

# Modeling and design of experiments

## Chapitre 5: Surface response

Dr Jean-Marie Fürbringer

École Polytechnique Fédérale de Lausanne

Fall 2025

## Surface response

1. Lack of fit
2. Classical response surface designs
3. Canonical analysis

Surface response      Lack of Fit  
Classical designs  
Canonical analysis

5.2 Is a quadratic function necessary?

Surface response      Lack of Fit  
Classical designs  
Canonical analysis

5.3 Central composite design ( $N_0 = 1, \alpha = 1.22$ )

Surface response      Lack of Fit  
Classical designs  
Canonical analysis

5.14 Isosurfaces of a quadratic function

Ellipsoid

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = y$$

Hyperboloid

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = \pm y$$

Dr. Jean-Marie Fürbringer      Modeling and design of experiments

# 5.1 Lack of Fit

## 5.1.1 Two objectives for performing experiments

1. Quantify the effects of the factors  $x_i$  on the response  $y$  in a minimal number of experiments to :
  - ▶ Select significant factors
  - ▶ Perform a Pareto analysis (sort effects by order of importance)

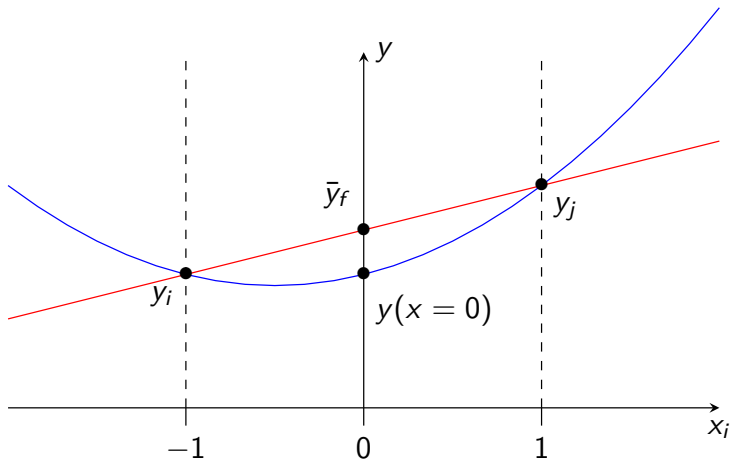
For this objective a first degree model with or without interactions is sufficient :

$$a_0 + \sum_{i=1}^n a_i x_i + \sum_{i < j}^n a_{ij} x_i x_j$$

2. Determine the combination of the factors that allows us to optimize the response, also with a minimal number of experiments  
For this objective a quadratic model is necessary :

$$a_0 + \sum_{i=1}^n a_i x_i + \sum_{i \leq j}^n a_{ij} x_i x_j$$

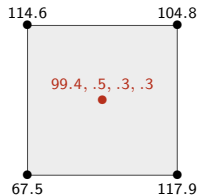
## 5.1.2 Is a quadratic function necessary ?



## 5.1.3 Test the curvature with a central point

- ▶ An experimental situation with two factors  $x_1$  and  $x_2$ , 4 factorial measurements,  $y_f(i)$ , 4 measurements at the center of the experimental space,  $y_c(j)$
- ▶ Is the linear model with interactions  $y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2$  sufficient ?

Run	Factorial	Center
1	67.5	99.4
2	114.6	99.5
3	117.9	99.3
4	104.8	99.3
$\bar{y}$	101.2	99.4
$s^2$		0.009
$\delta$	1.83	



Tester  $H_0 : \alpha_{11} + \alpha_{22} + \dots = 0$

$$t_o = \frac{\bar{y}_f - \bar{y}_c}{\sqrt{s^2 \left( \frac{1}{n_f} + \frac{1}{n_c} \right)}}$$

if  $|t_o| > t_{\alpha/2, n_c - 1}$  then  $H_0$  is rejected.

