

多项式输出中相关变量的结构贡献和相关贡献分析*

吴丹清, 吕震宙, 胡吉祥

(西北工业大学 航空学院, 陕西 西安 710072)

摘要:以二次不含交叉项的多项式为例,解析推导了正态相关输入变量对输出响应量方差贡献的结构贡献部分和相关贡献部分。通过算例验证了解析结果的正确性。将所研究的指标和已有的进行对比,归纳出结构贡献部分和相关贡献部分的侧重与统一。解析解可直接用于极限状态函数不超过两次且不含交叉项的结构和相关贡献的识别,为其他新的数值算法提供了参考对照。

关键词:结构贡献;相关贡献;二次不含交叉项;解析解

中图分类号:TB114.3 **文献标志码:**A **文章编号:**1001-2486(2014)05-026-07

Analysis of structural and correlative contributions by correlated variables in polynomial output

WU Danqing, LYU Zhenzhou, HU Jixiang

(School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract:Based on the example of the quadratic polynomial without cross terms outputs, the structural and correlative contributions of the correlated normal input variables were derived out, which had impact on the variance contributions of output response variables. Numerical results of several examples validate the correctness of analytical solutions. Comparing the presented indices with the existent indices, some general conclusions were drawn from the structural contributions and the correlative contributions, and the differences between them were identified. The proposed analytical solutions can be directly used to identify the structural contributions and the correlative contributions of the quadratic polynomial without cross terms, which offer a reference to other numerical algorithms.

Key words: structural contribution; correlative contribution; quadratic polynomial without cross terms; analytical solution

重要性测度分析模型可分为3类:非参模型^[1-2]、矩独立模型^[3-4]和基于方差的模型^[5-7]。工程实际中很多输入变量之间存在相互影响,不能将各个变量简单地处理成独立变量来研究其方差贡献,因而研究相关输入变量的重要性测度十分必要。Xu提出将相关输入变量的方差贡献 S_i 分解为独立贡献部分 S_i^U 和相关贡献部分 S_i^C ^[8],该指标的模型基础为Sobol等证明的方差分解式,只适用于线性可加模型。Mara提出相关变量正交独立化变换方法^[9],转换之后的变量独立部分去除了该变量与其他变量的相关部分和交叉部分,将变量的方差贡献分离得更彻底,但该变换每次只对一个变量变换,并会产生意义不明确的中间变量增加计算量。Li提出将相关变量方差贡献分为结构贡献和相关贡献^[10],该方法不但适用于非线性模型和非线性相关情况,而且能够清晰

反映输入变量各部分方差贡献。

1 Li 指标的定义

为分清相关变量的方差贡献的来源, Li 提出将相关输入变量的方差贡献分解为结构贡献 $S_{p_j}^a$ 和相关贡献 $S_{p_j}^b$ ^[10]:

$$S_{p_j} = S_{p_j}^a + S_{p_j}^b \quad (1)$$

Li 将 n 维输入变量 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 的输出响应量 $f(\mathbf{x})$ 的表达式分解为:

$$f(\mathbf{x}) = f_0 + \sum_{i=1}^n f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) + \dots + f_{12 \dots n}(x_1, x_2, \dots, x_n) = f_0 + \sum_{j=1}^{2^n-1} f_{p_j}(\mathbf{x}_{p_j}) \quad (2)$$

其中 $f_0 = E(Y)$, $f_i(x_i)$ 表示变量 x_i 的输出贡献,

* 收稿日期:2014-04-01

基金项目:国家自然科学基金资助项目(51175425),航空基金资助项目(2011ZA53015),博士学科点专项科研基金资助项目(20116102110003)

作者简介:吴丹清(1990—),女,江苏南京人,博士研究生,E-mail:wudanqing86@163.com;

吕震宙(通信作者),女,教授,博士,博士生导师,E-mail:zhenzhou@nwpu.edu.cn

$f_{ij}(x_i, x_j)$ 以及更高阶的分解项反映了输入变量间由于响应量表达式的形式而产生的交互作用的贡献,并在此基础上给出灵敏度指标的定义:

$$\begin{cases} S_{p_j} = Cov(f_{p_j}, Y) / V(Y) \\ S_{p_j}^a = Var(f_{p_j}) / V(Y) \\ S_{p_j}^b = Cov(f_{p_j}, \sum_{k=1, k \neq j}^{n_p} f_{p_k}) / V(Y) \end{cases} \quad (3)$$

本文以二次不含交叉项的输出响应表达式为研究对象,故只需考虑输出响应量 $f(x)$ 表达式中的一阶分解项,式(2)可化简为:

$$f(x) = f_0 + \sum_{i=1}^n f_i(x_i) \quad (4)$$

因而 Li 各项指标的计算公式可写为:

$$\begin{cases} S_i = Cov(f_i, Y) / V(Y) \\ S_i^a = Var(f_i) / V(Y) \\ S_i^b = Cov(f_i, \sum_{k=1, k \neq i}^n f_k) / V(Y) \end{cases} \quad (5)$$

