解説

グローバル感度解析と 高次元モデル表現（HDMR）

香田　正人

1.　はじめに

感度解析は，システムモデルにおける入力変数（パラメータ）の現象論的な不確実さが，出力変数の認識論的な不確実さに与える影響を定量化する技術である．感度解析の近年における技術展開については，本誌第55巻第10号の特集[1]に詳しい．本稿では，グローバル感度解析[2]の開発において定式化された，高次元モデル表現（HDMR：High Dimensional Model Representation）について補足を行い，いくつかの解析的な例を示す．

2.　グローバル感度解析

グローバル感度解析が，従来の感度解析と大きく異なるのは次の3点である．

① 着目する入力変数は，一定あるいは無限のパラメータ空間で，その不確実さを表現する確率密度関数で示される分布に従って調べられる．

② 入力変数によって生起されたモデル出力は，全パラメータ空間における他のすべての入力変数のグローバルな変化も同時に勘案して評価される．

③ モデル独立である．

ここにモデル独立（model independent）とは，特定のモデル形式（例えば，線型システムか否か）や個別のモデル構造（加法性や単調性など）に制約されることなく，あるいは，入力変数間の相互作用の有無などにもよらず，感度解析が実行可能なことである．

すなわち，グローバル感度解析では，対象への入出力関係などのシステム機能に着目し，モデルが内包する不確実さを確率密度関数の分布形状として組み込んでいる．さらに，現象に対して演繹的に導かれたモデル構造に形式的に依存するのではなく，メタモデルによる帰納的でモデル独立な定式化がなされているので一般性が高い．具体的には，分散分析の適用により，モデル出力の総分散（すなわち不確実さ）を関連するモデル因子の各次元の偏分散（partial variances）に分解して，総分散への寄与度を評価する[2]．

3.　分散分析に基づく感度指標

モデル関数を $y=f(x_1, x_2, \cdots, x_k)=f(\mathbf{x})\in L_2$ とする．ここに，y はモデル出力，$x_i(i=1, 2, \cdots, k)$ は k 次元の入力変数（パラメータ）である．簡単のため，入力変数が単位超立方体 $\Omega^k=\{\mathbf{x}|0\leq x_i\leq 1\ (i=1, \cdots, k)\}$ 上で一様分布する場合を取り扱う．

着目する入力変数 x_i の不確実さ低減によるモデル出力の分散低減は，$E(\cdot)$ を期待値，$E(y|x_i)$ を条件付期待値として，次式で定義される V_i によって評価できる．

$$V_i=V(y)-E[V(y|x_i)]=V[E(y|x_i)] \tag{1}$$

ここに，$V(y)$ はモデル出力の分散，$V(y|x_i)$ は着目する入力変数 x_i を固定したときの条件付分散を表し，(1)式右辺の分散 $V[E(y|x_i)]$ は以下で定義される．

$$V[E(y|x_i)]=E[\{E(y|x_i)-E(y)\}^2] \tag{2}$$

(1)式は x_i の不確実さが除かれた場合のモデル出力の分散低減の期待値を表す．なお，(1)式の V_i と(2)式は等価であるが，それらは後述する1次の偏分散に他ならない．

(1)式の V_i を分散 $V(y)$ で除して，1次感度指標を次式で定義する．

$$S_i=\frac{V[E(y|x_i)]}{V(y)}=\frac{V_i}{V(y)} \tag{3}$$

明らかに $0\leq S_i\leq 1$ である．

着目する入力変数の次元の増加に従って(1)式の V_i

こうだ　まさと
筑波大学 大学院システム情報工学研究科
〒305-8573 つくば市天王台 1-1-1

を拡張することにより，高次の感度指標を定義できる．例えば，2入力変数 $x_i, x_j (i \neq j)$ に対する条件付分散は，それぞれの単独の1次効果の和 $V_i + V_j$ に，残差に相当する2次オーダーの項として V_{ij} を付加して次式で表せる．

$$V[E(y|x_i, x_j)] = V_i + V_j + V_{ij} \tag{4}$$

(4)式から，V_{ij} が以下で求まる．

$$V_{ij} = V[E(y|x_i, x_j)] - V_i - V_j \tag{5}$$

上記の分散分析の数学的整合性は，後述する Sobol' 法[3]によって示すことができる．

(5)式を $V(y)$ で除して2次感度指標を次式で定義する．

$$S_{ij} = \frac{V_{ij}}{V(y)} = \frac{V[E(y|x_i, x_j)]}{V(y)} - S_i - S_j \tag{6}$$

