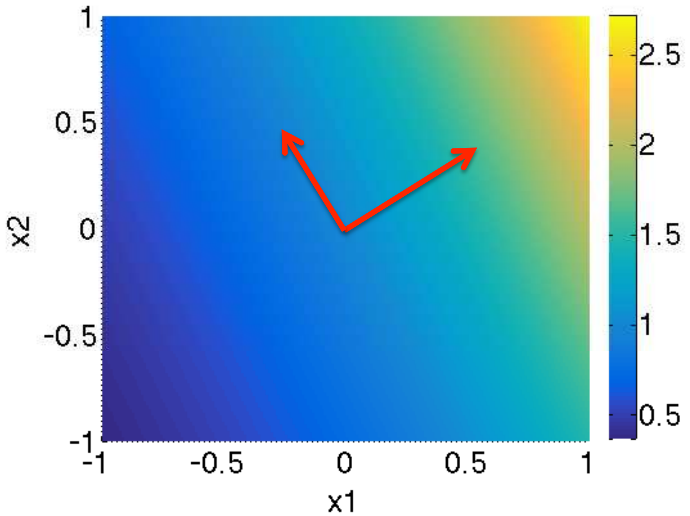
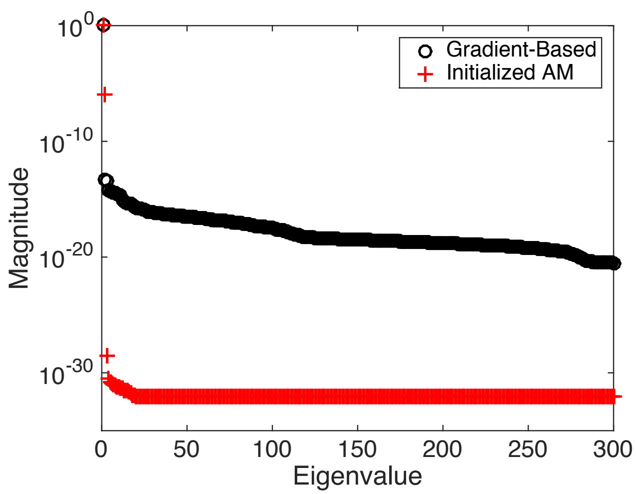
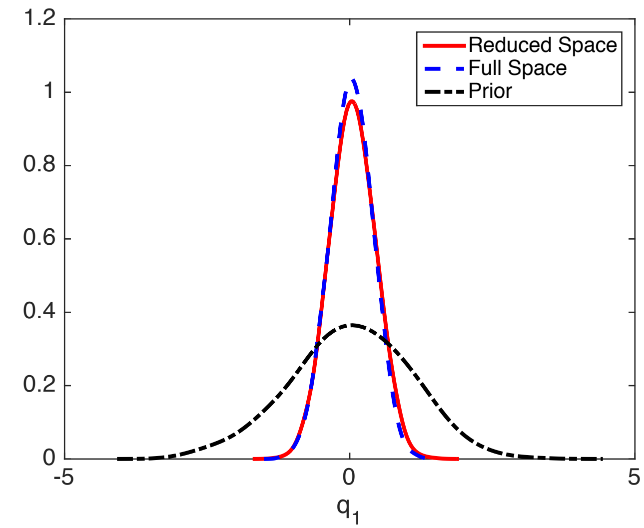
Ties between Sensitivity Analysis, Active Subspaces, Identifiability Analysis, and Parameter Subset Selection

Ralph C. Smith
Department of Mathematics
North Carolina State University







Support: DOE Consortium for Advanced Simulation of LWR (CASL)
NNSA Consortium for Nonproliferation Enabling Capabilities (CNEC)
NSF Data-Enabled Science and Engineering of Atomic Structure (SEAS)
Air Force Office of Scientific Research (AFOSR)

Example 1: SIR Model for Disease Dynamics

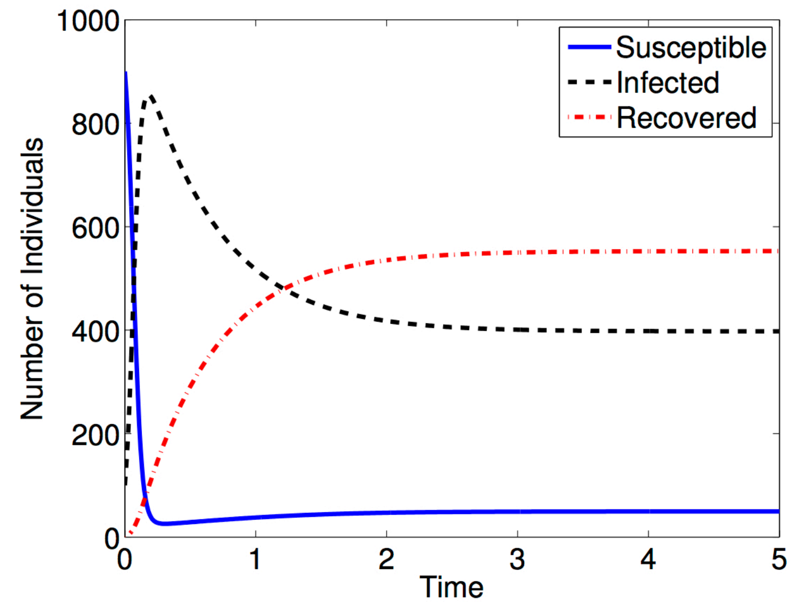
SIR Model: Population N

$$\frac{dS}{dt} = \delta N - \delta S - \gamma kIS \quad , \quad S(0) = S_0$$

$$\frac{dI}{dt} = \gamma kIS - (r + \delta)I \quad , \quad I(0) = I_0$$

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

Typical Realization



Parameters:

- γ: Infection coefficient
- k: Interaction coefficient
- r: Recovery rate
- δ: Birth/death rate

Note: Parameters $\theta = [\gamma, k, r, \delta]$ not uniquely determined by data

Goal: Use sensitivity analysis to isolate subset of influential or identifiable parameters

Example 2: Bone Model from Pharmacology

Bone Model: Employed for quantitative system pharmacology (QSP)

Subset of Equations: 8 for simplified model

Note:

- Use to study osteoporosis
- L1, L2 are lumped states representing RANK,RANKL (nuclear factor kappa-B ligand)

$$\frac{dL1}{dt} = R_{L1} + k_{OB\to L1}\cdot(\mathrm{FOB}+\mathrm{SOB}) + k_{L2\to L1}\cdot L2 + k_{CMX\to L}\cdot \mathrm{CMX} - d_{L1}\cdot L1 \tag{21}$$

$$\frac{dL2}{dt} = R_{L2} + k_{OB\to RANKL}\cdot(\mathrm{FOB}+\mathrm{SOB}) + k_{CMX\to L}\cdot \mathrm{CMX} - \left(d_{L2} + \frac{1}{3}\cdot\frac{(k_{int}-d_{RANKL})\cdot C}{K_{SS}+C}\right)\cdot L2 \tag{22}$$

$$\frac{d\mathrm{CMX}}{dt} = k_{L2\to CMX}\cdot L2 - k_{CMX\to L}\cdot \mathrm{CMX} \tag{23}$$