式(5)中 S_i 反映变量 x_i 的总贡献, S_i^a 反映了 x_i 的结构贡献, S_i^b 反映了 x_i 与除去 x_i 之外的其他所有变量 x_{-i} 的相关贡献。从定义式来看,总贡献是变量的一阶分解项与总输出响应量之间协方差与总方差的比值,由于协方差反映了两者的线性相关关系,因而该指标衡量两者之间的总的关系;由于一阶分解项是含有该变量的组成功能函数,因而反映了该变量在输出响应表达式中的结构,结构贡献是变量一阶分解项的方差与总方差的比值。相关贡献是变量一阶分解项与其余变量的一阶分解项的协方差在总方差中所占的比例,是二者相关性的度量。

2 结构和相关贡献的解析解

2.1 输出响应量总方差

设含正态相关变量的二次不含交叉项的响应量表达式为:

$$y = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n b_i x_i^2 \quad (6)$$

式(6)中, a_0 为常数项, a_i 为第 i 个一阶项的系数, b_i 为第 i 个二阶项的系数, $x_i \sim N(\mu_i, \sigma_i^2)$ 。为得到 Li 提出的各项指标 (S_i, S_i^a, S_i^b) 的解析表达,需计算输出响应量的总方差 $V(Y)$, 第 i 个变量的组成分量函数 $f_i(x_i)$ 与输出响应量 Y 的协方差 $Cov(f_i, Y)$, 以及第 i 个变量的组成功能函数 $f_i(x_i)$ 自身的方差 $Var(f_i)$ 。

输出响应量的总方差贡献可以表示为

$$\begin{aligned} V(y) = & \sum_{i=1}^n [a_i^2 V(x_i) + b_i^2 V(x_i^2)] + \\ & \sum_{i=1}^n \sum_{j=1}^n 2a_i b_j Cov(x_i, x_j^2) + \\ & \sum_{i=1}^n \sum_{j=i+1}^n 2[a_i a_j Cov(x_i, x_j) + \\ & b_i b_j Cov(x_i^2, x_j^2)] \end{aligned} \quad (7)$$

在进一步推导之前,给出一些由正态分布变量的概率密度函数推导的公式^[11]:

$$\begin{cases} E(x^3) = 3\mu_x \sigma_x^2 + \mu_x^3 \\ E(x^4) = 3\sigma_x^4 + 6\mu_x^2 \sigma_x^2 + \mu_x^4 \\ E(x_i x_j) = \mu_i \mu_j + \rho_{ij} \sigma_i \sigma_j \\ E(x_i x_j^2) = \mu_i \mu_j^2 + (\mu_i \sigma_j + 2\rho_{ij} \mu_j \sigma_i) \sigma_j \\ E(x_i^2 x_j^2) = \mu_i^2 \mu_j^2 + \mu_j^2 \sigma_i^2 + \mu_i^2 \sigma_j^2 + \\ 4\rho_{ij} \mu_i \mu_j \sigma_i \sigma_j + (2\rho_{ij}^2 + 1) \sigma_i^2 \sigma_j^2 \end{cases} \quad (8)$$

ρ_{ij} 表示两个变量 x_i 和 x_j 之间的 Pearson 相关系数,利用方差公式和协方差的性质及式(8),输出响应量的总方差为^[11]:

$$\begin{aligned} V(y) = & \sum_{i=1}^n [a_i^2 + 2b_i^2(\sigma_i^2 + 2\mu_i^2)] \sigma_i^2 + \\ & \sum_{i=1}^n \sum_{j=1}^n 4a_i b_j \mu_j \rho_{ij} \sigma_i \sigma_j + \\ & \sum_{i=1}^n \sum_{j=i+1}^n 2\rho_{ij} \sigma_i \sigma_j [a_i a_j + \\ & 2b_i b_j (2\mu_i \mu_j + \rho_{ij} \sigma_i \sigma_j)] \end{aligned} \quad (9)$$

2.2 变量的各方差贡献

求解 Li 的指标,关键是求解变量相关情况下高维分解模型中的各个分量。本文采用 Li 提出的高维模型替代对其进行分析^[12]:

$$\begin{cases} f_0 = \int f(\mathbf{x}) \omega(\mathbf{x}) d\mathbf{x} \\ f_i(x_i) = \int f(\mathbf{x}) \omega_{-i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i} - f_0 - h_i(x_i) \\ f_{ij}(x_i, x_j) = \int f(\mathbf{x}) \omega_{-ij}(\mathbf{x}_{-ij}) d\mathbf{x}_{-ij} - f_0 - \\ f_i(x_i) - f_j(x_j) - h_{ij}(x_i, x_j) \end{cases} \quad (10)$$