In the present case :  
 $t_o \approx 27 > t_{0.025, 3} \approx 3.18$

## 5.1.4 With Matlab

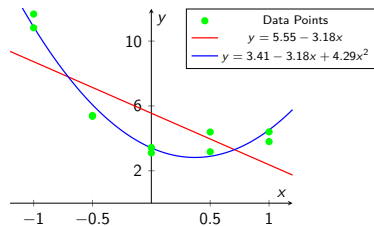
If there are sufficient degrees of freedom (replicates and more than the minimal scheme), the routine `anova` can compute the lack of fit. `anova mdl2, 'summary'`

	SumSq	DF	MeanSq	F	pValue
<b>Total</b>	1613.8	7	230.54		
<b>Model</b>	1607.1	3	535.7	320.36	3.2166e-05
<b>. Linear</b>	701.09	2	350.55	209.63	8.9309e-05
<b>. Nonlinear</b>	906.01	1	906.01	541.81	2.019e-05
<b>Residual</b>	6.6888	4	1.6722		
<b>. Lack of fit</b>	6.6613	1	6.6613	726.68	0.00011202
<b>. Pure error</b>	0.0275	3	0.0091667		

## 5.1.5 Is there a lack of fit ?

Situation : 10 data points  $y_i$  have been acquired for 5 values of  $x$ .  
Is this data better presented by a linear response or a quadratic model ?

	$x$	$y$
1	-1	11.67
2	-1	10.82
3	-0.5	5.41
4	-0.5	5.36
5	0	3.10
6	0	3.43
7	0.5	3.17
8	0.5	3.39
9	1	4.40
10	1	3.80



## 5.1.6 Estimates, residues and sums of squares

$i$	$y$	$\bar{y}$	$y - \bar{y}$	$\hat{y}_1$	$\epsilon_1$	$\hat{y}_1 - \bar{y}$	$\hat{y}_2$	$\epsilon_2$	$\hat{y}_2 - \bar{y}$
1	9.5	10.6	-1.08	9.4	0.12	1.2	10.7	-1.2	-0.12
1	11.7	10.6	1.08	9.4	2.29	1.2	10.7	0.97	-0.12
2	7.9	7.7	0.25	7.7	0.23	-0.02	7.0	0.89	0.64
2	7.4	7.7	-0.25	7.7	-0.28	-0.02	7.0	0.38	0.64
3	3.0	3.5	-0.41	6.0	-2.93	-2.53	4.7	-1.61	-1.21
3	3.9	3.5	0.41	6.0	-2.12	-2.53	4.7	-0.8	-1.21
4	4.9	4.6	0.36	4.3	0.67	0.31	3.6	1.33	0.97
4	4.2	4.6	-0.36	4.3	-0.04	0.31	3.6	0.62	0.97
5	3.0	3.6	-0.6	2.6	0.44	1.03	3.9	-0.88	-0.29
5	4.2	3.6	0.6	2.6	1.63	1.03	3.9	0.31	-0.29
SS	436.8	433.0	3.8	415.0	21.7	18.0	427.2	9.6	5.8

$$SS_{PE} = \sum_i (y_{ij} - \bar{y}_i)^2$$

$$SS_{LOF} = \sum_i (\hat{y}_i - \bar{y}_i)^2$$

## 5.1.7 Differentiate first degree to quadratic model

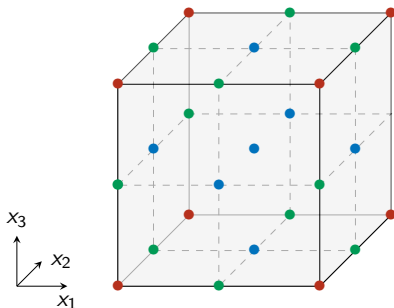
Source	ss	df	ms	F	p
$a_o$	356.8	1			
Linear	58.3	1	58.3	21.4	0.0017
$\epsilon_1$	21.7	8	2.7		
LoF	18.0	3	6.0	7.9	0.02
Pure error	3.8	5	0.75		

