Global Sensitivity Analysis:

E.g.) Portfolio model Amounts invested

$$Y = c_1 Q_1 + c_2 Q_2$$

Portfolios

Take $c_1 = 2$

$c_2 = 1$

Note: $Y \sim N(0, \sigma_Y^2)$ with

$$\sigma_Y^2 = c_1\sigma_1^2 + c_2\sigma_2^2 = 13$$

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

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

More volatile

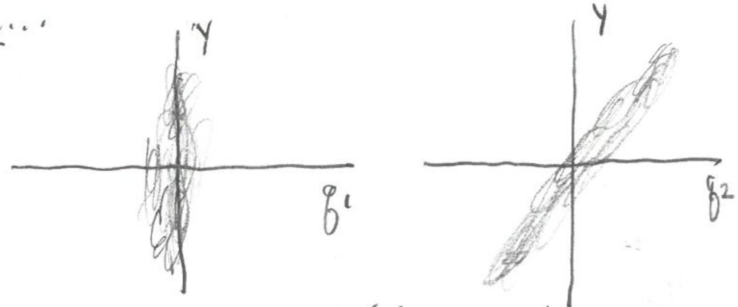
Local Sensitivities:

$$S_1 = \frac{\partial Y}{\partial Q_1} = c_1 = 2$$

$$S_2 = \frac{\partial Y}{\partial Q_2} = c_2 = 1$$

Conclusion: Y more sensitive wrt Q_1

But...



Conclusion: More of variability attributed to Q_2

GSA: 1) Analysis of Variance (ANOVA)

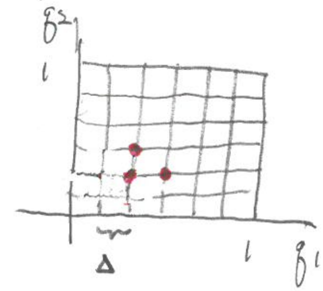
2) Morris screening

Morris Screening: Section 15.2 $Y = f(Q)$

Initial Assumption: Independent uniformly distributed parameters

$$Q = [Q_1, \ldots, Q_p] \sim \mathcal{U}([0,1]^p)$$

e.g.) $p = 2$



Elementary Effects: Coarse derivative approximations

$$d_i = \frac{f(q_1, \ldots, q_{i-1}, q_i + \Delta, q_{i+1}, \ldots, q_p) - f(q)}{\Delta}$$

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

$\Delta \in \left\{ \frac{1}{l-1}, \ldots, 1 - \frac{1}{l-1} \right\}$, l is level ; e.g.) $\Delta = \frac{1}{40}$