$$\frac{d\mathrm{OC}}{dt} = R_{OC} - d_{OC}\cdot\left(\rho_1 + (a_1-\rho_1)\frac{\mathrm{TG}F^{\gamma_1}}{\delta_1^{\gamma_1}+\mathrm{TG}F^{\gamma_1}}\right)\cdot\left(a_2 - (a_2-\rho_2)\frac{(\mathrm{CMX}/10)^{\gamma_2}}{\delta_2^{\gamma_2}+(\mathrm{CMX}/10)^{\gamma_2}}\right)\cdot \mathrm{OC} \tag{24}$$

$$\frac{d\mathrm{TGF}}{dt} = k_{OC\to TGF}\cdot \mathrm{OC} - d_{TGF}\cdot \mathrm{TGF} \tag{25}$$

$$\frac{d\mathrm{ROB}}{dt} = R_{ROB}\cdot\left(\rho_3 + (a_3-\rho_3)\frac{\mathrm{TG}F^{\gamma_3}}{\delta_3^{\gamma_3}+\mathrm{TG}F^{\gamma_3}}\right) - k_{ROB\to OB}\cdot \mathrm{ROB} \tag{26}$$

Example 2: Bone Model from Pharmacology

Subset of Parameters: 34 for simplified model

Note: Estimated values often obtained “empirically"

Goal: Determine parameter identifiability/influence for estimation/inference/UQ

Parameter	Description	Value	
		Before estimation	**After estimation (%RSE)**
R_{L1}	Production rate of *L*1	75.0	
$k_{OB \to L1}$	Rate constant expressing the effect of OB to the production of *L*1	55.3	
$k_{L2 \to L1}$	Rate constant from *L*2 to *L*1	160	
$k_{CMX \to L}$	Rate constant from CMX to lumped state (*L*1 or *L*2)	0.112	
d_{L1}	Degradation rate constant of *L*1	0.970	
R_{L2}	Production rate of *L*2	0.000160	0.00337 (9.1%)
$k_{OB \to RANKL}$	Rate constant expressing the effect of OB to the production of RANKL	0.234	
d_{L2}	Degradation rate constant of *L*2	0.00110	
d_{RANKL}	Degradation rate constant of RANKL	0.00290	
k_{int}	Elimination rate constant of the denosumab-RANKL complex	0.00795[a]	
K_{SS}	Steady-state constant for denosumab-RANKL binding affinity (ng/ml)	138[a]	63.4 (63.7%)
$k_{L2 \to CMX}$	Rate constant from *L*2 to CMX	0.0000190	
R_{OC}	Production rate of OC	0.00000298	
d_{OC}	Degradation rate constant of OC	0.0292[b]	0.0898 (5.4%)
a_1	Maximum anticipated response of TGF to the degradation of OC	2.18[b]	
ρ_1	Minimum anticipated response of TGF to the degradation of OC	0.200[b]	
δ_1	Amount of TGF that produces the half-maximal response to the degradation of OC	16.2[b]	

Motivation for Sensitivity Analysis

Motivation:

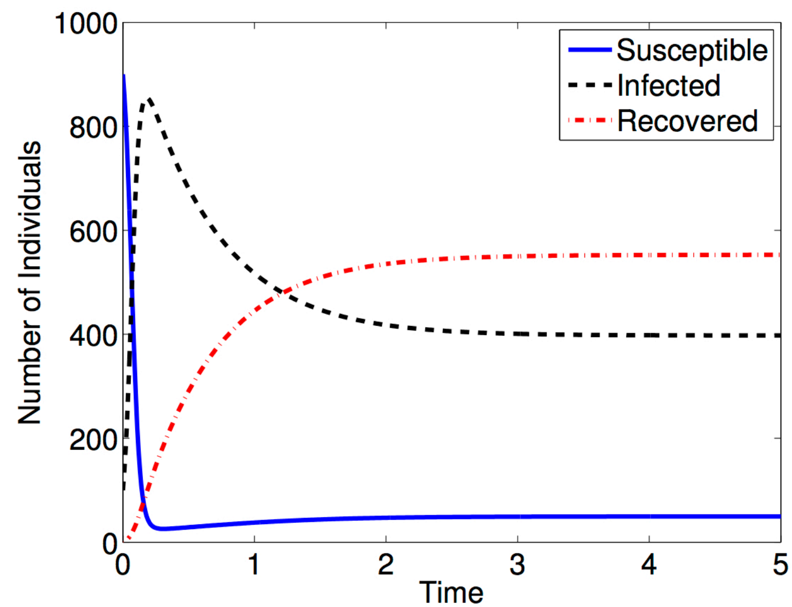
- Ascertain whether the model is robust or overly fragile with regard to certain parameters
- Determine whether the model can be simplified by fixing or freezing insensitive parameters
- Specify regions in the parameter space that optimally impact responses or their uncertainties
- Guide experimental design to determine measurement regimes that have the greatest impact on parameter or response sensitivity

SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \underline{\gamma k} IS \quad , \quad S(0) = S_0$$

$$\frac{dI}{dt} = \underline{\gamma k} IS - (r + \delta) I \quad , \quad I(0) = I_0$$

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



Local Sensitivity Analysis

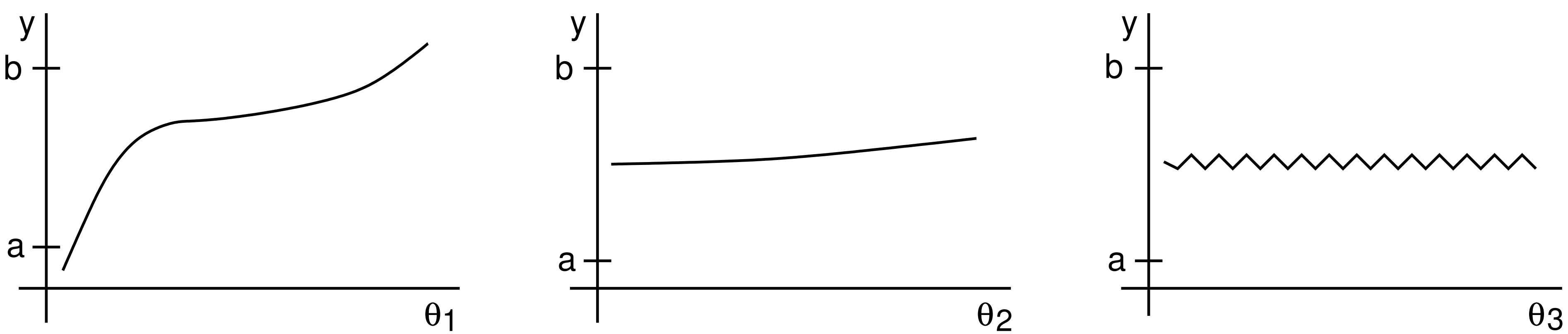
Strategy: Compare derivatives of the response with respect to parameters at a nominal value; i.e., $\frac{\partial y}{\partial \theta_i}(\theta^*)$

Advantages:

- Often fairly easy to implement using adjoints, sensitivity equations, or complex-step derivative approximations

Disadvantages:

- It is local and can be misleading for determining parameter influence
- It does not account for response and parameter uncertainties



