

# 输入变量相关情况下重要性测度指标关系<sup>1)</sup>

宋静文 吕震宙<sup>2)</sup>

(西北工业大学航空学院, 西安 710072)

**摘要** 为了根据不同需求找到输入变量相关情况下最优重要性测度分析方法, 非常有必要对现有重要性测度方法之间的关系进行探究比较. 分别以不含交叉项和包含交叉项的二次多项式输出模型为例, 解析推导了相关正态输入变量情况下基于协方差分解的重要性测度指标, 包括总贡献、结构贡献和相关贡献. 进一步以所推导的解析结论为基础, 理论推导了传统基于方差的重要性测度指标与基于协方差分解的重要性测度指标之间的关系. 并从特殊二次多项式模型的研究结果对一般模型做出推断, 然后从高维模型分解的角度验证所推断的结论, 并详细阐述了不同重要性测度指标的优缺点. 最后结合具体算例深入地分析了各指标之间的关系, 为输入变量相关情况下结构性能对输入特性的重要性分析与工程设计提供指导.

**关键词** 重要性测度, 相关变量, 方差分解, 协方差分解, 二次多项式

中图分类号: V211, TB114.3 文献标识码: A doi: 10.6052/0459-1879-13-369

## 引言

重要性测度分析又称为全局灵敏度分析, 是结构系统不确定分析问题中的一个重要方面, 它旨在研究结构系统输出性能不确定性向输入变量的分配问题, 进而识别影响结构系统输出性能的主、次输入变量, 为可靠性设计和模型简化提供丰富信息<sup>[1]</sup>. 目前应用较多的是基于方差的重要性测度分析方法, Sobol 和 Saltelli 等<sup>[2-3]</sup> 提出并完善了输入变量独立情况下基于方差的重要性测度体系及求解方法. 在实际工程问题中, 很多情况下输入变量是相关的, 由于相关性的影响, 这一重要性测度体系的内含将发生较大变化<sup>[4-5]</sup>, 而目前对于相关变量条件下基于方差的重要性测度研究还十分有限.

Xu 等<sup>[5]</sup> 指出, 对于含相关输入变量的非线性程度不高的模型 (通常针对可加模型), 可以将某相关输入变量对输出响应性能的方差贡献分解为两部分: 独立贡献和相关贡献. 文献 [6] 针对不含交叉项二次多项式模型推导了某一变量的主贡献、独立贡献及相关贡献解析解. 这种分解方法适用于变量之间不含交互作用或者交互作用较小以致可以忽略的情况. Hao 等<sup>[7]</sup> 进一步给出了考虑交互作用时相关输入变量方差贡献的新解释, 并解析推导了含交叉项二次

多项式模型的相关总贡献和自身独立贡献的解析解. Li 等<sup>[8]</sup> 通过对高维替代模型的协方差分解, 提出了一种对独立和相关输入变量均适用的重要性测度指标体系. 对于单个输入变量或者输入变量的子集, 定义了总贡献、结构贡献和相关贡献 3 个新的重要性测度指标.

现有文献中对 Li 等<sup>[8]</sup> 所提出的重要性测度指标求解方法的研究还十分有限, 同时也未将相关输入变量情况下传统的基于方差的重要性测度指标与基于协方差分解的重要性测度指标进行分析比较. 因此本文先针对相关正态输入变量情况下不含交叉项和包含交叉项的二次多项式模型, 研究 Li 等提出的重要性测度体系的解析求解方法. 又进一步以这些解析解为基础, 参考文献 [6] 和 [7] 中的传统重要性测度指标解析解, 探究传统基于方差的重要性测度指标与 Li 等提出的基于协方差的重要性测度指标之间的关系. 最后从高维模型分解的角度对各重要性测度指标的关系进行一般性验证, 旨在得出有工程指导意义的结论.

为便于行文, 在本论文中将几个概念的简称进行约定:

(1) Li 指标: 对独立或相关输入变量下的一般性

2013-10-31 收到第 1 稿, 2014-01-09 收到修改稿.

1) 国家自然科学基金 (51175425)、博士学科点专项科研基金 (20116102110003) 和 2013 年西北工业大学本科毕业设计 (论文) 重点扶持资助项目.

2) 吕震宙, 教授, 研究方向: 可靠性工程, 灵敏度分析, 多学科优化. E-mail: zhenzhoulu@nwpu.edu.cn

模型, 将单个输入变量或者输入变量的子集  $\mathbf{x}_{P_j}$  ( $P_j$  表示集合  $\{1, 2, \dots, n\}$  的一个非空子集) 对输出响应性能的方差总贡献  $V_{P_j}^L$  分解为结构贡献 (反映模型的结构特征的贡献, 用  $V_{P_j}^{La}$  表示) 和相关贡献 (反映输入变量之间概率分布的相关性贡献, 用  $V_{P_j}^{Lb}$  表示);

(2) Xu 指标: 对相关输入变量下非线性程度不高的模型, 将某相关输入变量  $x_k$  对输出响应性能的方差总贡献  $V_k^X$  分解为独立贡献 (仅由  $x_k$  独立部分产生, 用  $V_k^{XU}$  表示) 和相关贡献 (由  $x_k$  和其他输入变量的相关部分产生, 用  $V_k^{XC}$  表示);

(3) Hao 指标: 对相关输入变量下包含交互作用的模型, 将某相关输入变量  $x_k$  对输出响应的方差贡献解释为相关总贡献  $V_k^{HT}$ , 并将其分解为自身独立贡献  $V_k^{Hu}$  以及  $x_k$  与其他变量的相关性产生的方差贡献  $V_k^{Hc}$ .

### 1 二次多项式模型重要性测度解析分析

对一般形式的响应函数  $y = g(\mathbf{x})$  (其中黑体  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  表示  $n$  维输入矢量) 进行高维模型展开 [2,8-9]

$$g(\mathbf{x}) = g_0 + \sum_{i=1}^n g_i(x_i) + \sum_{i=1}^n \sum_{j=i+1}^n g_{ij}(x_i, x_j) + \dots + g_{1,2,\dots,n}(x_1, x_2, \dots, x_n) = g_0 + \sum_{j=1}^{2^n-1} g_{P_j}(\mathbf{x}_{P_j}) \quad (1)$$

式中下标  $P_j$  表示  $\{1, 2, \dots, n\}$  的真子集,  $P_j$  中所有元素代表的输入变量集合为  $\mathbf{x}_{P_j}$ . 分量函数  $g_{P_j}(\mathbf{x}_{P_j})$  表达式为

$$\left. \begin{aligned} g_0 &= \int g(\mathbf{x}) f_X(\mathbf{x}) d\mathbf{v}(\mathbf{x}) \\ g_i(x_i) &= \int g(\mathbf{x}) f_{X_{-i}}(\mathbf{x}_{-i}) d\mathbf{v}(\mathbf{x}_{-i}) - g_0 \\ g_{ij}(x_i, x_j) &= \int g(\mathbf{x}) f_{X_{-ij}}(\mathbf{x}_{-ij}) d\mathbf{v}(\mathbf{x}_{-ij}) - g_i(x_i) - g_j(x_j) - g_0 \\ &\dots \end{aligned} \right\} \quad (2)$$