分散分析では，モデル出力の総分散について以下の分散分解（variance decomposition）が知られている．

$$V(y) = \sum_{i=1}^{k} V_i + \sum_{1 \leq i < j \leq k} V_{ij} + \cdots + V_{12\cdots k} \tag{7}$$

ここで，感度指標を $S_{ij\ell mn} = V_{ij\ell mn}/V(y)$ などと分散比で定義し，(7)式の両辺を総分散で除せば次式を得る．

$$1 = \sum_{i-1}^{k} S_i + \sum_{1 \leq i < j \leq k} S_{ij} + \cdots + S_{12\cdots k} \tag{8}$$

上式は，感度指標の総和が1となるので便利である．

4. 高次元モデル表現（HDMR）

Sobol' 法[3]では入力変数の次元の増加に従って，モデル関数を以下のように展開する．

$$f(\mathbf{x}) = f(x_1, \cdots, x_k) = f_0 + \sum_{i=1}^{k} f_i(x_i) + \sum_{1 \leq i < j \leq k} f_{ij}(x_i, x_j) + \cdots + f_{12\cdots k}(x_1, x_2, \cdots x_k) \tag{9}$$

上式において，$f_0 = \int_{\Omega^k} f(\mathbf{x}) d\mathbf{x}$ は定数（平均値）である．また，各展開項に含まれる入力変数の組合せの数は次元の増加に対応して ${}_kC_1, {}_kC_2, \cdots, {}_kC_k$ となり，f_0 を含めると(9)式右辺の全項数は $1 + {}_kC_1 + {}_kC_2 + \cdots + {}_kC_k = 2^k$ となる．上の展開は，一般に高次元モデル表現（HDMR：High Dimensional Model Representation）と呼ばれる[2]．

ここで，(9)式右辺の定数項を除くすべての展開項に関して，各項に含まれる任意の入力変数についての積分が0となることを仮定すると，異なる2つの展開項間に Ω^k 上の積分（内積）に関する直交性が成立する．このとき，着目する x_i を除く $(k-1)$ 次元の入力変数について(9)式を積分すると，1次の展開項 $f_i(x_i)$ が以下のように求まる．

$$\int_0^1 \cdots \int_0^1 f(\mathbf{x}) d\mathbf{x}_{\sim i} = f_0 + f_i(x_i) \tag{10}$$

ここに，$\int \cdots \int (\cdot) d\mathbf{x}_{\sim i}$ は x_i を除くすべての入力変数についての積分を表す．次に，2次の展開項 $f_{ij}(x_i, x_j)$ についても，着目する $x_i, x_j (i \neq j)$ を除く $(k-2)$ 次元の入力変数について(9)式を積分すれば次式が得られる．

$$\int_0^1 \cdots \int_0^1 f(\mathbf{x}) d\mathbf{x}_{\sim ij} = f_0 + f_i(x_i) + f_j(x_j) + f_{ij}(x_i, x_j) \tag{11}$$

上式の $\int \cdots \int (\cdot) d\mathbf{x}_{\sim ij}$ は $x_i, x_j (i \neq j)$ を除くすべての入力変数についての積分を示す．同様にして，入力変数の次元の増加に対応して，積分次元を順次下げていけば(9)式の展開が得られる．

ここで，下の(12)式で与えられる総分散 D と，各展開項の2乗の積分（期待値）として(13)式により評価される偏分散（partial variance）$D_{i_1 \cdots i_s}$ を定義する．

$$D = \int_{\Omega^k} f^2(\mathbf{x}) d\mathbf{x} - f_0^2 \tag{12}$$

$$D_{i_1 \cdots i_s} = \int_0^1 \cdots \int_0^1 f_{i_1 \cdots i_s}^2 (x_{i_1}, \cdots, x_{i_s}) dx_{i_1} \cdots dx_{i_s} \tag{13}$$

(13)式中の添字は $1 \leq i_1 < \cdots < i_s \leq k (s = 1, \cdots, k)$ である．

(9)式の両辺を2乗して Ω^k 上で積分し，各展開項間の直交性を用いると，(12)式の総分散が入力変数の次元の増加に従った偏分散の和に展開できて次式が得られる．

$$D = \sum_{i-1}^{k} D_i + \sum_{1 \leq i < j \leq k} D_{ij} + \cdots + D_{12\cdots k} \tag{14}$$