Source	ss	df	ms	F	p
$a_o$	216.6	1			
Quadratic	210.6	2	105.3	77	0.000017
$\epsilon_2$	9.57	7	1.37		
LoF	5.8	2	2.9	3.9	0.1
Pure error	3.8	5	0.75		

## 5.2 Classical designs

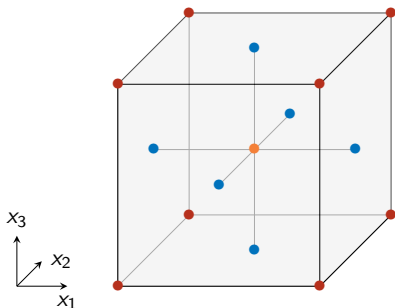
- ▶ Factorial  $3^k$  designs
- ▶ Composite design
- ▶ Doehlert design
- ▶ Box-Behnken

## 5.2.1 The $3^k$ designs



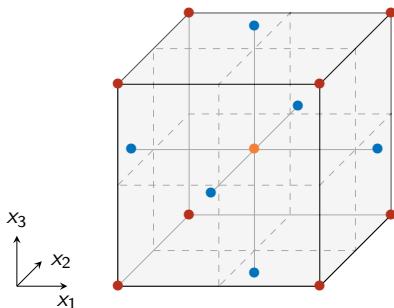
3 factors : 27 measurement points , 3 levels per factor

## 5.2.2 The central composite design ( $\alpha = 1$ )



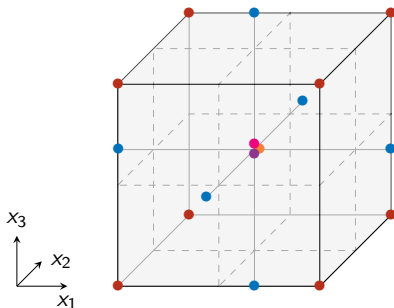
3 factors : 15 measurement points , 3 levels per factors

## 5.2.3 Central composite design ( $N_o = 15$ , $\alpha = 1.22$ )



3 factors : 15 measurement points , 5 levels per factors

## 5.2.4 Central composite design ( $N_o = 3, \alpha = 1.353$ )



3 factors : 15 measurement points , 3 levels per factors

## 5.2.5 Optimisation of the radius $\alpha$

Trade off between *isovariance* per rotation and *pseudo-orthogonality*

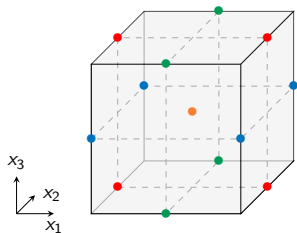
- ▶ The isovariance per rotation (rotability) :  

$$\text{var}_y(x_1, x_2, \dots) = \text{var}_y(\sqrt{x_1^2 + x_2^2 + \dots})$$
- ▶ The pseudo-orthogonality : property limiting the number of terms in the matrix of dispersion (improve the accuracy of the estimated coefficients).
- ▶ Matlab :  $E = \text{ccdesign}(n)$

Nb of factors ( $n$ )	2	3	4	5	5	6	6
Factorial design	$2^2$	$2^3$	$2^4$	$2^{5-1}$	$2^5$	$2^{6-1}$	$2^6$
Nb fact exp ( $2^{n-k}$ )	4	8	16	16	32	32	64
Nbr star pts ( $2n$ )	4	6	8	10	10	12	12
Nbr central pts ( $n_o$ )	1-3	1-3	1-3	1-3	1-3	1-3	1-3
Total ( $2^{n-k} + 2n + N_o$ )	9-11	15-17	25-27	27-29	43-45	45-45	77-79
$\alpha$ si $n_o = 1$	1	1.22	1.41	1.55	1.60	1.72	1.76
$\alpha$ si $n_o = 2$	1.08	1.29	1.48	1.61	1.66	1.78	1.82
$\alpha$ si $n_o = 3$	1.15	1.35	1.55	1.66	1.72	1.83	1.89