其中 $\omega(\mathbf{x}) = \prod_{i=1}^n \omega_i(x_i)$, $\omega_i(x_i)$ 表示 x_i 的边缘概率密度函数, $\omega_{-i}(\mathbf{x}_{-i})$ 表示除 x_i 外其他所有变量的联合边缘概率密度函数,类似地, $\omega_{-ij}(\mathbf{x}_{-ij})$ 表示除 x_i 和 x_j 外其他所有变量的联合边缘概率密度函数。 $h_i(x_i)$ 与所有包含 x_i 并高于一阶的组成分量函数有关,相应地, $h_{ij}(x_i, x_j)$ 与所有包含 x_i 和 x_j 并且高于二阶的组成分量函数有关。由于本文研究对象为二次不含交叉项的输出响应函数,因而 $h_i(x_i) = 0$, 高维模型替代结果中只含有

f_0 和 $f_i(x_i)$ 项, $f_{ij}(x_i, x_j)$ 以及更高阶项均不予考虑。故本文用到的模型替代为:

$$\begin{cases} f_0 = \int f(\mathbf{x}) \omega(\mathbf{x}) dx \\ f_i(x_i) = \int f(\mathbf{x}) \omega_{-i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i} - f_0 \end{cases} \quad (11)$$

f_0 和 $f_i(x_i)$ 的解析表达为:

$$f_0 = E(Y) = a_0 + \sum_{i=1}^n a_i u_i + \sum_{i=1}^n b_i (\mu_i^2 + \sigma_i^2) \quad (12)$$

$$f_k = f_k(x_k) = \int f(\mathbf{x}) \omega_{-k}(\mathbf{x}_{-k}) d\mathbf{x}_{-k} - f_0 = a_0 + \sum_{i=1, i \neq k}^n a_i u_i + a_k x_k + \sum_{i=1, i \neq k}^n b_i (\mu_i^2 + \sigma_i^2) + b_k x_k^2 \quad (13)$$

变量 x_k 的总方差贡献为:

$$\begin{aligned} Cov(f_k, Y) &= \sum_{i=1}^n a_i a_k \rho_{ik} \sigma_i \sigma_k + \\ &\sum_{i=1}^n 2a_i b_k \rho_{ik} u_i \sigma_i \sigma_k + \sum_{i=1}^n 2a_k b_i \rho_{ik} u_i \sigma_i \sigma_k + \\ &\sum_{i=1}^n b_i b_k (4\rho_{ik} \mu_i \mu_k \sigma_i \sigma_k + 2\rho_{ik}^2 \sigma_i^2 \sigma_k^2) \end{aligned} \quad (14)$$

变量 x_k 的结构方差贡献:

$$V(f_k) = a_k^2 \sigma_k^2 + b_k^2 (2\sigma_k^4 + 4\mu_k^2 \sigma_k^2) + 4a_k b_k \mu_k \sigma_k^2 \quad (15)$$

式(5) 可计算 S_k 和 S_k^a, S_k^b 则用式(16) 计算:

$$S_k^b = S_k - S_k^a \quad (16)$$

式(14) 、式(15) 的推导过程如下:

由协方差计算公式可知: $Cov(f_i, Y) = E(f_i Y) - E(f_i)E(Y)$ 。

$$f_k Y = (a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n b_i x_i^2) [a_0 + a_k x_k + \sum_{i=1, i \neq k}^n a_i u_i + \sum_{i=1, i \neq k}^n b_i (\mu_i^2 + \sigma_i^2) + b_k x_k^2] \quad (17)$$

推导期望 $E(f_k Y)$:

$$\begin{aligned} E(f_k Y) &= a_0^2 + 2a_0 \sum_{i=1}^n a_i u_i + 2a_0 \sum_{i=1}^n b_i (\mu_i^2 + \sigma_i^2) + \\ &\sum_{i=1}^n \sum_{j=1}^n a_i a_j \mu_i \mu_j + 2 \sum_{i=1}^n \sum_{j=1}^n a_i b_j \mu_i (\mu_j^2 + \sigma_j^2) + \\ &\sum_{i=1}^n 2a_i b_k \rho_{ik} u_k \sigma_i \sigma_k + \sum_{i=1}^n 2a_k b_i \rho_{ik} u_i \sigma_i \sigma_k + \\ &\sum_{i=1}^n \sum_{j=1}^n b_i b_j (\mu_i^2 + \sigma_i^2) (\mu_j^2 + \sigma_j^2) + \sum_{i=1}^n a_i a_k \rho_{ik} \sigma_i \sigma_k + \\ &\sum_{i=1}^n b_i b_k (4\rho_{ik} \mu_i \mu_k \sigma_i \sigma_k + 2\rho_{ik}^2 \sigma_i^2 \sigma_k^2) \end{aligned} \quad (18)$$

f_k 和 Y 的期望的乘积为:

$$E(f_k)E(Y) = [a_0 + \sum_{i=1}^n a_i u_i + \sum_{i=1}^n b_i (\mu_i^2 + \sigma_i^2)] [a_0 + \sum_{i=1}^n a_i u_i + \sum_{i=1}^n b_i (\mu_i^2 + \sigma_i^2)] \quad (19)$$