其中  $\mathbf{x}_{-i}$  和  $\mathbf{x}_{-ij}$  分别表示所有输入变量中除去  $x_i$  和除去  $(x_i, x_j)$  以外的所有剩余输入变量。  $d\mathbf{v}(\mathbf{x})$  为广义体积微分。  $f_X(\mathbf{x})$  为所有输入向量的联合概率密度函数。  $f_{X_{-i}}(\mathbf{x}_{-i})$  和  $f_{X_{-ij}}(\mathbf{x}_{-ij})$  分别表示  $\mathbf{x}_{-i}$  和  $\mathbf{x}_{-ij}$  的边缘概率密度函数, 它们可由  $f_X(\mathbf{x})$  分别对  $x_i$  及  $x_i$  和  $x_j$

积分得到. 对式 (1) 求方差得到协方差分解式如下 [8]

$$V(y) = Cov(y, y) = Cov\left(g_0 + \sum_{j=1}^{2^n-1} g_{P_j}(\mathbf{x}_{P_j}), y\right) = \sum_{j=1}^{2^n-1} \left[ V(g_{P_j}(\mathbf{x}_{P_j})) + Cov\left(g_{P_j}(\mathbf{x}_{P_j}), \sum_{k=1, k \neq j}^{2^n-1} g_{P_k}(\mathbf{x}_{P_k})\right) \right] \quad (3)$$

其中,  $V(\cdot)$  表示方差,  $Cov(\cdot, \cdot)$  表示协方差. 那么 Li 等提出的重要性测度指标总贡献  $S_{P_j}^L$ 、结构贡献  $S_{P_j}^{La}$  和相关贡献  $S_{P_j}^{Lb}$  分别定义为

$$\left. \begin{aligned} S_{P_j}^L &= V_{P_j}^L / V(y) = Cov(g_{P_j}(\mathbf{x}_{P_j}), y) / V(y) \\ S_{P_j}^{La} &= V_{P_j}^{La} / V(y) = V(g_{P_j}(\mathbf{x}_{P_j})) / V(y) \\ S_{P_j}^{Lb} &= V_{P_j}^{Lb} / V(y) = S_{P_j}^L - S_{P_j}^{La} \end{aligned} \right\} \quad (4)$$

由于  $P_j$  表示的是  $\{1, 2, \dots, n\}$  的子集, 因此 Li 指标能够反映单个或者多个变量共同作用下对性能响应量方差的影响. 在结构系统不确定性分析中, 常通过替代模型近似隐式响应函数, 而由替代模型得到的显式函数的形式常采用二次多项式 [10]. 因此这里分别针对二次不含交叉项多项式和二次含交叉项多项式输出模型, 推导相关正态输入变量对响应量方差贡献的 Li 指标解析解.

#### 1.1 不含交叉项二次多项式输出模型 Li 指标解析解

对于含  $n$  维相关正态输入变量的不含交叉项二次多项式输出模型

$$y = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n b_i x_i^2 \quad (5)$$

设输入变量服从正态分布, 即  $x_i \sim N(\mu_i, \sigma_i^2)$ , 且  $x_i$  与  $x_j$  间的 Pearson 相关系数为  $\rho_{ij}$ . 由于不含交叉项的二次多项式输出模型中没有变量的交互作用, 因此, 其高维模型展开只有一阶项, 且按照式 (2) 求解分量函数如下

$$g_0 = \int g(\mathbf{x}) f_X(\mathbf{x}) d\mathbf{v}(\mathbf{x}) = a_0 + \sum_{i=1}^n [a_i \mu_i + b_i (\mu_i^2 + \sigma_i^2)] \quad (6)$$

$$g_k(x_k) = \int g(\mathbf{x}) f_{X_{-k}}(\mathbf{x}_{-k}) d\mathbf{v}(\mathbf{x}_{-k}) - g_0 = a_k (x_k - \mu_k) + b_k (x_k^2 - \mu_k^2 - \sigma_k^2) \quad (7)$$

其中  $k = 1, 2, \dots, n$ . 将  $g_k(x_k)$  代入 Li 指标定义式 (4), 分别可求得相关输入变量  $x_k$  的总贡献  $V_k^L$ 、结构贡献  $V_k^{La}$  和相关贡献  $V_k^{Lb}$  如下

$$V_k^L = Cov(g_k(x_k), y) = \sum_{i=1}^n [(a_i + 2b_i\mu_i)(a_k + 2b_k\mu_k) + 2\rho_{ik}\sigma_i\sigma_k b_i b_k] \rho_{ik}\sigma_i\sigma_k \quad (8)$$

$$V_k^{La} = V(g_k(x_k)) = \sigma_k^2 [(a_k + 2b_k\mu_k)^2 + 2b_k^2\sigma_k^2] \quad (9)$$

$$V_k^{Lb} = Cov\left(g_k(x_k), \sum_{i=1, i \neq k}^n g_i(x_i)\right) = \sum_{\substack{i=1 \\ i \neq k}}^n [(a_i + 2b_i\mu_i)(a_k + 2b_k\mu_k) + 2\rho_{ik}\sigma_i\sigma_k b_i b_k] \rho_{ik}\sigma_i\sigma_k \quad (10)$$

其中  $\rho_{ik}$  在  $i = k$  时的值为 1. 从式 (3) 看出, 当高维模型展开仅包含一阶项时, 输出响应量的总方差为  $\sum_{k=1}^n Cov(g_k(x_k), y)$ , 经过验证其与直接求得的输出总方差<sup>[6]</sup>是相同的, 即

$$V(y) = \sum_{k=1}^n Cov(g_k, y) = \sum_{k=1}^n \sum_{i=1}^n [(a_i + 2b_i\mu_i)(a_k + 2b_k\mu_k) + 2\rho_{ik}\sigma_i\sigma_k b_i b_k] \rho_{ik}\sigma_i\sigma_k \quad (11)$$

比较式 (8)~(10) 可以看出, 输入变量  $x_k$  的结构贡献  $V_k^{La}$  为总贡献求和式 (8) 中  $i = k$  这一项, 而相关贡献  $V_k^{Lb}$  是总贡献中  $i \neq k$  的部分. 若输入变量相互独立, 式 (10) 中的各求和项均有  $\rho_{ik} = 0$ , 则  $V_k^{Lb} = 0$ , 从而  $V_k^L = V_k^{La}$ , 那么 Li 指标就退化为一个总贡献指标  $V_k^L$ .