## 5.2.6 Box-Behnken design

$$E = \begin{pmatrix} 0 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \\ -1 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

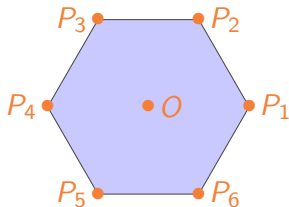


- ▶ 3 levels per factor
- ▶ No experiments at the vertices
- ▶ Isovariant per rotation : 4 ou 7 factors
- ▶ Blocking : factorial design  $2^2$
- ▶ Matlab :  $E = \text{bbdesign}(n)$

Factors	Coefficients	Run
3	10	13
4	15	25
5	21	41
6	28	49
7	36	57

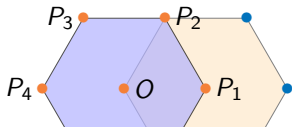
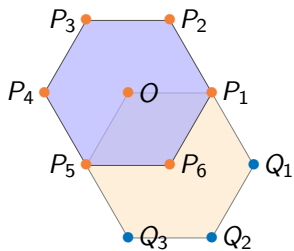
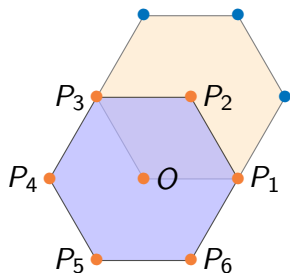
## 5.2.7 2D Doehlert design

$$E = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -1 & 0 \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} & -\frac{\sqrt{3}}{2} \end{pmatrix}$$

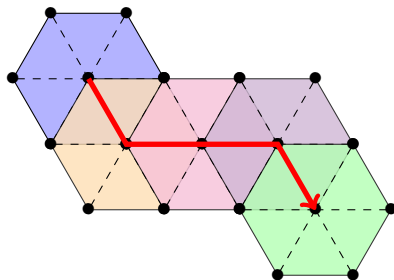


- ▶ 2 factors : 7 measurement points, 3 and 5 levels per factor,
- ▶ No experiments at the vertices
- ▶ Isovariant per rotation
- ▶ Matlab :  $E = \text{doehlert}(n)$  to download from Moodle

## 5.2.8 Shift of a 2D Doehlert design



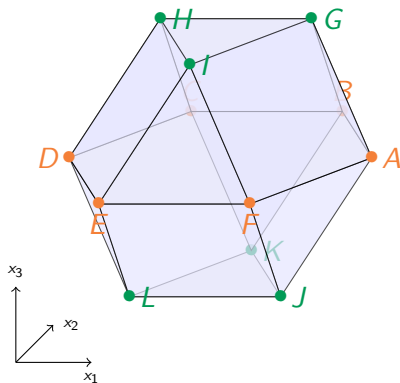
## 5.2.9 Hexagonal simplex



- Sequential exploration of a domain ( $7+3+3+3+3=19$  exp)

## 5.2.10 3D Doehlert design

$$E = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \\ -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ -\frac{1}{2} & -\frac{\sqrt{3}}{2} & -\frac{\sqrt{2}\sqrt{3}}{3} \\ \frac{1}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{2}\sqrt{3}}{3} \\ -\frac{1}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{2}\sqrt{3}}{3} \\ \frac{1}{2} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{2}\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}\sqrt{3}}{3} \\ 0 & \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}\sqrt{3}}{3} \end{pmatrix}$$