(a) Identifiable and influential

(b) Minimally influential

(c) Minimally influential with large derivatives

Global Sensitivity Analysis

Example: Linear constitutive relation

$$\sigma = Ee + c\frac{de}{dt}$$

Nominal Values: $E = 100, \; c = 0.1$

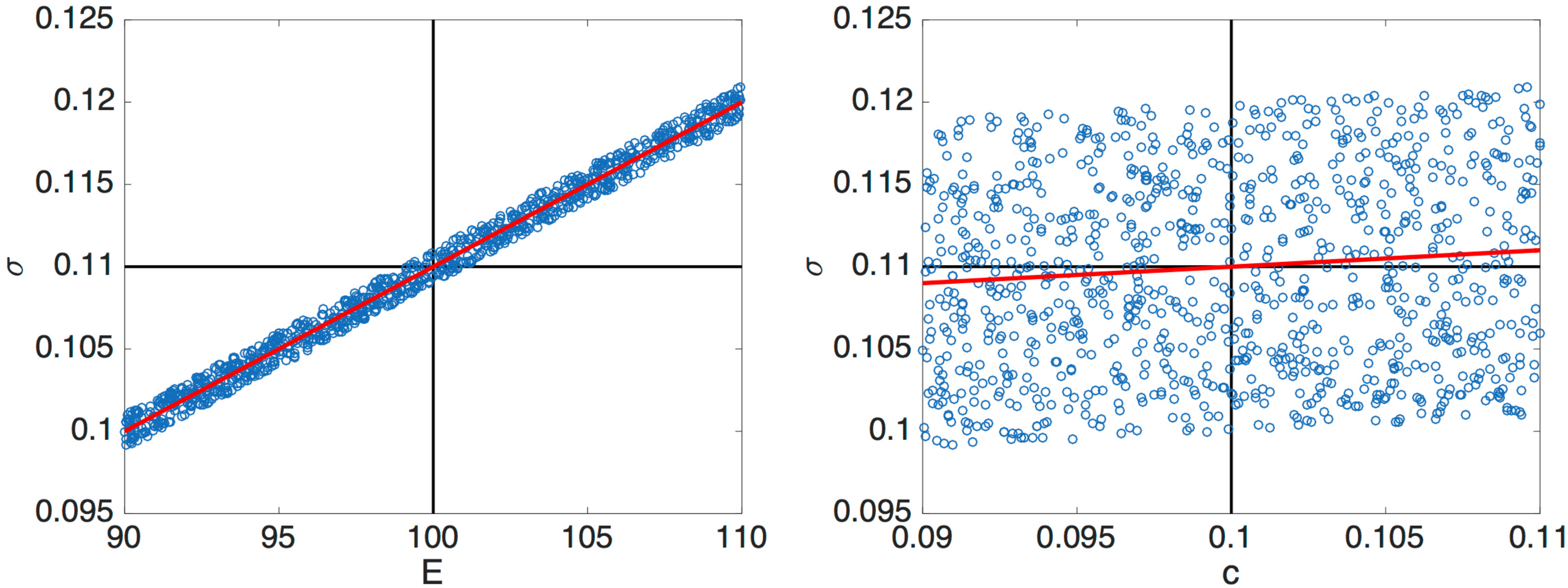
Uncertainty: 10% of nominal values

$$E \sim \mathcal{U}(90, 110) \; , \; c \sim \mathcal{U}(0.09, 0.11)$$

Local Sensitivities:

$$\frac{\partial \sigma}{\partial E} = e = 0.001$$

$$\frac{\partial \sigma}{\partial c} = \frac{de}{dt} = 0.1$$



Global Sensitivity: E is more influential

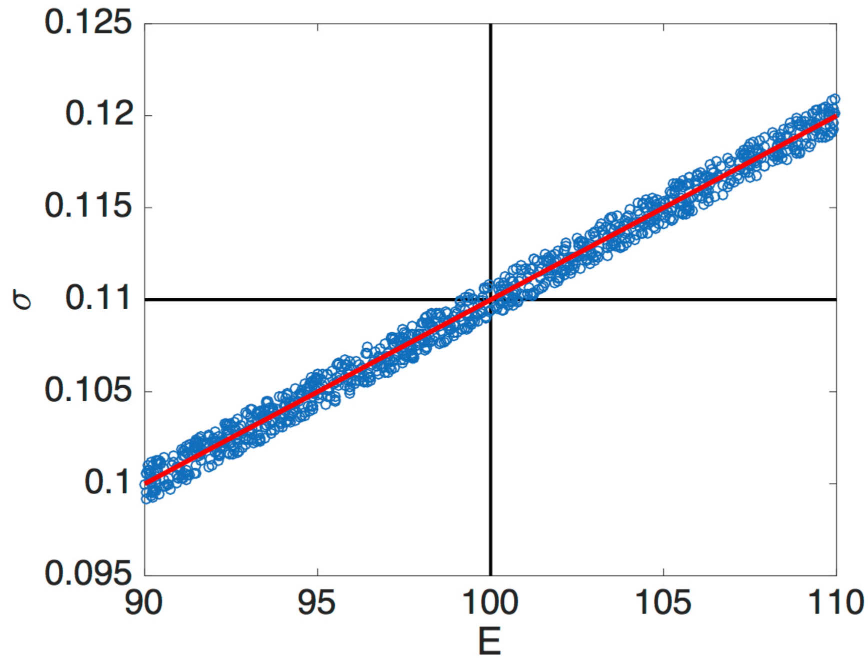
Variance-Based Methods

Sobol Representation: For now, take $q_i \sim \mathcal{U}(0,1)$ and $\Gamma = [0,1]^p$

Take

$$Y = f(q) = f_0 + \sum_{i=1}^{p} f_i(q_i) + \sum_{1 \leqslant i < j \leqslant p} f_{ij}(q_i, q_j)$$

Analogy: Taylor or Fourier series



Here

$$f_0 = \int_\Gamma f(q)dq$$

$$f_i(q_i) = \int_{\Gamma^{p-1}} f(q)dq_{\sim i} - f_0$$

Variances:

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

$$D = \text{var}(Y)$$

Sobol Indices: $S_i = \frac{D_i}{D}$

Statistical Interpretation:

$$D_i = \text{var}[\mathbb{E}(Y|q_i)] \Rightarrow S_i = \frac{\text{var}[\mathbb{E}(Y|q_i)]}{\text{var}(Y)}$$

Global Sensitivity Analysis: Analysis of Variance

Sobol' Representation: $Y = f(q)$

$$f(q) = f_0 + \sum_{i=1}^{p} f_i(q_i) + \sum_{i \leqslant i < j \leqslant p} f_{ij}(q_i, q_j) + \cdots + f_{12 \cdots p}(q_1, \ldots, q_p)$$