### 1.2 含交叉项的二次多项式输出模型 Li 指标解析解

对于相关正态输入变量情况下含交叉项二次多项式输出模型

$$y = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n b_i x_i^2 + \sum_{i=1}^n \sum_{j=i+1}^n c_{ij} x_i x_j \quad (12)$$

由于该多项式输出模型仅含二阶交互作用, 因而高维替代模型的三阶及以上的分量函数均为 0, 只需求解零阶、一阶和二阶分量函数即可. 式 (12) 的高维替

代模型形式为

$$g(\mathbf{x}) = g_0 + \sum_{k=1}^n g_k(x_k) + \sum_{k_1=1}^n \sum_{k_2=k_1+1}^n g_{k_1 k_2}(x_{k_1}, x_{k_2}) \quad (13)$$

按照式 (2) 可推导得到分量函数表达式如下

$$g_0 = \int g(\mathbf{x}) f_X(\mathbf{x}) d\mathbf{v}(\mathbf{x}) = a_0 + \sum_{i=1}^n [a_i \mu_i + b_i (\mu_i^2 + \sigma_i^2)] + \sum_{i=1}^n \sum_{j=i+1}^n c_{ij} \mu_i \mu_j \quad (14)$$

$$g_k(x_k) = \int g(\mathbf{x}) f_{X_{-k}}(\mathbf{x}_{-k}) d\mathbf{v}(\mathbf{x}_{-k}) - g_0 = a_k x_k + b_k x_k^2 + \sum_{\substack{i=1 \\ i \neq k}}^n c_{ik} x_k \mu_i - a_k \mu_k - b_k (\mu_k^2 + \sigma_k^2) - \sum_{\substack{i=1 \\ i \neq k}}^n c_{ik} \mu_i \mu_k \quad (15)$$

$$g_{k_1 k_2}(x_{k_1}, x_{k_2}) = \int g(\mathbf{x}) f_{X_{-k_1 k_2}}(\mathbf{x}_{-k_1 k_2}) d\mathbf{v}(\mathbf{x}_{-k_1 k_2}) - g_{k_1}(x_{k_1}) - g_{k_2}(x_{k_2}) - g_0 = c_{k_1 k_2} (x_{k_1} - \mu_{k_1})(x_{k_2} - \mu_{k_2}) \quad (16)$$

求得各阶分量函数后, 即可求得相关输入变量  $x_k$  的总贡献  $V_k^L$  和结构贡献  $V_k^{La}$

$$V_k^L = Cov(g_k(x_k), y) = \sum_{i=1}^n [(a_i + 2b_i\mu_i)(a_k + 2b_k\mu_k) + 2\rho_{ik}\sigma_i\sigma_k b_i b_k] \rho_{ik}\sigma_i\sigma_k + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq k}}^n c_{ik} \mu_i \rho_{ik} \sigma_i \sigma_k (a_i + 2b_i\mu_i) + \sum_{i=1}^n \sum_{j=i+1}^n \sum_{\substack{l=1 \\ l \neq k}}^n c_{ij} c_{lk} \mu_l \sigma_k (\mu_j \rho_{ki} \sigma_i + \mu_i \rho_{kj} \sigma_j) + \sum_{i=1}^n \sum_{j=i+1}^n [a_k c_{ij} (\mu_i \rho_{kj} \sigma_k \sigma_j + \mu_j \rho_{ki} \sigma_k \sigma_i) + 2b_k c_{ij} (\mu_k \mu_j \rho_{ki} \sigma_k \sigma_i + \mu_k \mu_i \rho_{kj} \sigma_k \sigma_j + \rho_{kj} \rho_{ki} \sigma_k^2 \sigma_j \sigma_i)] \quad (17)$$

$$V_k^{La} = V(g_k(x_k)) = \sigma_k^2 \left[ \left( a_k + \sum_{\substack{i=1 \\ i \neq k}}^n c_{ik} \mu_i + 2b_k \mu_k \right)^2 + 2b_k^2 \sigma_k^2 \right] \quad (18)$$

两个相关输入变量  $x_{k_1}$  和  $x_{k_2}$  的总贡献  $V_{k_1 k_2}^L$  和结构贡献  $V_{k_1 k_2}^{La}$

$$V_{k_1 k_2}^L = Cov(g_{k_1 k_2}(x_{k_1}, x_{k_2}), y) = \sum_{i=1}^n 2c_{k_1 k_2} b_i \rho_{ik_1} \rho_{ik_2} \sigma_i^2 \sigma_{k_1} \sigma_{k_2} + \sum_{i=1}^n \sum_{j=i+1}^n c_{k_1 k_2} c_{ij} \sigma_i \sigma_j \sigma_{k_1} \sigma_{k_2} (\rho_{ik_1} \rho_{jk_2} + \rho_{jk_1} \rho_{ik_2}) \quad (19)$$

$$V_{k_1 k_2}^{La} = V(g_{k_1 k_2}(x_{k_1}, x_{k_2})) = c_{k_1 k_2}^2 \sigma_{k_1}^2 \sigma_{k_2}^2 (\rho_{k_1 k_2}^2 + 1) \quad (20)$$

与式(11)的求法相同, 输出响应模型(12)的总方差就是所有变量的一阶、二阶总贡献之和, 即

$$V(y) = \sum_{k=1}^n Cov(g_k(x_k), y) + \sum_{k_1=1}^n \sum_{k_2=k_1+1}^n Cov(g_{k_1 k_2}(x_{k_1}, x_{k_2}), y) \quad (21)$$

与不含交叉项二次多项式结论类似, 当变量独立时, 经过验证  $x_k$  的总贡献等于结构贡献  $V_k^{La}$ ,  $x_{k_1}$  和  $x_{k_2}$  的总贡献等于结构贡献  $V_{k_1 k_2}^{La}$ . 因而所推导的解析表达式对独立和相关正态输入变量的情况均是适用的, 既能作为独立正态输入变量主贡献的参照解<sup>[3]</sup>, 又提供了一种新的相关变量重要性测度的高效求解方法.

## 2 输入变量相关情况下重要性测度指标的比较

在相关输入变量对响应量方差影响的研究中, 主要采用的重要性测度指标是 Xu 指标、Hao 指标和 Li 指标, 而 Hao 指标是对 Xu 指标内含的进一步完善. 本文先将 Xu 指标和 Li 指标的方差贡献解析解进行对比, 再将 Hao 指标与 Li 指标进行对比.

### 2.1 二次不含交叉项多项式模型 Li 指标与 Xu 指标的比较

#### (1) Li 指标总贡献与 Xu 指标总贡献的比较

对于式(5)所示的不含交叉项的二次多项式模型, Xu 指标中  $x_k$  的总贡献解析表达式为<sup>[6]</sup>

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

为便于从表达式角度比较其关系, 将 Xu 指标总贡献  $V_k^X = V[E(y|x_k)]$  写成协方差的形式  $V_k^X = Cov(E(y|x_k), E(y|x_k))$ , 其中条件期望  $E(y|x_k)$  的表达式如下

$$E(y|x_k) = \sum_{i=1}^n \left\{ a_i \left[ \mu_i + \rho_{ik} (x_k - \mu_k) \frac{\sigma_i}{\sigma_k} \right] + b_i \left[ \sigma_i^2 (1 - \rho_{ik}^2) + \mu_i^2 + 2\mu_i \rho_{ik} (x_k - \mu_k) \frac{\sigma_i}{\sigma_k} + \rho_{ik}^2 (x_k - \mu_k)^2 \frac{\sigma_i^2}{\sigma_k^2} \right] \right\} \quad (23)$$

记上式求和项中  $i = k$  时的结果为  $E(y|x_k)|_{i=k} = a_k x_k + b_k x_k^2$ . 那么  $V_k^X$  的部分协方差为

$$V_k^X|_{i=k} = Cov(E(y|x_k)|_{i=k}, E(y|x_k)) = \sum_{i=1}^n [(a_i + 2b_i \mu_i)(a_k + 2b_k \mu_k) + 2b_i b_k \rho_{ik} \sigma_i \sigma_k] \rho_{ik} \sigma_i \sigma_k \quad (24)$$

对照式(24)和式(8), 可以看出关系式  $V_k^X|_{i=k} = V_k^L$  成立, 说明对于不含交叉项的二次多项式输出模型, Li 指标总贡献  $V_k^L$  是 Xu 指标总贡献  $V_k^X$  中  $i = k$  项, 体现了某一变量  $x_k$  与所有变量的协方差关系. 而 Xu 指标的总贡献定义式中, 当  $x_k$  被固定, 其他变量中与  $x_k$  相关的部分也被固定, 因此还计入了除  $x_k$  以外的其他所有变量  $x_{-k}$  与  $x_k$  相关的部分对响应方差的贡献.

