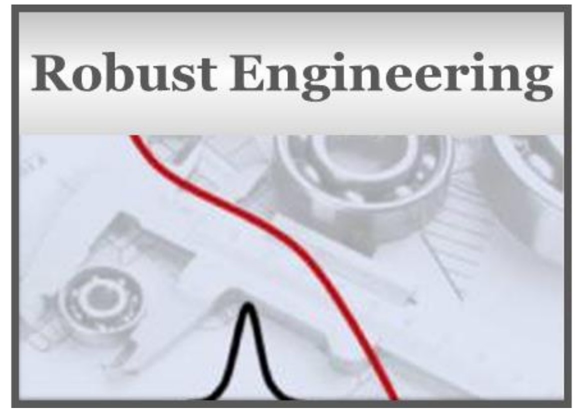
UNIVERSITY of BRADFORD

Statistiques Appliquées Formation Industrielle



Session 9: Developing A Multi-Level Design

In This Session We Will…

- Show how an experiment can be divided into blocks, using an extended model matrix.
- Discuss the idea of randomizing the run order.
- Return to Central Composite designs and look at some criteria for;
  - choosing the number of runs at the centre.
  - placing the axial points.

Blocks

- In Session 5 we introduced the idea of running an experiment in blocks
  - the blocks might be different days, different machines, different batches of material, etc.
- We gave this example of a 2^3 design run in two blocks:

Block	x_1	x_2	x_3
1	–1	–1	–1
1	+1	–1	+1
1	–1	+1	+1
1	+1	+1	–1
2	–1	–1	+1
2	+1	–1	–1
2	–1	+1	–1
2	+1	+1	+1

We said that this pattern could be worked out 'on the back of an envelope' but other cases are more challenging

Constructing A Design In Two Blocks

The steps:

- Extend the model matrix M to include a two-level blocking factor x_b.
- Re-allocate rows to blocks to make x_b as orthogonal as possible to the other columns of M.

Exact orthogonality is the ideal, but not usually achievable

Note: the method can be extended to more than 2 blocks

Example: Face-centred CC Design, Three Factors

Design

x_1	x_2	x_3
-1	-1	-1
+1	-1	-1
-1	+1	-1
+1	+1	-1
-1	-1	+1
+1	-1	+1
-1	+1	+1
+1	+1	+1
-1	0	0
+1	0	0
0	-1	0
0	+1	0
0	0	-1
0	0	+1
0	0	0
0	0	0

M

e.g. the $x_1{}^2$ column is the square of the x_1 column

Intercept	x_1	x_2	x_3	$x_1{}^2$	$x_2{}^2$	$x_3{}^2$	x_1x_2	x_1x_3	x_2x_3
1	-1	-1	-1	1	1	1	1	1	1
1	1	-1	-1	1	1	1	-1	-1	1
1	-1	1	-1	1	1	1	-1	1	-1
1	1	1	-1	1	1	1	1	-1	-1
1	-1	-1	1	1	1	1	1	-1	-1
1	1	-1	1	1	1	1	-1	1	-1
1	-1	1	1	1	1	1	-1	-1	1
1	1	1	1	1	1	1	1	1	1
1	-1	0	0	1	0	0	0	0	0
1	1	0	0	1	0	0	0	0	0
1	0	-1	0	0	1	0	0	0	0
1	0	1	0	0	1	0	0	0	0
1	0	0	-1	0	0	1	0	0	0
1	0	0	1	0	0	1	0	0	0
1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0

The Extended Model Matrix For An Arbitrary Split

Inter-cept	x_1	x_2	x_3	$x_1{}^2$	$x_2{}^2$	$x_3{}^2$	x_1x_2	x_1x_3	x_2x_3	x_b
1	-1	-1	-1	1	1	1	1	1	1	-1
1	1	-1	-1	1	1	1	-1	-1	1	-1
1	-1	1	-1	1	1	1	-1	1	-1	-1
1	1	1	-1	1	1	1	1	-1	-1	-1
1	-1	-1	1	1	1	1	1	-1	-1	-1
1	1	-1	1	1	1	1	-1	1	-1	-1
1	-1	1	1	1	1	1	-1	-1	1	-1
1	1	1	1	1	1	1	1	1	1	-1
1	-1	0	0	1	0	0	0	0	0	1
1	1	0	0	1	0	0	0	0	0	1
1	0	-1	0	0	1	0	0	0	0	1
1	0	1	0	0	1	0	0	0	0	1
1	0	0	-1	0	0	1	0	0	0	1
1	0	0	1	0	0	1	0	0	0	1
1	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	1