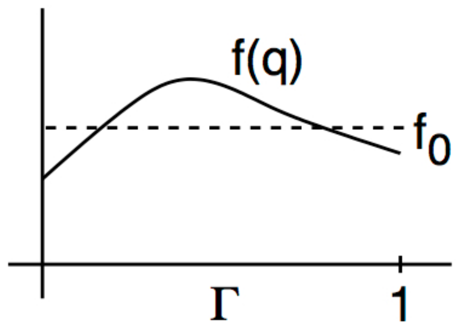
$$= f_0 + \sum_{i=1}^{p} \sum_{|u|=i} f_u(q_u)$$

where

$$f_0 = \int_\Gamma f(q) \rho(q) dq = \mathbb{E}[f(q)]$$

$$f_i(q_i) = \mathbb{E}[f(q)|q_i] - f_0$$

$$f_{ij}(q_i, q_j) = \mathbb{E}[f(q)|q_i, q_j] - f_i(q_i) - f_j(q_j) - f_0$$



Typical Assumption: $q_1, q_2, \ldots, q_p$ independent. Then

$$\int_\Gamma f_u(q_u) f_v(q_v) \rho(q) dq = 0 \quad \text{for } u \neq v$$

$$\Rightarrow \text{var}[f(q)] = \sum_{i=1}^{p} \sum_{|u|=i} \text{var}[f_u(q_u)]$$

Sobol' Indices:

$$S_u = \frac{\text{var}[f_u(q_u)]}{\text{var}[f(q)]} \quad , \quad T_u = \sum_{v \subseteq u} S_v$$

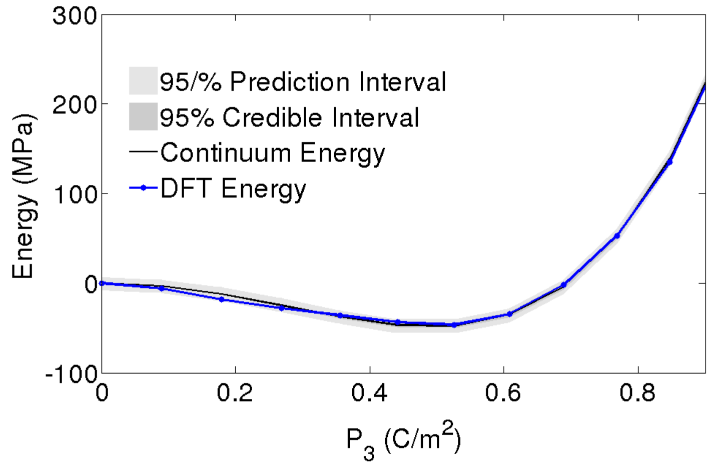
Note: Magnitude of S_i, T_i quantify contributions of q_i to $\text{var}[f(q)]$

Global Sensitivity Analysis

Example: Continuum energy

Question: Do we use 4th or 6th-order Landau energy?

$$\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$



Landau energy

Parameters:

$$q = [\alpha_1, \alpha_{11}, \alpha_{111}]$$

Global Sensitivity Analysis:

	α_1	α_{11}	α_{111}
S_k	0.62	0.39	0.01
T_k	0.66	0.38	0.06
μ_k^*	0.17	0.07	0.03

Conclusion:

α_{111} insignificant and can be fixed

Global Sensitivity Analysis

Example: Quantum-informed continuum model

Question: Do we use 4th or 6th-order Landau energy?

$$\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$

Parameters:

$$q = [\alpha_1, \alpha_{11}, \alpha_{111}]$$

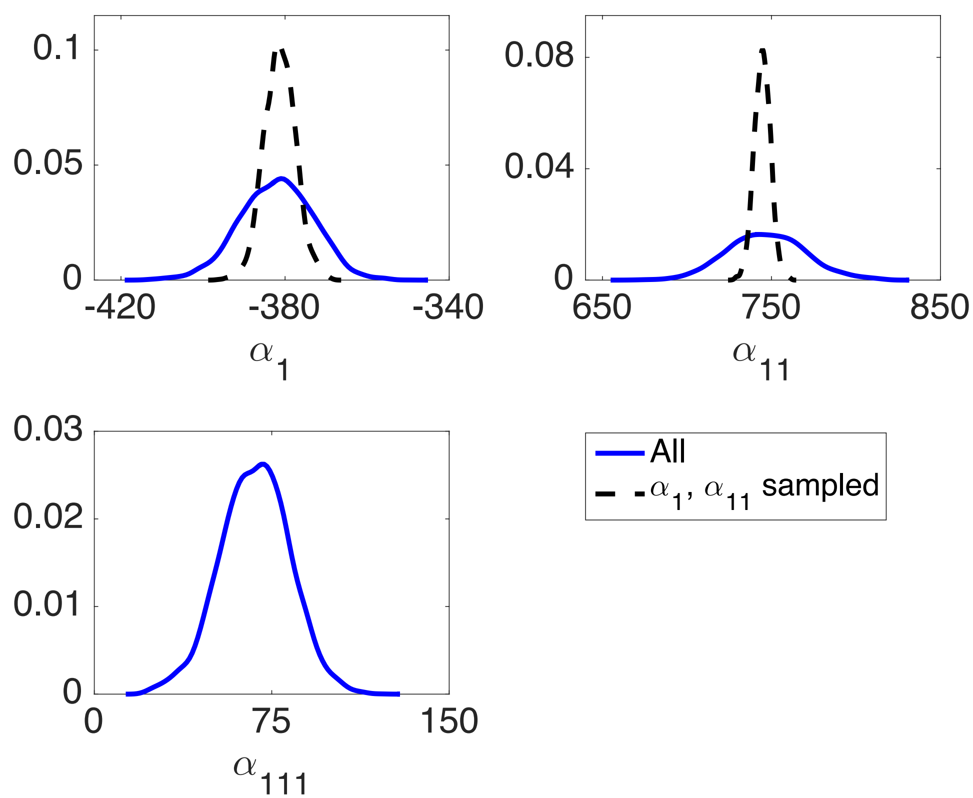
Global Sensitivity Analysis:

	α_1	α_{11}	α_{111}
S_k	0.62	0.39	0.01
T_k	0.66	0.38	0.06
μ_k^*	0.17	0.07	0.03

Conclusion:

α_{111} insignificant and can be fixed

Problem: We obtain different distributions when we perform Bayesian inference with fixed non-influential parameters



Global Sensitivity Analysis

Example: Quantum-informed continuum model

Question: Do we use 4th or 6th-order Landau energy?

$$\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$

Parameters:

$$q = [\alpha_1, \alpha_{11}, \alpha_{111}]$$

Global Sensitivity Analysis:

	α_1	α_{11}	α_{111}
S_k	0.62	0.39	0.01
T_k	0.66	0.38	0.06
μ_k^*	0.17	0.07	0.03

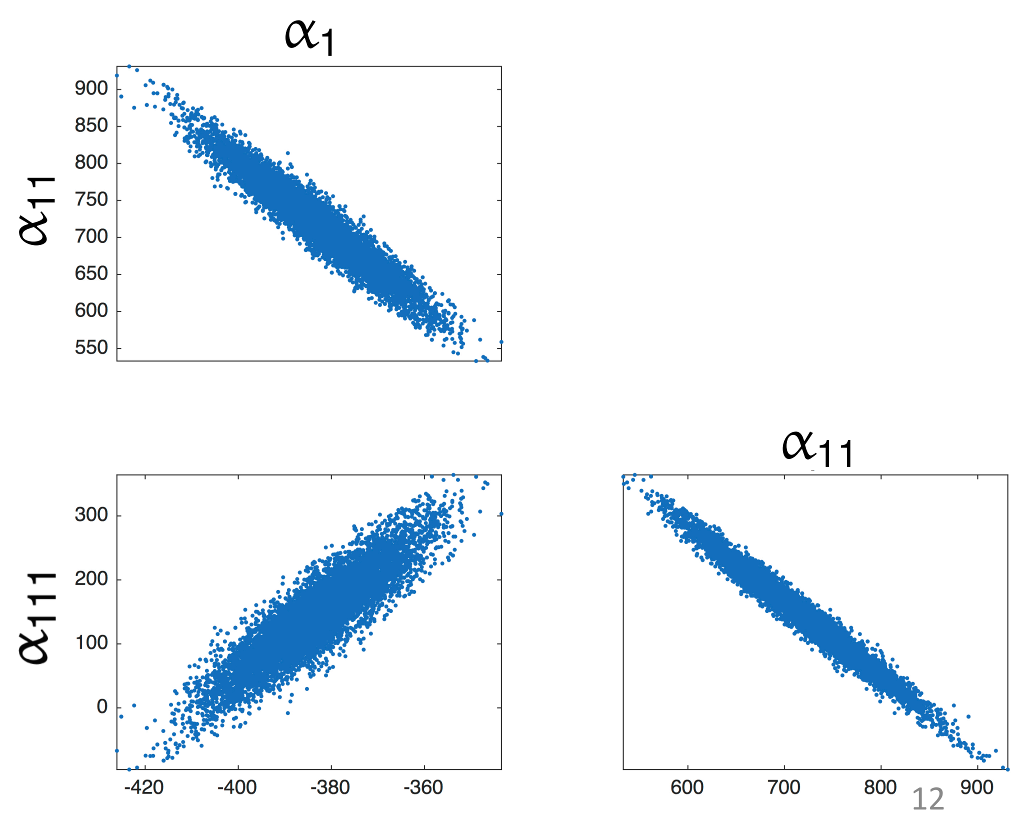
Note:

- Global sensitivity methods generally require parameter distribution, which is typically not known *a priori*.

Alternative: Active subspaces

Problem:

- Parameters correlated
- Cannot fix α_{111}



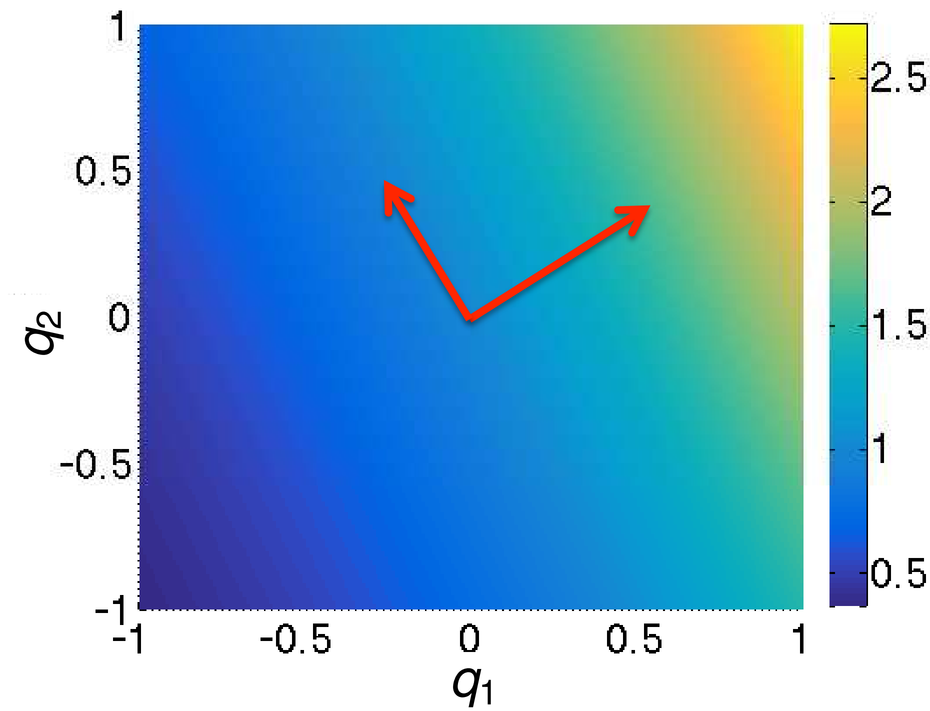
Active Subspaces

Note:

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

Example: $y = \exp(0.7q_1 + 0.3q_2)$

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



A Bit of History:

- Often attributed to Russi (2010).
- Global version of *identifiable subspaces* from systems and control; e.g., Reid (1977).
- For linearly parameterized problems, active subspace given by SVD or QR; Beltrami (1873), Jordan (1874), Sylvester (1889), Schmidt (1907), Weyl (1912). See 1993 *SIAM Review* paper by Stewart.

Gradient-Based Active Subspace Construction

Active Subspace: Consider

$$f = f(q) \;,\; q \in \mathbb{Q} \subseteq \mathbb{R}^p$$

- E.g., see [Constantine, SIAM, 2015; Stoyanov & Webster, *IJUQ*, 2015]

and

$$\nabla_q f(q) = \left[\frac{\partial f}{\partial q_1}, \cdots, \frac{\partial f}{\partial q_p} \right]^T$$

Construct outer product

$$C = \int (\nabla_q f)(\nabla_q f)^T \rho dq$$

$\rho(q)$: Distribution of input parameters q

Question: Can we avoid dependence on an unknown parameter density??

Partition eigenvalues: $C = W \Lambda W^T$

$$\Lambda = \begin{bmatrix} \Lambda_1 & \\ & \Lambda_2 \end{bmatrix} \;,\; W = [W_1 \quad W_2]$$

Rotated Coordinates:

$$y = W_1^T q \in \mathbb{R}^n \quad \text{and} \quad z = W_2^T q \in \mathbb{R}^{p-n}$$

Active Variables

Active Subspace: Range of eigenvectors in W_1

Back to Local Sensitivity

Statistical Observation Model:

$$y_i = f(t_i, \theta^*) + \varepsilon_i \;,\; i = 1, \ldots, n \qquad \varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

Strategy:

- Generate y_i using nominal parameter value θ^*
- Goal: Investigate structure of the model for various values of θ^*

Gradient and Sensitivity Matrix:

$$\nabla_\theta f(t_i, \theta^*) = \left[\frac{\partial f}{\partial \theta_1}(t_i, \theta^*), \ldots, \frac{\partial f}{\partial \theta_p}(t_i, \theta^*) \right]^T$$

$$\mathcal{X}(\theta^*) = \begin{bmatrix} \frac{\partial f}{\partial \theta_1}(t_1, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(t_1, \theta^*) \\ \vdots & & \vdots \\ \frac{\partial f}{\partial \theta_1}(t_n, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(t_n, \theta^*) \end{bmatrix}$$

Parameter Subset Selection

Taylor Expansion: Consider

$$f(t_i, \theta^* + \Delta\theta) \approx f(t_i, \theta^*) + \nabla_\theta f(t_i, \theta^*) \cdot \Delta\theta$$

about nominal value θ^* which minimize

$$J(\theta) = \frac{1}{n}\sum_{i=1}^{n}[y_i - f(t_i, \theta)]^2$$

$$\mathcal{X}(\theta^*) = \begin{bmatrix} \frac{\partial f}{\partial \theta_1}(t_1, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(t_1, \theta^*) \\ \vdots & & \vdots \\ \frac{\partial f}{\partial \theta_1}(t_n, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(t_n, \theta^*) \end{bmatrix}$$