上式は，分散分析で得られた(7)式に他ならず，(14)式右辺の全項数は $(2^k - 1)$ となる．これは，分散分析（ANOVA）に対応する総分散の HDMR 展開として，一般に ANOVA-HDMR と呼ばれる．上式中の1次の偏分散 D_i は入力変数 x_i に関する主効果を表し，2次の偏分散 $D_{ij} (i \neq j)$ は，入力変数単独の主効果の和 $D_i + D_j$ では説明できない2変数間の交互効果（interaction effect）を表す．高次の偏分散についても同様である．

条件付期待値を用いれば，(9)式右辺における HDMR の2次までの展開項が以下のように表される．

$$f_0 = E(f) = \int_{\Omega^k} f(\mathbf{x}) d\mathbf{x} \tag{15}$$

$$f_i = E(f|x_i) - f_0 = \int_0^1 \cdots \int_0^1 f(\mathbf{x}) d\mathbf{x}_{\sim i} - f_0 \tag{16}$$

$$f_{ij} = E(f|x_i, x_j) - f_0 - f_i - f_j$$

$$=\int_0^1\cdots\int_0^1 f(\mathbf{x})d\mathbf{x}_{\sim ij}-f_0-f_i-f_j \qquad (17)$$

上の(16)式と(17)式中の積分は，着目する入力変数を除いて評価した条件付期待値を表し，両式はそれぞれ(10)式と(11)式に対応する．高次の展開項についても同様な関係式が成立する．

1次の偏分散は，(16)式を2乗して期待値をとることにより次式となる．

$$D_i=E[f_i^2(x_i)]=E[\{E(f|x_i)-f_0\}^2]$$
$$=V[E(f|x_i)] \qquad (18)$$

これは(1)式および(2)式に他ならない．また，(17)式より得られる関係式，$E(f|x_i, x_j)-f_0=f_i+f_j+f_{ij}$ の両辺を2乗して期待値をとり，各展開項の直交性を用いれば，対応する2次の偏分散が下のように得られる．

$$D_{ij}=E[f_{ij}^2(x_i, x_j)]$$
$$=V[E(f|x_i, x_j)]-D_i-D_j \qquad (19)$$

上式は(5)式あるいは(4)式と等価で，(5)式の数学的な整合性を示すものである．同様に，高次の偏分散も再帰的に計算することができる．

1970年代に最初のグローバル感度解析手法として開発されたFAST (Fourier Amplitude Sensitivity Test) [4][5]もANOVA-HDMRの一例である．

5. 例

グローバル感度解析においては，全パラメータ空間にわたる入力変数の不確実さの分布に沿って，モデル出力の総分散が一度に包括的に評価される．一般的な解析には，確率密度関数に従った入力変数のサンプリングに基づいて，モンテカルロ法を適用するのが普通である．以下では，高次元モデル表現（HDMR）の標準的なケースとして，入力変数が一様分布や正規分布をする場合に，HDMRへの有用な洞察を得ることができるような例をいくつか示す．

5.1 Ishigami Function；x_i は3次元の一様分布

次のIshigami functionはグローバル感度解析における代表的なテスト関数である[6]．

$$f(x_1, x_2, x_3)=\sin(x_1)+a\sin^2(x_2)+bx_3^4\sin(x_1)$$

ここで，3次元の入力変数は区間 $[-\pi, \pi]$ 上の一様分布とする．

Ishigami functionの関数形から $f_3=f_{12}=f_{23}=f_{123}=0$ が明らかなので，HDMRは次式となる．

$$f(x_1, x_2, x_3)=f_0+f_1(x_1)+f_2(x_2)+f_{13}(x_1, x_3)$$
$$=a/2+(1+b\pi^4/5)\sin(x_1)$$
$$+a(\sin^2(x_2)-1/2)+b(x_3^4-\pi^4/5)\sin(x_1)$$

ここに，(15)式から下の定数項を得る．

$$f_0=E(f)=\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f(x_1, x_2, x_3)dx_1dx_2dx_3/8\pi^3$$
$$=a/2$$

