

A Simplified Model of Quadratic Polynomials RSM And Its Design of Experiment

Jinwen Wei*

College of Mechanical Engineering, Guangxi University, Nanning, Guangxi, China, 530004

*Corresponding author

Abstract—Considering the inefficiency of classical DOE (Design Of Experiment) based polynomials RSM (Response Surface Method), this paper proposes a simplified experimental modeling method to approximate a quadratic polynomials RSM model (called Archetype in this paper). This simplified model contains the same number of coefficients as that of modeling variables, and uses orthogonal vectors of modeling variables as DOE. Such related problems as basis function, DOE, etc. are analyzed theoretically and demonstrated with examples. When applied to approximate an Archetype of positive define, central distribution eigenvalue and high dimension modeling variables, experiment times can be greatly decreased while enough precision is obtained in this simplified RSM model.

Keywords—quadratic polynomials; Response Surface Method (RSM); Design Of Experiment (DOE); simplified model

I. INTRODUCTION

Response surface method (RSM) is a test modeling method, which has been widely used in structural optimization and dynamic analysis, etc^[1]. The key of RSM modeling lies in two aspects: base function selection and experiment design, which aim to obtain accurate object model with fewer experiment times^[2]. Among the RSM basis functions, polynomials, especially quadratic polynomials, are the most commonly used basis functions with the advantages of simple calculation, multi constraint problems and sensitivity analysis. Two polynomial model usually adopts two forms with and without cross terms:

$$a_0 + \sum a_i x_i + \sum a_{ij} x_i^2 \quad \text{and} \quad a_0 + \sum a_i x_i + \sum \sum a_{ij} x_i x_j$$

There are few parameters in the model of without cross terms. When the parameters are fitted by Design Of Experiment (DOE), the sampling mean is usually taken as the sampling Center^[3]. This leads to larger fitting errors at the non sampling center points, so many iterative improvement methods have been proposed^[4]. However, the accuracy of this model is not high, due to the omitting of too many cross terms.

More coefficients are contained in a model with cross terms, which can improve its accuracy. But it needs more complex processing process, including introducing error correction terms, interpolation point instead of limit state method^[8], and so on^[5]. In addition, the idea of gradually introducing cross terms is put forward^[3]. Due to the need to consider the intersection degree of each cross terms, these

methods are tedious and inefficient, effective only for individual objects, rather than universal.

The classical DOE of polynomial RSM includes central composite design^[6], orthogonal design, Latin hypercube design, etc., and their various improvements. The sampling method and the number of samples vary with different methods, for example, an orthogonal design is determined by a designing tables; a Latin hypercube design is determined according to the actual variable number and level number. Experiment efficiency is low in these DOEs due to many factors. First of all the range of value are usually given while the specific sample value of an object is not given^[7]. And generally the more common sample points, the higher accuracy, but in many cases, a large number of samples cannot obtain high precision model^[8]. For example, the local modeling of a complex archetype usually rely only on regional samples, while scatter sample points decrease instead of increase accuracy and efficiency.

In a common polynomial based RSM model, the number of coefficients increases rapidly with the number of variables, which makes the modeling method and experiment design tedious and inefficient. In this paper, a simplified form of quadratic polynomial RSM is proposed. Cross terms are naturally contained in the model without special treatment; when an experimental design is carried out, experimental times are guaranteed to be the least under a certain accuracy of modeling, and the modeling efficiency is improved. Then this method is applied in the cookery design of an auto cooker, to show its high modeling efficiency.

