$$\boxed{\text{Parameter Subset Selection}}$$

- Given $\vec{S}_j = \frac{\partial f}{\partial \theta_j}$, we know that given $J(\theta)$ to be minimized

$$\Rightarrow J(\theta) = \frac{1}{n}\sum (y-f)^2 \Rightarrow J(\vec{\theta}+\Delta\theta) = \frac{1}{n}\left[\underset{\sim}{S}\Delta\theta\right]^T\left[\underset{\sim}{S}\Delta\theta\right]$$

So $$J(\vec{\theta}+\Delta\theta) = \frac{1}{n}\Delta\vec{\theta}^T\left(\underset{\sim}{S}^T\underset{\sim}{S}\right)\Delta\vec{\theta}, \quad \vec{\theta}\in\mathbb{R}^p$$

- If $\Delta\vec{\theta}$ is an eigenvector of $\underset{\sim}{S}$ $\Rightarrow$ $\underset{\sim}{S}^T\underset{\sim}{S}\,\Delta\vec{\theta} = \vec{\lambda}\Delta\vec{\theta}$, $\vec{\lambda}>0$, $\vec{\lambda}$ are eig. values of $\underset{\sim}{S}^T\underset{\sim}{S}$

$$\Rightarrow \boxed{J(\vec{\theta}+\Delta\theta) \approx \frac{1}{n}\vec{\lambda}\,\|\Delta\vec{\theta}\|_2^2}$$

- If any $\lambda_i \approx 0$, then θ_i has relatively little impact $J(\vec{\theta})$

Global Sensitivity Analysis (GSA) (Chapter 9)

- GSA quantifies the uncertainty in Y that is apportioned to $\vec{\theta}$.

Variance-Based Methods

- Consider a scalar, nonlinear model

independent

$$Y = F(\vec{\theta}), \quad \vec{\theta}\in\Gamma \ \& \ \theta_i \sim \mathcal{U}(0,1) \quad \left(\text{so } \Gamma = [0,1]^p\right)$$

- We can decompose $f(\vec{\theta})$ into a hierarchical form

$$F(\vec{\theta}) = f_0 + \sum_{i=1}^{p} f_i(\theta_i) + \sum_{1\le i<j\le p} f_{ij}(\theta_i,\theta_j) + \cdots$$

where

$f_0 \equiv$ mean over Γ

$f_i \equiv$ ... order response from θ_i

where

$f_0 \equiv$ mean over Γ

$f_i \equiv$ first order response from θ_i

$f_{ij} \equiv$ Second order interactions from θ_i + θ_j

- We assume that the functions satisfy

$$\int_0^1 f_i(\theta_i)\, d\theta_i = 0$$

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

i.e. each f is orthogonal.

- Each f can be expressed as:

$$f_0 = \int_\Gamma F(\vec{\theta})\, d\vec{\theta} = E[F(\vec{\theta})]$$

$$f_i(\theta_i) = \int_{\Gamma^{P-1}} F(\vec{\theta})\, d\vec{\theta}_{\sim i} - f_0 = E[F(\vec{\theta}) \mid \theta_i] - f_0$$

$$f_{ij}(\theta_i, \theta_j) = \int_{\Gamma^{P-2}} F(\vec{\theta})\, d\vec{\theta}_{\sim i,j} - f_i - f_j - f_0 = E[F(\vec{\theta}) \mid \theta_i, \theta_j] - f_i - f_j - f_0$$

where $\vec{\theta}_{\sim i} = \vec{\theta} \setminus \theta_i$, $\vec{\theta}_{\sim ij} = \vec{\theta} \setminus \{\theta_i, \theta_j\}$

- This is a (functional) analysis of variance (ANOVA) decomposition of $F(\vec{\theta})$.
  - Hoeffding-Sobol decomposition.

- If θ_i are independent

$$\Rightarrow \int_\Gamma f_i(\theta_i) f_j(\theta_j)\, d\vec{\theta} = 0$$

$$\Rightarrow \int_{\Gamma} f_i(\theta_i)\, f_j(\theta_j)\, d\vec{\theta} = 0$$

- Let the variance of $F(\vec{\theta}) = Y$

$$D = Var[Y] = \int_{\Gamma} \left(F(\vec{\theta})\right)^2 d\vec{\theta} - f_0^2$$

- Since each f term is orthogonal, then

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

Because $Var[f_i] = E[f_i^2] - (E[f_i])^2$

$+ \; E[f_i] = \int_0^1 f_i(\theta_i)\, d\theta_i = 0$

where
$$D_i = \int_0^1 f_i^2(\theta_i)\, d\theta_i = Var[f_i(\theta_i)]$$

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

Sobol Indices)