Extra column in M

Rows 1-8 are in Block 1,
9-16 in Block 2

The blocking factor is coded using –1 / +1 as usual

Now re-arrange this column to minimize the multiple correlation (R^2) between x_b and the other columns

An Optimal Solution And A Nearly Optimal Solution

x_1	x_2	x_3	Block
-1	-1	-1	-1
+1	-1	-1	+1
-1	+1	-1	+1
+1	+1	-1	-1
-1	-1	+1	+1
+1	-1	+1	-1
-1	+1	+1	-1
+1	+1	+1	+1
-1	0	0	+1
+1	0	0	+1
0	-1	0	+1
0	+1	0	+1
0	0	-1	+1
0	0	+1	+1
0	0	0	-1
0	0	0	-1

Optimal

6 runs in Block '-1', 10 in Block '+1'

All the star points are in Block '+1'

x_1	x_2	x_3	Block
-1	-1	-1	+1
+1	-1	-1	-1
-1	+1	-1	-1
+1	+1	-1	+1
-1	-1	+1	-1
+1	-1	+1	+1
-1	+1	+1	+1
+1	+1	+1	-1
-1	0	0	+1
+1	0	0	-1
0	-1	0	+1
0	+1	0	-1
0	0	-1	-1
0	0	+1	+1
0	0	0	+1
0	0	0	-1

Nearly Optimal

8 runs in each block

3 star points in each Block

Background Trends

- In many experimental set-ups there is concern that an uncontrolled variable will drift during the experiment.

  ... causing a trend in the response, independent of our factor changes.

- If the form of the trend can be anticipated (e.g. roughly linear in time), a run order can be selected to 'neutralise' the trend.

- We treat 'time' (really run number) as a covariate x_t and use the same ideas as in designing blocked experiments.

Randomizing The Run Order

- Randomization is strongly recommended in most DoE literature.
  - if full randomization is too costly, stratified randomization can be used (randomization within different groups of runs).
- Can be thought of as an 'insurance policy';
  - the idea is that for any specific column in the model matrix M, there is only a small chance that it will be strongly correlated with the trend or other covariate,
  - even when the form of the covariate is entirely unknown.

In a experiments it usually costs more to use a random run order

Randomization is not very effective in small experiments because there aren't enough different run orders

Another Look At Central Composite Designs

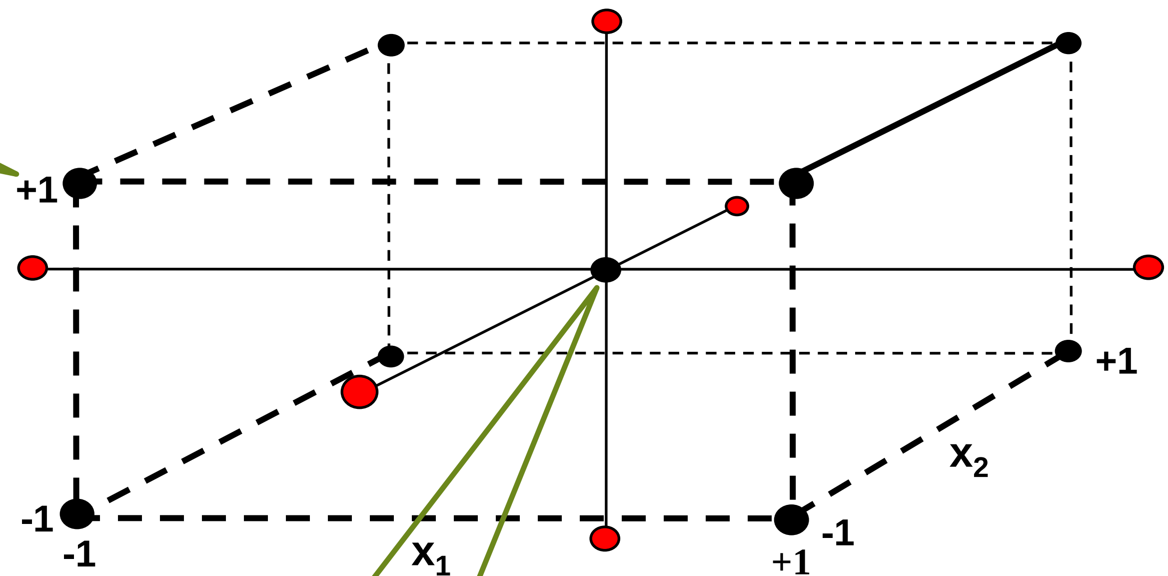
The most general CC design has 5 levels of each factor