(2) Li 指标结构贡献与 Xu 指标独立贡献的比较  
由文献[6]可知 Xu 指标独立贡献  $V_k^{XU}$  表达式为

$$V_k^{XU} = \sigma_k^2 (a_k^2 + 4a_k b_k \mu_k + 4b_k^2 \mu_k^2 + 2b_k^2 \sigma_k^2) - \mathbf{B}_{k(-k)} \mathbf{B}_{(-k)(-k)}^{-1} \mathbf{B}_{k(-k)}^T (a_k^2 + 4a_k b_k \mu_k + 4b_k^2 \mu_k^2 - 2b_k^2 \mathbf{B}_{k(-k)} \mathbf{B}_{(-k)(-k)}^{-1} \mathbf{B}_{k(-k)}^T) \quad (25)$$

其中  $\mathbf{B}$  表示协方差矩阵.

比较式(18)和式(25)可知, Xu 指标的独立贡献  $V_k^{XU}$  是  $x_k$  中去除了  $x_k$  与其他变量相关性的贡献, 它是 Li 指标结构贡献  $V_k^{La}$  的一部分, 即有  $V_k^{XU} < V_k^{La}$ . 而 Li 指标结构贡献  $V_k^{La}$  是  $x_k$  在模型结构中总的贡献, 由于  $x_k$  与其他变量相关, 因而必然包含变量  $x_k$  与其他变量相关部分的方差贡献. 此外 Li 指标相关贡献  $V_k^{Lb}$  和 Xu 指标相关贡献  $V_k^{XC}$  可根据公式:  $V_k^{Lb} = V_k^L - V_k^{La}$  和  $V_k^{XC} = V_k^X - V_k^{XU}$  求得. 结合前面已



得出  $V_k^L < V_k^X$  的结论，显然 Li 指标的相关贡献  $V_k^{Lb}$  必然是 Xu 指标相关贡献  $V_k^{Xc}$  的一部分。

分析二次不含交叉项多项式模型的解析解可知，Xu 指标能够更严格的将变量的相关部分和独立部分贡献区分开，而 Li 指标的结构贡献和相关贡献则不能很好地反映相关性关系。同时，Xu 指标总贡献中将每个变量与  $x_k$  的相关部分贡献都计入，那么各变量的总指标会存在重叠部分，从该角度来看，则 Li 指标能够更好的反映变量的结构作用。

### 2.2 二次含交叉项多项式模型 Li 指标与 Hao 指标的比较

对式 (12) 所示的含交叉项二次多项式模型，Hao 等提出了相关变量重要性测度指标的新解释 [7]。本节以前面推导的 Li 指标解析解与文献 [7] 中的解析结论为基础，深入分析两套指标之间的关系，以便揭示两套指标的内含关系，并为输出性能的设计提供更明确的指导信息。

#### (1) Li 指标总贡献与 Hao 指标相关总贡献的比较

输入变量相关情况下输出响应模型 (12) 的条件期望  $E(y|x_k)$  为

$$E(y|x_k) = \sum_{i=1}^n [a_i E(x_i|x_k) + b_i E(x_i^2|x_k)] + \sum_{i=1}^n \sum_{j=i+1}^n c_{ij} E(x_i x_j|x_k) \quad (26)$$

其中  $E(x_i|x_k)$ ,  $E(x_i^2|x_k)$ ,  $E(x_i x_j|x_k)$  可由条件期望的性质推导 [11]，此处不再详述。采用 2.1.1 小节类似的方法，取上式中  $i = k$  项并记为  $E(y|x_k)|_{i=k}$ ，那么

$$E(y|x_k)|_{i=k} = a_k x_k + b_k x_k^2 + \sum_{\substack{l=1 \\ l \neq k}}^n c_{lk} \left[ \rho_{lk} \frac{\sigma_l}{\sigma_k} x_k^2 + \left( \mu_l - \rho_{lk} \frac{\sigma_l}{\sigma_k} \mu_k \right) x_k \right] \quad (27)$$

则 Hao 指标的相关总贡献  $V_k^{HT}$  中的部分协方差为

$$V_k^{HT}|_{i=k} = Cov(E(y|x_k)|_{i=k}, E(y|x_k)) = \sum_{i=1}^n [(a_i + 2b_i \mu_i)(a_k + 2b_k \mu_k) + 2\rho_{ik} \sigma_i \sigma_k b_i b_k] \rho_{ik} \sigma_i \sigma_k + \sum_{i=1}^n \sum_{\substack{l=1 \\ l \neq k}}^n c_{lk} \mu_l \rho_{lk} \sigma_l \sigma_k (a_i + 2b_i \mu_i) + \sum_{i=1}^n \sum_{j=i+1}^n \sum_{\substack{l=1 \\ l \neq k}}^n c_{ij} c_{lk} \mu_l \sigma_k (\mu_j \rho_{ki} \sigma_i + \mu_i \rho_{kj} \sigma_j) +$$

$$\sum_{i=1}^n \sum_{j=i+1}^n [a_k c_{ij} (\mu_i \rho_{kj} \sigma_k \sigma_j + \mu_j \rho_{ki} \sigma_k \sigma_i) + 2b_k c_{ij} (\mu_k \mu_j \rho_{ki} \sigma_k \sigma_i + \mu_k \mu_i \rho_{kj} \sigma_k \sigma_j + \rho_{kj} \rho_{ki} \sigma_k^2 \sigma_j \sigma_i)] \quad (28)$$

比较式 (28) 与式 (17) 可知， $V_k^{HT}|_{i=k} = V_k^L$ ，该结论与不含交叉项二次多项式情况下的结论一致。

#### (2) Li 指标结构贡献与 Hao 指标自身独立贡献的比较

Hao 指标体系在求  $x_k$  的自身独立贡献  $V_k^{Hu}$  时需将  $x_k$  分解为独立部分  $x_{ku}$  和相关部分  $x_{kc}$  (即  $x_k = x_{ku} + x_{kc}$ )，此时若二次多项式中包括平方项，则会产生附加交叉作用，因而影响各分解部分的正确性 [7]。为此本文只讨论不含平方项的模型

$$y = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n c_{ij} x_i x_j \quad (29)$$

由文献 [7] 可知，Hao 指标自身独立贡献  $V_k^{Hu}$  为

$$V_k^{Hu} = (\sigma_k^2 - \mathbf{B}_{k(-k)} \mathbf{B}_{(-k)(-k)}^{-1} \mathbf{B}_{k(-k)}^T) \left( a_k + \sum_{\substack{i=1 \\ i \neq k}}^n c_{ik} \mu_i \right)^2 \quad (30)$$