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

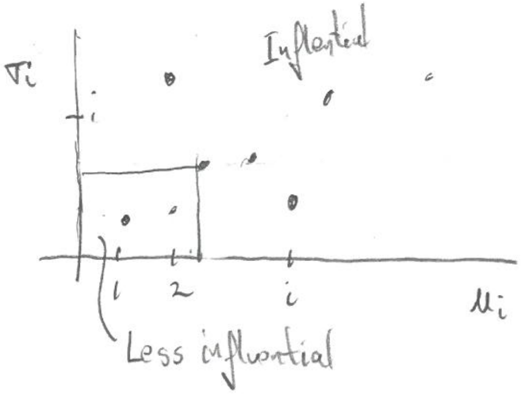
i^{th} component

Global Sensitivity Measures: , $i = 1, \ldots, p$

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

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

Objective: Use μ_i^* and σ_i^2 to quantify relative influence of parameters



e.g.) $p = 8$

Issues: 1.) Relative rather than absolute ranking

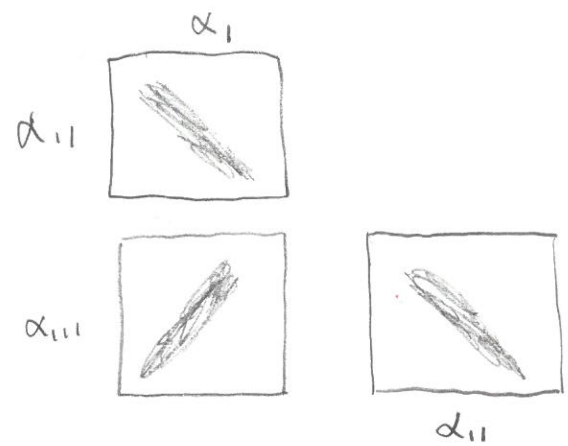
2.) Parameters often correlated and hence not independent. Can make incorrect conclusions based on incorrect assumption of independence.

3.) How do we construct indices for time- or space-dependent responses – more generally infinite-dimensional? Vector valued responses?

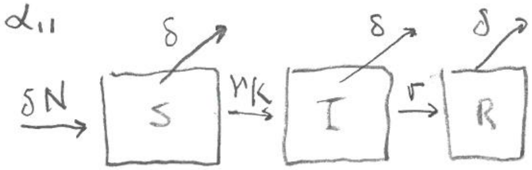
e.g.) $y(p,q) = \alpha_1 p^2 + \alpha_{11} p^4 + \alpha_{111} p^6$

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

	α_1	α_{11}	α_{111}
μ_i^*	0.17	0.07	0.03



Example: SIR Model



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

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

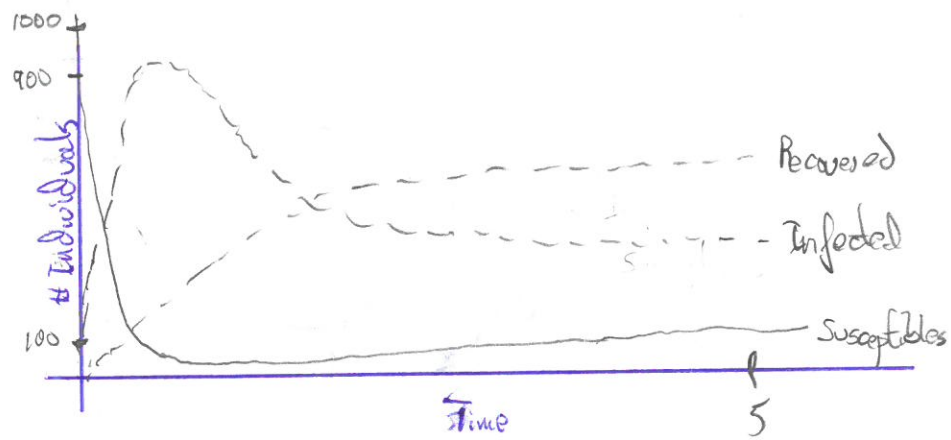
$$\frac{dR}{dt} = rI - \delta R \quad , \; R(0) = R_0$$

Assumed Parameter Distributions

$\gamma \sim \mathcal{U}(0,1)$	$k \sim \text{Beta}(\alpha, \beta)$	$r \sim \mathcal{U}(0,1)$	$\delta \sim \mathcal{U}(0,1)$
Infection Coef	Interaction Coef	Recovery Rate	Birth/Death Rate

Note: $Q = [\gamma, k, r, \delta]$ not jointly identifiable

Typical Realization: $S(0) = 900,\ I(0) = 100,\ R(0) = 0$



Note: $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$

$\Rightarrow S(t) + I(t) + R(t) = N$ Population constant

Quantity of Interest:

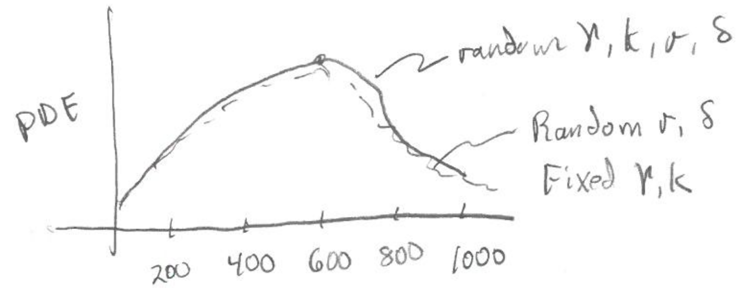
$$y(q) = \int_0^5 R(t,q)\,dt$$

Morris Indices:

	γ	k	r	δ
$\mu_i^* \ (\times 10^3)$	.25	.28	2.02	1.23
$\sigma_i \ (\times 10^3)$	.95	1.62	6.67	3.99

Conclusion: γ, k less influential and can be fixed at nominal values

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



Question: How do we construct time-dependent indices?

Analysis of Variance (Sobol):

- Initial Assumptions: $Q_i \sim \mathcal{U}(0,1)$, independent, $\Gamma = [0,1]^p$
- Sobol Representation: Truncated at 2^{nd} order – Exact if p^{th} order

$$f(q) \simeq f_0 + \underbrace{\sum_{i=1}^{p} f_i(q_i)}_{\text{1st order interactions}} + \underbrace{\sum_{1 \le i < j \le p} f_{ij}(q_i, q_j)}_{\text{2nd order interactions}}$$

Notes: i) Analogies: Taylor & Fourier series

ii) Need constraints to construct unique representation

- Derivatives: Taylor
- Orthogonality: Fourier

e.g. $f(q) = \sin(\pi q)$

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

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

Sobol Constraints:

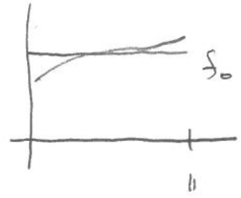
$$\int_0^1 f_i(q_i)\,dq_i = \int_0^1 f_{ij}(q_i, q_j)\,dq_i = \int_0^1 f_{ij}(q_i, q_j)\,dq_j = 0$$

Ensures:

$$\int_\Gamma f_i(q_i) f_j(q_j)\,dq_i\,dq_j = \int_\Gamma f_i(q_i) f_{ij}(q_i, q_j)\,dq_i\,dq_j = 0$$

Then

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



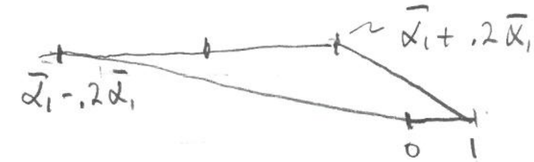
$$f_i(q_i) = \int_{\Gamma^{p-1}} f(q)\,dq_{\sim i} - f_0$$

$$f_{ij}(q_i, q_j) = \int_{\Gamma^{p-2}} f(q)\,dq_{\sim\{ij\}} - f_i(q_i) - f_j(q_j) - f_0$$

Note: $q_{\sim i} = [q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_p]$

Example: $y = \int_0^1 [\alpha_1 P^2 + \alpha_{11} P^4]\,dP = \frac{1}{3}\alpha_1 + \frac{1}{5}\alpha_{11}$

Note: $\alpha_1 \sim \mathcal{U}([\bar{\alpha}_1 - .2\bar{\alpha}_1, \bar{\alpha}_1 + .2\bar{\alpha}_1])$



$\bar{\alpha}_1 = -389.4$

$q_1 \in [0,1]$

Consider $y = aq_1 + bq_2$.

