

文章编号: CSTAM2014-P19-0041

# 不同参数对气动性能影响程度的量化评价方法

罗长童, 姜宗林, 王 春, 胡宗民

(中国科学院力学研究所高温气体动力学国家重点实验室, 北京 100190)

**摘 要:** 影响气动力/热的因素往往比较多, 包括来流(如静温、静压、马赫数)、飞行姿态(如攻角)、试验条件(如模型尺度、壁温)等。正确评价这些因素的相对影响程度对于飞行器的设计和控制都有重要意义。然而目前广泛采用的都是定性化或准定性化评价方法, 缺乏对各因素相对影响程度的量化指标。本文通过分析传统相关系数背后的思想, 引入适用范围更广的概念——相对影响率(RIR)。RIR 在多维空间上采用缺位法(思想类适于反证法)定义, 具有全局优化特性, 适用于线性和非线性情形的多参数分析, 为各参数对气动力/热性能的影响程度提供了科学的定量化评价依据。

**关键词:** 气动性能; 量化评价; 缺位法; 相对影响率

**中图分类号:** O355, O224

**文献标识码:** A

## 0 引 言

影响气动力/热的因素往往比较多, 包括来流(如静温、静压、马赫数)、飞行姿态(如攻角、侧滑角)、试验条件(如模型尺度、壁温)等。正确评价这些因素的相对影响程度对于飞行器的设计和控制都有重要意义。然而, 对于多个影响因素的情况, 准确描述其相对影响并非易事。比如在研究不同因素对某飞行器典型位置无因次热流影响规律时, 通常可以用表1描述。

表1中的结论往往是通过数据曲线对比观察得到的, 换句话说, 这是一种通过人工方式得到的离散量化指标。显然结论具有一定模糊性和人为因素。

事实上, 目前广泛采用的都是定性化或准定性化评价方法, 缺乏一个科学的对各影响因素相对影响程度进行刻画的量化指标。本文以数据的相关性分析方法为基础, 分析其背后蕴含的简化思想, 并

据此将相关概念拓广到非线性、多参数问题的研究中。

表 1.不同因素对某飞行器典型位置无因次热流影响

影响因素	典型位置无因次热流影响规律	敏感性程度评价
静压	基本不随来流静压变化	☆☆☆
静温	随来流静温增加而增加	★★☆
马赫数	随来流马赫数增加而增加	★☆☆
高度	随飞行高度增加而减小	★★☆
攻角	随来流攻角增加而明显增加	★★★
模型尺度	基本不随模型尺度变化	☆☆☆
壁温	随壁温增加而减小	★★☆

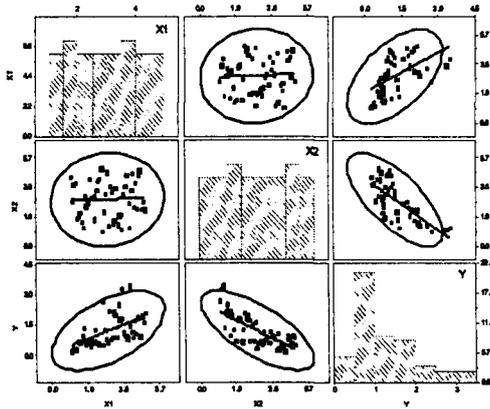
## 1 研究基础

对于多参数问题, 如果某一个参数的影响程度很小, 那么该参数和目标之间在统计上应当是不相关的或相关性很小。据此, 我们曾经给出过相关阵

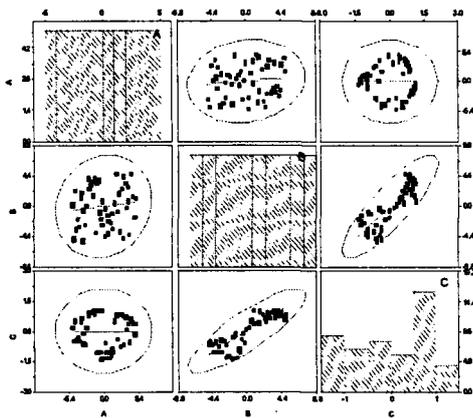
基金项目: 国家自然科学基金(编号 90916028)