Nomenclature:

[A] — transform matrix of an archetype quadratic polynomials;

λ_i , [P] — eigenvalue and eigen-matrix of [A] respectively;

$y(X)$, $y_H(X)$ —quadratic polynomials archetype and its simplified model respectively;

Q — model vector of simplified model y_H , i.e. $y_H(X) = (X^T Q)^2$;

Z_i — sample vector of sample set (matrix [Z]), i.e. Z_i is the i^{th} vector of [Z], [Z] = [Z_1, \dots, Z_n];

z — modeling variable of vector Z in normalized space S_n ; $|z| \leq 1$;

z_a — experimental variable of vector Z_a in actual value space $S_a=[z_{amin}, z_{amax}]$;

II. A QUADRATIC RSM AND ITS SIMPLIFIED MODEL

For a common quadratic RSM fitting model

$$y(X) = a_{1,1}x_1^2 + \dots + a_{n,n}x_n^2 + 2a_{1,2}x_1x_2 + \dots + 2a_{n-1,n}x_{n-1}x_n \quad (1)$$

$$= X^T [A] X$$

where: $X = \{x_1, \dots, x_n\}^T$;

Supposing that the symmetric matrix

$$[A] = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ & \ddots & \\ a_{n1} & & a_{nn} \end{bmatrix}$$

is a positive definite matrix;

$\lambda_1 > \lambda_2 > \dots > \lambda_n \geq 0$, composing a diagonal matrix $[A] = \text{diag}(\lambda_1, \dots, \lambda_n)$, are the eigenvalues of matrix $[A]$, whose corresponding eigenvectors are $[P_1, \dots, P_n] = [P]$ respectively, i.e. $[A] = [P][A][P]^T$.

Make a transformation on the quadratic model of Eq. (1) as

$$y(X) = \lambda_1(X^T P_1)^2 + \dots + \lambda_n(X^T P_n)^2 = (X^T P_1 \sqrt{\lambda_1})^2 + \dots + (X^T P_n \sqrt{\lambda_n})^2 \quad (1)$$

Thus due to $\Sigma(X^T P_i)^2 = \|X\|^2$, a quadratic function $y(X)$ can be regarded as the weighted mean of its n eigenvalues, λ_i , $i=1, \dots, n$. The weight of eigenvalue λ_i is the square of the corresponding coordinate component (projection) of vector X on P_i of the orthogonal system $[P]$. On the other hand quadratic polynomial $y(X)$ can also be regarded as a quadratic sum of projection of vector X on an orthogonal vector serial $\sqrt{\lambda_i} P_i$, $i=1, \dots, n$.

An n -order quadratic contains $(n^2+n)/2$ coefficients. If the coefficients are examined by experiments, at least the same number of experiments are required. When the n is large, the examination is almost infeasible. Therefore only the projection squares of X on one vector Q is expected to take, that is, to replace the RSM of Eq.(1) (called as archetype model) with a simplified one

$$y_H(X) = (X^T Q)^2 = X^T [QQ^T] X \quad (3)$$

to approximate an archetype $y(X)$. Only n coefficients are contained in the simplified model $y_H(X)$, in which the n -dimensional model vector Q can be determined by n experiments only.

The n trials are carried out, and the experiment sample set are $[Z]_{n \times n} = [Z_1, \dots, Z_n]$. Supposing that

$$y(Z_1) = (Z_1^T Q)^2$$

$$\vdots$$

$$y(Z_n) = (Z_n^T Q)^2$$

when $[Z]$ is nonsingular, there exist

$$Q = [Z^T]^{-1} \left\{ \sqrt{y(Z_1)}, \dots, \sqrt{y(Z_n)} \right\}^T \quad (4)$$

If the sample set is $[Z]_{n \times m} = [Z_1, \dots, Z_m]$, where the sample number m is larger than n , the dimension of examined vector, then Q is calculated by the least square method

$$Q = [ZZ^T]^{-1} [Z] \left\{ \sqrt{y(Z_1)}, \dots, \sqrt{y(Z_m)} \right\}^T$$

III. MODEL EXPECTATION

A. Expectation of Archetype Model, $E[y(X)]$

Let X be a vector with components of independent and identically distribution Gaussian random variable: $x_j \sim N(0, \sigma^2)$, $j=1, \dots, n$. P_i be a unit vector, then $X^T P_i \sim N(0, \sigma^2)$, therefore

$$E[(X^T P_i)^2] = D(X^T P_i) + [E(X^T P_i)]^2 = \sigma^2$$

From Eq.(2), it's obtained

$$E[y(X)] = \sigma^2 \Sigma \lambda \quad (5)$$

B. Expectation of Simplified Model, $E[y_H(X)]$

The transformation matrix of the simplified model obtained by the sample set $[Z_1, \dots, Z_n] = [Z]$ is $[QQ^T]$. The characteristic equation of matrix $[QQ^T]$ is