Then

$$f_0 = \int_0^1 \int_0^1 [aq_1 + bq_2]\,dq_1\,dq_2 = \frac{a+b}{2}$$

$$f_1(q_1) = \int_0^1 [aq_1 + bq_2]\,dq_2 - f_0 = aq_1 - \frac{a}{2}$$

$$f_2(q_2) = \int_0^1 [aq_1 + bq_2]\,dq_1 - f_0 = bq_2 - \frac{b}{2}$$

Note: $\int_0^1 f_1(q_1)\,dq_1 = \int_0^1 f_2(q_2)\,dq_2 = 0$

Statistical Interpretations:

Notation: $\mathbb{E}(Y|q_i) = \int_{\Gamma^{p-1}} f(q)\, dq_{\sim i}$

$$\mathbb{E}(Y|q_i,q_j) = \int_{\Gamma^{p-2}} f(q)\, dq_{\sim\{ij\}}$$

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

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

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

$$f_{ij}(q_i,q_j) = \mathbb{E}(Y|q_i,q_j) - f_i(q_i) - f_j(q_j) - f_0$$

Total Variance:

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

Partial Variances:

$$D_i = \int_0^1 f_i^2(q_i)\, dq_i \qquad \text{since } \int_0^1 f_i(q_i)\, dq_i = 0$$

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

Sobol Indices:

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

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

Variance Interpretations: Verified shortly

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

and

$$S_{T_i} = 1 - \frac{var[\mathbb{E}(Y|q_{\sim i})]}{var(Y)}$$

Example: $Y = c_1 Q_1 + c_2 Q_2$ $\qquad Q_1 \sim N(0,1)$, $Q_2 \sim N(0,9)$

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

Then

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

$$= \int_0^1 \left[\int_{\Gamma^{p-1}} f(q)\, dq_{\sim i} - f_0\right]^2 dq_i$$

$$= \int_0^1 \left[\int_{\Gamma^{p-1}} f(q)\, dq_{\sim i}\right]^2 dq_i - f_0^2 \quad *$$

$$= \mathbb{E}[\mathbb{E}(Y|q_i)]^2 - [\mathbb{E}[\mathbb{E}(Y|q_i)]]^2 = var[\mathbb{E}(Y|q_i)]$$

since

$$\mathbb{E}[\mathbb{E}(Y|q_i)] = \int_0^1 \Big[\int_{\Gamma^{p-1}} \underset{*}{f(q)}\, dq_{\sim i}\Big] dq_i = f_0$$

Example: SIR Model

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

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

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

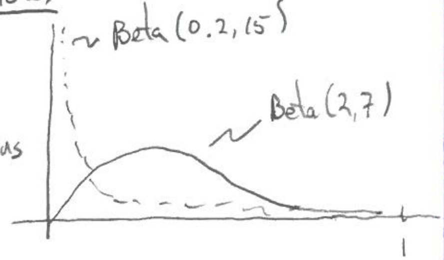
with

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

and

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

Note:



Results: Beta(2,7) – Many interactions

	γ	k	r	δ
S_i	0.1	0.03	0.79	0.18
S_{T_i}	-0.06	-0.05	0.56	0.20
$\mu_i^* \,(\times 10^3)$	0.25	0.28	2.02	1.23
$\sigma_i \,(\times 10^3)$	0.95	1.62	6.67	3.99

Note: 1) See previous plot of pdf at $t_f = 5$
2) See book for Beta(0.2, 15)

Example: $Y = c_1 Q_1 + c_2 Q_2$ with $Q_1 \sim N(0,\sigma_1^2)$

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

and

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

$c_1 = 2,\ c_2 = 1$

$$\Rightarrow p(q_1) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-q_1^2/2\sigma_1^2}$$

$$p(q_2) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-q_2^2/2\sigma_2^2}$$

Then

$$f_0 = \iint_{\mathbb{R}^2} [c_1 q_1 + c_2 q_2]\, p(q_1)\, p(q_2)\, dq_1\, dq_2 = 0$$

$$f_1(q_1) = \int_{\mathbb{R}} [c_1 q_1 + c_2 q_2]\, p(q_2)\, dq_2 = c_1 q_1$$

$$f_2(q_2) = \int_{\mathbb{R}} [c_1 q_1 + c_2 q_2]\, p(q_1)\, dq_1 = c_2 q_2$$

$$f_{12}(q_1, q_2) = 0$$

and

$$D_i = \int_{\mathbb{R}} f_i^2(q_i)\, p(q_i)\, dq_i = \int_{\mathbb{R}} c_i^2 q_i^2\, p(q_i)\, dq_i = c_i^2 \sigma_i^2$$

