Parameter Selection Techniques

Motivation: Consider spring model

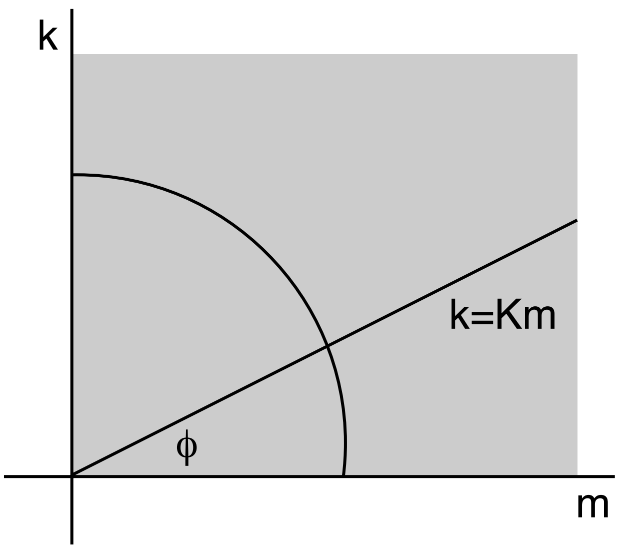
$$m\frac{d^2z}{dt^2} + kz = 0,$$

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

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

Observation: Parameters $\theta = [k, m]$ not uniquely determined by displacement data

Admissible Parameter Space: $\Theta = (0, \infty) \times (0, \infty)$



Note: Determination of slope equivalent to specifying ϕ

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

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

Note: $\Theta = I(\theta) \oplus NI(\theta)$

Parameter Selection Techniques

HIV Model:

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1-\varepsilon)k_1 V T_1$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1-f\varepsilon)k_2 V T_2$$

$$\dot{T}_1^* = (1-\varepsilon)k_1 V T_1 - \delta T_1^* - m_1 E T_1^*$$

$$\dot{T}_2^* = (1-f\varepsilon)k_2 V T_2 - \delta T_2^* - m_2 E T_2^*$$

$$\dot{V} = N_T \delta (T_1^* + T_2^*) - cV - [(1-\varepsilon)\rho_1 k_1 T_1 + (1-f\varepsilon)\rho_2 k_2 T_2]V$$

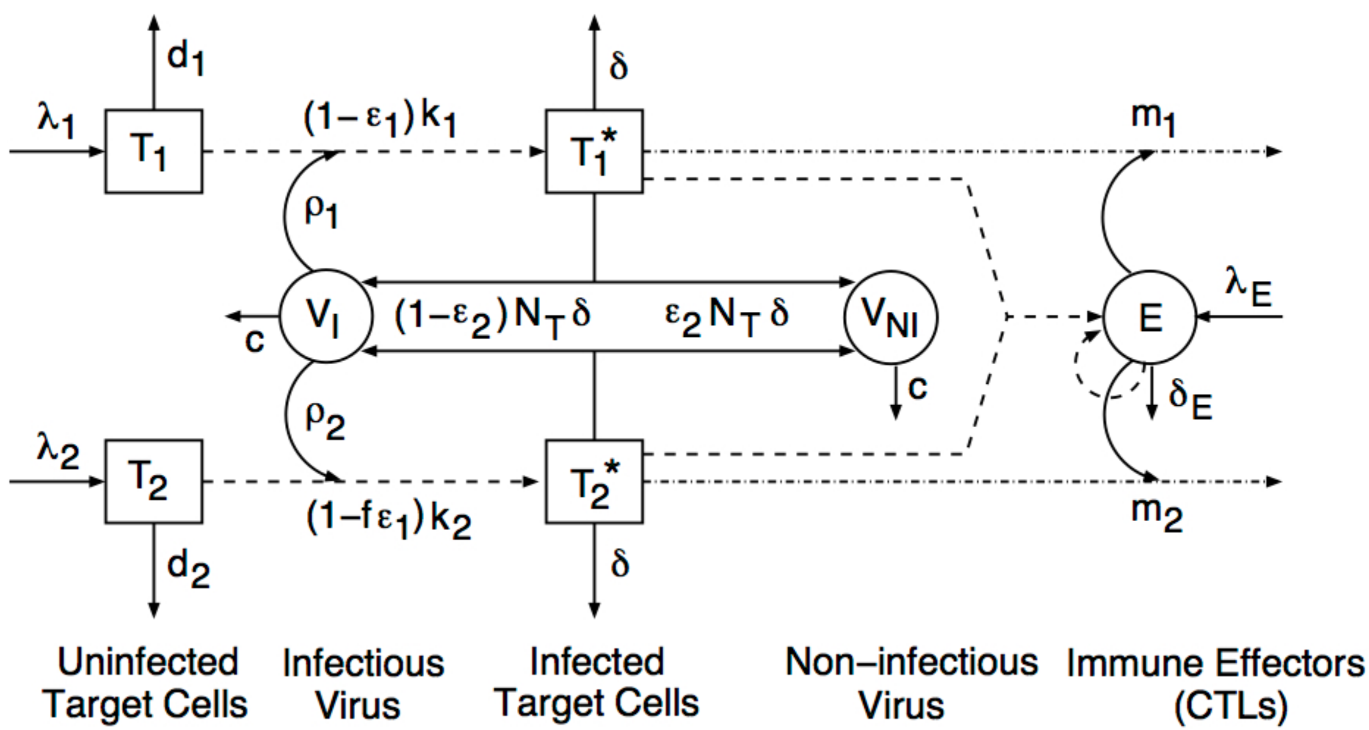
$$\dot{E} = \lambda_E + \frac{b_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_b}E - \frac{d_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_d}E - \delta_E E.$$

Notes: 21 parameters

[Adams, Banks et al., 2005]

Notation: $\dot{E} \equiv \dfrac{dE}{dt}$

Compartments:



Parameter Selection Techniques

HIV Model: Used for characterization and control treatment regimes.

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1-\varepsilon) k_1 V T_1$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1-f\varepsilon) k_2 V T_2$$

$$\dot{T}_1^* = (1-\varepsilon) k_1 V T_1 - \delta T_1^* - m_1 E T_1^*$$

$$\dot{T}_2^* = (1-f\varepsilon) k_2 V T_2 - \delta T_2^* - m_2 E T_2^*$$

$$\dot{V} = N_T \delta (T_1^* + T_2^*) - cV - [(1-\varepsilon)\rho_1 k_1 T_1 + (1-f\varepsilon)\rho_2 k_2 T_2] V$$

$$\dot{E} = \lambda_E + \frac{b_E (T_1^* + T_2^*)}{T_1^* + T_2^* + K_b} E - \frac{d_E (T_1^* + T_2^*)}{T_1^* + T_2^* + K_d} E - \delta_E E.$$

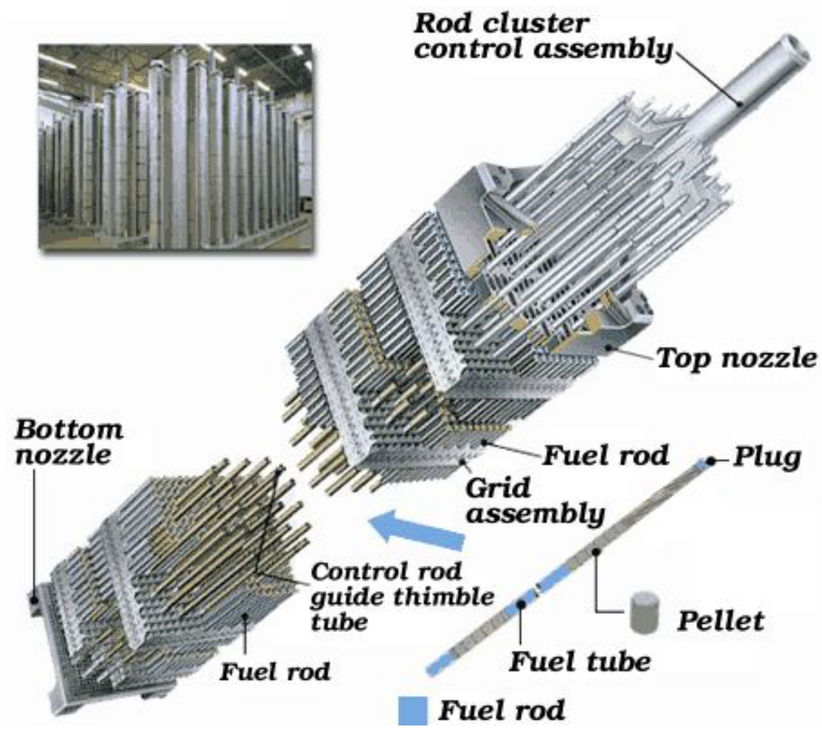
Parameters: Most are unknown and must be estimated from data

λ_1	Target cell 1 production rate	ρ_1	Ave. virions infecting type 1 cell
λ_2	Target cell 2 production rate	ρ_2	Ave. virions infecting type 2 cell
d_1	Target cell 1 death rate	b_E	Max. birth rate immune effectors
d_2	Target cell 2 death rate	d_E	Max. death rate immune effectors
k_1	Population 1 infection rate	K_b	Birth constant, immune effectors
k_2	Population 2 infection rate	K_d	Death constant, immune effectors
c	Virus natural death rate	λ_E	Immune effector production rate
δ	Infected cell death rate	δ_E	Natural death rate, immune effectors
ε	Population 1 treatment efficacy	N_T	Virions produced per infected cell
m_1	Population 1 clearance rate	f	Treatment efficacy reduction
m_2	Population 2 clearance rate		

Pressurized Water Reactors (PWR)

3-D Neutron Transport Equations:

$$\frac{1}{|v|}\frac{\partial \varphi}{\partial t} + \Omega \cdot \nabla \varphi + \Sigma_t(r,E)\varphi(r,E,\Omega,t)$$
$$= \int_{4\pi} d\Omega' \int_0^\infty dE' \Sigma_s(E' \to E, \Omega' \to \Omega)\varphi(r,E',\Omega',t)$$
$$+ \frac{\chi(E)}{4\pi} \int_{4\pi} d\Omega' \int_0^\infty dE' \nu(E')\Sigma_f(E')\varphi(r,E',\Omega',t)$$



Challenges:

- Linear in the state but function of 7 independent variables:

$$r = x, y, z; E; \Omega = \theta, \phi; t$$

- Very large number of inputs; e.g., 100,000; Active subspace construction is critical.
- ORNL Code SCALE: can take minutes to hours to run.
- SCALE TRITON has adjoint capabilities via TSUNAMI-2D and NEWT.