1次の展開項は，(16)式に従い次のように計算される．

$$f_1(x_1)=E(f|x_1)-f_0$$
$$=\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f(x_1, x_2, x_3)dx_2dx_3/4\pi^2-a/2$$
$$=\sin(x_1)+\int_{-\pi}^{\pi}a\sin^2(x_2)dx_2/2\pi$$
$$+\sin(x_1)\int_{-\pi}^{\pi}bx_3^4dx_3/2\pi-a/2$$
$$=(1+b\pi^4/5)\sin(x_1)$$
$$f_2(x_2)=E(f|x_2)-f_0=a(\sin^2(x_2)-1/2)$$

(17)式より，2次の項 $f_{13}=b(x_3^4-\pi^4/5)\sin(x_1)$ を得る．さらに，任意の異なる2項がすべて直交することが分かる．

次に，Ishigami functionのANOVA-HDMRを求める．まず，総分散が(12)式から下のように得られる．

$$D=1/2+a^2/8+b\pi^4/5+b^2\pi^8/18$$

1次の偏分散は，(18)式に従い次のように計算される．

$$D_1=\int_{-\pi}^{\pi}f_1^2(x_1)dx_1/2\pi=1/2+b\pi^4/5+b^2\pi^8/50$$
$$D_2=\int_{-\pi}^{\pi}f_2^2(x_2)dx_2/2\pi=a^2/8$$

2次の偏分散は，(19)式より $D_{13}=b^2\pi^8/18-b^2\pi^8/50$ となる．また，$D_3=D_{12}=D_{23}=D_{123}=0$ なので，分散分解 $D=D_1+D_2+D_{13}$ が成立することを確認できる．感度指標は(8)式に従い計算すればよい．

5.2 $y=\prod_{i=1}^{k}x_i$；x_i は区間 $[0, X_i]$ 上の一様分布

まず，(15)式より下の定数項が得られる．

$$f_0=E(y)=\prod_{i=1}^{k}\left(\int_0^{X_i}x_idx_i/X_i\right)=\prod_{i=1}^{k}(X_i/2)$$

さらに，(16)式と(17)式に従って，HDMRの1次と2次の各展開項が次のように計算できる．

$$f_i=E(y|x_i)-f_0=x_i\prod_{j\neq i}\left(\int_0^{X_j}x_jdx_j/X_j\right)-f_0$$
$$=(x_i-X_i/2)\prod_{j\neq i}(X_j/2)$$
$$f_{ij}=E(y|x_i, x_j)-f_i-f_j-f_0$$
$$=x_ix_j\prod_{\ell\notin(i,j)}\left(\int_0^{X_\ell}x_\ell dx_\ell/X_\ell\right)-f_i-f_j-f_0$$
$$=(x_i-X_i/2)(x_j-X_j/2)\prod_{\ell\notin(i,j)}(X_\ell/2)$$

HDMRの一般項は，下のように求められる．

$$f_{i_1\cdots i_s}=\prod_{j\in I_s}(x_j-X_j/2)\prod_{\ell\notin I_s}(X_\ell/2)$$

ここに，I_s は入力変数を示す添字の集合 $I_s=$

$(i_1, \cdots, i_s)$ を表し，入力変数間の交互効果が2次以上のすべての展開項に及ぶことを示している．また，任意の異なる2展開項がすべて直交することも分かる．

総分散は，⑿式から次のように計算される．

$$D=\prod_{i=1}^{k}\left(\int_0^{X_i} x_i^2 dx_i/X_i\right)-f_0^2=(3^{-k}-4^{-k})\prod_{i=1}^{k}X_i^2$$

また，n 次 $(1\leq n\leq k)$ の偏分散は，n 個の入力変数の組合せによらず，同一値 $D(n)=4^{-k}\prod_{i=1}^{k}X_i^2/3^n$ となる．したがって，n 次の感度指標は $D(n)/D=3^{-n}/\{(4/3)^k-1\}$ となり，感度は次数の増加に従って1/3のべき（冪）で減少する．そうして，次数の増加に対応して偏分散に含まれる入力変数の組合せの数が ${}_kC_1, {}_kC_2, \cdots, {}_kC_k$ となることから，下のように⒁式の分散分解が成立する．