$$D_{ij} = \iint_{\mathbb{R}^2} f_{ij}^2\, p(q_i)\, p(q_j)\, dq_i\, dq_j = 0$$

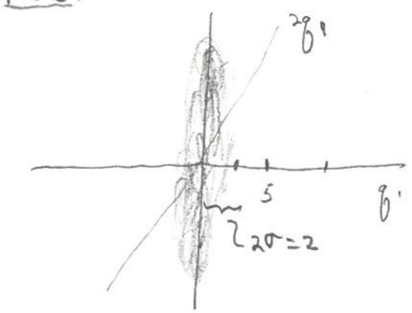
$$\Rightarrow D = D_1 + D_2 + \sum_{1 \le i < j \le 2} D_{ij} = c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2$$

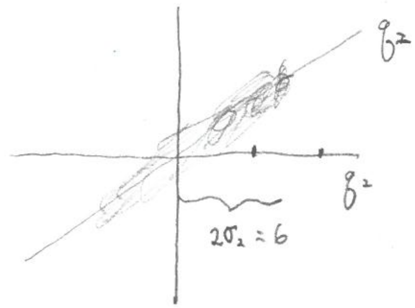
Finally,

$$S_i = \frac{c_i^2 \sigma_i^2}{c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2}$$

$$\Rightarrow \boxed{S_1 = \frac{4}{13},\quad S_2 = \frac{9}{13}}$$

Plots:





$D_1 = 4 \cdot 1 = 4$ $D_2 = 1 \cdot 9 = 9$

Note: $S_{T_i} = \dfrac{\mathbb{E}[\mathrm{var}(Y|q_{\sim i})]}{\mathrm{var}(Y)}$

Then

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

$$\Rightarrow \mathrm{var}(Y|q_i) \simeq 0 \quad \text{since } \mathrm{var}(Y|q_i) \geq 0$$

$\Rightarrow$ Noninfluential

Example: Consider

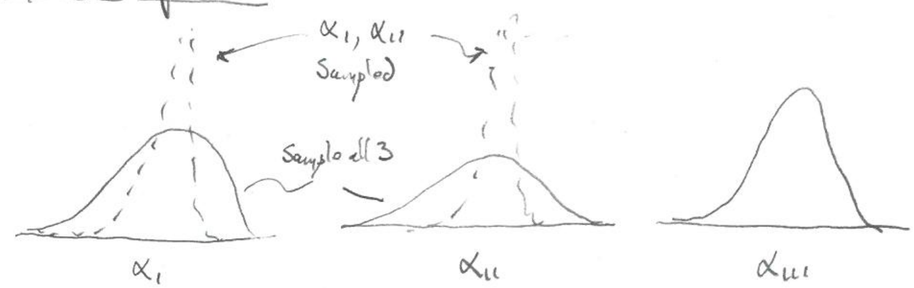
$$Y(P,q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$

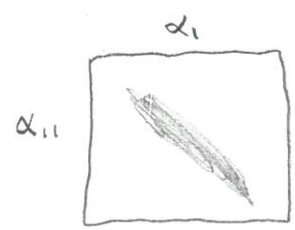
with $q = [\alpha_1, \alpha_{11}, \alpha_{111}]$.