Parameter Subspaces

Definition: Consider

$$y = f(\theta) \quad , \; \theta = [\theta_1, \dots, \theta_p]$$

The parameters are identifiable at θ^* if $f(\theta) = f(\theta^*)$ implies that $\theta = \theta^*$ for all admissible $\theta \in \Theta$. The parameters are identifiable with respect to a space $I(\theta)$, termed the identifiable subspace, if this holds for all $\theta^* \in I(\theta)$. The nonidentifiable subspace $NI(\theta)$ is the orthogonal complement of $I(\theta)$ with respect to Θ

Example: Consider $\theta = [\theta_1, \theta_2]$ in $\Theta = \mathbb{R}^2$ and $y = \theta_1$. Then

$$NI(\theta) = \{\theta_2 \in \mathbb{R}\} \; , \; I(\theta) = \{\theta_1 \in \mathbb{R}\}$$

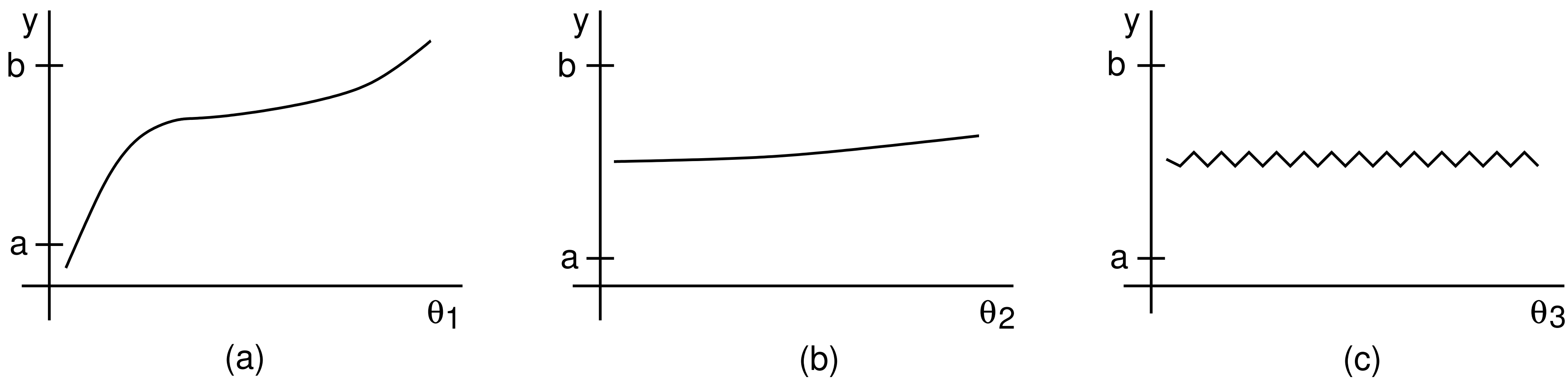
Example: Take $y = \theta_1 - \theta_2$. Then

$$NI(\theta) = \{(\theta_1, \theta_2) \in \mathbb{R}^2 \,|\, \theta_1 = \theta_2\}$$

$$I(\theta) = \{(\theta_1, \theta_2) \in \mathbb{R}^2 \,|\, \theta_1 = -\theta_2\}$$

Noninfluential Parameters

Definition: The parameters $\theta = [\theta_1, \ldots, \theta_p]$ are functionally noninfluential on the manifold $\mathcal{NI}(\theta)$ if $|f(\theta) - f(\theta^*)| < \epsilon$ for all $\theta, \theta^* \in \mathcal{NI}(\theta)$



Identifiable and influential

Minimally influential

Minimally influential with large derivatives

Parameter Selection Techniques

Techniques: $y = f(\theta)$

1. Local sensitivity analysis: Based on derivatives $\frac{\partial y}{\partial \theta_i}$

2. Global sensitivity analysis: Quantifies how uncertainties in model outputs are apportioned to uncertainties in model inputs; e.g., ANOVA

3. Parameter subset selection (PSS) techniques

4. Active subspace techniques based on QR or SVD

Note: 1, 2 and 3 determine subsets of parameters whereas 4 determines subspace

Local Sensitivity Analysis

Strategy: Approximate derivatives

$$s_i = \frac{\partial f}{\partial \theta_i}(\theta^*)$$

Issues:

- Does not quantify uncertainties
- Local at θ^*

Example: Spring model

$$\frac{d^2 z}{dt^2} + C\frac{dz}{dt} + Kz = 0$$

$$z(0) = 2\;,\;\; \frac{dz}{dt}(0) = -C$$

Displacement Observations:

$$y = [1 \;\; 0]\begin{bmatrix} z \\ \dot{z} \end{bmatrix} = z$$

Then $y(t) = 2e^{-Ct/2}\cos\left(\sqrt{K - C^2/4}\cdot t\right)$

Techniques to Compute Local Sensitivities:

1. Analytic
2. Sensitivity equations
3. Finite-difference or complex step
4. Automatic differentiation

Techniques for Local Sensitivity Analysis

1. Analytic: Use symbolic package; e.g., Maple, Mathematica

$$\frac{\partial y}{\partial K} = \frac{-2t}{\sqrt{4K - C^2}} e^{-Ct/2} \sin\left(\sqrt{K - C^2/4} \cdot t\right)$$

$$\frac{\partial y}{\partial C} = e^{-Ct/2} \left[\frac{Ct}{\sqrt{4K - C^2}} \sin\left(\sqrt{K - C^2/4} \cdot t\right) - t \cos\left(\sqrt{K - C^2/4} \cdot t\right) \right]$$

Sensitivity Matrix: $\theta = [C, K]$

$$\boldsymbol{\mathcal{X}}(\boldsymbol{\theta}^*) = \begin{bmatrix} \frac{\partial y}{\partial K}(t_1, \boldsymbol{\theta}^*) & \frac{\partial y}{\partial C}(t_1, \boldsymbol{\theta}^*) \\ \vdots & \vdots \\ \frac{\partial y}{\partial K}(t_n, \boldsymbol{\theta}^*) & \frac{\partial y}{\partial C}(t_n, \boldsymbol{\theta}^*) \end{bmatrix}$$

Information Matrix: $\mathcal{F} = \mathfrak{X}^T \mathfrak{X}$

Techniques for Local Sensitivity Analysis

2. Sensitivity Equations:

$$\frac{d}{dK}\left[\frac{d^2z}{dt^2} + C\frac{dz}{dt} + Kz\right] = 0$$

$$\Rightarrow \frac{d^2z_K}{dt^2} + C\frac{dz_K}{dt} + Kz_K = -z \quad , \; z_K \equiv \frac{\partial z}{\partial K}$$

System:

$$\frac{d^2z_K}{dt^2} + C\frac{dz_K}{dt} + Kz_K = -z \quad , \; z_K(0) = \frac{dz_K}{dt}(0) = 0$$

$$\frac{d^2z}{dt^2} + C\frac{dz}{dt} + Kz = 0 \quad z(0) = 2, \frac{dz}{dt}(0) = 0$$

Similarly:

$$\frac{d^2z_C}{dt^2} + C\frac{dz_C}{dt} + Kz_C = -\frac{dz}{dC} \quad , \; z_C \equiv \frac{\partial z}{\partial C}$$

$$z_C(0) = 0 \; , \; \frac{dz_C}{dt}(0) = -1$$

Techniques for Local Sensitivity Analysis

3. Finite-Difference or Complex Step:

$$\frac{\partial y}{\partial K}(t) \approx \frac{z(t, K + h_K, C) - z(t, K, C)}{h_K}$$

$$\frac{\partial y}{\partial C}(t) \approx \frac{z(t, K, C + h_C) - z(t, K, C)}{h_C}$$

Issues:

1) Stepsizes h_K, h_C must reflect magnitudes of coefficients; e.g., $h_K = 10^{-6}|K|$

2) $\frac{\text{small}}{\text{small}}$ can be inaccurate

Solution: Complex steps

4. Automatic Differentiation:

• Perform differentiation of basic operations – e.g., addition, subtraction, multiplication, division, composition – at the compiler level;