$$|QQ^T - \lambda[I]| = \left| [Z^T]^{-1} \begin{Bmatrix} y(Z_1) \\ \vdots \\ y(Z_n) \end{Bmatrix} \{y(Z_1), \dots, y(Z_n)\} [Z]^{-1} - \lambda[I] \right| = 0$$

$$\text{From } \left\{ \begin{Bmatrix} \sqrt{y(Z_1)} \\ \vdots \\ \sqrt{y(Z_n)} \end{Bmatrix} \left\{ \sqrt{y(Z_1)}, \dots, \sqrt{y(Z_n)} \right\} - \lambda [Z]^T [Z] \right\} = 0, \text{ it}$$

is obvious that if $[Z]^T [Z] = [I]$ (i.e. $[Z]$ is orthogonal) the eigenvalues of $[QQ^T]$ are $\{\Sigma y, 0, \dots, 0\}$.

When $[Z_1, \dots, Z_n] = [Z]$ is a unit orthogonal basis, the sum of the archetype function of an arbitrary orthogonal system, e.g. $[Z]$, is

$$y(Z_1) + \dots + y(Z_n) = \left[\lambda_1 (Z_1^T P_1)^2 + \dots + \lambda_n (Z_n^T P_n)^2 \right] + \dots$$

$$+ \left[\lambda_1 (Z_n^T P_1)^2 + \dots + \lambda_n (Z_n^T P_n)^2 \right] = \sum_i \lambda_i$$

So the eigenvalues of $[QQ^T]$ are $\{\Sigma\lambda, 0, \dots, 0\}$, and $\|Q\|^2 = Q^T Q = \Sigma y = \Sigma \lambda$, which means that the eigenvalues of the simplified model are determined by its archetype one. Then it is obtained that $X^T Q \sim N(0, \Sigma \lambda \sigma^2)$, and that

$$E[y_H(X)] = D(X^T Q) + [E(X^T Q)]^2 = \Sigma \lambda \sigma^2 \quad (6)$$

A simplified model has the same eigenvalues as those of its archetype one if the sample set matrix $[Z]$ is orthogonal.

C. The Function Values of Simplified Model at a Sample Point, $y_H(Z_i)$

If sample set $[Z]$ is orthogonal, then at any sample vector Z_i , there exists

$$y_H(Z_i) = y(Z_i) \quad (7)$$

Eq.(7) indicates that a simplified model can obtain high precision near the sample vectors. This feature guarantees the accuracy of a simplified model (a closed approximation to its archetype $y_H(X)$) in local domain of sampling area.

Let P_0 be the eigenvector corresponding to the nonzero eigenvalues, $\Sigma\lambda$, of $[QQ^T]$. From Eq.(2) and Eq.(3), it's obtained

$$y_H(X) = (X^T Q)^2 = \Sigma \lambda (X^T P_0)^2 = (X^T \sqrt{\Sigma \lambda} P_0)^2,$$

and $P_0 = Q / (\Sigma \lambda)^{0.5}$, i.e. P_0 is a unit vector. Therefore the simplified model can be regarded as an average of equal weight $(X^T P_0)^2$.

IV. ANALYSIS ON THE INFLUENCE FACTORS OF THE SIMPLIFIED MODEL ERROR

Based on the quadratic form archetype, the error is defined as: $y - y_H$. It is easy to see that the error expectation is $E(y - y_H) = 0$. And the variance is

$$D(y - y_H) = E(y - y_H)^2 - [E(y - y_H)]^2 = E(y^2) + E(y_H^2) - 2E(y y_H).$$

The variance of the error $y - y_H$ is the function $D(\sigma)$. The larger the variance is, the more easily the error is dispersed, which means that the precision and reliability of the simplified model are worse. The calculation of variance polynomial is as follows:

A. Calculation of $E(y^2)$

From Eq.(2), there exists

$$y^2 = [\lambda_1 (X^T P_1)^2 + \dots + \lambda_n (X^T P_n)^2] \times [\lambda_1 (X^T P_1)^2 + \dots + \lambda_n (X^T P_n)^2]$$

and

$$E[(X^T P_j)^2 (X^T P_k)^2] = \begin{cases} \sigma^4 & j \neq k \\ 3\sigma^4 & j = k \end{cases}$$

therefore

$$E[y^2(X)] = 3\sigma^4 \lambda_1^2 + \sigma^4 \lambda_1 \lambda_2 \dots + \sigma^4 \lambda_1 \lambda_n + \dots + \sigma^4 \lambda_n \lambda_1 + \dots$$

$$+ \sigma^4 \lambda_{n-1} \lambda_n + 3\sigma^4 \lambda_n^2 = \sigma^4 (\Sigma \lambda)^2 + 2\sigma^4 \Sigma \lambda^2 \quad (8)$$

and

$$D(y) = E(y^2) - [E(y)]^2 = 2\sigma^4 \Sigma \lambda^2$$

B. The Calculation of $E(y y_H)$

From $E[(X^T P_i)^2 (X^T Q)^2] = \sigma^4 [\Sigma \lambda + 2(P_i^T Q)^2]$, there exists

$$E(y y_H) = E[\lambda_1 (X^T P_1)^2 (X^T Q)^2 + \dots + \lambda_n (X^T P_n)^2 (X^T Q)^2]$$

$$= \sigma^4 [(\Sigma \lambda)^2 + 2\Sigma (P_i^T Q)^2 \lambda_i] \quad (10)$$

In summary Eq.(8) ~ (10), the variance of the error between the simplified model and the archetype is

$$D(y - y_H) = 2\sigma^4 [(\Sigma \lambda)^2 + \Sigma \lambda^2 - 2\Sigma (P_i^T Q)^2 \lambda_i] \quad (11)$$

According to Eq.(2)~(4), it's known that $\lambda_i = (P_i^T Q_{[P]})^2$, and that

$$\Sigma (P_i^T Q)^2 \lambda_i = \{(P_1^T Q_{[Z]})^2, \dots, (P_n^T Q_{[Z]})^2\} \cdot \{(P_1^T Q_{[P]})^2, \dots, (P_n^T Q_{[P]})^2\}$$

where:

$Q_{[Z]}$, $Q_{[P]}$ symbol the simplified model vectors from sample set $[Z]$ and $[P]$ respectively, while $Q_{[Z]}$ also symbols Q mentioned above.

It can be seen that the closer the sample set $[Z]$ is to $[P]$, the closer simplified model $Q_{[Z]}$ is to simplified model $Q_{[P]}$, and the larger the inner product $\Sigma (P_i^T Q)^2 \lambda_i$ is, and also the smaller the variance $D(y - y_H)$ is; When $[Z] = [P]$, the inner product is the largest and the variance is minimum: $D(y - y_H) = 2\sigma^4 [(\Sigma \lambda)^2 - \Sigma \lambda^2]$; Furthermore if λ_1 is much larger than other eigenvalues, then $D(y - y_H) \rightarrow 0$.

V. TRANSFORMATION BETWEEN MODEL VARIABLES AND EXPERIMENT VARIABLES

The above analysis shows that the experiment sample set $[Z]$ must be orthogonal. However, as the components of a sample vector Z_i , any experiment variables such as z_a generally cannot be actually taken as negative values in actual experiments, and the values of each variable may vary greatly. It is difficult to implement an orthogonal sample set, so

normalization transformation on $[Z]$ must be carried out. This paper uses the following transformation method:

A. Experiment Variable and Modeling Variable

An experiment variable, symbolized by z_a , corresponds to experiment actual value in actual value space, S_a ; while a modeling variable, symbolized by z , is used in normalized space, S_n : $|z| \leq 1$, for modeling. There is a transformation relationship between them:

$$z = (z_a - \mu) / s \tag{12}$$

where: μ and s are the expectation and mean squared deviation (MSD) of the experimental variables respectively. When μ and s are unknown, $(z_{amax} + z_{amin}) / 2$ and radius $(z_{amax} - z_{amin}) / 2$ are used as μ and s respectively, between the actual value range, $S_a = [z_{amin}, z_{amax}]$, of an experimental variable;