## 5.2.11 Sequentiality with Doehlert designs

Number of runs :  
 $N_{exp} = k^2 + k + 1$

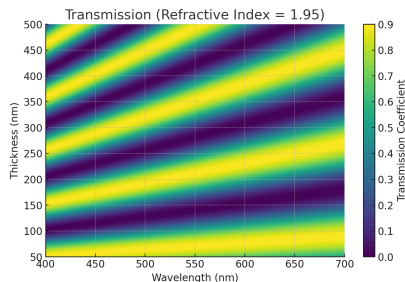
$$E = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & -\frac{\sqrt{3}}{2} & -\frac{\sqrt{2}\sqrt{3}}{3} & 0 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & \frac{\sqrt{2}\sqrt{3}}{3} & 0 & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & \frac{\sqrt{2}\sqrt{3}}{3} & 0 & 0 \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} & -\frac{\sqrt{2}\sqrt{3}}{3} & 0 & 0 \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}\sqrt{3}}{3} & 0 & 0 \\ 0 & \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}\sqrt{3}}{3} & 0 & 0 \\ -\frac{1}{2} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{6}}{12} & -\frac{\sqrt{2}\sqrt{5}}{4} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{12} & \frac{\sqrt{2}\sqrt{5}}{4} & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{12} & \frac{\sqrt{2}\sqrt{5}}{4} & 0 \\ \frac{1}{2} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{6}}{12} & -\frac{\sqrt{2}\sqrt{5}}{4} & 0 \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{12} & \frac{\sqrt{2}\sqrt{5}}{4} & 0 \\ 0 & \frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{12} & -\frac{\sqrt{2}\sqrt{5}}{4} & 0 \\ 0 & 0 & -\frac{\sqrt{2}\sqrt{3}}{4} & \frac{\sqrt{2}\sqrt{5}}{4} & 0 \\ 0 & 0 & \frac{\sqrt{2}\sqrt{3}}{4} & -\frac{\sqrt{2}\sqrt{5}}{4} & 0 \\ -\frac{1}{2} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{6}}{12} & -\frac{\sqrt{10}}{20} & -\frac{\sqrt{3}\sqrt{5}}{5} \\ \frac{1}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{12} & \frac{\sqrt{10}}{20} & \frac{\sqrt{3}\sqrt{5}}{5} \\ -\frac{1}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{12} & \frac{\sqrt{10}}{20} & \frac{\sqrt{3}\sqrt{5}}{5} \\ \frac{1}{2} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{6}}{12} & -\frac{\sqrt{10}}{20} & -\frac{\sqrt{3}\sqrt{5}}{5} \end{pmatrix}$$

## 5.2.12 Application : light transmission

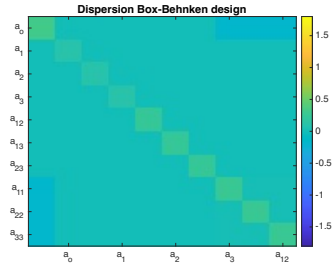
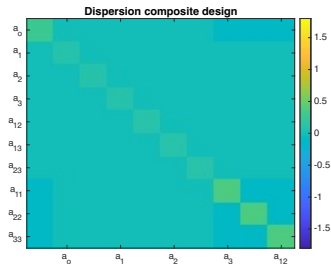
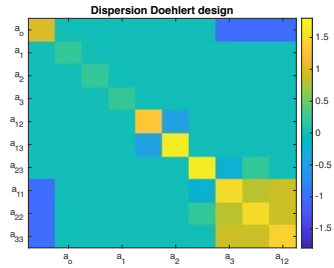
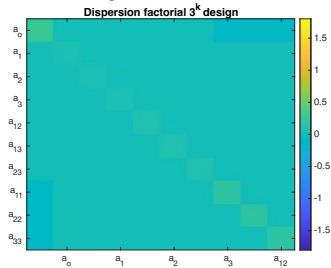
- ▶ The transmission or reflection of light by a thin film is a complex phenomenon.
- ▶ Based on Snell and Fresnel equations, for a perpendicular non-polarized beam, the transmission coefficient  $T$  giving the fraction of intensity which is transmitted is function of the wave length  $\lambda$ , of the refractive index  $n$  and the film thickness  $t$  such as

$$T = \left( 1 - \left( \frac{n-1}{n+1} \right)^2 \right) \sin^2 \left( \frac{2\pi n t}{\lambda} \right)$$