• Good software for ODE and some for PDE

Information Matrix

Relate Sensitivities to Taylor Expansion: Consider

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

about nominal value θ^* obtained my minimizing

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

Here

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

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

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

Sensitivity Matrix:

$$\mathcal{X}(\theta^*) = \begin{bmatrix} \frac{\partial f}{\partial \theta_1}(s_1, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(s_1, \theta^*) \\ \vdots & & \vdots \\ \frac{\partial f}{\partial \theta_1}(s_n, \theta^*) & \cdots & \frac{\partial f}{\partial \theta_p}(s_n, \theta^*) \end{bmatrix}_{n \times p}$$

Fisher Information Matrix

Note:

$$J(\theta^* + \Delta\theta) \approx \frac{1}{n}\Delta\theta^T \mathcal{X}^T \mathcal{X} \Delta\theta$$

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

$$\Rightarrow \mathcal{X}^T \mathcal{X} \nabla\theta = \lambda\Delta\theta$$

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

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

$\Rightarrow$ Nonidentifiable

Note: Estimator for covariance matrix

$$V = s^2[\mathcal{X}^T\mathcal{X}]^{-1} = \begin{bmatrix} \mathrm{var}(\theta_1) & \mathrm{cov}(\theta_1, \theta_2) & \cdots & \mathrm{cov}(\theta_1, \theta_n) \\ \mathrm{cov}(\theta_2, \theta_1) & \mathrm{var}(\theta_2) & \mathrm{cov}(\theta_2, \theta_3) & \\ \vdots & & \vdots & \\ \mathrm{cov}(\theta_n, \theta_1) & \cdots & \cdots & \mathrm{var}(\theta_n) \end{bmatrix}$$

Fisher Information Matrix

Note:

$$J(\theta^* + \Delta\theta) \approx \frac{1}{n}\Delta\theta^T \mathcal{X}^T \mathcal{X} \Delta\theta$$

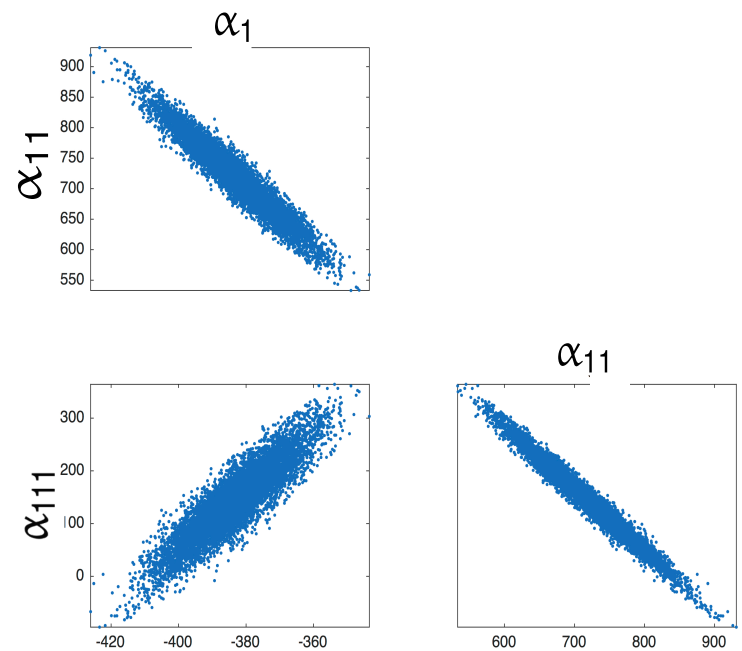
Strategy: Take $\Delta\theta$ to be eigenvector of $\boxed{\mathcal{X}^T \mathcal{X}}$ Information Matrix

$$\Rightarrow \mathcal{X}^T \mathcal{X} \nabla\theta = \lambda \Delta\theta$$

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

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

$\Rightarrow$ Nonidentifiable



Example:

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

Parameters:

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

Result: rank($\chi^T\chi$) = 3 so all parameters identifiable

Fisher Information Matrix

Parameter Subset Selection (PSS) Algorithm:

1. Set $n = p$ and threshhold ε

2. Compute eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors $v_1, \ldots, v_n$ of $\chi^T\chi$ and order the eigenvalues by magnitude:

$$|\lambda_1| \leqslant |\lambda_2| \leqslant \ldots \leqslant |\lambda_n|$$

3. If $|\lambda_1| > \varepsilon$, stop

4. If $|\lambda_1| < \varepsilon$, one or more parameters is not identifiable

- Identify component of v_1 with largest magnitude. This corresponds to least identifiable parameter
- Remove column of $\mathcal{X}$ that corresponds to this component and set $n = n - 1$

Global Sensitivity Analysis

Example: Portfolio model

$$Y = c_1\theta_1 + c_2\theta_2$$

Take

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

$$\theta_1 \sim N(0,1) \ , \ \theta_2 \sim N(0,9)$$

Note:

- θ_1 and θ_2 represent hedged portfolios
- c_1 and c_2 amounts invested in each portfolio

Local Sensitivities:

$$\frac{\partial Y}{\partial \theta_1} = 2 \ , \ \frac{\partial Y}{\partial \theta_2} = 1$$

Conclusion: Investment is more sensitive to Portfolio 1 than to Portfolio 2

Limitations:

- Does not accommodate potential uncertainty in parameters.
- Does not accommodate potential correlation between parameters.
- Sensitive to units and magnitudes of parameters.

Global Sensitivity Analysis

Example: Portfolio model

$$Y = c_1\theta_1 + c_2\theta_2$$

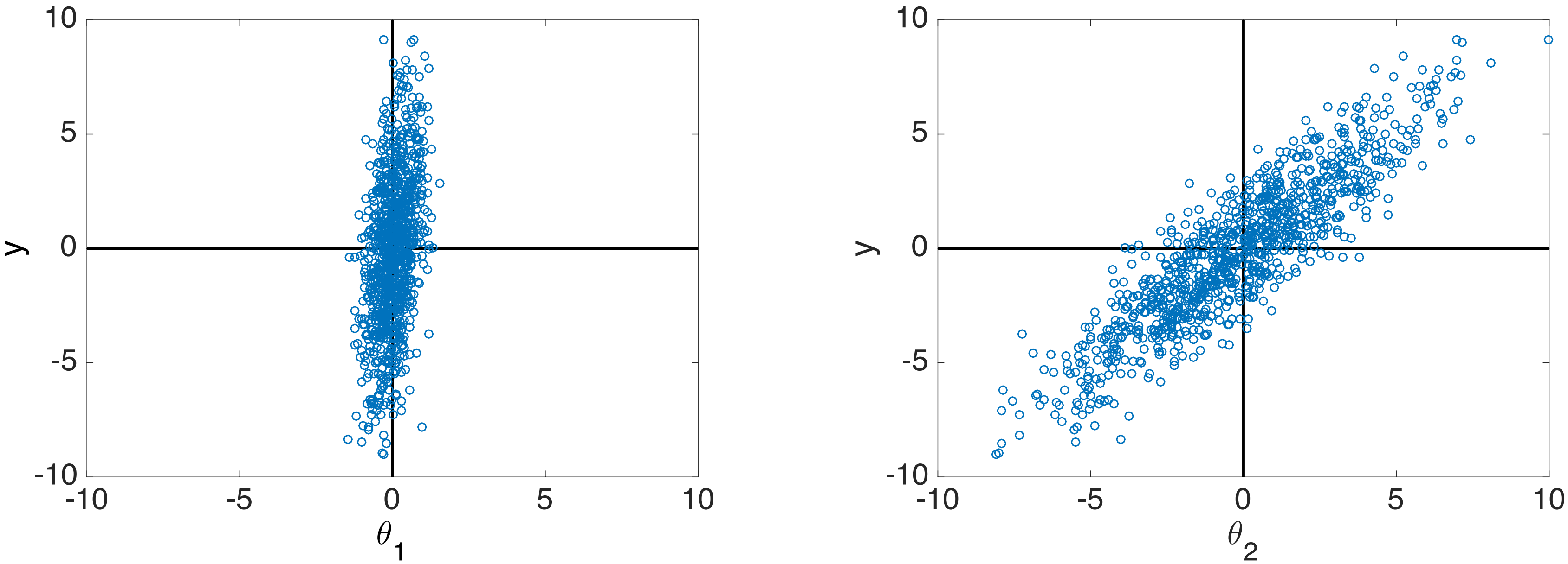
Take

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

$$\theta_1 \sim N(0,1)\ ,\ \theta_2 \sim N(0,9)$$

Note:

- θ_1 and θ_2 represent hedged portfolios
- c_1 and c_2 amounts invested in each portfolio



Local Sensitivities:

$$\frac{\partial Y}{\partial \theta_1} = 2\ ,\ \frac{\partial Y}{\partial \theta_2} = 1$$

Solutions:

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

Global Sensitivity Analysis

Example: Portfolio model

$$Y = c_1\theta_1 + c_2\theta_2$$

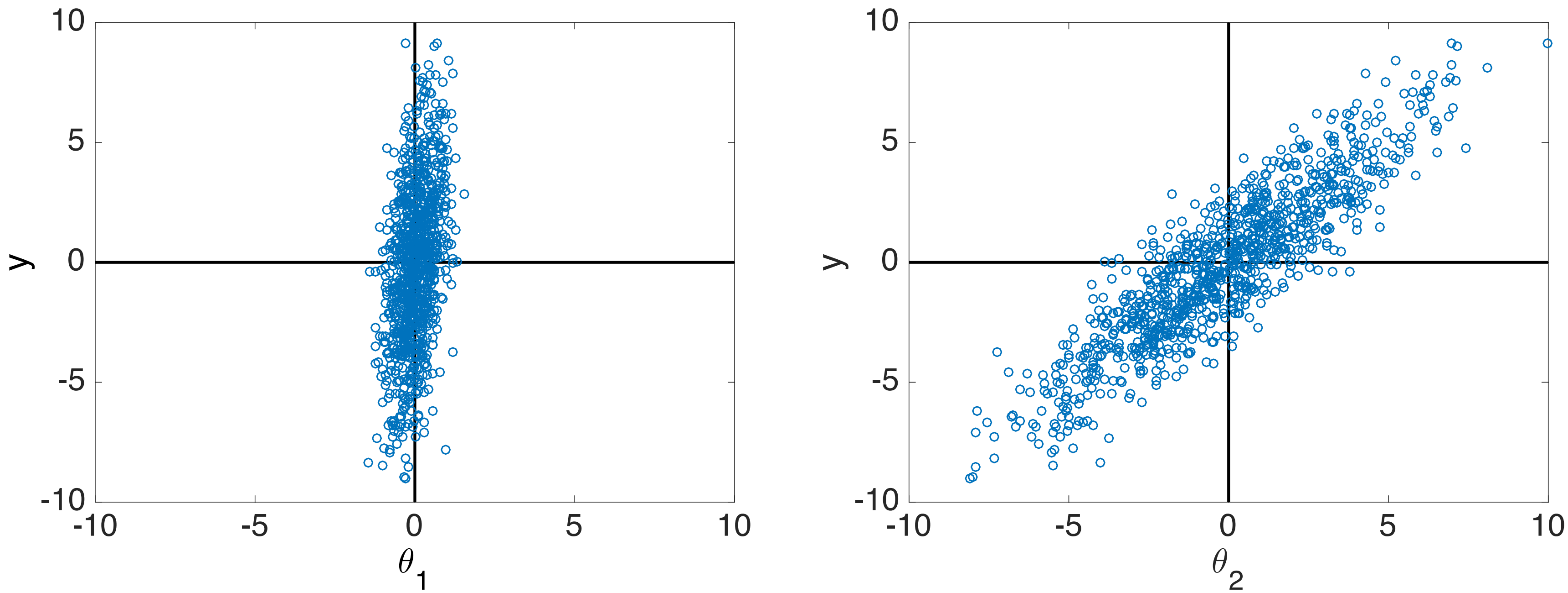
Take

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

$$\theta_1 \sim N(0,1) \ , \ \theta_2 \sim N(0,9)$$

Statistical Motivation: Consider variability of expected values

$$D_i = \text{var}[\mathbb{E}(Y|\theta_i)]$$



Note: Here $D_2 > D_1$

Analysis of Variance (ANOVA): Sobol Analysis

Initial Assumption: Independent uniformly distributed parameters

$$\theta = [\theta_1, \ldots, \theta_p] \sim \mathcal{U}([0,1]^p)$$

Sobol Representation: Truncate at 2nd order – exact if pth order

$$f(\theta) = f_0 + \sum_{i=1}^{p} f_i(\theta_i) + \sum_{1 \leqslant i < j \leqslant p} f_{ij}(\theta_i, \theta_j)$$

Notes:

- Analogies: Taylor or Fourier series
- Need constraints to construct unique representation
    - Derivatives: Taylor
    - Orthogonality: Fourier

Fourier: $f(q) = \sum_{m=1}^{\infty} B_m \sin(m\pi q) = \sin(\pi q)$

Example: $f(\theta) = \sin(\pi\theta)$

Taylor: $f(\theta) = \pi\theta - \frac{(\pi\theta)^3}{3!} + \frac{(\pi\theta)^5}{5!} + \cdots \approx \pi\theta$

Fourier: $f(\theta) = \sum_{m=1}^{\infty} B_m \sin(m\pi\theta) = \sin(\pi\theta)$

Analysis of Variance (ANOVA): Sobol Analysis

Sobol Representation: Truncate at 2nd order – exact if pth order

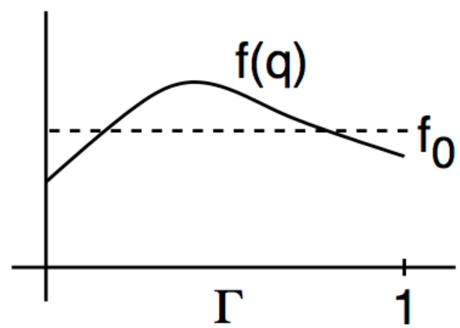
$$f(\theta) = f_0 + \sum_{i=1}^{p} f_i(\theta_i) + \sum_{1 \leqslant i < j \leqslant p} f_{ij}(\theta_i, \theta_j)$$

Sobol Constraints:

$$\int_0^1 f_i(\theta_i) d\theta_i = \int_0^1 f_{ij}(\theta_i, \theta_j) d\theta_i = \int_0^1 f_{ij}(\theta_i, \theta_j) d\theta_j = 0$$

Coefficients:

$$f_0 = \int_\Gamma f(\theta) d\theta$$



$$f_i(\theta_i) = \int_{\Gamma^{p-1}} f(\theta) d\theta_{\sim i} - f_0$$

$$f_{ij}(\theta_i, \theta_j) = \int_{\Gamma^{p-2}} f(\theta) d\theta_{\sim \{ij\}} - f_i(\theta_i) - f_j(\theta_j) - f_0$$

Note: $\theta_{\sim i} = [\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_p]$

Analysis of Variance (ANOVA)

Example: $y = a\theta_1 + b\theta_2$

Then

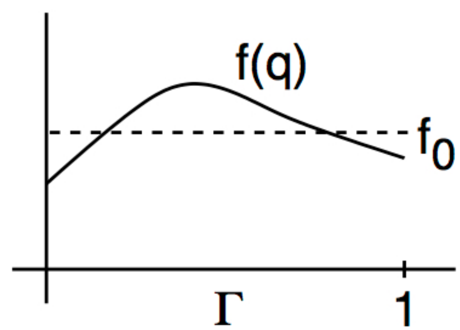
$$f_0 = \int_0^1 \int_0^1 [a\theta_1 + b\theta_2] d\theta_1 d\theta_2 = \frac{a+b}{2}$$

$$f_1(\theta_1) = \int_0^1 [a\theta_1 + b\theta_2] d\theta_2 - f_0 = a\theta_1 - \frac{a}{2}$$

$$f_2(\theta_2) = \int_0^1 [a\theta_1 + b\theta_2] d\theta_1 - f_0 = a\theta_2 - \frac{b}{2}$$

Coefficients:

$$f_0 = \int_\Gamma f(\theta) d\theta$$



$$f_i(\theta_i) = \int_{\Gamma^{p-1}} f(\theta) d\theta_{\sim i} - f_0$$

$$f_{ij}(\theta_i, \theta_j) = \int_{\Gamma^{p-2}} f(\theta) d\theta_{\sim\{ij\}} - f_i(\theta_i) - f_j(\theta_j) - f_0$$

Analysis of Variance (ANOVA)

Statistical Interpretations:

$$\mathbb{E}(Y|\theta_i) = \int_{\Gamma^{p-1}} f(\theta) d\theta_{\sim i}$$

$$\mathbb{E}(Y|\theta_i, \theta_j) = \int_{\Gamma^{p-2}} f(\theta) d\theta_{\sim \{ij\}}$$

Recall: $f_{X_1}(x_1) = \int_{\mathbb{R}} f_X(x_1, x_2) dx_2$

Note:

$$f_0 = \mathbb{E}(Y)$$

$$f_i(\theta_i) = \mathbb{E}(Y|\theta_i) - f_0$$

$$f_{ij}(\theta_i, \theta_j) = \mathbb{E}(Y|\theta_i, \theta_j) - f_i(\theta_i) - f_j(\theta_j) - f_0.$$

Total Variance:

$$D = \mathrm{var}(Y) = \int_\Gamma f^2(\theta) d\theta - f_0^2$$

$$= \sum_{i=1}^{p} D_i + \sum_{1 \leqslant i < j \leqslant p} D_{ij}$$

Partial Variances:

$$D_i = \int_0^1 f_i^2(\theta_i) d\theta_i \quad \text{since} \int_0^1 f_i(\theta_i) d\theta_i = 0$$

$$D_{ij} = \int_0^1 \int_0^1 f_{ij}^2(\theta_i, \theta_j) d\theta_i d\theta_j.$$

Analysis of Variance (ANOVA)

Sobol Indices:

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

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

Variance Interpretations: Verified shortly

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

and

$$S_{T_i} = \frac{\mathbb{E}[\mathrm{var}(Y|\theta_{\sim i})]}{\mathrm{var}(Y)}$$

