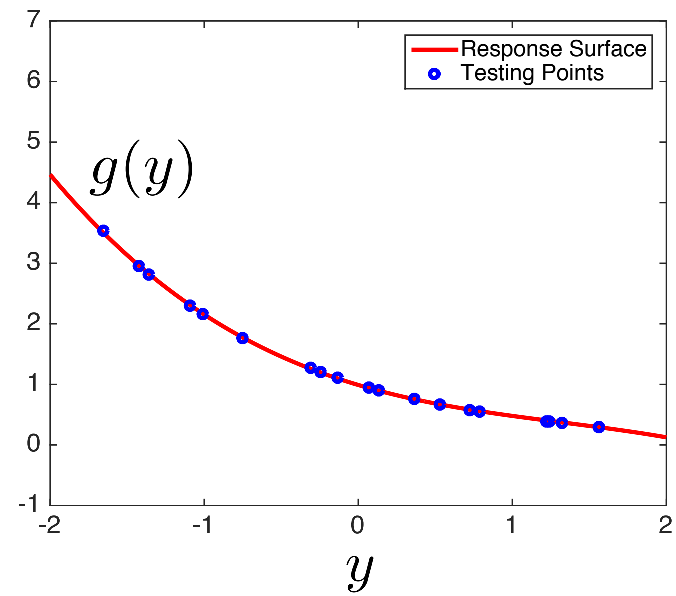
Example: $y = \exp(0.7x_1 + 0.3x_2)$

Active Subspace: For gradient matrix G, form SVD

$$G = U\Lambda V^T$$

Eigenvalue spectrum indicates 1-D active subspace with basis

$$U(:,1) = [0.91 \ , \ 0.39]$$



Exercise 4:

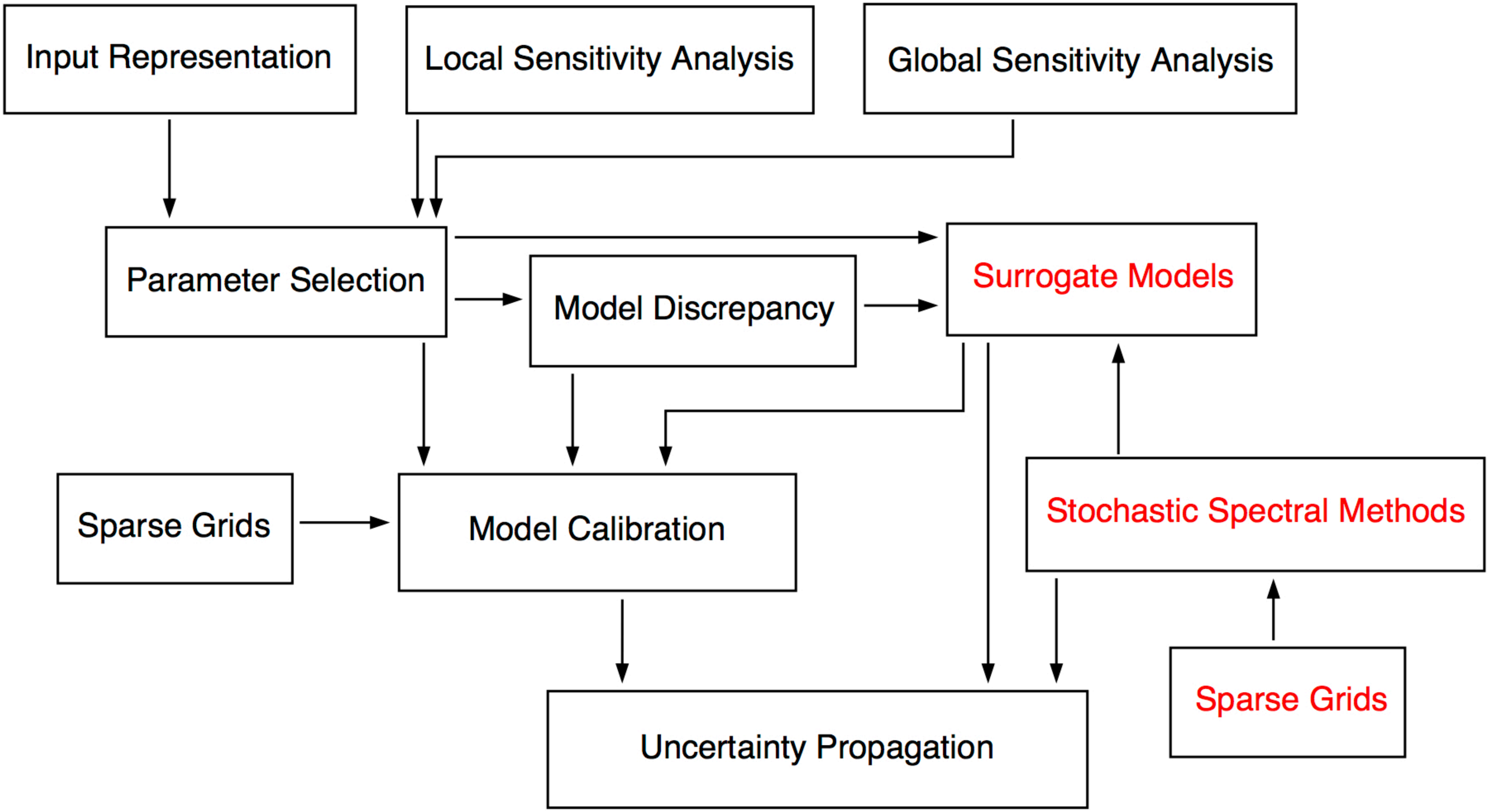
- Construct a gradient matrix G and evaluate it at random points in the interval

  $$[-1,1] \times [-1,1]$$

  U(-1,1). Now take the SVD to determine a basis for the active subspace.