Results:

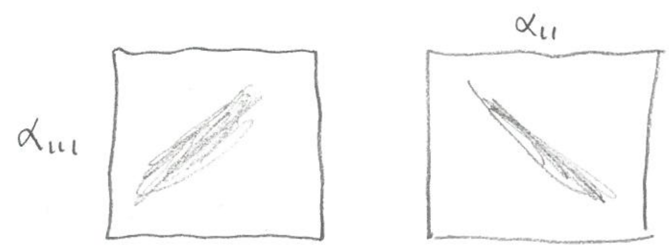
	α_1	α_{11}	α_{111}
S_i	.62	.39	.01
S_{T_i}	.66	.38	.06
μ_i^*	.17	.07	.03

Bayesian Inference:





Note:
- Not independent!!
- Extensions exist to incorporate distribution but difficult



Implementation Algorithms: If one uses M Monte Carlo evaluations to approximate $\mathbb{E}(Y|q_i)$ and repeats M times to approximate variance, this requires $M^2 p$ evaluations, which is typically computationally prohibitive. Various algorithms reduce this to $M(p+2)$; e.g., see work by Saltelli.

Example: Algorithm 15.10 in Book

Recall:

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

where

$$D = \int_\Gamma f^2(q)dq - f_0^2$$

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

$$= \int_0^1 \left[\int_{\Gamma^{p-1}} f(q)dq_{\sim i} - f_0 \right]^2 dq_i$$

$$= \int_0^1 \left[\int_{\Gamma^{p-1}} f(q)dq_{\sim i} \right]^2 dq_i - f_0^2$$

Algorithm 15.10. 1. Create two $M \times p$ sample matrices

$$A = \begin{bmatrix} q_1^1 & \cdots & q_i^1 & \cdots & q_p^1 \\ \vdots & & & & \vdots \\ q_1^M & \cdots & q_i^M & \cdots & q_p^M \end{bmatrix} , \; B = \begin{bmatrix} \hat{q}_1^1 & \cdots & \hat{q}_i^1 & \cdots & \hat{q}_p^1 \\ \vdots & & & & \vdots \\ \hat{q}_1^M & \cdots & \hat{q}_i^M & \cdots & \hat{q}_p^M \end{bmatrix},$$

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

2. Create $M \times p$ matrices

$$C_i = \begin{bmatrix} \hat{q}_1^1 & \cdots & q_i^1 & \cdots & \hat{q}_p^1 \\ \vdots & & & & \vdots \\ \hat{q}_1^M & \cdots & q_i^M & \cdots & \hat{q}_p^M \end{bmatrix},$$

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

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

$$y_A = f(A) \; , \; y_B = f(B) \; , \; y_{C_i} = f(C_i)$$

by evaluating the model at the input values in A, B, and C_i. The evaluation of y_A and y_B requires $2M$ model evaluations, whereas the evaluation of y_{C_i}, $i = 1, \ldots, p$, requires pM evaluations. Hence the total number of model evaluations is $M(p+2)$.

4. The estimates for the first-order sensitivity indices are

$$S_i = \frac{\mathrm{var}[\mathbb{E}(Y|q_i)]}{\mathrm{var}(Y)} = \frac{\frac{1}{M} y_A^T y_{C_i} - f_0^2}{\frac{1}{M} y_A^T y_A - f_0^2} = \frac{\frac{1}{M}\sum_{j=1}^M y_A^j y_{C_i}^j - f_0^2}{\frac{1}{M}\sum_{j=1}^M (y_A^j)^2 - f_0^2}, \tag{15.24}$$

where the mean is approximated by

$$f_0^2 = \left(\frac{1}{M}\sum_{j=1}^M y_A^j\right)\left(\frac{1}{M}\sum_{j=1}^M y_B^j\right). \tag{15.25}$$

The estimates for the total effects indices are

$$S_{T_i} = 1 - \frac{\mathrm{var}[\mathbb{E}(Y|q_{\sim i})]}{\mathrm{var}(Y)} = 1 - \frac{\frac{1}{M} y_B^T y_{C_i} - f_0^2}{\frac{1}{M} y_A^T y_A - f_0^2} = 1 - \frac{\frac{1}{M}\sum_{j=1}^M y_B^j y_{C_i}^j - f_0^2}{\frac{1}{M}\sum_{j=1}^M (y_A^j)^2 - f_0^2}. \tag{15.26}$$