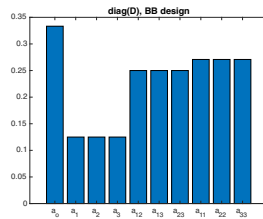
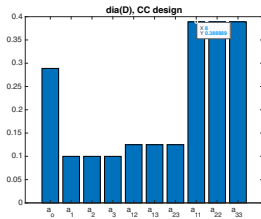
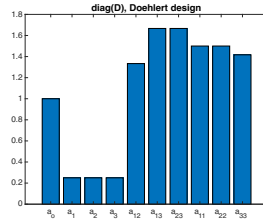
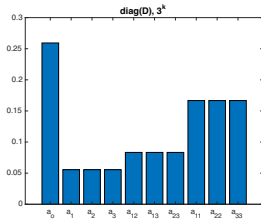
- ▶ The objective is to experimentally determine  $T$  around the point ( $\lambda = 475 \text{ nm}$  (blue),  $t = 425 \text{ nm}$ ,  $n = 1.95$ )



## 5.2.13 Dispersion matrices

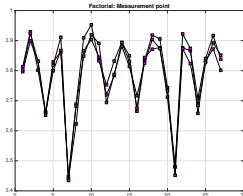


## 5.2.14 Diagonal elements of dispersion

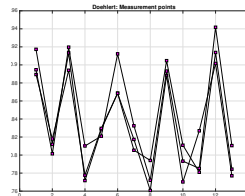


## 5.2.15 Data points

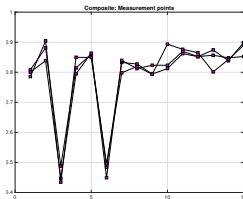
### Factorial $3^k$



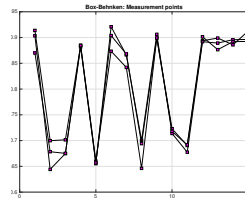
### Doehlert



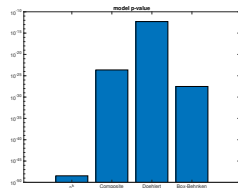
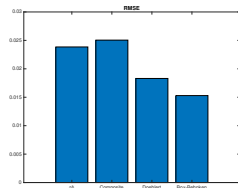
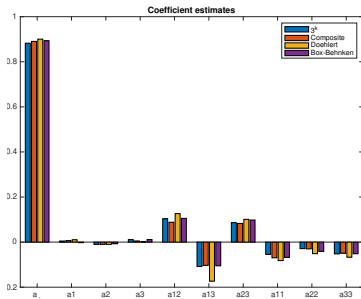
### Composite



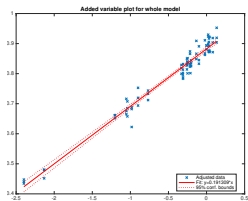
### Box-Behnken



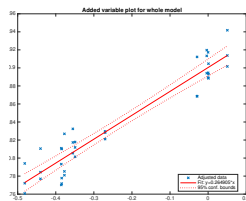
## 5.2.16 Fit



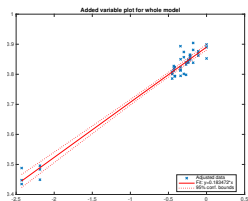
## 5.2.17 Plot added

Factorial  $3^k$ 

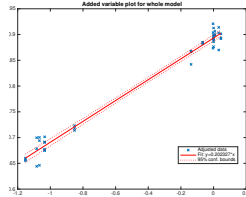
Doehlert



Composite



Box-Behnken



## 5.3 Canonical analysis

## 5.3.1 Geometry of the second degree

- ▶ The function  $a_o + \sum_{i=1}^n a_i x_i + \sum_{i \leq j}^n a_{ij} x_i x_j$  can be written as

$$y = a_o + (x_1, \dots, x_n) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + (x_1, \dots, x_n) \begin{pmatrix} a_{11} & & \frac{1}{2} a_{1n} \\ & \ddots & \\ \frac{1}{2} a_{1n} & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