作者简介: 罗长童(1974-), 男, 河南固始人, 副研究员, 流体力学、计算数学; E-mail: luo@imcch.ac.cn

图的概念,即将数据投影到两两配对的二维平面上,再根据数据点的分布判断各参数之间以及各参数与目标之间的相关关系。



(a)



(b)

图1 多维离散数据的相关阵图示例

比如有两个参数 $X_1, X_2$ ,共同影响目标 $Y$ ,现有50个样本点 $(X_{1i}, X_{2i}, Y_i), i = 1, 2, \dots, 50$ . 可以将这50个点分别投影到 $X_1 - X_2, X_1 - Y$ 和 $X_2 - Y$ 平面,如图1(a)所示。在图1(a)中, $X_1$ 和 $X_2$ 没有相关性(或相关性小),其相关椭圆接近圆形; $X_1$ 和 $Y$ 是正相关的,而 $X_2$ 和 $Y$ 是负相关的,它们的相关椭圆都很“扁”,而且 $X_1 - Y$ 对应的椭圆长轴斜率为正,而 $X_2 - Y$ 对应的椭圆长轴斜率为负。事实上,三者

的关系为 $Y = X_1/X_2, (X_1, X_2) \in [1, 5]^2$ 。

一般情况下,通过相关阵图可以看出定性的相关关系。甚至在其他参数影响不大情况下可近似得到定量关系。然而,相关阵图在高度非线性情况下可能无法得到定性关系。例如,对于非线性函数关系

$$Y = (5 * X_2 + 0.2 * X_1^2) / (3 + X_1^2 + X_2^2),$$

$(X_1, X_2) \in [-5, 5]^2$ ,其50个随机样本的相关阵图如图1(b)所示,从右上角的椭圆接近圆形,说明在本例中, $X_1$ 和 $Y$ 相关关系探测失败。事实上,还有些情况是由于试验条件限制等原因,本来不相关的变量,其相关椭圆也可能很“扁”。因而,对于多参数非线性问题,必须引入新的概念,才能准确进行定量描述。

## 2 相对影响率的概念

### 2.1 概念的提出

早在1895年,皮尔逊(K. Pearson)就提出相关系数的概念[1],其定义可描述为:

$$r = \frac{E[(X - E(X))(Y - E(Y))]}{\sigma_X \sigma_Y},$$

它本质上将多参数问题降维到二维空间,即将样本点投影到二维平面,进行线性拟合,根据斜率正负和残差大小判断相关关系:斜率为正,则正相关;斜率为负,则负相关;斜率为零,则不相关;相关情况下,残差为零则相关系数为+1或-1,否则介于-1到+1之间。皮尔逊相关系数对线性依赖性太强,于是,斯皮尔曼(C. Spearman)根据序关系引入秩(rank)的概念,将相关性概念由线性推广到单调情形[2]。随后,肯德尔(M. Kendall)引入相容关系(concordant, discordant),对相关关

系作进一步推广[3]。斯皮尔曼相关系数 $\rho$ 和肯德尔 $\tau$ 都能部分描述非线性相关关系，只要两个变量之间存在某种序关系即可。但是对于一般的非线性相关问题，前述的三个相关系数都可能无法正确表达。

不妨以 $Y = X^m$ ， $m = 1, 2, 10, 100$ 为例，其在不同样本区间上的相关系数如表2所示。可见，随着非线性增强，皮尔逊相关系数越来越小于1，偏离其定义本身的意图。斯皮尔曼相关系数和肯德尔相关系数在一定程度上弥补了其缺陷，如其中的 $X \in [0, 1]$ 情形，这两个相关系数就准确表达了其相关性的含义。但在 $X \in [-1, 1]$ 情形，它们的缺陷就显示出来了：虽然 $X-Y$ 在非线性意义下是相关的，但除线性关系 $Y=X$ 外，其他非线性情形，三个相关系数取值均为0，这显然背离了相关关系的内涵。

表2.传统相关系数的局限性

	Pearson's r	Spearman's $\rho$	Kendall $\tau$	
[0,1]	$Y = X$	1	1	1
	$Y = X^2$	0.97	1	1
	$Y = X^{10}$	0.66	1	1
	$Y = X^{100}$	0.25	1	1
[-1,1]	$Y = X$	1	1	1
	$Y = X^2$	0	0	0
	$Y = X^{10}$	0	0	0
	$Y = X^{100}$	0	0	0