It is often useful to explore a bit further out along the axes

Cube levels still coded -1 and +1 (usually)

Star (axial) points outside the cube

Distance to centre point $= \alpha\ (> 1)$



Let # runs at the centre $= n_C$

Example

$k = 4, \quad \alpha = \sqrt{2}, \quad n_C = 2$

factors

The cube is a 2^4 design

Star points

Centre point is replicated

Row	x_1	x_2	x_3	x_4
1	+1	+1	+1	+1
2	+1	+1	+1	–1
3	+1	+1	–1	+1
4	+1	+1	–1	–1
5	+1	–1	+1	+1
6	+1	–1	+1	–1
7	+1	–1	–1	+1
8	+1	–1	–1	–1
9	–1	+1	+1	+1
10	–1	+1	+1	–1
11	–1	+1	–1	+1
12	–1	+1	–1	–1
13	–1	–1	+1	+1
14	–1	–1	+1	–1
15	–1	–1	–1	+1
16	–1	–1	–1	–1
17	–1.414	0	0	0
18	+1.414	0	0	0
19	0	–1.414	0	0
20	0	+1.414	0	0
21	0	0	–1.414	0
22	0	0	+1.414	0
23	0	0	0	–1.414
24	0	0	0	+1.414
25	0	0	0	0
26	0	0	0	0

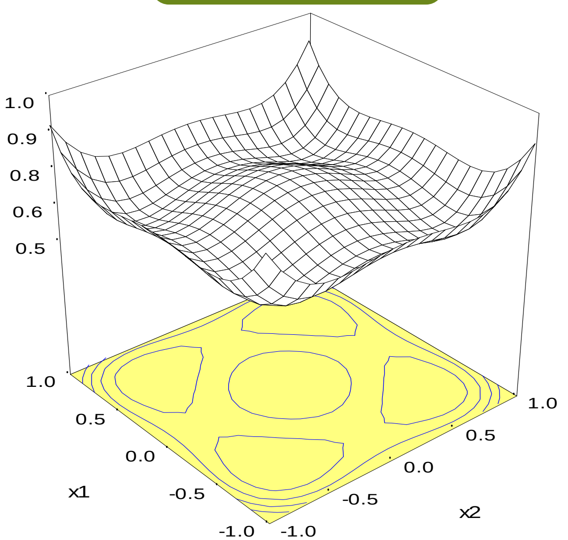
Choosing α and n_C In A CC Design

- These criteria have been recommended for choosing α and n_C :
  - the shape of the prediction standard deviation (PSD) function,
  - the shape of the region of interest,
  - orthogonal blocking.
- We'll discuss these criteria, then make practical recommendations.