而 Li 指标结构贡献  $V_k^{La} = \sigma_k^2 \left( a_k + \sum_{\substack{l=1 \\ l \neq k}}^n c_{lk} \mu_l \right)^2$ ，显然  $V_k^{Hu}$  是  $V_k^{La}$  的一部分，与 Xu 指标独立贡献与 Li 指标结构贡献的关系类似。Hao 指标其他分解部分 (如独立交叉贡献) 的解析推导较为复杂，这些方差分解量与 Li 指标的比较将在算例 2 中简单分析。

### 2.3 一般形式响应模型 Li 指标与 Hao 指标的比较

由于 Xu 指标是 Hao 指标在不含交互作用时的特殊情况，因此本文只分析 Hao 指标与 Li 指标的关系。根据 2.1 和 2.2 节的结论推测：一般形式输出响应模型下，相关输入变量  $x_k$  的 Li 指标总贡献  $V_k^L$  是 Hao 指标相关总贡献  $V_k^{HT}$  解析表达式中的某一项；Li 指标结构贡献  $V_k^{La}$  完全包含了 Hao 指标的自身独立贡献  $V_k^{Hu}$ ，同时还包含了一部分与相关性有关的贡献。

以上结论可用示意图 1 表示，其中小圆内符号表示该小圆所连接的大圆包围的区域。以下将针对一般形式的响应模型对该结论进行证明。

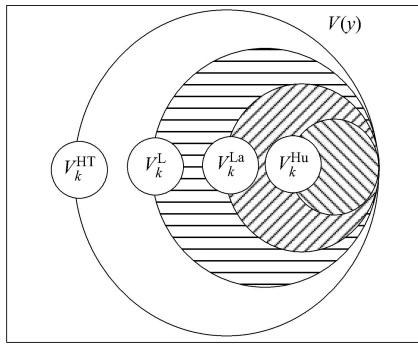


图1 Li 指标和 Hao 指标关系示意图

Fig. 1 The relation between Li's and Hao's indices

由于一般形式输出响应模型均可进行高维模型展开(如式(1)所示),本节从高维分解的角度研究一般响应模型下 Li 指标与 Hao 指标的关系,同时也作为对上述结论的证明.

(1) 可加模型 Li 指标与 Hao 指标的比较

可加模型的高维模型展开形式为  $y = g(\mathbf{x}) = g_0 + \sum_{i=1}^n g_i(x_i)$ . 相关输入变量  $x_k$  对应的分量函数为  $g_k(x_k)$ , 则 Li 指标中  $x_k$  的总贡献为

$$V_k^L = Cov\left(g_k(x_k), \sum_{i=1}^n g_i(x_i)\right) = \sum_{i=1}^n [E(g_k(x_k)g_i(x_i)) - E(g_k(x_k))E(g_i(x_i))] \tag{31}$$

由于高维模型展开的各分量函数期望值为 0<sup>[8]</sup>(后面推导均用到此结论), 上式简化为

$$V_k^L = \sum_{i=1}^n \int g_k(x_k)g_i(x_i)f_{X_k, X_i}(x_k, x_i)dx_k dx_i \tag{32}$$

可加模型的条件期望  $E(y|x_k) = g_0 + \sum_{i=1}^n \hat{g}_i(x_k)$ ,

其中  $\hat{g}_i(x_k)$  表示分量函数  $g_i(x_i)$  在条件概率密度函数  $f_{x_{-k}|X_k}(\mathbf{x}_{-k})$  下对除  $x_k$  以外其他所有变量  $\mathbf{x}_{-k}$  的积分<sup>[12]</sup>

$$\hat{g}_i(x_k) = \int g_i(x_i)f_{X_{-k}|X_k}(\mathbf{x}_{-k})d\mathbf{x}_{-k} = \int g_i(x_i)\frac{f_X(\mathbf{x})}{f_{X_k}(x_k)}d\mathbf{x}_{-k} \tag{33}$$

特别的, 当  $i = k$  时  $\hat{g}_i(x_k) = g_k(x_k)$ , 则有  $E(y|x_k)|_{i=k} = g_k(x_k)$ . 那么  $x_k$  的 Hao 指标相关总贡献  $V_k^{HT}$  的部分协方差为

$$V_k^{HT}|_{i=k} = Cov(E(y|x_k)|_{i=k}, E(y|x_k)) = \sum_{i=1}^n \int g_k(x_k)\left(\int g_i(x_i)\frac{f_X(\mathbf{x})}{f_{X_k}(x_k)}d\mathbf{x}_{-k}\right)f_{X_k}(x_k)dx_k =$$

$$\sum_{i=1}^n \int g_k(x_k)g_i(x_i)f_{X_k, X_i}(x_k, x_i)dx_k dx_i \tag{34}$$

对比式(31)和式(34)有  $V_k^L = V_k^{HT}|_{i=k}$ , 说明对于一般形式的可加模型, 图1所示的 Li 指标总贡献  $V_k^L$  与 Hao 指标相关总贡献  $V_k^{HT}$  的关系成立.

条件期望  $E(y|\mathbf{x}_{-k}) = g_0 + \hat{g}_k(\mathbf{x}_{-k}) + \sum_{i \neq k}^n g_i(x_i)$ , 其中  $\hat{g}_k(\mathbf{x}_{-k})$  表示分量函数  $g_k(x_k)$  在条件概率密度函数  $f_{X_k|\mathbf{x}_{-k}}(x_k)$  下对  $x_k$  的积分. 那么 Hao 指标自身独立贡献为

$$V_k^{Hu} = V(y) - V(E(y|\mathbf{x}_{-k})) = V(g_k(x_k)) - V(\hat{g}_k(\mathbf{x}_{-k})) + 2\left[Cov\left(g_k(x_k), \sum_{i \neq k}^n g_i(x_i)\right) - Cov\left(\hat{g}_k(\mathbf{x}_{-k}), \sum_{i \neq k}^n g_i(x_i)\right)\right] \tag{35}$$

由式(31)可知

$$Cov\left(g_k(x_k), \sum_{i \neq k}^n g_i(x_i)\right) = \sum_{i \neq k}^n \int g_k(x_k)g_i(x_i)f_{X_k, X_i}(x_k, x_i)dx_k dx_i \tag{36}$$

且有

$$Cov\left(\hat{g}_k(\mathbf{x}_{-k}), \sum_{i \neq k}^n g_i(x_i)\right) = \sum_{i \neq k}^n \int \left(\int g_k(x_k)\frac{f_X(\mathbf{x})}{f_{X_{-k}}(\mathbf{x}_{-k})}d\mathbf{x}_k\right) \cdot g_i(x_i)f_{X_{-k}}(\mathbf{x}_{-k})d\mathbf{x}_{-k} = \sum_{i \neq k}^n \int g_k(x_k)g_i(x_i)f_{X_k, X_i}(x_k, x_i)dx_k dx_i \tag{37}$$

所以  $V_k^{Hu} = V(g_k(x_k)) - V(\hat{g}_k(\mathbf{x}_{-k}))$ , 而  $V(g_k(x_k))$  恰好是变量  $x_k$  的 Li 指标结构贡献, 从而验证了图1中 Hao 指标自身独立贡献  $V_k^{Hu}$  与 Li 指标结构贡献  $V_k^{La}$  关系的正确性, 并且  $V_k^{La}$  比  $V_k^{Hu}$  多包含的部分的一般求解式为  $V(\hat{g}_k(\mathbf{x}_{-k}))$ .