Since $y_i \approx f(t_i, \theta^*)$,

$$\begin{aligned} J(\theta^* + \Delta\theta) &\approx \frac{1}{n}\sum_{i=1}^{n}[\nabla_\theta f(t_i, \theta^*) \cdot \Delta\theta]^2 \\ &= \frac{1}{n}\Delta\theta^T \mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)\Delta\theta \end{aligned}$$

Strategy: Take $\Delta\theta$ to be eigenvector of $\mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)$ Information Matrix

$$\Rightarrow \mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)\Delta\theta = \lambda\Delta\theta$$

$$\Rightarrow J(\theta^* + \Delta\theta) \approx \frac{\lambda}{n}\|\Delta\theta\|_2^2$$

Note:

$\lambda \approx 0 \Rightarrow$ Perturbations $J(\theta^* + \Delta\theta) \approx 0$

$\Rightarrow$ Nonidentifiable

Parameter Subset Selection

Strategy: Take $\Delta\theta$ to be eigenvector of $\mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)$ Information Matrix

$$\Rightarrow \mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)\Delta\theta = \lambda\Delta\theta$$

$$\Rightarrow J(\theta^* + \Delta\theta) \approx \frac{\lambda}{n}\|\Delta\theta\|_2^2$$

Note: $\lambda \approx 0 \Rightarrow$ Perturbations $J(\theta^* + \Delta\theta) \approx 0$

$\Rightarrow$ Nonidentifiable

Algorithm:

1. Set threshold η, set $j = p$, and specify nominal input vector θ^*.
2. Construct $\mathcal{X}(\theta^*)$.
3. Compute ordered eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_j$ of $F(\theta^*) = \mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)$.
4. If $\lambda_1 > \eta$, stop. All remaining inputs are identifiable. Elseif $\lambda_1 < \eta$, at least one parameter is not identifiable.
   a. For computed eigenvector v_1 associated with λ_1, identify component having the largest magnitude. This corresponds to the least identifiable parameter.
   b. Remove the corresponding column in $\mathcal{X}(\theta^*)$, set $j = j - 1$, and repeat 3.

Parameter Subset Selection

Strategy: Take $\Delta\theta$ to be eigenvector of $\boxed{\mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)}$ Information Matrix

$$\Rightarrow \mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)\Delta\theta = \lambda\Delta\theta$$

$$\Rightarrow J(\theta^* + \Delta\theta) \approx \frac{\lambda}{n}\|\Delta\theta\|_2^2$$

Note: $\lambda \approx 0 \Rightarrow$ Perturbations $J(\theta^* + \Delta\theta) \approx 0$

$\Rightarrow$ Nonidentifiable

Note: Covariance matrix estimator [Seber and Wild, 2003]

$$V(\theta^*) = s^2[\mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)]^{-1} = \begin{bmatrix} \text{var}(\hat{\theta}_1) & \text{cov}(\hat{\theta}_1, \hat{\theta}_2) & \cdots & \text{cov}(\hat{\theta}_1, \hat{\theta}_n) \\ \text{cov}(\hat{\theta}_2, \hat{\theta}_1) & \text{var}(\hat{\theta}_2) & \text{cov}(\hat{\theta}_2, \hat{\theta}_3) & \\ \vdots & & \vdots & \\ \text{cov}(\hat{\theta}_n, \hat{\theta}_1) & \cdots & \cdots & \text{var}(\hat{[}\theta_n) \end{bmatrix}$$

Relation to Sensitivity and Statistical Identifiability

Statistical Observation Model:

$$y_i = f(t_i, \theta^*) + \varepsilon_i \ , \ i = 1, \ldots, n \qquad \varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

Sensitivity Matrix:

$$\mathcal{X}(\theta^*) = \begin{bmatrix} \frac{\partial f}{\partial \theta_1}(t_1, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(t_1, \theta^*) \\ \vdots & & \vdots \\ \frac{\partial f}{\partial \theta_1}(t_n, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(t_n, \theta^*) \end{bmatrix}$$

Definition: Parameters θ sensitivity identifiable at θ^* if and only if $\mathcal{X}(\theta^*)$ is one-to-one
Unidentifiable Subspace (Reid 1977): $\mathcal{N}(\mathcal{X}(\theta^*)) = \mathcal{N}(\mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*))$

Definition: For $y = [y_1, \ldots, y_n]$, parameters are statistically unidentifiable if θ and θ^* yield same likelihood

$$L(\theta|y) = L(\theta^*|y) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum_{i=1}^{n}[y_i - f(t_i, \theta)]^2/2\sigma^2}$$

Result: If $F(\theta) = \mathcal{X}^T(\theta)\mathcal{X}(\theta)$ is sufficiently regular in neighborhood of θ, then θ is locally statistically identifiable if and only if $F(\theta)$ is nonsingular.

Question: Can we make this quasi-global?

Thought: Borrow Ideas from Active Subspaces

Statistical Model:

$$y_i = f(t_i, \theta^*) + \varepsilon_i \ , \ \ i = 1, \dots, n$$

Gradient and Sensitivity Matrix:

$$\nabla_\theta f(t_i, \theta^*) = \left[\frac{\partial f}{\partial \theta_1}(t_i, \theta^*), \dots, \frac{\partial f}{\partial \theta_p}(t_i, \theta^*) \right]^T \qquad \mathcal{X}(\theta^*) = \begin{bmatrix} \frac{\partial f}{\partial \theta_1}(t_1, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(t_1, \theta^*) \\ \vdots & & \vdots \\ \frac{\partial f}{\partial \theta_1}(t_n, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(t_n, \theta^*) \end{bmatrix}$$

Observations:

- Parameter Subset Selection: Negligible eigenvalues of $F(\theta^*) = \mathcal{X}^T(\theta^*)\mathcal{X}(\theta^*)$
- Active subspace: Consider

$$C = \int_\Gamma (\nabla_\theta f)(\nabla_\theta f)^T \rho(\theta) d\theta \quad \text{Note: n = 1}$$

- Potential Strategy: Average the Information matrix!

$$F_{glob} = \int_\Gamma \mathcal{X}^T(\theta)\mathcal{X}(\theta)\rho(\theta) d\theta \approx \frac{1}{M}\sum_{k=1}^{M} \mathcal{X}^T(\theta^k)\mathcal{X}(\theta^k)$$

Averaged Parameter Subset Selection

Example: Consider

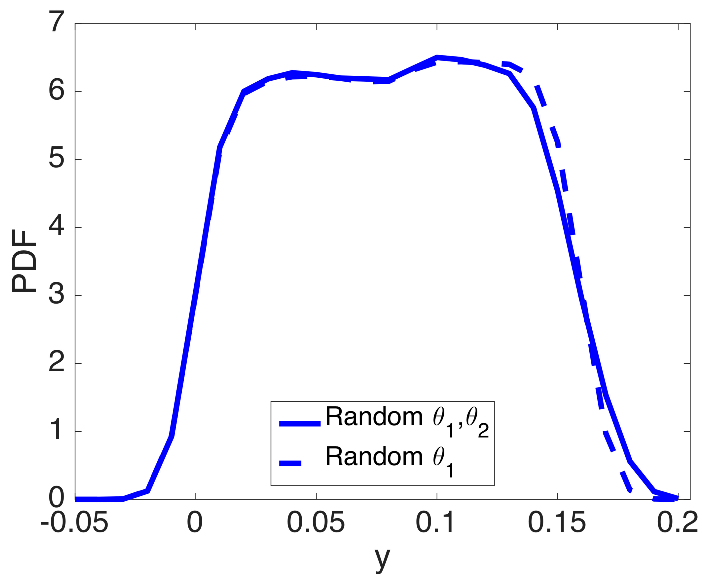
$$y_i = \theta_1\theta_2 t_i + \varepsilon_i \ , \ i = 1, ..., n$$