式(18) 减去式(19) 即得 $f_i(x_i)$ 与输出响应量 Y 的协方差 $Cov(f_i, Y)$ 。

由计算式 $V(f_k) = E(f_k^2) - E(f_k)^2$, 先推导被减数项:

$$\begin{aligned} E(f_k^2) &= a_0^2 + b_k^2 (3\sigma_k^4 + 6\mu_k^2 \sigma_k^2 + \mu_k^4) + \\ &a_k^2 (\mu_k^2 + \sigma_k^2) + \sum_{i=1, i \neq k}^n \sum_{j=1, j \neq k}^n b_i b_j (\mu_i^2 + \sigma_i^2) (\mu_j^2 + \sigma_j^2) + \\ &2 \sum_{i=1, i \neq k}^n b_i b_k (\mu_i^2 + \sigma_i^2) (\mu_k^2 + \sigma_k^2) + \\ &2a_0 (\sum_{i=1}^n a_i u_i + \sum_{i=1}^n b_i (\mu_i^2 + \sigma_i^2)) + 2 \sum_{i=1, i \neq k}^n a_i u_i (a_k u_k + \\ &\sum_{i=1}^n b_i (\mu_i^2 + \sigma_i^2)) + \sum_{i=1, i \neq k}^n \sum_{j=1, j \neq k}^n a_i a_j u_i u_j + \\ &2a_k [\sum_{i=1}^n b_i (\mu_i^2 + \sigma_i^2) u_k + 2b_k \mu_k \sigma_k^2] \end{aligned} \quad (20)$$

而组成分量函数 $f_k(x_k)$ 期望的平方为:

$$\begin{aligned} E(f_k)^2 &= a_0^2 + \sum_{i=1}^n \sum_{j=1}^n a_i a_j u_i u_j + \\ &\sum_{i=1}^n \sum_{j=1}^n b_i b_j (\mu_i^2 + \sigma_i^2) (\mu_j^2 + \sigma_j^2) + \\ &2 \sum_{i=1}^n \sum_{j=1}^n a_i b_j u_i (\mu_j^2 + \sigma_j^2) + \\ &2a_0 (\sum_{i=1}^n a_i u_i + \sum_{i=1}^n b_i (\mu_i^2 + \sigma_i^2)) \end{aligned} \quad (21)$$

式(20) 减去式(21) 即得 x_i 的结构方差贡献 $V(f_k)$ 。

3 两种理论解的比较

Li 和 Xu 都是基于方差模型提出的重要性测度指标, 二者研究对象相同, 因此本文对这两种指标的理论解的关系加以研究。文献[13] 中给出了 Xu 指标的理论解。由两类指标的定义可知, 两类解析解的定义基础不同, Xu 的定义是将输出响应量的总方差进行模型分解后讨论各阶方差对主方差的贡献, 而 Li 的定义是将输出响应量的功能函数进行模型分解后讨论各组成分量函数对总的功能函数的贡献, 因而二者对应的指标内涵会有所不同。

Xu 指标体系下 x_k 的一阶方差贡献 V_k :

$$\begin{aligned} V_k &= \sum_{i=1}^n [a_i^2 \rho_{ik}^2 \sigma_i^2 + b_i^2 (4\mu_i^2 \rho_{ik}^2 \sigma_i^2 + 2\rho_{ik}^4 \sigma_i^4)] + \\ &\sum_{i=1}^n \sum_{j=1}^n 4a_i b_j \rho_{ik} \rho_{jk} \mu_j \sigma_i \sigma_j + \\ &\sum_{i=1}^n \sum_{j=i+1}^n 2[a_i a_j \sigma_i \sigma_j \rho_{ik} \rho_{jk} + \\ &b_i b_j (4\rho_{ik} \rho_{jk} \mu_i \mu_j \sigma_i \sigma_j + 2\rho_{ik}^2 \rho_{jk}^2 \sigma_i^2 \sigma_j^2)] \end{aligned} \quad (22)$$

$V_k = V[E(Y| x_k)] = V(Y) - E[V(Y| x_k)]$ 是指固定变量 x_k 时输出响应总方差 V_y 的减少量, 该部分贡献中不仅包含 x_k 自身的影响, 还包括所有与 x_k 相关的其他变量的影响。式(17) 和式(14) 中的各加和因子的意义如表 1 所示。

表 1 两类指标 V_k 因子的含义

Tab. 1 Meaning of components of V_k in two indices

式(22) 加和因子	含 义
$\sum_{i=1}^n [a_i^2 \rho_{ik}^2 \sigma_i^2 + b_i^2 (4\mu_i^2 \rho_{ik}^2 \sigma_i^2 + 2\rho_{ik}^4 \sigma_i^4)]$	输出模型函数中与变量 x_k 相关(包括含有 x_k 的项)的各项方差贡献之和
$\sum_{i=1}^n \sum_{j=1}^n 4a_i b_j \rho_{ik} \rho_{jk} \mu_j \sigma_i \sigma_j$	输出函数中与变量 x_k 相关的一次项和二次项之间的协方差之和
$\sum_{i=1}^n \sum_{j=i+1}^n 2[a_i a_j \rho_{ik} \rho_{jk} \sigma_i \sigma_j + b_i b_j (4\rho_{ik} \rho_{jk} \mu_j \sigma_i \sigma_j + 2\rho_{ik}^2 \rho_{jk}^2 \sigma_i^2 \sigma_j^2)]$	输出函数中与变量 x_k 相关的不同一次项之间、不同二次项之间的协方差之和
式(14) 加和因子	含 义
$\sum_{i=1}^n a_i a_k \rho_{ik} \sigma_i \sigma_k$	输出函数中含有变量 x_k 的一次项和与其相关的一次项(包括含有 x_k 的一次项)的协方差之和
$\sum_{i=1}^n 2a_k b_i \rho_{ik} u_i \sigma_i \sigma_k$	输出函数中含有变量 x_k 的一次项和与其相关的二次项的协方差之和
$\sum_{i=1}^n 2a_i b_k \rho_{ik} u_i \sigma_i \sigma_k$	输出函数中含有变量 x_k 的二次项和与其相关的一次项的协方差之和
$\sum_{i=1}^n b_i b_k (4\rho_{ik} \mu_i \mu_k \sigma_i \sigma_k + 2\rho_{ik}^2 \sigma_i^2 \sigma_k^2)$	输出函数中含有变量 x_k 的二次项和与其相关的二次项(包括含有 x_k 的二次项)的协方差之和