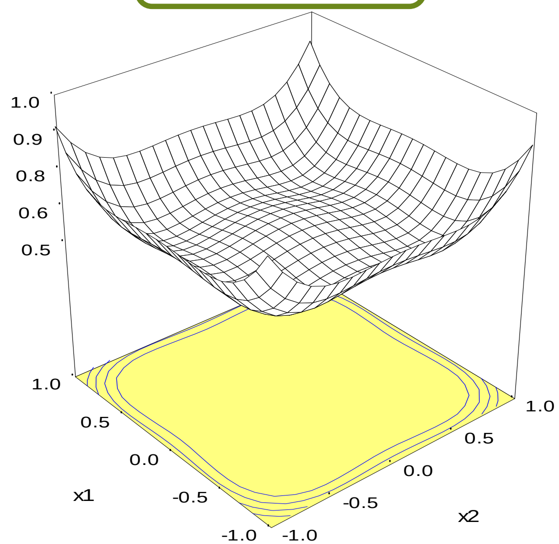
The PSD Function

- If we don’t replicate the centre run, the PSD has a ‘bump’ in the middle.
  - e.g. face-centred design with $k = 2$.

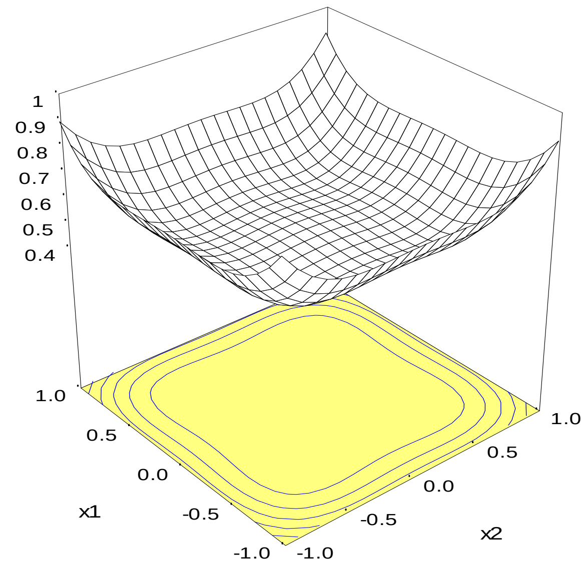
$n_c = 1$



$n_c = 2$

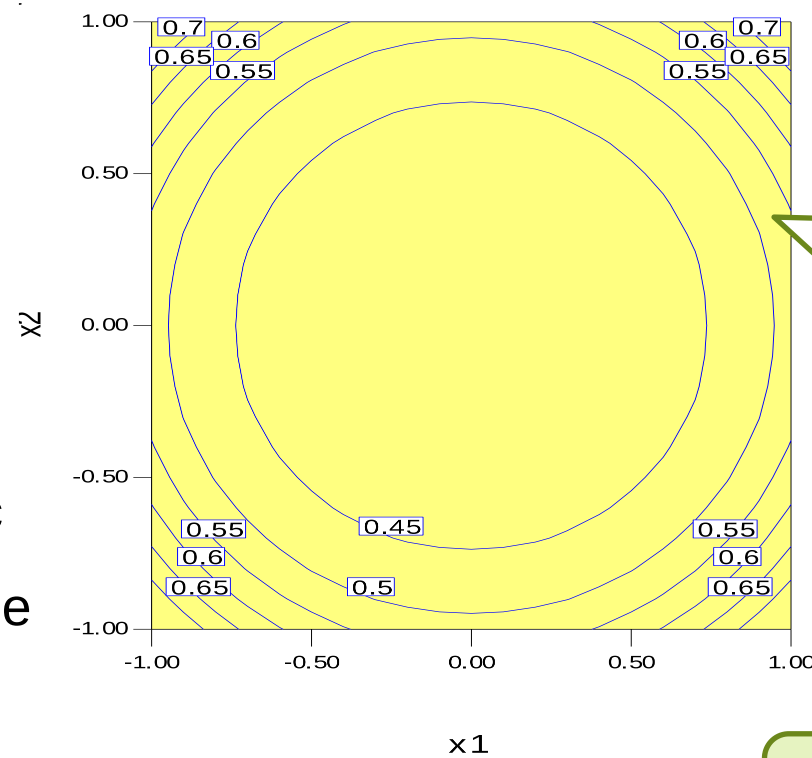


$n_c = 3$



Rotatability

- α can be chosen to make the design rotatable.
- This means that the PSD function is spherically symmetric (in coded units) i.e. the contours are circular:



Uncertainty is a function of distance from the centre (in coded units)

But we will want to predict with actual engineering units

… and this usually destroys rotatability

Rotatability requires large α e.g. 1.682 for $k = 3$

The Region of Interest: Spherical CC Designs

- The cube points lie at a distance $\sqrt{k}$ from the centre.
- So if $\alpha = \sqrt{k}$, all points except the centre lie on a hypersphere (radius $\sqrt{k}$),
  - there is extensive theory about spherical CC designs.
- If the experimenters really want to explore this shape of region, the design is obviously appropriate.
- **But** except for small k, the volume of this hypersphere is very big compared with the cube.
  - e.g. for $k = 4$, volume of hypersphere is about 5 times bigger than the cube.
  - for $k = 6$, volume of hypersphere is about 17 times bigger.

It is very ambitious to try to explore such a large region

Orthogonal Blocking

- For example, with $k = 3$, a CC design can be run in two blocks as follows:
  - Block 1: cube + 4 centre runs;
  - Block 2: star at $\alpha = 1.633$ + 2 centre runs.

These blocks are orthogonal to all the other effects in a 2nd order model

6 runs at the centre (out of 20 in all)

Practical Choices For A CC design

- Ignore rotatability.
- If the region of interest is strictly a hypercube in coded units, use the face-centred design with $n_C = 2$.
- In other cases consider pushing the star points out to (say) $\alpha = 1.3$, with $n_C = 2 \; or \; 3$.
- Blocking: if two or more blocks are needed, use a computer algorithm to choose the blocks to give approximate orthogonality.

Don't use software defaults unless you really want them!

In This Session We Have…

- Shown how an experiment can be divided into blocks, using an extended model matrix.
- Explained that although most DoE literature recommends randomizing the run order, this is not effective in small experiments.
- Returned to Central Composite designs and looked at some criteria for:
  - choosing the number of runs at the centre,
  - placing the axial points.

Session 9: Developing a Multi Level Design

Tutorial and Exercise

Tutorial

- **Session TS08+09: Three Level Designs**
- **Objective:**
  - Develop skills for generating and evaluating three-level designs.
  - This tutorial is based on the Technical Sessions TS08 and TS09 - see the Technical Session slides for details.
- **Python Environment**

  A self-guided tutorial has been created as a Colab notebook with pre-designed Python code and notes. For this tutorial, follow the instructions in the notes, upload data files and run the code. No modification of code is required. Interpret the results in accordance with the Technical session.
- **Tutorial Tasks**
  1. Generate a face-centred Central Composite design for three factors.
  2. Make a 3D scatter plot of the design in factor space.
  3. Generate a D-optimal design in 15 runs for fitting a 2nd order response surface in three factors.
  4. Compare the CC and D-optimal designs.

Exercise

- **Sessions TS08+09 & 10: Three Level Experiment**
- **Objective**
- To design, analyse and discuss the results of a follow up three level experiment for the Virtual Catapult.

  https://sigmazone.com/catapult/

Exercise

- **Sessions TS08+09 & 10: Three Level Experiment**

- **Python Environment**

  The exercise has been created as a Colab notebook with notes. Follow the instructions in the notes, and create your own code using tutorials 08 to 11 as a guide. Interpret the results in accordance with the Technical Sessions.

- **Objectives:**
  - To plan, run and analyse a three-level experiment on the catapult, using **three** factors selected on the basis of your screening results
  - To make predictions for factor combinations that you have not already tested
  - To test your predictions by firing the catapult.

- **Guidelines**

  Amongst other things, you need to decide:
  - Which design to use; which three factors to use; which levels to use for each factor; how many runs to make.
  - If you decide to run a CC design you need to decide how many runs to make at the centre point. If you choose a D-optimal design you should consider whether to add one or more runs at the centre.