(2) 一般形式输出响应模型 Li 指标与 Hao 指标的比较

在实际应用中, 通常将性能响应量函数展开到二阶就能满足所需的精度<sup>[9]</sup>

$$y = g(\mathbf{x}) \approx g_0 + \sum_{i=1}^n g_i(x_i) + \sum_{i=1}^n \sum_{j=i+1}^n g_{ij}(x_i, x_j) \tag{38}$$

$x_k$  的 Li 指标总贡献为

$$V_k^L = Cov\left(g_k(x_k), \sum_{i=1}^n g_i(x_i) + \sum_{1 \leq i < j \leq n} g_{ij}(x_i, x_j)\right) = \sum_{i=1}^n Cov(g_k(x_k), g_i(x_i)) + \sum_{i=1}^n \sum_{j=i+1}^n Cov(g_k(x_k), g_{ij}(x_i, x_j)) \quad (39)$$

式 (38) 的条件期望为

$$E(y|x_k) = g_0 + g_k(x_k) + \sum_{i \neq k} \hat{g}_i(x_k) + \sum_{i=1}^n \sum_{j=i+1}^n \hat{g}_{ij}(x_k) \quad (40)$$

其中  $\hat{g}_{ij}(x_k)$  表示分量函数  $g_{ij}(x_i, x_j)$  在条件概率密度函数  $f_{X_{-k}|X_k}(\mathbf{x}_{-k})$  下对  $\mathbf{x}_{-k}$  的积分. 取  $E(y|x_k)$  中  $g_k(x_k)$  这一项, 那么相关总贡献  $V_k^{HT}$  的部分协方差为

$$V_k^{HT}|_{i=k} = Cov(g_k(x_k), E(y|x_k)) = \sum_{i=1}^n Cov(g_k(x_k), \hat{g}_i(x_k)) + \sum_{i=1}^n \sum_{j=i+1}^n Cov(g_k(x_k), \hat{g}_{ij}(x_k)) \quad (41)$$

前面已经说明, 式 (41) 中  $\sum_{i=1}^n Cov(g_k(x_k), \hat{g}_i(x_k))$

与  $\sum_{i=1}^n Cov(g_k(x_k), g_i(x_i))$  相等, 对式 (41) 的后一部分

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=i+1}^n Cov(g_k(x_k), \hat{g}_{ij}(x_k)) = \\ & \sum_{i=1}^n \sum_{j=i+1}^n \int g_k(x_k) \left( \int g_{ij}(x_i, x_j) \cdot \frac{f(\mathbf{x})}{f(x_k)} d\mathbf{v}(\mathbf{x}_{-k}) \right) f(x_k) dx_k = \\ & \sum_{i=1}^n \sum_{j=i+1}^n Cov(g_k(x_k), g_{ij}(x_i, x_j)) \end{aligned} \quad (42)$$

因此,  $V_k^L = V_k^{HT}|_{i=k}$  同样成立. 同理,  $x_k$  的 Hao 指标的自身独立贡献  $V_k^{Hu}$  也是 Li 指标结构贡献  $V_k^L$  的一部分, 证明过程与前面相同.

### 3 重要性测度指标优缺点分析

Xu 指标通常适用于可加模型, Hao 指标在 Xu 指标的基础上考虑了交互作用的影响, 这两套指标都能够严格的将各相关输入变量中的独立和相关部分对总方差的贡献区分开. 但是它们均是针对单个

变量方差贡献进行分解, 未给出二阶及以上的重要性测度指标的定义及求解方法. 比如在输入变量  $x_k$  的相关总贡献  $V_k^{HT}$  中, 相关总贡献只包含了  $x_i$  与其他所有变量的相关性有关的那部分交叉贡献, 而对于不相关的两个变量之间交互作用的方差贡献并没有涉及 (算例 2 能够清晰体现).

Li 指标从高维模型展开的角度定义了总贡献、结构贡献和相关贡献, 包含一阶、二阶以及更高阶的指标, 因此能够很好的将变量的单独作用和交互作用分离开, 并实现对响应量总方差的完整分解, 从而避免 Xu 指标和 Hao 指标中信息重叠的情况. 但是该指标的结构贡献和相关贡献没有将变量的独立作用和和相关作用完全区分开. 可以推测, Li 指标的二阶求解结果能够帮助求解 Hao 指标中分散在相关部分和独立部分的交叉贡献, 这还有待进一步的探索.

通过第 2 部分详细比较 Xu 指标、Hao 指标与 Li 指标的解析表达式, 定量地明确了各指标中所包含内在信息之间的关系. 这些结论能够帮助工程人员更好的了解现有理论研究中的成果, 对于需要考虑输入变量相关性的问题, 合理选取所需要的重要性测度分析方法.

## 4 算例

**算例 1** 含交叉项二次多项式模型形式如下

$$y = 5 + x_1 + 2x_2 + 3x_3 + 2x_1^2 + 4x_2^2 + 2x_3^2 + 3x_1x_2 + x_1x_3 + 3x_2x_3 \quad (43)$$

其中  $x_i (i = 1, 2, 3) \sim N(\mu_i, \sigma_i^2)$ ,  $\boldsymbol{\mu} = (1, 2, 2)$ ,  $\boldsymbol{\sigma} = (2, 1, 2)$ , 相关系数  $\rho_{12} = 0.3$ ,  $\rho_{13} = 0.4$ ,  $\rho_{23} = 0.2$ . 运用 1.2 节的结论计算 Li 指标, 得到的结果列于表 1 的解析栏中 (用 ANA 表示), 用正交基函数法<sup>[13]</sup> (样本量  $N = 10000$ ) 的计算结果作为对照 (用 NUM 表示).

表 1 中计算结果表明, 本文推导的相关正态输入变量情况下二次含交叉项多项式模型重要性测度解析解是正确的, 另外一阶和二阶总贡献求和为 1, 进一步说明重要性测度指标  $S_{p_i}^L$  是对方差的完整分解. 3 个变量的一阶方差贡献都远大于二阶方差贡献, 因此在基于协方差分解的重要性测度体系中, 变量之间交互作用的影响相比变量自身的影响是比较小的, 这与传统基于方差的重要性测度指标相似. 总指标  $S_{p_i}^L$  一列中  $x_3$  的一阶指标最大, 说明变量  $x_3$  是相对最重要的变量, 那么要想减小输出响应模型

表 1 含交充满项二次多项式 Li 指标计算结果

Table 1 Computational results of Li's indices for example 1

	$S_{P_j}^{La}$		$S_{P_j}^{Lb}$		$S_{P_j}^L$	
	ANA	NUM	ANA	NUM	ANA	NUM
$x_1$	0.1590	0.1592	0.1339	0.1338	0.2930	0.2931
$x_2$	0.1505	0.1507	0.0869	0.0865	0.2374	0.2372
$x_3$	0.2816	0.2820	0.1274	0.1271	0.4091	0.4091
$x_1x_2$	0.0078	0.0078	0.0148	0.0149	0.0226	0.0227
$x_1x_3$	0.0037	0.0037	0.0138	0.0138	0.0175	0.0175
$x_2x_3$	0.0074	0.0074	0.0131	0.0131	0.0205	0.0205
sum	0.6100	0.6089	0.3900	0.3911	1.0000	1.0000