因而，2012年郑等人提出广义相关测度(GMC)的概念<sup>[4]</sup>。

$$\{GMC(Y|X), GMC(X|Y)\} = \left\{ 1 - \frac{E\{(Y - E(Y|X))^2\}}{\text{var}(Y)}, 1 - \frac{E\{(X - E(X|Y))^2\}}{\text{var}(X)} \right\}$$

真正将相关系数推广到一般的非线性情形。

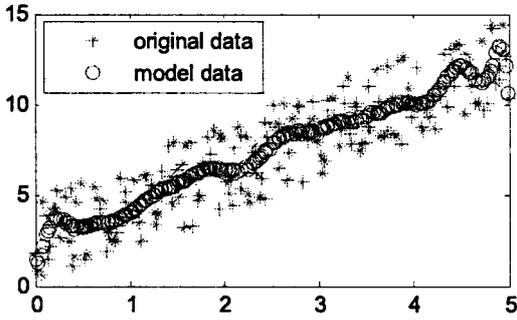
注意到广义相关测度仍然蕴含降维（到二维）的思想，即对于多因素问题，仍然需要投影到二维平面以后进行处理，因而，只能分别考察各自对目标函数的影响。要综合考察多参数对目标函数的相对影响大小，必须引入新的概念。

## 2.2 相对影响率的定义

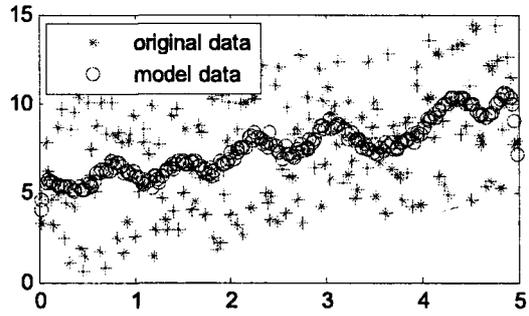
已有相关系数都蕴含降维思想，把多维问题简化为二维，信息被过度压缩，因而在本文中，我们放松维度要求，采取缺位法定义，仅将问题从 $n+1$ 维降到 $n$ 维，以确保样本数据包含的信息得以充分保留。缺位法的定义方法类似于反证法，其基本思想如下：假如因素 $X_i$ 在多因素 $(X_1, X_2, \dots, X_n)$ 中相对影响率大，那么在 $X_i$ 缺位后，相对于模型 $m_i(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$ 而言，数据发散性大；反之，缺位后数据仍然能够保持相对集中。这里的集中在二维空间中表现为共线，在三维空间中表现为共面或共线，在 $n+1$ 维空间中表现为数据点集中于 $k$ 维超曲面( $k \leq n$ )。下面我们以一个浅显的例子来说明缺位法、发散度和相对影响的关系。设 $Y_1 = X_1 + 2X_2$ ，在区域 $[0, 5]^2$ 上取200个样本点。当然，本例中各影响因素的关系为已知，在实际应用中是没有这种显式关系的，这里仅仅为了说明概念的含义。显然，如果将 $Y_1$ 看成气动性能指标，那么

其中的 $X_1$ 影响较小, 而 $X_2$ 的影响较大, 影响率应当是2倍的关系。利用缺位法, 在 $X_1$ 缺位的条件下建立某种逼近模型 $m_1(X_2)$ , 将200个样本点的 $(X_2, Y_1)$ 和 $(X_2, m_1(X_2))$ 绘制与统一坐标系中对比如图2(a), 所

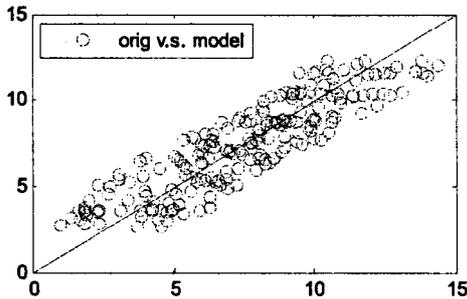
示; 同理将 $X_2$ 缺位后的点 $(X_1, Y_2)$ 和 $(X_1, m_2(X_1))$ 绘制于图2(b). 可见发散度和相对影响率一致。