由表 1 可知, V_k 中包含的信息是较多的, 将功能函数中任一包含变量 x_k 或与其相关的项均囊括在内, 而 $Cov(f_k, Y)$ 中两项之间的协方差中至少有一项含有 x_k 。因而 Xu 的指标 V_k 较 Li 的指标 $Cov(f_k, Y)$ 多出的信息正是功能函数中不含有 x_k 且与其相关的各项之间的协方差。若 x_k 不与其他变量相关时, 前者较后者的多余部分就不存在, 两类指标就是相同的。

Xu 指标体系下独立部分方差贡献 V_k^U :

$$V_k^U = (a_k^2 + 4a_k b_k \mu_k + 4b_k^2 \mu_k^2) (\sigma_k^2 - B_{k(-k)} B_{(-k)(-k)}^{-1} B_{k(-k)}^T) + 2b_k^2 (\sigma_k^4 - B_{k(-k)} B_{(-k)(-k)}^{-1} B_{k(-k)}^T B_{(-k)(-k)}^{-1} B_{k(-k)}^T) \quad (23)$$

其中 $B_{k(-k)} = (\sigma_{k1} \sigma_{k2} \cdots \sigma_{k(k-1)} \sigma_{k(k+1)} \cdots \sigma_{kn})$, $\sigma_{ki} = \rho_{ki} \sigma_k \sigma_i$, $B_{(-k)(-k)}$ 是除去变量 x_k 以外的其他所有变量的协方差阵。 $B_{k(-k)} B_{(-k)(-k)}^{-1} B_{k(-k)}^T = C_{kkn} \sigma_k^2$, 其中 C_{kkn} 由变量之间的相关系数决定。

$V_k^U = V(Y) - V[E(Y|x_k)] = E[V(Y|x_k)]$ 表示固定除 x_k 外其他所有变量时方差的平均变化量, 由于变量之间的相关性, x_k 与其他变量相关的部分也被固定, 剩余可以发生变化的即 Xu 定义的独立部分。 V_k^U 表达式中加和因子 $(a_k^2 + 4a_k b_k \mu_k + 4b_k^2 \mu_k^2) \sigma_k^2 + 2b_k^2 \sigma_k^4$ 是功能函数中含有 x_k 的一次项和二次项自身的方差, V_k^U 中的剩余项则由与 x_k 相关的相关系数决定。相应的 $V(f_k)$ 表达式包含的是功能函数中含有 x_k 的一次项和二次项自身的方差。

对比分析可知, Xu 的指标 V_k^U 较 Li 的指标 $V(f_k)$ 减去了 x_k 与其他变量相关部分的方差贡献。若变量 x_k 不与其他变量相关时, 前者较后者减去的部分就不存在, 两类指标就是相同的。两类指标的相关方差贡献是上述两个方差项之差, 故二者对照在此就不予赘述。

变量独立时, Xu 的指标与 Li 指标的关系:

$$V_k = Cov(f_k, Y) = a_k^2 \sigma_k^2 + 4a_k b_k \mu_k \sigma_k^2 + b_k^2 (4\mu_k^2 \sigma_k^2 + 2\sigma_k^4) \\ V_k^U = V(f_k) = 2b_k^2 \sigma_k^4 + (a_k^2 + 4a_k b_k \mu_k + 4b_k^2 \mu_k^2) \sigma_k^2 \quad (24)$$

当输入变量相互独立时, Xu 指标和 Li 指标是完全相同的。这也与所有变量均独立时变量对输出响应量的方差贡献理论体系相统一。此外, 在输出响应函数为二次不含交叉项时, 不论是否考虑输入变量相关, Li 指标的结构方差贡献 $V(f_k)$ 的表达式不变, 它反映了分量函数 f_k 对总方差的结构贡献。

4 算例

算例 1: 响应函数为 $y = x_1 + x_2 + x_3 + x_1^2 + x_2^2 + 2x_3^2$, $x_i \sim N(0, 1)$, $\rho_{12} = 0.3$, 其余变量间两两独立, 计算结果如图 1 所示(其中数值方法为大样本 Monte Carlo 数字模拟法^[14])。