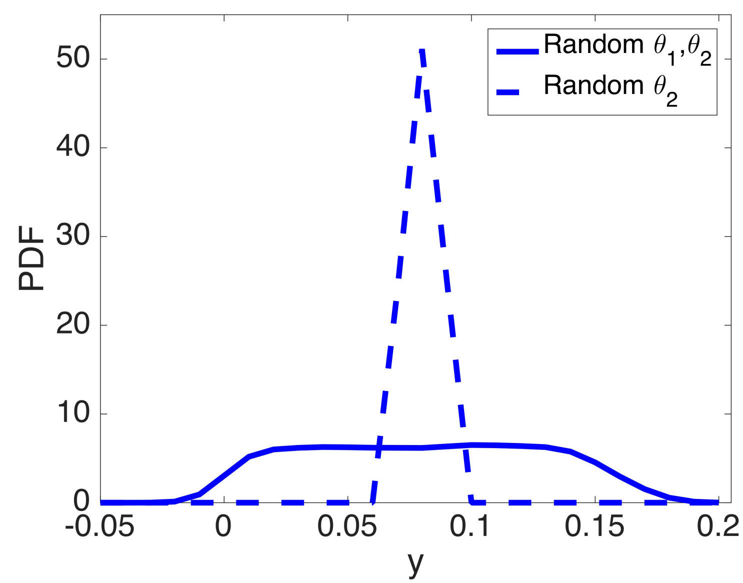
with nominal values $\theta_1 = 0.1, \theta_2 = 0.8$. Here $F = \sum_{i=1}^n t_i^2 A$ where

$$A = \begin{bmatrix} \theta_2^2 & \theta_1\theta_2 \\ \theta_2\theta_1 & \theta_1^2 \end{bmatrix} = \begin{bmatrix} 0.64 & 0.08 \\ 0.08 & 0.01 \end{bmatrix}$$

$$\Rightarrow D = \begin{bmatrix} 0 & 0 \\ 0 & 0.6500 \end{bmatrix} , \ V = \begin{bmatrix} 0.1240 & -0.9923 \\ -0.9923 & -0.1240 \end{bmatrix}$$

Conclusion: Fix θ_2 at nominal value for subsequent analysis.





Averaged Parameter Subset Selection

Example: Consider

$$y_i = \theta_1 \theta_2 t_i + \varepsilon_i \ , \ \ i = 1, \ldots, n$$

Global: Take $\theta_1, \theta_2 \sim \mathcal{U}(0, 1)$, which yields $F_{glob} = \sum_{i=1}^{n} t_i^2 A_{glob}$ with

$$A_{glob} = \begin{bmatrix} 1/3 & 1/4 \\ 1/4 & 1/3 \end{bmatrix}$$

$$\Rightarrow D_{glob} = \begin{bmatrix} 0.0833 & 0 \\ 0 & 0.5833 \end{bmatrix} \ , \ \ V_{glob} = \begin{bmatrix} -0.7071 & 0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}$$

Results:

- This yields incorrect conclusion that parameters are identifiable!
- Furthermore, θ_1 is noninfluential for nominal values

  $\theta_1 = 0.8, \theta_2 = 0.1$

Conclusions:

- PSS provides different information from active subspace analysis;
- Subset of identifiable parameters can vary through parameter space.

Quasi-Global Parameter Subset Selection

Strategy:

1. Randomly sample $q^m \sim \mathcal{U}(\Gamma)$;
2. Determine identifiable parameter set θ_{id}^m;
3. Take quasi-global set to be $\theta_{id} = \cap_m \theta_{id}^m$.

Example: Consider $\theta = [\theta_1, \theta_2, \theta_3] \in \Gamma = [0,1]^3$ and

$$y_i = \theta_1 \theta_2 t_i + \theta_3 \ , \ i = 1, \ldots, n$$

Result: $\theta_{id} = [\theta_3]$

Strategy:

- Determine quasi-global identifiable parameter set by considering intersection of local results;
- Map regions where unidentifiable parameter sets may be identifiable.
- Fix unidentifiable parameters for subsequent optimization, Bayesian or frequentist inference, and uncertainty propagation.

Quasi-Global Parameter Subset Selection

Example: SIR with population N

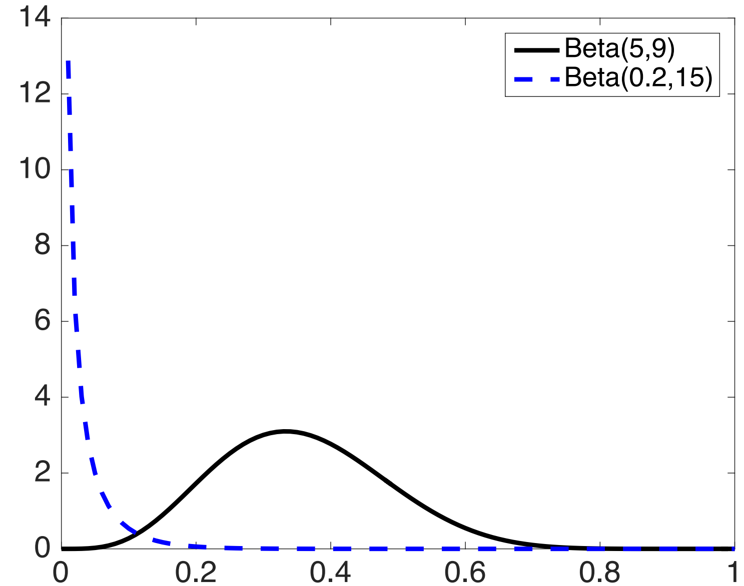
$$\frac{dS}{dt} = \delta N - \delta S - \gamma \underline{k} IS \quad , \; S(0) = S_0$$

$$\frac{dI}{dt} = \underline{\gamma} kIS - (r + \delta) I \quad , \; I(0) = I_0$$

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

Response:

$$f(t_i, \theta) = R(t_i, \theta)$$



Parameters:

- γ: Infection coefficient
- k: Interaction coefficient
- r: Recovery rate
- δ: Birth/death rate

Parameter Distributions:

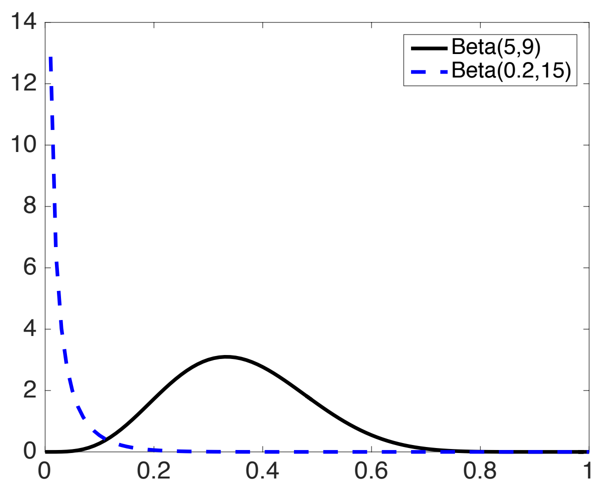
$$\gamma \sim \mathcal{U}(0, 0.1) \;,\; k \sim \mathrm{Beta}(\alpha, \beta) \;,\; r \sim \mathcal{U}(0, 1.2) \;,\; \delta \sim \mathcal{U}(0, 0.5)$$

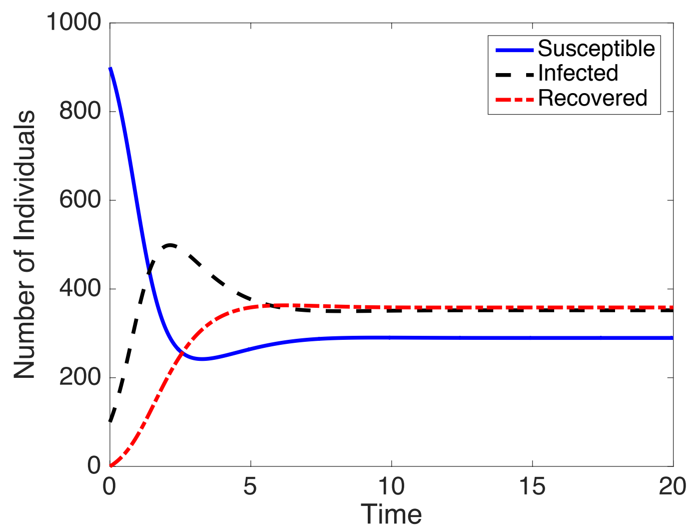
Quasi-Global Parameter Subset Selection

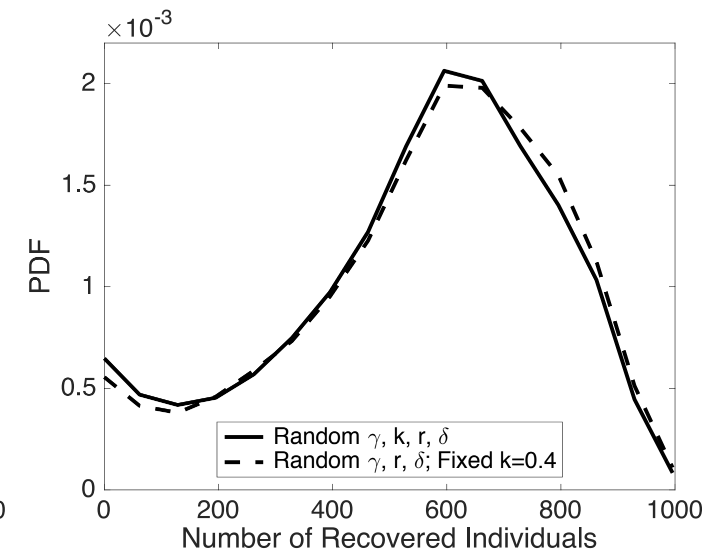
Case i: $k \sim \text{Beta}(5, 9)$

- Significant interactions
- $\theta_{id} = [\gamma, r, \delta]$
- Consistent with GSA

	γ	k	r	δ
S_i	0.2207	0.0106	0.5104	0.2415
S_{T_i}	0.2744	0.0297	0.5365	0.2460
$\mu_i^* \ (\times 10^4)$	5.6325	0.4919	0.9112	1.4347
$\sigma_i \ (\times 10^5)$	1.2175	0.0830	0.1399	0.0986
$v_i \ (\times 10^8)$	1360.2	0.7710	4.1408	4.4926







Quasi-Global Parameter Subset Selection

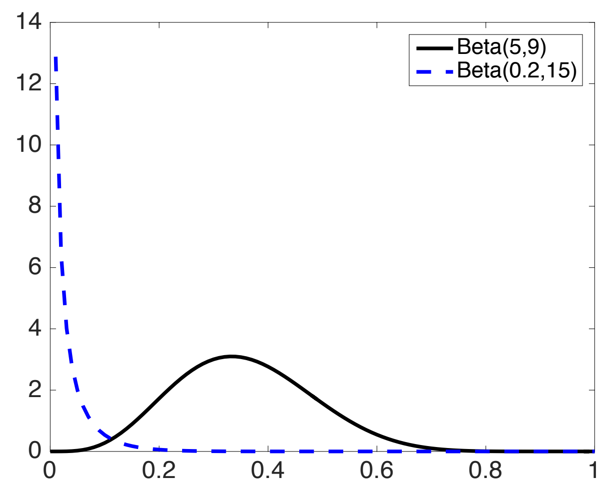
Case ii: $k \sim \text{Beta}(0.2, 15)$

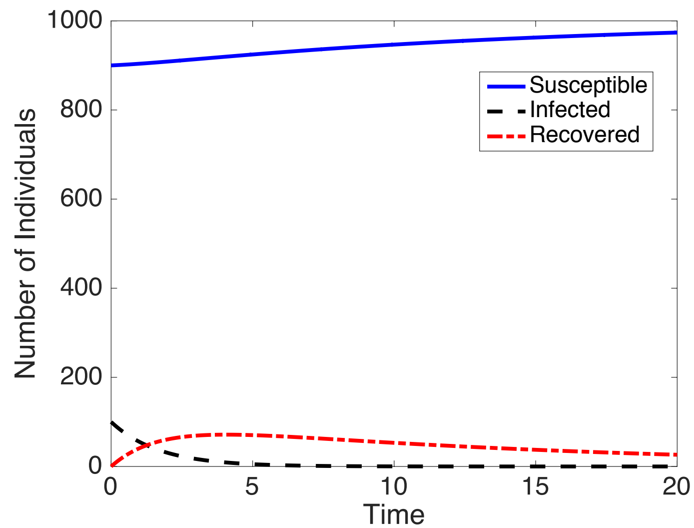
- Limited interactions
- $\theta_{id} = [r, \delta]$
- NOT consistent with GSA

	γ	k	r	δ
S_i	0.0424	0.5873	0.0231	0.0571
S_{T_i}	0.2216	0.8686	0.0936	0.1568
μ_i^* $(\times 10^5)$	0.1988	1.3001	0.0161	0.0622
σ_i $(\times 10^5)$	0.4452	1.3819	0.0449	0.0954
v_i $(\times 10^8)$	42.0180	210.2049	0.7099	1.8157

Note:

- Relative influence of k, γ depends on which has larger nominal value





Remark:

- Parameter subset selection addresses parameter identifiability
- GSA addresses statistical/functional parameter influence

Concluding Remarks

Notes:

• Parameter selection can isolate identifiable and influential parameters. It incorporates parameter correlation structure.

• Objectives complement but are different from active subspace analysis.

• Complex-step often provides robust derivative approximations.

• Quasi-global algorithm needs to be tested on much larger-scale models.

• Parameter Subset Selection: Addresses parameter identifiability.

• Global Sensitivity Analysis: Quantifies relative statistical or functional parameter influence.

• Prediction is very difficult, especially if it's about the future, Niels Bohr.