Note:

$$S_{T_i} \approx 0 \Rightarrow \mathbb{E}[\mathrm{var}(Y|\theta_{\sim i})] \approx 0$$

$$\Rightarrow \mathrm{var}(Y|\theta_{\sim i}) \approx 0 \;\; \text{since} \;\; \mathrm{var}(Y|\theta_{\sim i}) \geqslant 0$$

$\Rightarrow$ Parameter is noninfluential

Analysis of Variance (ANOVA)

Sobol Indices:

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

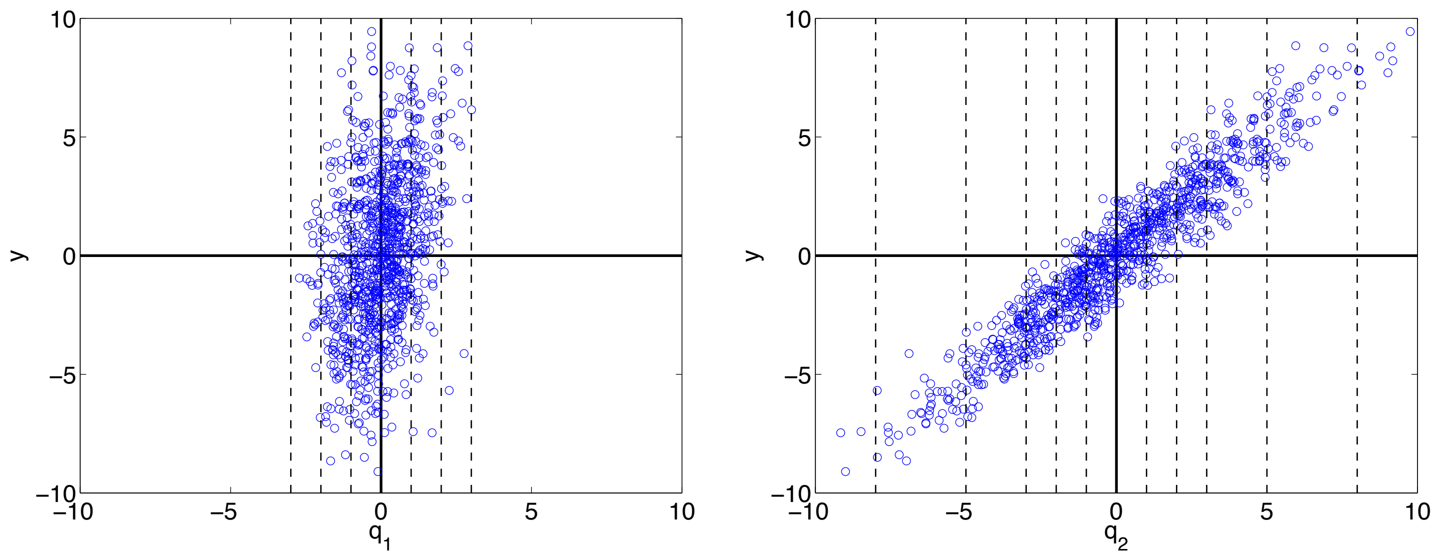
$$S_{T_i} = S_i + \sum_{j=1}^{p} S_{ij}$$

Variance Interpretations: Verified shortly

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

Example: Portfolio model

$$Y = c_1\theta_1 + c_2\theta_2$$



Analysis of Variance (ANOVA)

Example: Portfolio model

$$Y = c_1\theta_1 + c_2\theta_2$$

Take

$$\theta_1 \sim N(0, \sigma_1^2) \qquad \Rightarrow \qquad \rho(\theta_1) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\theta_1^2/2\sigma_1^2}$$

$$\theta_2 \sim N(0, \sigma_2^2) \qquad\qquad \rho(\theta_2) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\theta_2^2/2\sigma_2^2}$$

and

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

$$\sigma_1 = 1 \; , \; \sigma_2 = 3$$

Then

$$f_0 = \iint_{\mathbb{R}^2} [c_1\theta_1 + c_2\theta_2]\rho(\theta_1)\rho(\theta_2)d\theta_1 d\theta_2 = 0$$

$$f_1(\theta_1) = \int_{\mathbb{R}} [c_1\theta_1 + c_2\theta_2]\rho(\theta_2)d\theta_2 = c_1\theta_1$$

$$f_2(\theta_2) = \int_{\mathbb{R}} [c_1\theta_1 + c_2\theta_2]\rho(\theta_1)d\theta_1 = c_2\theta_2$$

$$f_{12}(\theta_1, \theta_2) = 0$$

Analysis of Variance (ANOVA)

Example: Portfolio model

$$Y = c_1\theta_1 + c_2\theta_2 \qquad c_1 = 2\,,\ c_2 = 1 \qquad \sigma_1 = 1\,,\ \sigma_2 = 3$$

Variances:

$$D_i = \int_{\mathbb{R}} f_i^2(\theta_i)\rho(\theta_i)d\theta_i = \int_{\mathbb{R}} c_i^2\theta_i^2\rho(\theta_i)d\theta_i = c_i^2\sigma_i^2$$

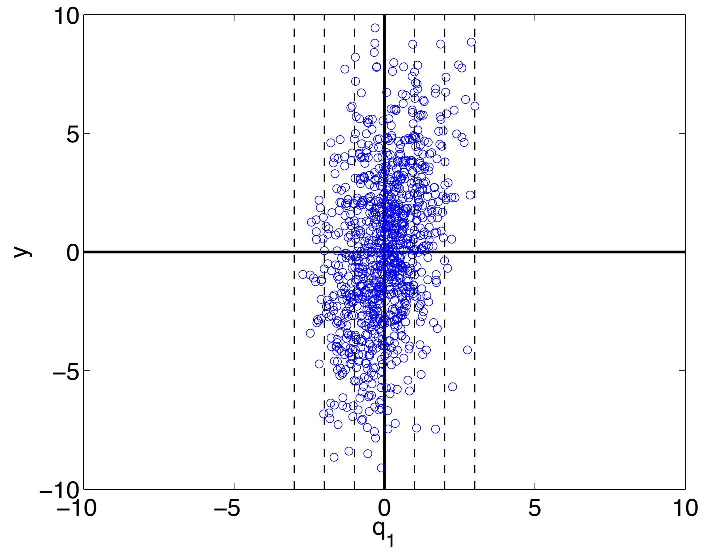
$$D_{12} = \iint_{\mathbb{R}^2} f_{12}^2\rho(\theta_1)\rho(\theta_2)d\theta_1 d\theta_2 = 0$$

so

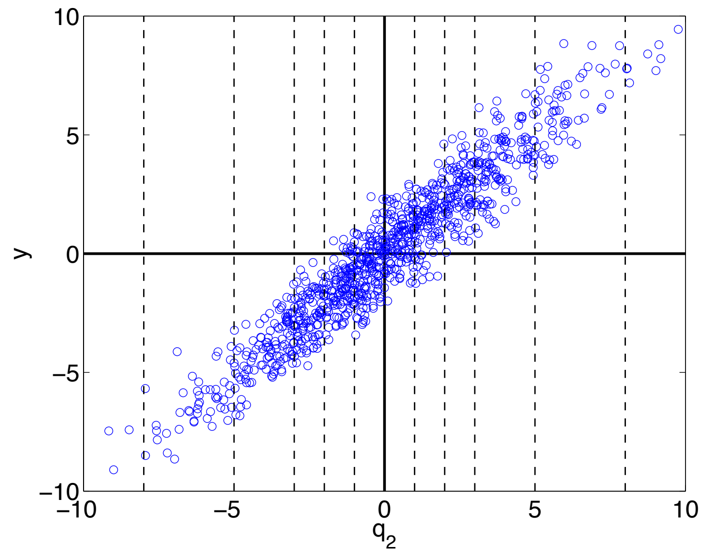
$$D = D_1 + D_2 + D_{12} = c_1^2\sigma_1^2 + c_2^2\sigma_2^2$$

Sobol Indices:

$$S_i = \frac{c_i^2\sigma_i^2}{c_1^2\sigma_1^2 + c_2^2\sigma_2^2} \Rightarrow \boxed{S_1 = \frac{4}{13}\,,\ S_2 = \frac{9}{13}}$$



$D_1 = 4$



$D_2 = 9$

Analysis of Variance (ANOVA)

Verification: Recall that $\text{var}(f) = \mathbb{E}(f^2) - [\mathbb{E}(f)]^2$

Then

$$\begin{aligned}
D_i = & \int_0^1 f_i^2(\theta_i) d\theta_i \\
= & \int_0^1 \left[\int_{\Gamma^{p-1}} f(\theta) d\theta_{\sim i} - f_0 \right]^2 dq_i \\
= & \int_0^1 \left[\int_{\Gamma^{p-1}} f(q\theta) d\theta_{\sim i} \right]^2 d\theta_i - f_0^2 \quad \star \\
= & \mathbb{E} \left[\mathbb{E}(Y|\theta_i) \right]^2 - \left[\mathbb{E}[\mathbb{E}(Y|\theta_i)]^2 \right] \\
= & \text{var}[\mathbb{E}(Y|\theta_i)]
\end{aligned}$$

since

$$\mathbb{E}[\mathbb{E}(Y|\theta_i)] = \int_0^1 \left[\int_{\Gamma^{p-1}} f(\theta) d\theta_{\sim i} \right] d\theta_i = f_0$$

$\star$

Saltelli Algorithm

Algorithm 9.8:

1. Create two $M \times p$ sample matrices

$$\boldsymbol{A} = \begin{bmatrix} \theta_1^1 & \cdots & \theta_i^1 & \cdots & \theta_p^1 \\ \vdots & & & & \vdots \\ \theta_1^M & \cdots & \theta_i^M & \cdots & \theta_p^M \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \hat{\theta}_1^1 & \cdots & \hat{\theta}_i^1 & \cdots & \hat{\theta}_p^1 \\ \vdots & & & & \vdots \\ \hat{\theta}_1^M & \cdots & \hat{\theta}_i^M & \cdots & \hat{\theta}_p^M \end{bmatrix},$$

where θ_i^j and $\hat{\theta}_i^j$ are quasi-random numbers drawn from the respective densities.