的不确定性首先应该从变量  $x_3$  入手。

**算例 2** 对二次含交叉项多项式模型  $y = x_1x_3 + x_2x_4$ , 其中  $(x_1, x_2, x_3, x_4) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ,  $\boldsymbol{\mu} = (0, 0, \mu_3, \mu_4)$ , 相关系数  $\rho_{13} = \rho_{14} = \rho_{23} = \rho_{24} = 0$ , 从而相关性矩阵为

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & 0 & 0 \\ \sigma_{21} & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & \sigma_{34} \\ 0 & 0 & \sigma_{43} & \sigma_4^2 \end{pmatrix}$$

按照式 (21) 得到输出总方差为

$$V(y) = \sigma_1^2(\sigma_3^2 + \mu_3^2) + \sigma_2^2(\sigma_4^2 + \mu_4^2) + 2\rho_{12}\sigma_1\sigma_2(\rho_{34}\sigma_3\sigma_4 + \mu_3\mu_4)$$

各变量的 Li 指标结构贡献  $V_k^{La}$ 、相关贡献  $V_k^{Lb}$  和总贡献  $V_k^L$  解析解列入表 2 中。根据文献 [7] 中推导的结论将各变量的相关总贡献  $V_k^{HT}$ 、自身独立贡献  $V_k^{Hu}$  和相关关系贡献  $V_k^{Hc}$  的解析解列入表 3 中。

表 2 各输入变量的 Li 指标解析结果

Table 2 Analytical results of Li's indices for example 2

$x_i$	$V_k^{La}$	$V_k^{Lb}$	$V_k^L$
$x_1$	$\mu_3^2\sigma_1^2$	$\mu_3\mu_4\rho_{12}\sigma_1\sigma_2$	$\mu_3^2\sigma_1^2 + \mu_3\mu_4\rho_{12}\sigma_1\sigma_2$
$x_2$	$\mu_4^2\sigma_2^2$	$\mu_3\mu_4\rho_{12}\sigma_1\sigma_2$	$\mu_4^2\sigma_2^2 + \mu_3\mu_4\rho_{12}\sigma_1\sigma_2$
$x_3$	0	0	0
$x_4$	0	0	0
$x_1x_3$	$\sigma_1^2\sigma_3^2$	$\sigma_1\sigma_2\sigma_3\sigma_4\rho_{12}\rho_{34}$	$\sigma_1^2\sigma_3^2 + \sigma_1\sigma_2\sigma_3\sigma_4\rho_{12}\rho_{34}$
$x_2x_4$	$\sigma_2^2\sigma_4^2$	$\sigma_1\sigma_2\sigma_3\sigma_4\rho_{12}\rho_{34}$	$\sigma_2^2\sigma_4^2 + \sigma_1\sigma_2\sigma_3\sigma_4\rho_{12}\rho_{34}$

表 3 各输入变量的 Hao 指标解析结果

Table 3 Analytical results of Hao's indices for example 2

$x_i$	$V_k^{HT}$	$V_k^{Hu}$	$V_k^{Hc}$
$x_1$	$(\sigma_1\mu_3 + \mu_4\rho_{12}\sigma_2)^2$	$\mu_3^2\sigma_1^2(1 - \rho_{12}^2)$	$\rho_{12}^2(\mu_3^2\sigma_1^2 + \mu_4^2\sigma_2^2) + 2\mu_3\mu_4\rho_{12}\sigma_1\sigma_2$
$x_2$	$(\sigma_2\mu_4 + \mu_3\rho_{12}\sigma_1)^2$	$\mu_4^2\sigma_2^2(1 - \rho_{12}^2)$	$\rho_{12}^2(\mu_3^2\sigma_1^2 + \mu_4^2\sigma_2^2) + 2\mu_3\mu_4\rho_{12}\sigma_1\sigma_2$
$x_3$	0	0	0
$x_4$	0	0	0

分析表 2 和表 3 可以看出, 一、二阶 Li 指标总贡献求和与总方差相等, 可保证分解结果的正确性。根据本文得到的结论, Li 指标总贡献  $V_k^L$  包含在 Hao 指标相关总贡献  $V_k^{HT}$  中, 且 Li 指标相关贡献包含在 Hao 指标相关关系贡献  $V_k^{Hc}$  中, 而  $x_3$  和  $x_4$  的  $V_k^{HT}, V_k^{Hc}$  均为零, 因此  $x_3$  和  $x_4$  的 Li 指标均为零, 也说明了这两个变量对响应模型没有一阶方差贡献。在相关总贡献  $V_k^{HT}$  中没有包含如  $\sigma_1^2\sigma_3^2\rho_{13}^2$  的相关交叉贡献, 这是由于相关总贡献中只包含与相关性有关的那部分交叉作用贡献, 而本算例中  $\rho_{13} = 0$  使得  $\sigma_1^2\sigma_3^2\rho_{13}^2$  未计入  $V_k^{HT}$  中。另需注意,  $V(y)$  中的  $\rho_{12}\rho_{34}\sigma_1\sigma_2\sigma_3\sigma_4$  项虽然包含在 Li 指标 (如表 2) 二阶方差贡献里, 却不属于 Hao 指标 (如表 3) 中的任意一个分解量, 这是由于 Hao 指标只是针对某一变量的方差贡献 (即一阶重要性测度指标), 而未给出二

阶及以上的重要性测度指标的定义及求解方法。

**算例 3** 用一个机翼三盒段来模拟机翼结构, 如图 2 所示, 此三盒段包含 28 个杆和 16 块板。设 3 个方向杆的长度分别为  $L_1, L_2$  和  $L_3$ , 所有杆的横截面积均相同, 用  $A$  表示。弹性常数用  $E$  表示, 泊松比均为

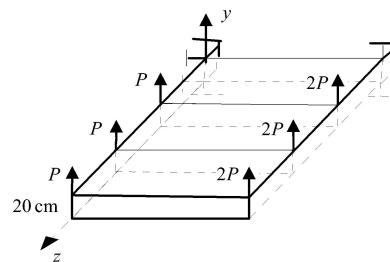


图 2 机翼三盒段结构简图

Fig. 2 Diagram of 3-box wing simulation structure



0.3, 字母  $P$  表示载荷,  $TH$  表示所有板的厚度 (均相同). 这些输入量均服从正态分布, 分布参数如表 4 所示. 相关系数  $\rho_{L_1L_2} = 0.1$ ,  $\rho_{L_1L_3} = 0.1$ ,  $\rho_{E(TH)} = 0.2$ , 其余输入量间相关系数为 0. 输出响应量为该结构在载荷  $P$  作用下的最大位移.

表 4 三盒段模型输入变量分布参数

Table 4 The distribution parameters of inputs for example 3

Variables	Index	Mean	Standard deviation
$A/\text{mm}^2$	$x_1$	100	10
$L_1/\text{mm}$	$x_2$	600	60
$L_2/\text{mm}$	$x_3$	200	20
$L_3/\text{mm}$	$x_4$	400	40
$E/\text{MPa}$	$x_5$	71 000	7 100
$P/\text{N}$	$x_6$	1 500	150
$TH/\text{mm}$	$x_7$	3	0.3

若采用不含交叉项二次响应面拟合得到

$$y = 0.0005A - 0.0042L_1 - 0.0733L_2 - 0.0084L_3 - 0.0001E + 0.0012P - 1.9584TH + 0.0001L_2^2 + 0.2291TH^2$$

检验得到  $R^2 = 0.9486$  (方差比  $R^2$  越接近 1, 表示拟合模型与原模型近似程度越高), 说明三盒段结构可近似看做可加模型且忽略交互作用的影响.  $Li$  指标计算结果如表 5,  $Xu$  指标计算结果如表 6.