由图 1 可发现对于二次不含交叉项的响应量函数, 解析解和数值解非常接近, 证明了所推导解

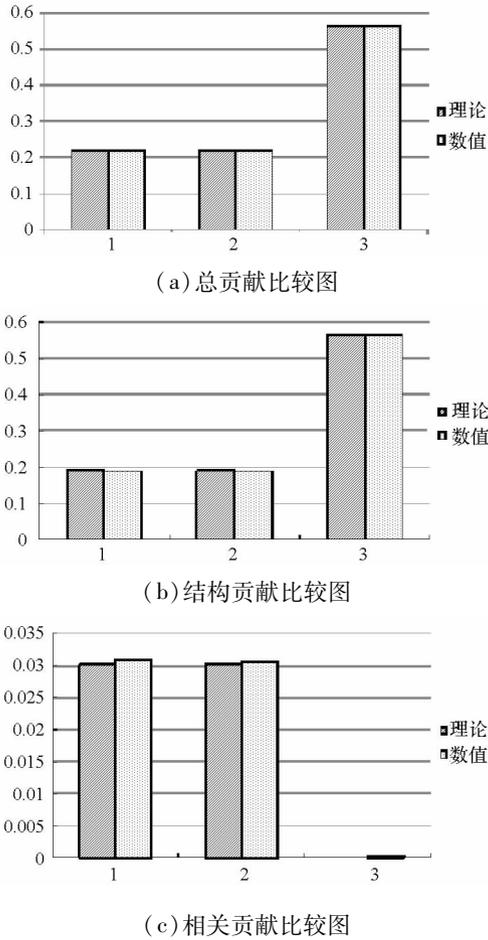


图 1 理论值与数值结果对照图

Fig. 1 Comparison of theoretical and numerical results

析解的正确性。由图 1 中(b)、(c)两图,发现 x_1 和 x_2 的结构贡献和相关贡献均相同,结构贡献相同是由于这两个变量的分布参数相同,并且二者在输出响应表达式中所占的位置也是相同的,相关贡献相同是因为二者只互相相关,与其他变量无关。对于 x_3 ,由于独立,因而相关贡献为 0,相应地,它的总贡献全部由结构贡献提供。

以算例 1 来验证前文得到的两类指标之间的关系,求得 Xu 指标结果如表 2 所示。表中 S_i 是变量 x_i 的一阶方差总贡献, S_i^U 是 x_i 的独立部分贡献, S_i^C 是 x_i 的相关部分贡献,各指标计算公式为: $S_i = V_i/V(\gamma)$, $S_i^U = V_i^U/V(\gamma)$, $S_i^C = S_i - S_i^U$ 。

表 2 算例 1 的 Xu 指标计算结果

Tab.2 Results of example 1 in Xu's indices

变量	S_i	S_i^U	S_i^C
x_1	0.2548	0.1813	0.0735
x_2	0.2548	0.1813	0.0735
x_3	0.5639	0.5639	0

对比表 2 与图 1 中各图,发现对于独立变量 x_3 ,两种定义下的各项指标对应相同。对于相关

变量 x_1, x_2 ,由于相关系数为正数,因而有 Xu 指标下的 S_i 较 Li 指标下的 S_i 偏大, Xu 指标下的 S_i^U 较 Li 指标下的 S_i^a 偏小,验证了文章第 3 部分给出的两类指标关系的正确性。

算例 2:图 2 给出了九盒段模拟结构的几何形状和受载方式。九盒段根部的前后缘 4 个节点为固定端,内部 4 个节点简化为只有一个自由度的支座。翼盒上表面的前后缘 6 个节点受载为 P ,中间部分的 6 个节点的受载为 $2P$ 。所有杆的长度均为 L 的整数倍,假设所有杆面积 A 都相同, E 表示杆元和板元的弹性模量,泊松比取 0.3, H 表示板(即蒙皮)厚度。各变量分布如表 3 所示(为表示方便,将 L, A, E, P, H 分别表示为 x_1, x_2, x_3, x_4, x_5)。相关系数 $\rho_{12} = 0.2, \rho_{13} = 0.1, \rho_{23} = 0.05$,其余变量之间两两独立。

表 3 九盒段结构中随机变量的分布参数

Tab.3 Distribution parameters of variables of nine-box

变量	$L(\text{m})$	$A(\text{cm}^2)$	$E(\text{GP})$	$P(\text{kN})$	$H(\text{mm})$
μ	0.2	1	71	10	3
σ	0.01	0.05	3.55	0.5	0.15

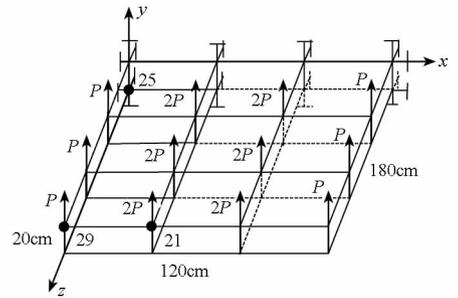


图 2 九盒段结构示意图

Fig. 2 Scheme of nine-box structure

对九盒段有限元分析后得极限状态方程:

$$g(x_1, x_2, x_3, x_4, x_5) = 510 - \sigma_{\max}$$

将输出响应量 g 与输入变量进行非线性回归,得到响应量的功能函数为:

$$g = 30.4x_1 - 45.5x_1^2 - 0.39x_2 + 0.15x_2^2 + 0.506x_3 - 0.003x_3^2 + 5.18x_4 - 0.035x_4^2 - 110.9x_5 + 13.4x_5^2 + 371$$