2. Create $M \times p$ matrices

$$\boldsymbol{C}_i = \begin{bmatrix} \theta_1^1 & \cdots & \hat{\theta}_i^1 & \cdots & \theta_p^1 \\ \vdots & & & & \vdots \\ \theta_1^M & \cdots & \hat{\theta}_i^M & \cdots & \theta_p^M \end{bmatrix},$$

which are identical to $\boldsymbol{A}$ with the exception that the i^{th} column is taken from $\boldsymbol{B}$.

3. Create the $2M \times p$ matrix

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{A} \\ \hline \boldsymbol{B} \end{bmatrix}$$

by appending $\boldsymbol{B}$ to $\boldsymbol{A}$.

Saltelli Algorithm

Algorithm 9.8: Continued

4. Compute $M \times 1$ vectors of model outputs

$$\boldsymbol{y}_{\boldsymbol{A}} = f(\boldsymbol{A})\ ,\ \boldsymbol{y}_{\boldsymbol{B}} = f(\boldsymbol{B})\ ,\ \boldsymbol{y}_{\boldsymbol{C}_i} = f(\boldsymbol{C}_i)\ ,\ \boldsymbol{y}_{\boldsymbol{D}} = f(\boldsymbol{D})$$

by evaluating the model at the input values in $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}_i$ and $\boldsymbol{D}$. Denote the j^{th} element of $f(D)$ by $\boldsymbol{y}_D^j$ with similar notation for the other vectors. The evaluation of $\boldsymbol{y}_A$ and $\boldsymbol{y}_B$ requires $2M$ model evaluations, whereas the evaluation of $\boldsymbol{y}_{\boldsymbol{C}_i}$, $i = 1, \ldots, p$, requires pM evaluations. The total number of model evaluations is thus $M(p+2)$.

5. Use Monte Carlo integration to approximate the first-order sensitivity indices

$$S_i = \frac{\mathrm{var}[\mathbb{E}(Y|\theta_i)]}{\mathrm{var}(Y)} \approx \frac{\frac{1}{M}\left[\boldsymbol{y}_{\boldsymbol{B}}^T\boldsymbol{y}_{\boldsymbol{C}_i} - \boldsymbol{y}_{\boldsymbol{B}}^T\boldsymbol{y}_{\boldsymbol{A}}\right]}{\frac{1}{2M}\boldsymbol{y}_{\boldsymbol{D}}^T\boldsymbol{y}_{\boldsymbol{D}} - [\mathbb{E}(\boldsymbol{y}_{\boldsymbol{D}})]^2} \tag{1}$$

and total indices

$$S_{T_i} = \frac{\mathbb{E}[\mathrm{var}(Y|\theta_{\sim i})]}{\mathrm{var}(Y)} \approx \frac{\frac{1}{2M}\left[\boldsymbol{y}_{\boldsymbol{A}}^T\boldsymbol{y}_{\boldsymbol{A}} - 2\boldsymbol{y}_{\boldsymbol{A}}^T\boldsymbol{y}_{\boldsymbol{C}_i} + \boldsymbol{y}_{\boldsymbol{C}_i}^T\boldsymbol{y}_{\boldsymbol{C}_i}\right]}{\frac{1}{2M}\boldsymbol{y}_{\boldsymbol{D}}^T\boldsymbol{y}_{\boldsymbol{D}} - [\mathbb{E}(\boldsymbol{y}_{\boldsymbol{D}})]^2}, \tag{2}$$

where $\mathbb{E}(\boldsymbol{y}_{\boldsymbol{D}}) \approx \frac{1}{2M}\sum_{j=1}^{2M} \boldsymbol{y}_{\boldsymbol{D}}^j$.

Morris Screening

Model: $y = f(\theta)$

Initial Assumption: Independent uniformly distributed parameters

$$\theta = [\theta_1, \ldots, \theta_p] \sim \mathcal{U}([0, 1]^p)$$

Elementary Effects: Coarse derivative approximations

$$d_i = \frac{f(\theta_1, \ldots, \theta_{i-1}, \theta_i + \Delta, \theta_{i+1}, \ldots, \theta_p) - f(\theta)}{\Delta}$$

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

$$\Delta \in \left\{ \frac{1}{\ell - 1}, \ldots, 1 - \frac{1}{\ell - 1} \right\} \; , \; \ell \text{ is level; e.g., } \Delta = \frac{1}{100}$$

$$e_i = [0, \ldots, 0, 1, 0 \ldots, 0]$$

Global Sensitivity Measures: i=1, …, p

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

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

Morris Screening

Forward Difference Algorithm: See also Algorithm 9.20

1. Specify ℓ, $\Delta = \frac{1}{\ell}$ and r; e.g., $\Delta = 10^{-4}$ and $r = 40$.
2. For $j = 1, ..., r$
   (a) Sample random or quasi-random point $\theta^j \in \Gamma = [0, 1]^p$.
   (b) Employ the finite-difference relation to compute d_i^j.
3. Compute μ_i^* and σ_i.
4. Determine noninfluential inputs Q_i as those having small values of μ_i^* and σ_i and influential inputs as those having large indices.

Issues:

- Provides relative than absolute rankings
- Parameters often correlated and hence not independent. One can make incorrect conclusions based on incorrect assumption of independence.
- How does one construct indices for time or space-dependent responses or, more generally infinite-dimensional responses? Same question for vector-valued responses.

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 $\theta = [\gamma, k, r, \delta]$ 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:

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

SIR Disease Example

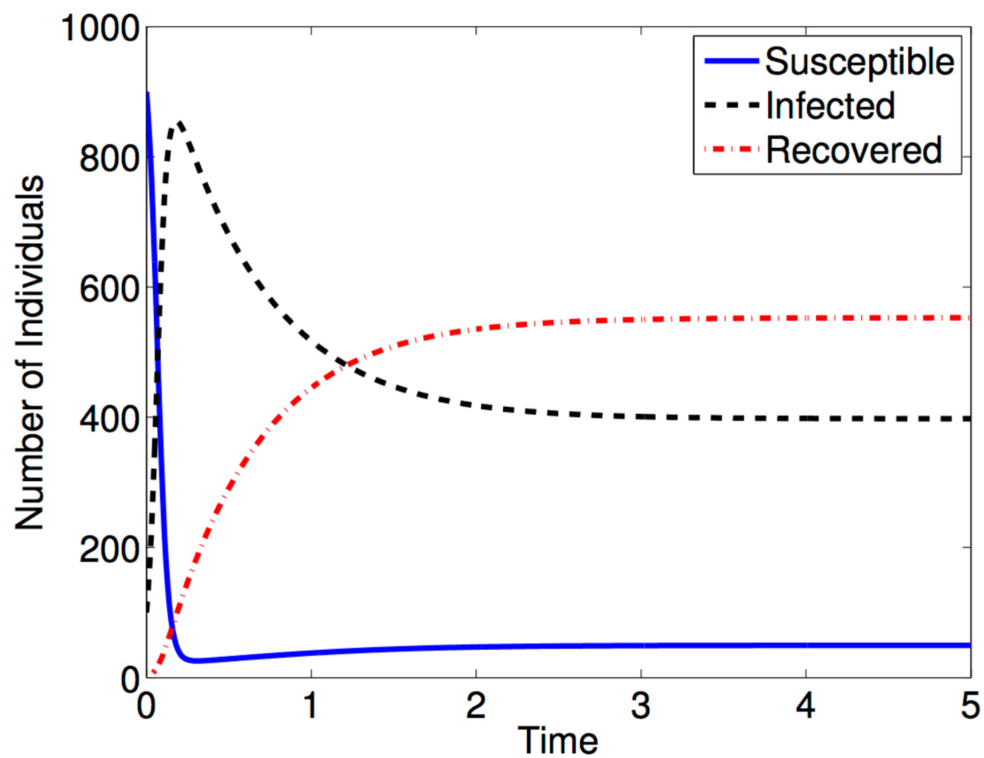
SIR Model:

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

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

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

Typical Realization:

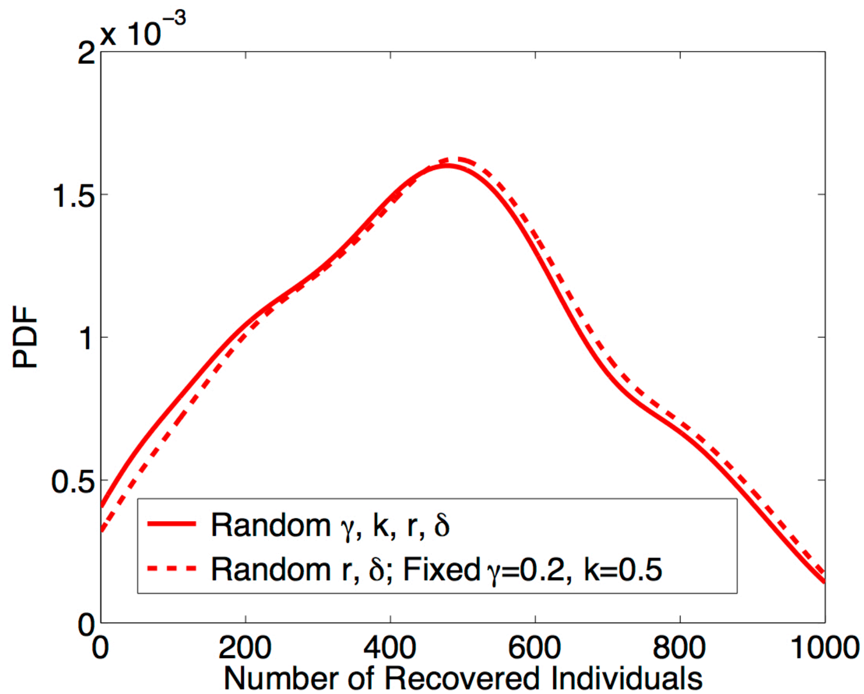


SIR Disease Example

Global Sensitivity Measures:

		γ	k	r	δ
Sobol	S_i	0.0331	0.0167	0.5335	0.4144
	S_{T_i}	0.0642	0.0189	0.5543	0.4257
Morris	μ_i^* $(\times 10^4)$	0.1448	0.2422	1.0257	1.0012
	σ_i $(\times 10^3)$	4.8329	5.7114	7.2911	5.5034
Time-Dependent	ν_i $(\times 10^9)$	4.6978	1.1471	0.3409	0.1962

Result: Densities for $R(t_f)$ at $t_f = 5$



Note: Can fix non-influential parameters

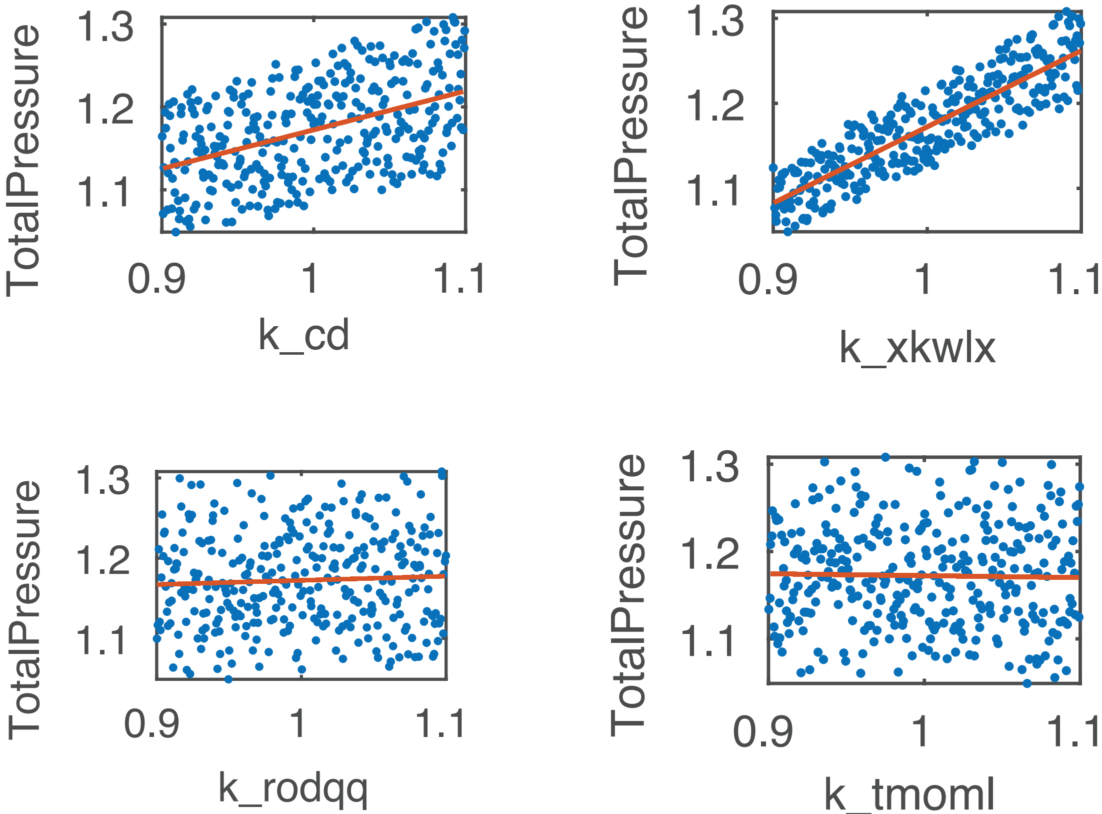
Subchannel Code COBRA-TF

33 Parameters:

Parameter	Partial correlation	Morris main	Morris interaction	Influence
k_sent	0.06			
k_sdent	0.05			
k_tmasv	-0.08			
k_tmasl	0.10			
k_moml		1.08×10^{-5}	1.52×10^{-5}	Low
k_xkes	-0.08			
k_xkge	0.07			
k_xkl	-0.12			
k_xkvls	-0.06			
k_xkwlw	-0.09			
k_tnrgv	0.05			
k_rodqq	0.46	1.26×10^{-2}	1.00×10^{-3}	Medium
k_sphts	0.08			
k_cond	-0.07			
k_xkwlx	0.99	1.79×10^{-1}	7.47×10^{-3}	High
k_cd	0.97	9.55×10^{-2}	7.98×10^{-3}	High
k_wkr	-0.06			

Subchannel Code COBRA-TF

33 Parameters:



Global Sensitivity Analysis

Example: Quantum-informed continuum model

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

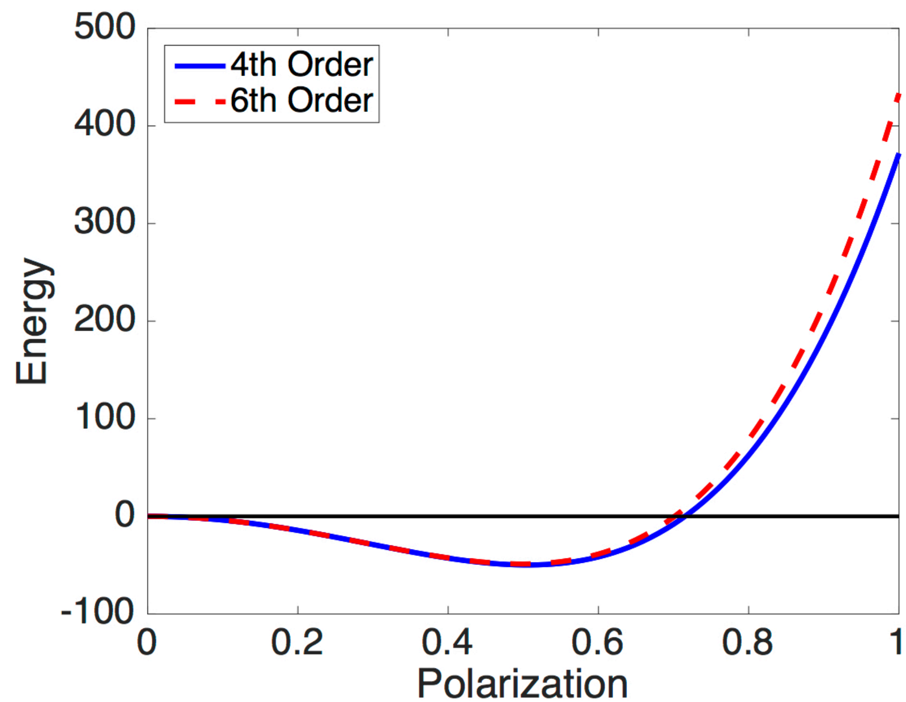
$$\psi(P,\theta) = \underline{\alpha_1} P^2 + \underline{\alpha_{11}} P^4 + \underline{\alpha_{111}} P^6$$

Parameters:

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

Global Sensitivity Analysis:

	α_1	α_{11}	α_{111}
S_i	0.62	0.39	0.01
S_{T_i}	0.66	0.38	0.06
μ_i^*	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?