- ▶ Equivalent to :

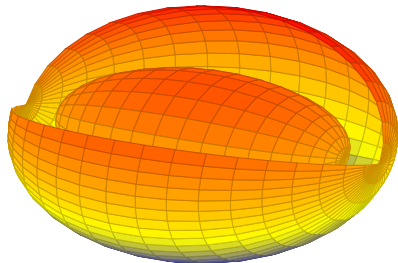
$$y = a_o + \vec{x} \cdot \vec{a} + \vec{x}^T A \vec{x}$$

- ▶ The *isosurfaces* of such a function are ellipsoids, or hyperboloids

## 5.3.2 Isosurfaces of a quadratic function

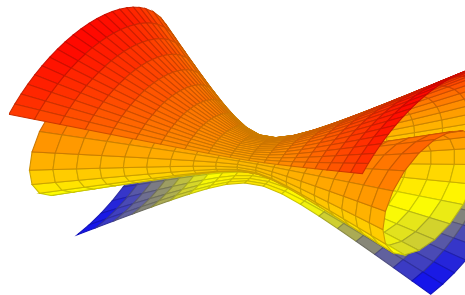
Ellipsoid

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = y$$



Hyperboloid

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = \pm y$$



## 5.3.3 Canonical analysis - fix point

- ▶ The center of the figure and the orientation of the axes, as well as the ratio of the axes is not known a priori!
- ▶ The canonical analysis consists in determining those informations
- ▶ First identify the center of the figure that can be an extremum or a saddle point
- ▶ We look for a point defined by  $\nabla y = 0$

$$\frac{\partial y}{\partial x_i} = a_i + a_{1i}x_1 + \dots + 2a_{ii}x_i + \dots + a_{in}x_n = 0$$

$$0 = \vec{a} + 2A\vec{x}$$

$$\vec{x}_s = -\frac{1}{2}A^{-1}\vec{a}$$

$$y_s = a_0 + \vec{x}_s \cdot \vec{a} + \vec{x}_s^T A \vec{x}_s$$

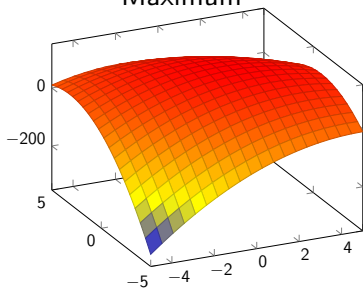
## 5.3.4 Canonical analysis - main axes

- ▶ Main axes are the eigen vectors  $A$ ,  $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$
- ▶ The increase of the function  $y = f(\vec{x})$  in the direction corresponding to the main axes is given by the eigen values,  $\lambda_1, \lambda_2, \lambda_3$
- ▶ In the directions where the eigen values are bigger, the contour lines are close to each one
- ▶ The function  $y$  can be re-written in a canonical form

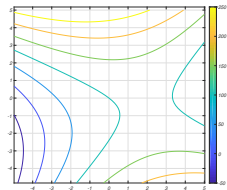
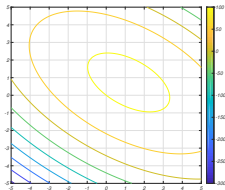
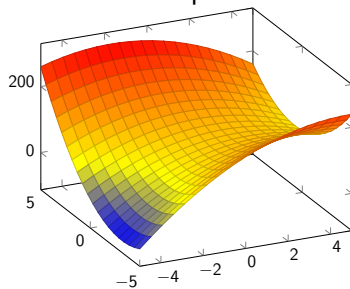
$$y = y_s + \sum_{i=1}^n \lambda_i \tilde{X}_i^2$$

- ▶ If all the eigen values have the same sign the figure is an ellipsoid, in the opposite case the figure is an hyperboloid