- How might you construct a response surface g(y) on the 1-D subspace y?

Steps in Uncertainty Quantification



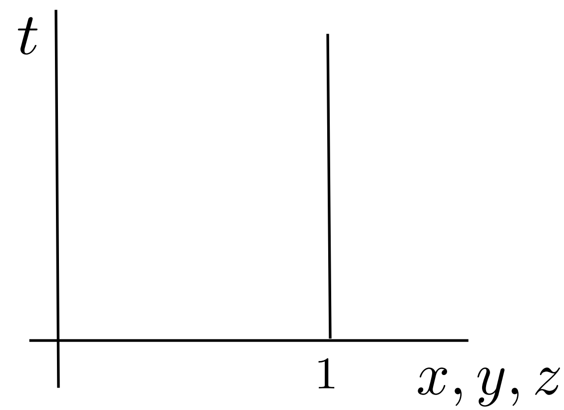
Surrogate Models

Recall: Consider the model

$$\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = f(q)$$

$$\text{Boundary Conditions}$$

$$\text{Initial Conditions}$$

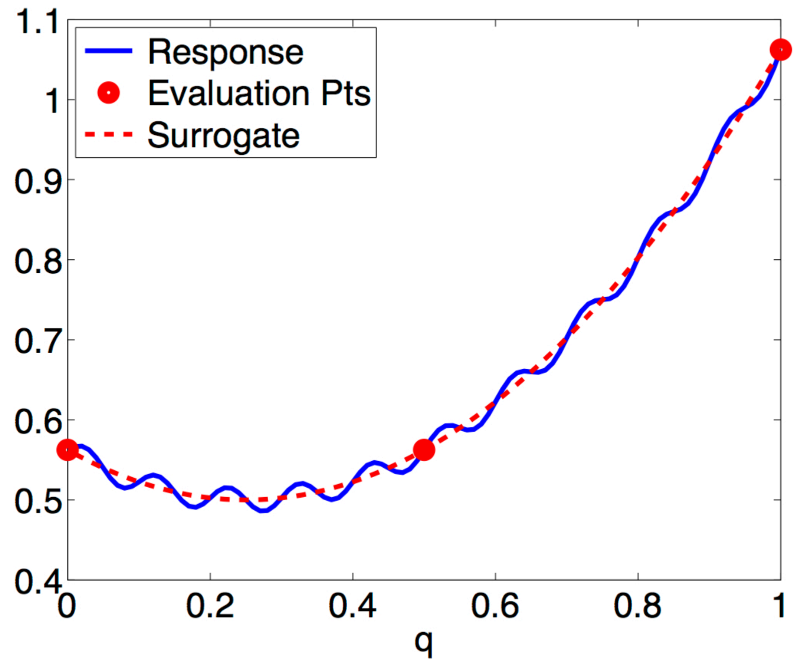


with the response

$$y(q) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 u(t, x, y, z; q) dx dy dz dt$$

Surrogate: Quadratic

$$y_s(q) = (q - 0.25)^2 + 0.5$$



Question: How do you construct a polynomial surrogate?

- Regression
- Interpolation

Statistical Model: $y_s(q)$: Emulator for $y(q)$

$$y_m = y_s(q^m) + \varepsilon_m \ , \ m = 1, \ldots, M$$

Surrogate Models

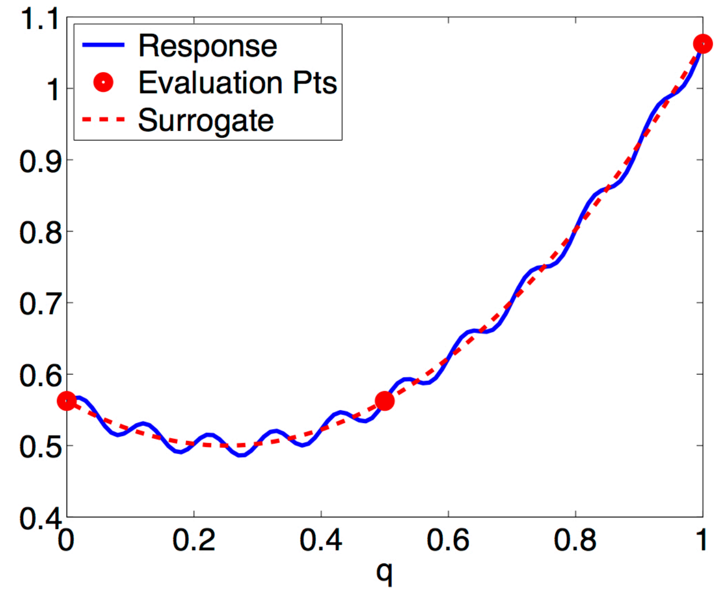
Question: How do we keep from fitting noise?

- Akaike Information Criterion (AIC)

$$AIC = 2k - 2\log\left(\frac{1}{M}\sum_{m=1}^{M}[y_m - y_s(q^m)]^2\right)$$

- Bayesian Information Criterion (AIC)

$$BIC = k\log(M) - 2\log\left(\frac{1}{M}\sum_{m=1}^{M}[y_m - y_s(q^m)]^2\right)$$



Example: $y = \exp(0.7x_1 + 0.3x_2)$

Exercise 4:

- Construct a polynomial surrogate using the code response_surface.m.
- What order seems appropriate?
- Modify the code to build a response surface for the 1-D subspace.