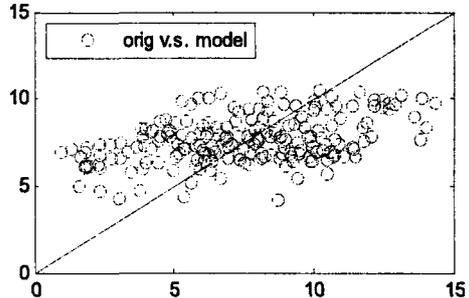
(a)  $X_1$ 缺位, 发散度较小, 表明 $X_1$ 相对影响率较小



(b)  $X_2$ 缺位, 发散度较大, 表明 $X_2$ 相对影响率较大



(c)  $X_1$ 缺位, 样本偏离对角线较小, 预测散度较小



(d)  $X_2$ 缺位, 样本偏离对角线较大, 预测散度较大

图 2. 缺位法、发散度和相对影响的关系

对于两个影响因素, 采用以上图示就能很直观地看出样本的发散程度, 对于三个影响因素, 也可以通过绘制三维散点图看出发散程度。限于空间表现形式, 四个以上的影响因素, 无法直接利用缺位后散点分布图表达。但无论维数多少, 我们总可以采用预测散度图, 即 $(Y, m_i(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n))$ 散点图, 比如对于上例, 其预测散度分布如图2(c),(d)所示, 同样表达了由于某因素缺位而引起的样本散度。

下面我们给出更确切的定义来描述这种由于缺

位而引起的样本散度。在一般的 $n$ 维实空间 $R^n$ 中, 影响目标函数 $f$ 的参数 $X_i$ 的相对影响率 $r_{X_i}$ 定义如下:

定义1. 设 $f: \Omega \mapsto R, (X_1, X_2, \dots, X_n) \in \Omega \subset R^n$ ,

$S_{X_i}$ 为参数 $X_i$ 的补空间( $n$ 维实空间),  $P_{f, X_i} \subset S_{X_i}$ 为

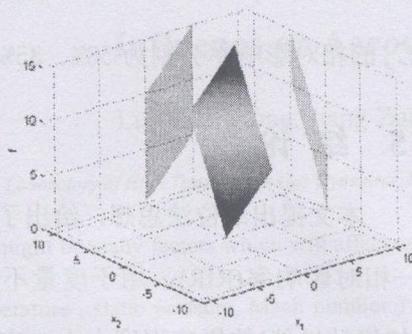
目标函数超曲面在的补空间 $S_{X_i}$ 上的投影,  $V_{f, X_i}$ 为

$P_{f, X_i}$ 在补空间 $S_{X_i}$ 上的体积,  $i=1, 2, \dots, n$ , 则参

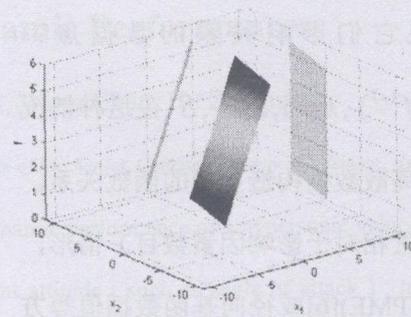
数 $X_i$ 对目标函数 $f$ 的相对影响率 $r_{X_i} = \frac{V_{f, X_i}}{\sum_{i=1}^n V_{f, X_i}}$ 。

按照以上定义,如果是双因素问题(仅有两个变量),那么可以很直观地看出相对影响率的几何意义,下面以几个例子说明(图3)。前两个例子是线性的,很容易得到各投影的体积(此时退化为面积),其中图3.(a)中,关于参数 $X_1$ 的补空间投影(右侧)

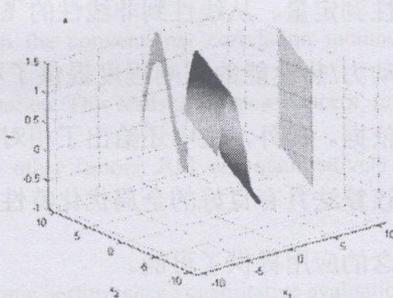
的面积是关于参数 $X_2$ 的补空间投影(左侧)面积的1/2,因而相对影响率 $r_{X_1} = 1/3$ ,  $r_{X_2} = 2/3$ 。同理可得,对于 $f_2$ ,有 $r_{X_1} = 1/3$ ,  $r_{X_2} = 2/3$  (图3.(b))。