B. Constructed an Orthogonal Matrix $[Z]$ of Model Variable in Normalized Space

Then according to Eq. (12) mapped it in actual value space into a sample set $[Z_a]$, used for experiments. So a simplified model as Eq.(3) can be established based on the variable matrix $[Z]$ and the experimental results $y([Z_a])$;

C. The Actual Variables be Transformed into Model Variables with Eq.(12) in the Calculation of a Function Value of Simplified RSM Model

VI. AN EXAMPLE

There are 2 quadratic archetypes, a and b , in form of $[P][A][P]^T$: both eigenvectors are

$$[P] = \begin{bmatrix} 0 & 1 & 0 \\ 0.707 & 0 & 0.707 \\ -0.707 & 0 & 0.707 \end{bmatrix},$$

while $[A_a] = \text{diag}(33, 2, 1)$, $[A_b] = \text{diag}(25, 6, 5)$, in which the eigenvalue distribution of the former is more concentrated than that of the latter.

Two orthogonal sample sets of normalized model variables are used to estimate the archetype with a simplified model. The first,

$$[Z1] = \begin{bmatrix} 0.14 & 1 & 0.14 \\ 0.7 & 0 & 0.7 \\ -0.7 & 0 & 0.7 \end{bmatrix},$$

which comes from the rotation around Z, Y and X coordinate axes $\pi/60, \pi/45, \pi/30$ respectively, of $[P]$.

The second,

$$[Z2] = \begin{bmatrix} 0.301 & 0.707 & -0.64 \\ 0.905 & 0 & 0.426 \\ 0.301 & -0.707 & -0.64 \end{bmatrix},$$

coming from the rotation around Z, Y and X coordinate axes $\pi/22, \pi/21, \pi/20$ respectively, of $[P]$. So $[Z1]$ is more closed to $[P]$ than $[Z2]$.

For archetype a , two simplified models from sample sets $[Z1]$ and $[Z2]$ are: $Q_{a1} = \{-0.5310, 4.5390, 3.8874\}$, and $Q_{a2} = \{-1.4841, 3.9287, 4.2848\}$;

While for archetype b , the 2 simplified models from the 2 sample sets are $Q_{b1} = \{-0.5310, 4.5390, 3.8874\}^T$ and $Q_{b2} = \{-1.4841, 3.9287, 4.2848\}^T$ respectively.

Four vectors, P_1 from $[P]$, $Z1_1$ from $[Z1]$, $Z2_1$ from $[Z2]$ each and $(Z1_1 + Z2_1) / 2$ are used as testing points (representing the domains of near the feature points, sample points and common point respectively) to evaluate the simplified models. The function values of the 2 archetype $y(X)$, and the corresponding simplified models $y_H(X)$ are calculated, and the results are shown in table 1 and Table 2 respectively. Some statistics (mean and variance) of model error ($e = y - y_H$), \bar{e} and S^2 , calculated below the 2 tables respectively, are used to evaluate the accuracy and reliability of the simplified models.

TABLE I. THE VALUES OF MODELS AT THE 4 TESTING POINTS

X	$X^T A_a X$	$X^T Q_{a1} Q_{a1}^T X$	$X^T Q_{a2} Q_{a2}^T X$
P_1	33	29.1741	21.6398
$Z1_1$	32.1804	32.1804	26.6792
$Z2_1$	29.6327	32.1842	29.6327
$(Z1_1 + Z2_1) / 2$	30.8832	32.1823	28.1369

$$\bar{e}_1 = -0.006175, S_1^2 = 5.7103; \bar{e}_2 = 4.9019, S_2^2 = 17.6860.$$

TABLE II. THE VALUES OF MODELS AT THE 4 TESTING POINTS

X	$X^T A_b X$	$X^T Q_{b1} Q_{b1}^T X$	$X^T Q_{b2} Q_{b2}^T X$
P_1	25	20.6025	15.4348
$Z1_1$	24.4878	24.4878	20.2813
$Z2_1$	22.8989	25.2828	22.8989
$(Z1_1 + Z2_1) / 2$	23.6436	24.8837	21.5701

$$\bar{e}_1 = 0.1934, S1^2 = 6.6023; \bar{e}_2 = 3.9613, S2^2 = 12.6799$$

This example has shown that both archetype themselves and sample sets have influences on the accuracy and reliability of simplified model. Three conclusions can be drawn from Table 1 and Table 2:

(1) The errors of the simplified models near the sample points are very small, even 0. And it is possibly to have higher

accuracy in the domain near the sample points than near the eigenvector points;

(2) Since the sample set [Z1] is closer to the eigenvector group [P] than [Z2], the corresponding simplified model Q_1 have higher accuracy than Q_2 around eigenvectors, and domains near their corresponding sample points;

(3) The eigenvalue distribution of archetype a is more concentrated than that of b , so the variance of the simplified model of a derived from the sample set [Z1] near [P] is smaller than that of b , that is, the simplified model of a is more reliable. However, this may be false for such testing points as [Z2], which is far away from [P].

VII. CONCLUSION

Compared with its quadratic archetype, a simplified modeling greatly decrease experiment times while a model is of certain precision. When the dimension of experiment variable vector is high, high modeling efficiency is obtained. Both the eigenvalue distribution of archetype and samples have effects on modeling precision. Because Eq.3 is positive definite, a precise modeling of this method applies to an positive definite archetype with eigenvalues of concentrated distribution, and requires that sample groups be orthogonal in normalized space S_n . And that the closer the sample group is to the eigenvector group [P] of archetype, the higher is the precision of the modeling. Eigenvector group [P] of an archetype can be substituted by that of similar model or model from experiments for a DOE because it's generally unknown actually. A simplification model can give enough precision when it is used to solve an optimization.

ACKNOWLEDGMENT

This research is supported by: (1) Science Foundation of Guangxi Education Department (No.2013LX015); (2) Scientific Research Project of Guangxi Education Department (No.20130351). (3) Application Research Project of Guangxi University (No.20121059). (4) Application Research Project of Guangxi University (No.20121720).

REFERENCES

- [1] Abul Kalam M, Khan A A, Khan S, et al. Optimizing indomethacin-loaded chitosan nanoparticle size, encapsulation, and release using Box-Behnken experimental design.[J]. International journal of biological macromolecules, 2016, 87: 329-340.
- [2] LI Chang, Han Xing. Analysis of reliability sensitivity for gear engagement based on response surface methods [J]. Journal of Aerospace Power, Mar. 2011, 26(3): 711-715.
- [3] LI Chang, Han Xing. Analysis of reliability sensitivity for gear engagement based on response surface methods [J]. Journal of Aerospace Power, Mar. 2011, 26(3): 711-715.
- [4] Guan X L, Melchers R E. Effect of response surface parameter variation on structural reliability estimates [J]. Structural Safety, 2001, 23: 429-444.
- [5] ZHANG Z, LI S-Y, TENG QJ, An improved response surface method for structural reliability analysis [J]. Journal of Dalian University of Technology, Jan. 2007, 47(1):57-60.
- [6] Asfaram A, Ghaedi M, Agarwal S, et al. Removal of basic dye Auramine-O by ZnS: Cu nanoparticles loaded on activated carbon: optimization of parameters using response surface methodology with central composite design[J]. RSC ADVANCES, 2015, 5(24): 18438-18450.
- [7] YAN M, SUN ZL, YANG Q. ANALYSIS METHOD OF RELIABILITY SENSITIVITY BASED ON RESPONSE SURFACE METHODS[J]. Journal of Mechanical Engineering, Oct. 2007, 43(10): 67-71.
- [8] WEI Jinwen, etc. Recognizing Data Fitting Based on the Robustness of Generalized Cross Product [J]. Journal of Mechanical Engineering, Jul. 2011, 47(14): 7-12.