对上式进行求解,结果如图 3、4 所示。

由图 3 可知 x_2 (杆的面积 A)的总贡献为 0,但从图 4 和图 5 来看它的结构和相关贡献却是不为 0 的, S_2^a 和 S_2^b 互为相反数才构成其总贡献为 0 的结果,故仅从总贡献为 0 不能判定变量对于输出响应没有影响。 x_3 (弹性模量 E)的相关贡献最大而其结构贡献是最小的。 x_4 (载荷 P)和 x_5 (蒙皮厚度 H)的总贡献较大,二者的结构贡献也较

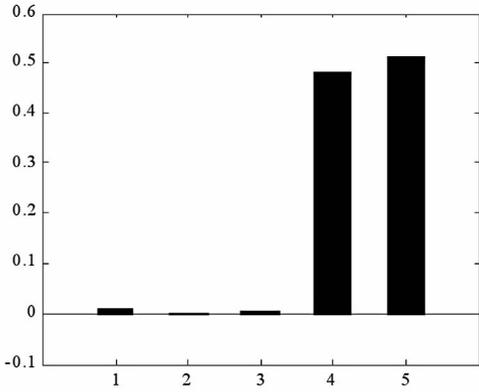


图3 变量总贡献图

Fig.3 Total contributions of variables

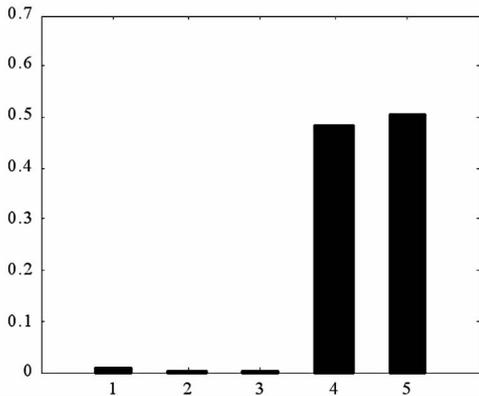


图4 变量结构贡献图

Fig.4 Structural contributions of variables

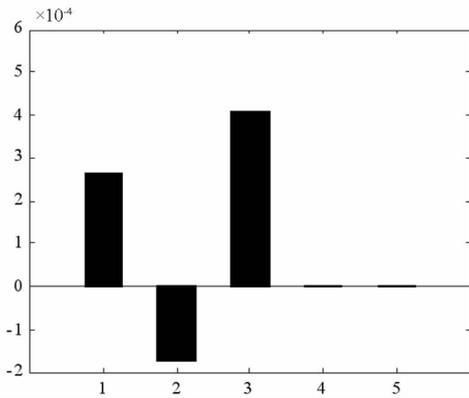


图5 变量相关贡献图

Fig.5 Correlative contributions of variables

大,相关贡献均为0,这是由于二者均为独立的变量。若要减小该极限状态函数的变异性,可以从减小 x_4 和 x_5 的结构作用或增大 x_2 的相关作用入手。

算例3:某飞机襟翼的有限元模型如图6,将襟翼所受的空气动力外载转化为作用在单元节点的集中载荷 F 。极限状态方程为:

$$g(x) = f(t_1, t_2, t_3, E_1, G_1, E_2, G_2, F)$$

其中 t_1, t_2, t_3 是襟翼3根梁的厚度, E_1 (E_2) 和 G_1 (G_2) 分别是弹性模量和剪切模量, F 代表施加

在节点上的外载,对于节点 i 处的外载 F_i ,有 $F_i = (1 + F) F_{i0}$, F_{i0} 是常值, F 是无量纲,所有变量正态分布,相关系数 $\rho_{t_1 t_2} = 0.05$, $\rho_{t_1 t_3} = 0.05$, $\rho_{t_2 E_2} = 0.1$,其余参数之间相关系数为0,变量的分布信息如表4所示。

表4 襟翼变量分布参数

Tab.4 Distribution parameters of variables of flap

变量	均值	标准差
t_1/cm	2	0.04
t_2/cm	2	0.04
t_3/cm	4	0.08
E_1/Mpa	70 380	1407.6
G_1/Mpa	26 458.6	529.172
E_2/Mpa	72 450	1449
G_2/Mpa	27 236.8	544.736
F	0	0.2

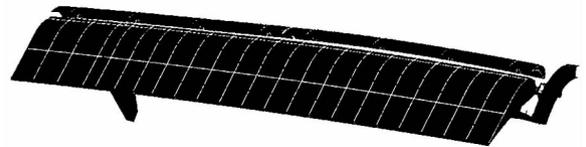


图6 某飞机襟翼图

Fig.6 A picture of the flap

运用非线性回归对输出响应和输入变量进行拟合,得到的拟合函数为:

$$y = 0.0003625E_1 + 0.001068G_1 + 0.000851G_2 + 28.6395F - 0.09316t_1^2 - 0.10742t_2^2 - 0.002722t_3^2 + 0.000352F^2$$