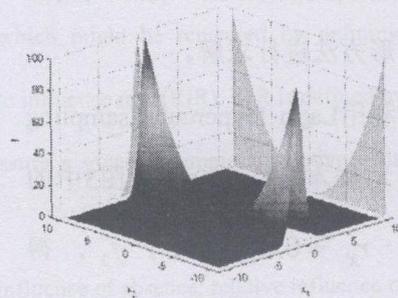
(a)  $f_1(X) = X_1 + 2X_2, X \in [0, 5]^2$



(b)  $f_2(X) = X_1 + 0.1X_2, X \in [0, 5]^2$



(c)  $f_3(X) = \sin(X_1) + 0.1X_2, X \in [0, 5]^2$



(d)  $f_4(X) = \exp(-X_1^2 - X_2^2) \cdot (X_1^2 + X_2^2), X \in [-10, 10]^2$

图3.相对影响率的几何意义(双因素问题)

然而对于非线性问题,求解补空间上投影的体积并没有那么简单,即使是双因素的二维问题,有时精确求解也很困难(如图3中的 $f_3, f_4$ )。因而,接下来我们给出求解相对影响率的数值算法。

### 3 相对影响率的数值算法

对于影响关系已知的情形(比如图3中的情形),我们利用均匀设计和数据建模,给出相对影响率的求解方法如下:

步1.在给定区域 $\Omega$ 上均匀采样(推荐采用Latin

Hypercube sampling或Orthogonal sampling方法),计算出有一定规模的样本点 $(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$ 及其对应目标值 $f^{(k)}, k = 1, 2, \dots, K$ 。

步2.对于每一个影响因素 $X_i$ ,去掉 $X_i$ 所在的列(使

$X_i$ 缺位),根据 $(x_1^{(k)}, \dots, x_{i-1}^{(k)}, x_{i+1}^{(k)}, \dots, x_n^{(k)})$ 和 $f^{(k)}$

的取值构造逼近模型(推荐采用DACEFIT方法

[5]) $m_i(x_1^{(k)}, \dots, x_{i-1}^{(k)}, x_{i+1}^{(k)}, \dots, x_n^{(k)})$ 。

步3.计算残差和

$$V_i = \sum_{k=1}^K |f^{(k)} - m_i(x_1^{(k)}, \dots, x_{i-1}^{(k)}, x_{i+1}^{(k)}, \dots, x_n^{(k)})|$$

步4.重复步2和步3得到所有  $V_i$ , 近似取  $V_{f, x_i} = V_i$ ,

$$\text{从而得到 } r_{x_i} = \frac{V_{f, x_i}}{\sum_{i=1}^n V_{f, x_i}}$$

在实际应用中, 影响因素和目标函数之间可能并没有显式的函数关系, 而只能根据实验或数值模拟得到一些反映它们影响关系的数据点:

$$(x_1^{(s)}, x_2^{(s)}, \dots, x_n^{(s)}, f^{(s)}), s = 1, 2, \dots, S.$$

在这种情况下, 可以首先利用离散数据构造显式的函数关系。

对于小样本(数据量相对于影响因素数目)情形, 推荐采用语法进化(PME)[6]或径向基函数插值等方法构造解析关系式; 对于大样本数据, 仍然可以通过上述的DACEFIT方法构造函数关系。然后利用上面的影响关系已知情形方法进行求解。

运用上述算法, 利用Latin Hypercube sampling方法进行均匀采样, 取样本数  $K=100$ , 对于图3中的  $f_1$ , 得  $r_{x_1} = 0.33$ ,  $r_{x_2} = 0.67$ ; 对于  $f_2$ , 得  $r_{x_1} = 0.09$ ,  $r_{x_2} = 0.91$ ; 对于  $f_3$ , 得  $r_{x_1} = 0.83$ ,  $r_{x_2} = 0.17$ ; 对于  $f_4$ , 得  $r_{x_1} = 0.57$ ,  $r_{x_2} = 0.43$ 。