比较表 5 和表 6 的最后一列, 各变量的  $S_k^X$  均大于  $S_k^L$ . 由于  $L_1$  与  $L_2, L_3$  相关,  $L_1$  的  $Xu$  指标总贡献比  $Li$  指标总贡献多包含了变量  $L_2, L_3$  中相关部分的方差贡献, 从而验证了本文得出的结论. 此外, 表 6 最后一列表明所有输入变量的  $Xu$  指标总贡献求和大于 1, 说明了各输入变量的总贡献中包含了大

表 5 机翼三盒段结构  $Li$  指标重要性测度结果

Table 5 Computational results of  $Li$ 's indices for example 3

	$S_k^{La}$	$S_k^{Lb}$	$S_k^L$
$A$	0.0004	0	0.0002
$L_1$	0.0221	-0.0021	0.0200
$L_2$	0.2618	0.0074	0.2692
$L_3$	0.4189	-0.0095	0.4094
$E$	0.0771	0.0135	0.0906
$P$	0.0685	0	0.0685
$TH$	0.0615	0.0135	0.0750
sum	0.9104	0.0228	0.9331

表 6 机翼三盒段结构  $Hao$  指标重要性测度结果

Table 6 Computational results of  $Hao$ 's indices for example 3

	$S_k^{XU}$	$S_k^{XC}$	$S_k^X$
$A$	0.0001	-0.0001	0
$L_1$	0.0112	0.2394	0.2506
$L_2$	0.1553	0.2732	0.4285
$L_3$	0.5002	0.2062	0.7063
$E$	0.0335	0.0489	0.0824
$P$	0.0367	0.0697	0.1063
$TH$	0.0509	0.1073	0.1582
sum	0.7879	0.9446	1.7325

量重叠部分. 因此当想要按照这 7 个变量得到对应总方差的完整分解, 采用  $Li$  指标中的数据要更加合理; 当需要严格将变量的相关部分和独立部分区分开时, 采用  $Xu$  指标更加合理.

### 5 结论

本文以多维正态变量的概率统计性质为基础, 针对不含交叉项与包含交叉项的二次多项式响应模型, 推导了相关正态输入变量的  $Li$  指标总贡献、结构贡献和相关贡献的解析解. 然后以所推导的解析结论为基础, 探究  $Li$  指标与  $Xu$  指标的关系以及  $Li$  指标与  $Hao$  指标的关系, 所得出的结论为:  $Li$  指标总贡献是  $Hao$  指标相关总贡献解析式中的某一项, 而  $Li$  指标结构贡献完全包含了  $Hao$  指标的自身独立贡献, 同时还包含一部分与相关性有关的贡献, 从而  $Hao$  指标中的相关贡献也就包含了  $Li$  指标的相关贡献. 最后从高维模型分解的角度验证以上结论. 算例部分验证了本文所推导结论的正确性, 并具体分析了不同重要性测度指标的关系.

### 参 考 文 献

- 1 Saltelli A, Ratto M. Global Sensitivity Analysis: The Primer. Chichester: Wiley, 2008
- 2 Sobol IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 2001, 55(1): 21-80
- 3 Satelli A, Sensitivity analysis for importance assessment. *Risk Analysis*, 2002, 22(3): 579-590
- 4 Iman RL, Johnson ME, Schroeder TA. Assessing hurricane effects. Part 1. Sensitivity analysis. *Reliability Engineering and System Safety*, 2002, 78(2): 131-145
- 5 Xu C, Gertner GZ. Uncertainty and sensitivity analysis for models with correlated parameters. *Reliability Engineering and System Safety*, 2008, 93(6): 1563-1573
- 6 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. 力学学报, 2012, 44(1): 167-173 (Hao Wenrui, Lü Zhenzhou, Wei Pengfei. Importance measure of correlated variables in polynomial output. *Chinese Journal of Theoretical and Applied Mechanics*, 2012, 44(1): 167-173 (in Chinese))
- 7 Hao W, Lu Z, Tian L. A new interpretation and validation of variance based importance measure analysis for model with correlated inputs. *Computer Physics Communications*, 2013, 184(5): 1401-1413
- 8 Li G, Rabitz H, Yelvington PE, et al. Global sensitivity analysis for systems with independent and/or correlated inputs. *The Journal of Physical Chemistry A*, 2010, 114(19): 6022-6032
- 9 Li G, Rabitz H. General formulation of HDMR component functions with independent and correlated variables. *Journal of Mathematical Chemistry*, 2012, 50(1): 99-130
- 10 吕震宙, 宋述芳, 李洪双等. 结构机构可靠性及可靠性灵敏度分析. 北京: 科学教育出版社, 2009 (Lü Zhenzhou, Song Sufang, Li Hongshuang, et al. The reliability and reliability sensitivity analysis for structure and machine system. Beijing: Science Press, 2009 (in Chinese))
- 11 赵选民, 徐伟, 师义民等. 数理统计 (第二版). 北京: 科学教育出版社, 2002 (Zhao Xuanmin, Xu Wei, Shi Yimin, et al. Mathematical statistics. 2nd edition. Beijing: Science Press, 2002 (in Chinese))
- 12 Li L, Lu Z. Importance analysis for models with correlated variables and its sparse grid solution. *Reliability Engineering and System Safety*, 2013, 119: 207-217
- 13 Li G, Hu J, Wang S, et al. Random sampling-high dimensional model representation and orthogonality of its different order component functions. *Journal of Physical Chemistry*, 2006, 110(7): 2474-2485

(责任编辑: 郭 旭)

(责任编辑: 刘希国)

## INVESTIGATION OF THE RELATION OF IMPORTANCE ANALYSIS INDICES FOR MODEL WITH CORRELATED INPUTS<sup>1)</sup>

Song Jingwen Lü Zhenzhou<sup>2)</sup>

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

**Abstract** Nowadays, several importance analysis methods have been developed for model with correlated inputs. For choosing the most appropriate analysis methods to meet different requirements, it is necessary to make differences among these existing methods. In this paper, the importance indices, including the total, the structural and the correlative contributions, derived from the covariance decomposition, are firstly derived for the quadratic polynomial without interaction terms and the one with interaction terms. Then, based on these derived analytical solutions, the relation between the traditional variance based method and the newly covariance based method is explored. The results derived from the quadratic polynomials are then extended to general cases, and validated from the point of high dimensional model representation. Three examples are introduced for investigating the relation between the two groups of importance indices, and relative merits of each. The conclusions are instructive and meaningful to importance analysis and engineering design when the model inputs are correlated.

**Key words** importance measure, correlated variable, variance decomposition, covariance decomposition, quadratic polynomial

Received 31 October 2013, revised 9 January 2014.

1) The project was supported by the National Natural Science Foundation of China (51175425), Research Found for the Doctoral Program of Higher Education of China (20116102110003), and the Undergraduate Thesis Project of Northwestern Polytechnical University in 2013.

2) Lü Zhenzhou, professor, research interests: reliability engineering, sensitivity analysis, multidiscipline optimization.

E-mail: zhenzhoulu@nwpu.edu.cn