$$\sum_{n=1}^{k}{}_kC_nD(n)=4^{-k}\{(1/3+1)^k-1\}\prod_{i=1}^{k}X_i^2=D$$

5.3 拡散プロセス

上の例に関連して，拡散プロセス（diffusion process）では，状態の初期値を y_0 として現時点 k における状態を $y_k=y_0\prod_{i=1}^{k}x_i$ とモデル化する場合がある．すなわち，x_i を時点 i におけるランダム係数として $(0, 1)$ 上の一様分布とし，状態変数に関して自己相似性（self-similarity）を有するランダム・インパルス $(0<x_i<1)$ の集積として現在値を表現する．このとき，$y_k=y_0\prod_{i=1}^{k}x_i$ に対数変換を適用して，$\log(y_k)=\log(y_0)+\sum_{i=1}^{k}\log(x_i)$ を考える．そうして，$\log(y_k/y_0)=\sum_{i=1}^{k}\log(x_i)$ の右辺各項が，大数の法則に基づいて正規分布をするとした加法的メタモデル，すなわち，対数正規分布のアンサンブルとして取り扱うと解析に便利である．こうしたモデルが金融工学や数理ファイナンスの分野に応用されることも多い．

5.4 $y=\prod_{i=1}^{k}x_i$；x_i は $N(0, \sigma_i)$ の正規分布

4節のSobol' 法に従って計算すれば，展開項は最高次の項を除いてすべて0となるので，HDMRが極めて簡単に $y=f_{12\cdots k}(x_1, x_2, \cdots, x_k)=\prod_{i=1}^{k}x_i$ となる．これは，純粋なノイズであって意味ある情報は得られない．そこで，モデルを2乗して，$y^2=\prod_{i=1}^{k}x_i^2$ についての対数変換により，$\log(y^2)=g(x_1, x_2, \cdots, x_k)=\sum_{i=1}^{k}\log(x_i^2)$ の加法的メタモデルを考える．このとき，主効果は $g_i=E[\log(y^2)|\log(x_i^2)]\propto\log(x_i^2)$ となり，変動 y^2 に対する x_i^2 の影響は，入力変数 x_i の変動が平均値0から離れるほど大きくなる．すなわち，標準偏差 $\sigma_i=\sqrt{E(x_i^2)}$ に比例して，入力変数の重要度をランク付けできる．

6. おわりに

作業仮説としてのモデルに関して，モデル因子を変化させたとき，異なる仮定に対して結果がどのくらい敏感に変動するかを検証し，その作業仮説の論理構造がどのように組み立てられているか，また代替的な帰結がどうなり得るかを分析する感度解析は，モデルの品質管理のためのツールとして理解可能である．特に，本稿で考察したグローバル感度解析は，メタモデルを含む様々なモデル構築に基づく研究方針や施策の優先度の検討など，社会的ソリューションの品質保証に活用できる．近年，科学技術を社会的文脈の中で新しく捉え直す，ポスト・ノーマルサイエンス（Post-Normal Science）[7]と呼ばれる考え方が提唱されている．グローバル感度解析はこうした考え方に沿った技術ともいえる．なお，ポスト・ノーマルサイエンスとグローバル感度解析の関連性については別稿[8]を参照されたい．

参考文献

[1] 香田正人編，特集 現代感度解析入門：FAST と Sobol' 法を中心として，オペレーションズ・リサーチ（日本オペレーションズ・リサーチ学会機関誌），Vol. 55, No. 10, 2010.

[2] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana and S. Tarantola : Global Sensitivity Analysis The Primer, John Wiley & Sons, 2008.

[3] I.M. Sobol' : "Sensitivity analysis for nonlinear mathematical models," Mathematical Modeling & Computational Experiment, Vol. 1, pp. 407–414, 1993.

[4] R.I. Cukier, H.B. Levine and K.E. Shuler : "Nonlinear sensitivity analysis of multiparameter model systems," Journal of Computational Physics, Vol. 26, pp. 1–42, 1978.

[5] M. Koda, G. McRae and J.H. Seinfeld : "Automatic sensitivity analysis of kinetic mechanisms," International Journal of Chemical Kinetics, Vol. 11, pp. 427–444, 1979.

[6] T. Ishigami and T. Homma : An importance quantification technique in uncertainty analysis for computer

models, Japan Atomic Energy Research Institute, Report JAERI-M 89-111, 1989.
[7] J.R. Ravetz : "What is Post-Normal Science," Futures, Vol. 31, pp. 647-653, 1999.
[8] 香田正人, "ポスト・ノーマルサイエンスとグローバル感度解析," 横幹 (横幹連合会誌), Vol. 5, No. 1, pp. 37-40, 2011.