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

Parameters:

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

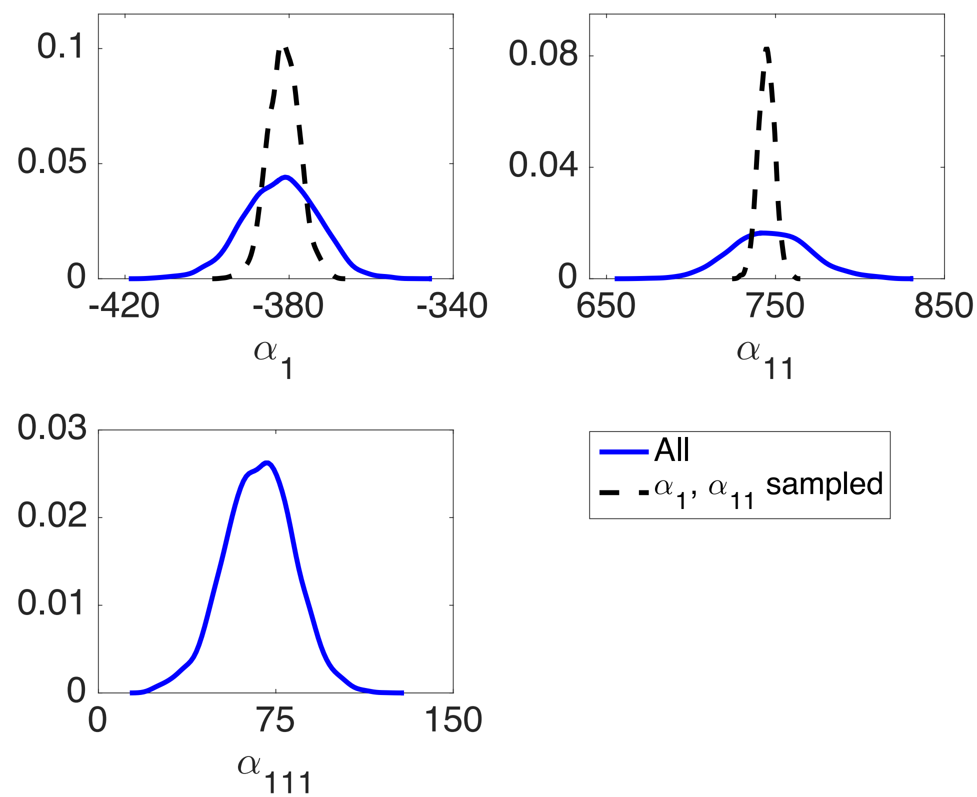
Global Sensitivity Analysis:

	α_1	α_{11}	α_{111}
S_i	0.62	0.39	0.01
S_{T_i}	0.66	0.38	0.06
μ_i^*	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?

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

Parameters:

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

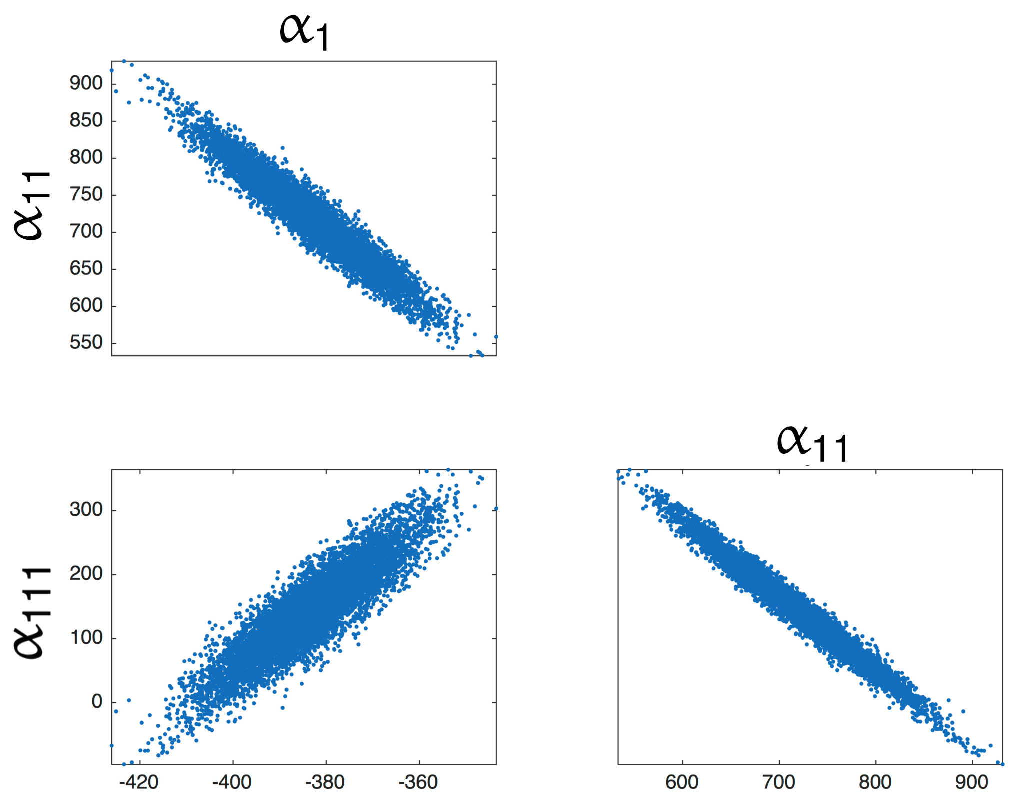
Global Sensitivity Analysis:

	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: Must accommodate correlation

Problem:

- Parameters correlated
- Cannot fix α_{111}



Global Sensitivity Analysis: Analysis of Variance

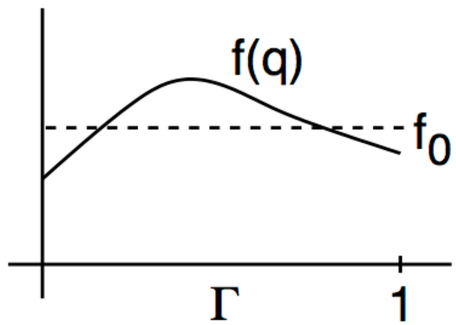
Sobol' Representation: $Y = f(\theta)$

$$f(q) = f_0 + \sum_{i=1}^{p} f_i(\theta_i) + \sum_{i\leqslant i<j\leqslant p} f_{ij}(\theta_i,\theta_j) + \cdots + f_{12\cdots p}(\theta_1,\ldots,\theta_p)$$

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

where

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



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

$$f_{ij}(\theta_i,\theta_j) = \mathbb{E}[f(\theta)|\theta_i,\theta_j] - f_i(\theta_i) - f_j(\theta_j) - f_0$$

Typical Assumption: $\theta_1, \theta_2, \ldots, \theta_p$ independent. Then

$$\int_\Gamma f_u(\theta_u) f_v(\theta_v)\rho(\theta)d\theta = 0 \quad \text{for } u \neq v$$

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

Sobol' Indices:

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

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

Global Sensitivity Analysis: Analysis of Variance

Sobol' Representation:

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

One Solution: Take variance to obtain

$$\text{var}[f(\theta)] = \sum_{i=1}^{p} \sum_{|u|=k} \text{cov}[f_u(\theta_u), f(\theta)]$$

Sobol' Indices:

$$S_u = \frac{\text{cov}[f_u(\theta_u), f(\theta)]}{\text{var}[f(\theta)]}$$

Pros:

- Provides variance decomposition that is analogous to independent case

Cons:

- Indices can be negative and difficult to interpret
- Often difficult to determine underlying distribution
- Monte Carlo approximation often prohibitively expensive.

Alternative: Construct active subspaces

- Can accommodate parameter correlation
- Often effective in high-dimensional space; e.g., p = 7700 for neutronics example

Additional Goal: Use Bayesian analysis on active subspace to construct posterior densities for physical parameters.

One Solution: Parameter Subset Selection

Note:

$$J(\theta^* + \Delta\theta) \approx \frac{1}{n}\Delta\theta^T \mathcal{X}^T \mathcal{X} \Delta\theta$$

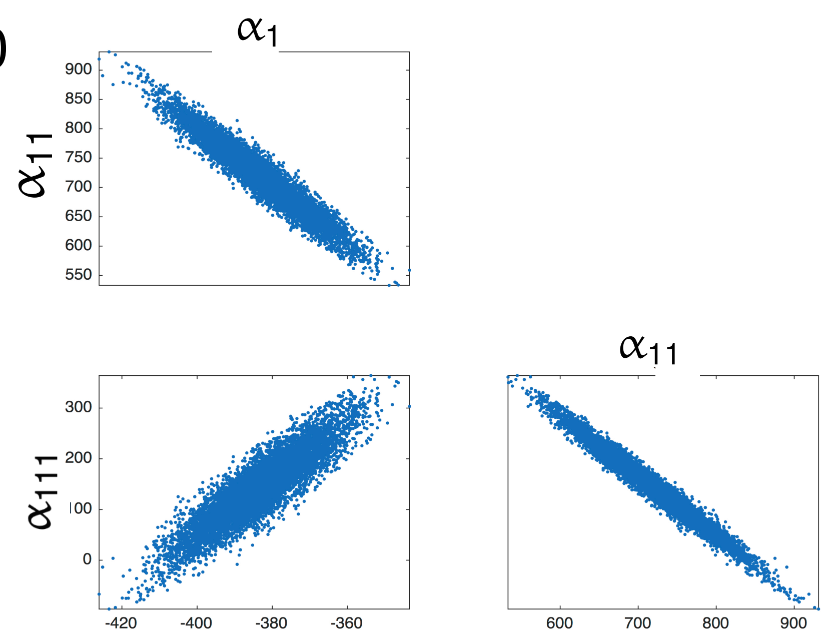
Strategy: Take $\Delta\theta$ to be eigenvector of $\mathcal{X}^T\mathcal{X}$ Information Matrix

$$\Rightarrow \mathcal{X}^T \mathcal{X} \nabla\theta = \lambda\Delta\theta$$

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

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

$\Rightarrow$ Nonidentifiable



Example:

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

Parameters:

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

Result: rank($\chi^T\chi$) = 3 so all parameters identifiable