对上式进行 Li 各项指标的求解,由于拟合函数中不含 E_2 的项,说明 E_2 不影响输出响应,故在求解 Li 各项指标时就不予考虑 E_2 的影响,结果如表5所示。

表5 飞机襟翼指标计算结果

Tab.5 Computation results of the flap

	$t_1(10^{-6})$	$t_2(10^{-6})$	$t_3(10^{-6})$	E_1	G_1	G_2	F
S	7.032	9.174	0.129	0.0077	0.0095	0.0064	0.9763
S^a	6.613	8.792	0.090	0.0077	0.0095	0.0064	0.9763
S^b	0.420	0.381	0.039	0	0	0	0

由表5知:对于结构贡献和相关贡献指标,独立变量 E_1, G_1, G_2 和 F 的相关贡献为0,相关变量 t_1, t_2, t_3 的相关贡献数值很小,说明变量之间的相关作用不明显。对于总贡献指标,飞机襟翼变量的总贡献指标按其相对大小可分为3类:数量级

很小的襟翼 3 根梁的厚度 t_1, t_2, t_3 ; 相对大小居中的弹性模量和剪切模量 E_1, G_1, G_2 以及数值最大的外载 F 。变量的总贡献大小排序可表明变量对输出响应变异性影响的重要性程度大小, 显然外载 F 对于输出响应的稳定性影响是最大的, E_1, G_1, G_2 次之, t_1, t_2, t_3 的影响最小。由于 F 的总贡献指标接近所有变量指标之和 1, 故增加 F 的稳定性对提高输出响应的稳定性十分有利。

5 结论

本文以二阶不含交叉项的多项式为例, 推导了正态相关情况下 Li 提出的基于方差重要性测度指标的解析解, 为求解该类情况下相关输入变量的重要性测度指标提供了有利的参考。文章对两类基于方差的重要性测度指标进行了简单地类比, 清楚地揭示了两类指标在描述相关变量重要性时的侧重与统一, 为工程实际选用合适的指标体系提供参考。

参考文献 (References)

- [1] Doksum K, Samarov A. Non-parametric estimation of global functionals and a measure of the explanatory power of covariates in regression [J]. *Annals of Statistics*, 1995, 23 (5): 1443 - 1473.
- [2] Satelli A. The role of sensitivity analysis in the corroboration of models and its link to model structural and parametric uncertainty [J]. *Reliability Engineering and System Safety*, 1997, 57(1): 1 - 4.
- [3] Boronovo E. A new uncertainty importance measure [J]. *Reliability Engineering & System Safety*, 2007, 92 (6): 771 - 784.
- [4] Cui L J, Lu Z Z, Zhao X P. Moment-independent importance measure of basic random variable and its probability density evolution solution [J]. *Science China Technological Sciences*, 2010, 53(4): 1138 - 1145.
- [5] Sobol I M. Global sensitivity indices for non-linear mathematical models and their Monte Carlo estimate [J]. *Mathematics and Computers in Simulation*, 2001, 55(1-3): 271 - 280.
- [6] Iman R L, Hora S C. A robust measure of uncertainty importance for use in fault tree system analysis [J]. *Risk Analysis*, 1990, 10(3): 401 - 406.
- [7] Satelli A. Sensitivity analysis for importance assessment [J]. *Risk Analysis*, 2002, 22(3): 579 - 590.
- [8] Xu C G, George Z G. Uncertainty and sensitivity analysis for models with correlated parameters [J]. *Reliability Engineering and System Safety*, 2008, 93(6): 1563 - 1573.
- [9] Mara T A, Tarantolab S. Variance-based sensitivity indices for models with dependent inputs [J]. *Reliability Engineering & System Safety*, 2012, 107: 115 - 121.
- [10] Li G Y, Rabitz H. Global sensitivity analysis for systems with independent and/or correlated inputs [J]. *Journal of Physical Chemistry A*, 2010, 114: 6022 - 6032.
- [11] Hao W R, Lu Z Z, Wei P F, et al. A new method on ANN for variance based importance measure analysis of correlated input variables [J]. *Structural Safety*, 2012, 38: 56 - 63.
- [12] Li G Y, Rabitz H. General formulation of HDMR component functions with independent and correlated variables [J]. *Journal of Mathematical Chemistry*, 2010, 50: 99 - 130.
- [13] 郝文锐, 吕震宙, 魏鹏飞. 多项式输出中相关变量的重要性测度分析 [J]. *力学学报*, 2012, 44(1): 167 - 173.
HAO Wenrui, LYU Zhenzhou, WEI Pengfei. Importance measure of correlated variables in polynomial output [J]. *Chinese Journal of Theoretical and Applied Mechanics*, 2012, 44(1): 167 - 173. (in Chinese)
- [14] Binder K. Applications of the Monte Carlo method in statistical physics [M]. Berlin: Springer-Verlag, 1984.