## 4 应用实例

对于高超声速飞行器, 影响其轴向力系数  $C_A$  的因素包括马赫数( $Ma$ )、雷诺数( $Re$ )和壁温比( $\hat{T}_w$ )等因素, 根据已知的试验和计算数据判断不同因素的影响大小对飞行器的设计和控制有重要的指导意义。对于某飞行器, 我们首先根据获得一定范围内的试验和计算数据, 利用解析矩阵进化算法(PME)得到了其解析关系(过程略):

$$C_A = 0.029 + 0.907 \left( \frac{1}{Ma^2} + \frac{1}{Re^{1/4}} + \frac{0.260T_0}{T_w} \right),$$

其中  $Ma \in [5, 8]$ ,  $Re \in [3.8 \times 10^5, 4.8 \times 10^7]$ , 壁温  $T_w \in [60, 250.3]K$ 。利用本文的缺位法分析, 可知马赫数  $Ma$ 、雷诺数  $Re$ 、壁温比  $\hat{T}_w$  对轴向力系数  $C_A$  的相对影响率分别为 58%、35%和 7%。

## 5 结论

本文提出缺位法思想, 给出了一种新的概念——相对影响率(RIR), 用于度量不同参数对目标值(如气动性能)的影响程度大小。它解决了以往方法信息被过度压缩等问题, 实现多因素影响分析从定性到定量、从线性到非线性的飞跃, 为各参数对气动力/热性能的影响程度提供了科学的定量化评价依据。另外, 文中还给出了相对影响率数值算法, 该算法具有良好的全局优化特性, 为相对影响率概念的应用提供了可能。

注意到相对影响率是一个区域相关量, 当影响因素的变化范围改变时, 相对影响率也会随之改变。这样就可以进一步将相对影响率的概念推广到局部区域上的相对影响率和极限意义下的相对影响率, 这些概念将在后续的研究中进一步探讨。另外, 关于文中提出的算法, 其相对稳定性也值得进一步验证。

## 参考文献

- [1] K. Pearson, Notes on regression and inheritance in the case of two parents, Proceedings of the Royal Society of London, 58: 240-242, 1895.
- [2] C. Spearman, The proof and measurement of association between two things, Amer. J. Psychol. 15: 72-101, 1904.
- [3] M. Kendall, A New Measure of Rank Correlation, Biometrika 30 (1-2): 81-89, 1938.

- [4] S. Zheng, N.-Z. Shi, Z. Zhang, Generalized Measures of Correlation for Asymmetry, Nonlinearity, and Beyond, *Journal of the American Statistical Association*, 107:499, 1239-1252, 2012.
- [5] Søren N. Lophaven, Hans Bruun Nielsen, Jacob Søndergaard, DACE - A Matlab Kriging Toolbox, Technical Report, Technical University of Denmark, IMM-REP-2002-12, 1-26, 2002.
- [6] C.T. Luo, S.-L. Zhang, Parse-matrix evolution for symbolic regression, *Engineering Applications of Artificial Intelligence*, 25: 1182–1193, 2012.

## **A New Method for Quantitative Measurement of Different Parameters' Influence on Aerodynamic Performances**

Luo Changtong, Jiang Zonglin, Wang Chun, Hu Zongmin

*(State Key Laboratory of High Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190)*

**Abstract:** There might be many factors which will affect aerodynamic performance of a hypersonic vehicle including free stream ( such as static temperature , static pressure, Mach number ) , flight attitude ( such as angle of attack ) , test conditions ( such as model scale , wall temperature ) and so on . Properly evaluate the relative influence of each factor is very important for aircraft design and flight control. However, the widely used methods are all qualitative or semi-qualitative. This paper find out that the information are over-compressed in the conventional correlation methods, which might be remedied by defining a new coefficient with less-compressed information. This leads to our new concept, relative influence rate (RIR). RIR is defined by considering its influence of absence, relative to other factors. RIR can quantitatively measure a given parameter's relative influence, both in linear and non-linear cases.

**Key words:** aerodynamic performance ; quantitative evaluation ; influence of absence; relative influence rate