Maximum



Saddle point



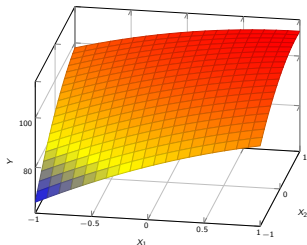
## 5.3.5 Canonical analysis - example

- ▶ Original model

$$y = 100 + 10x_1 + 12x_2 - 4x_1x_2 - 3x_1^2 - 5x_2^2$$

- ▶ Fix point determination

$$A = \begin{pmatrix} -3 & -2 \\ -2 & -5 \end{pmatrix} \quad \vec{a} = \begin{pmatrix} 10 \\ 12 \end{pmatrix} \Rightarrow \vec{x}_s = \begin{pmatrix} 1.18 \\ 0.73 \end{pmatrix}$$

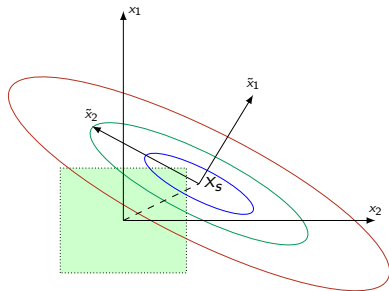


- ▶ Eigen values and eigen vectors

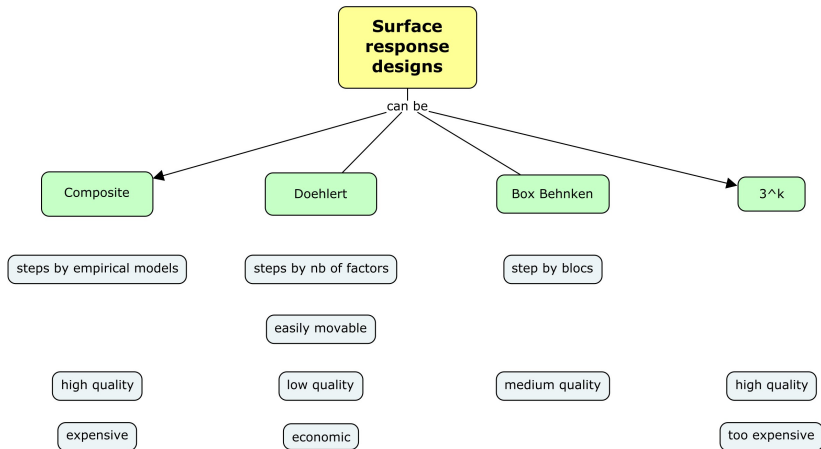
$$\begin{cases} \lambda_1 = -6.2 \\ \lambda_2 = -1.8 \end{cases} \quad \text{et} \quad \begin{cases} \tilde{x}_1 = 0.53\hat{x}_1 + 0.85\hat{x}_2 \\ \tilde{x}_2 = -.85\hat{x}_1 + 0.53\hat{x}_2 \end{cases}$$

- ▶ Canonical model

$$\tilde{y} = 110.3 - 6.2 \tilde{x}_1^2 - 1.8 \tilde{x}_2^2$$



## 5.3.6 Surface response designs : pro & cons



## 5.3.7 Conclusion

- ▶ Surface response methodology enables the efficient exploration and optimization of complex systems using mathematical modeling and experimental design.
- ▶ A quadratic model is often necessary to capture curvature in the response, and lack-of-fit tests (including central point analysis) are essential to assess model adequacy.
- ▶ Classical designs, such as  $3^k$ , Central Composite, Doehlert, and Box-Behnken, offer structured experimental frameworks that balance cost, information richness, and modeling power.
- ▶ Canonical analysis reveals key geometrical properties of the response surface, including the location and nature (maximum, minimum, saddle) of critical points.
- ▶ The choice of design should be informed by factors such as the number of variables, desired accuracy, experimental constraints, and software capabilities (e.g., MATLAB).
- ▶ Mastery of surface response techniques equips practitioners with a powerful toolbox for data-driven decision-making and process optimization in engineering and applied sciences.