Experimental design and treatment design

- Two separate aspects of an experimental study
- Treatment design:
  - What is done to an experimental unit
  - Examples include:
    - factorial treatment structures (2 way, 3 way)
    - choice of X values to fit a regression
- Experimental design:
- How treatments are randomized to experimental units
- Examples include:
  - Different blocking schemes (RCBD, Latin Square)
  - Subsampling
- Every study has a treatment design and an experimental design
- Can combine in all possibilities
  - e.g., 2 way factorial done in blocks

Response surface modeling

- Treatment design for fitting regression models
  - Usually quadratic polynomials in 2 or more variables
- Most frequently used in engineering
- Goals:
  - Most common: find the optimum level of all variables
  - Much less common: Does an X have an effect on Y?
  - While keeping the number of experimental runs small
- Why quadratic?
  - Over a small area of covariate space, any function can be approximated by a 1st order model
    $$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \varepsilon$$
  - Optimal choice of X's is not defined, on the edge of the covariate space
  - Over a larger area of covariate space, any function can be approximated by a 2nd order model
    $$Y = \beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3+\beta_4 X_1^2+\beta_5 X_2^2+\beta_6 X_3^2+\beta_7 X_1 X_2+\beta_8 X_1 X_3+\beta_9 X_2 X_3+\cdots+\varepsilon$$
  - Note: includes quadratic terms and all pairwise interactions (no 3 way interaction)
  - Optimal choice of $X's$ is a function of the β's.
- Need 3 levels of each X to fit the 2nd order model
- Full factorial for 3 factors would require $3^3 = 27$ treatments

- How can we reduce the number of experimental runs?

Motivating example:

- Organic material can absorb lead, Pb, as Pb(II), from aqueous solutions
- How effective are pistachio shells?
  - Currently just waste
- Evaluate 3 variables
  - initial pH of solution: 2.0 to 5.5
  - concentration of Pb, 5 to 50 ppm
  - contact time, 5 to 120 min
- Want to find values of these variables at which extraction of Pb is the greatest

"classical RSM":

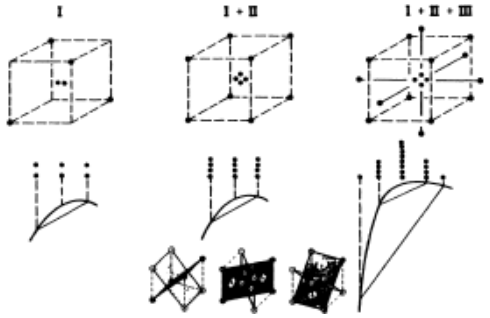
- Seminal paper: Box and Wilson, 1951
  - Developed RSM for chemical processes,
  - Wrote: "we believe that the methods will be of value in other fields where experimentation is sequential and the error fairly small"
  - Proposed the Central Composite Design
- experimental runs conducted sequentially
- Low variability
- So, relatively little replication
- Frequently no replication of many treatments
  - Many other designs developed over the last 70 years
  - Focus is on reducing the number of runs required
  - Or arranging treatments into incomplete blocks

Overall plan:

- Do not replicate all treatments - just one or a few
  - Key is small error variance
  - So observed value is close to true value
  - Some work on RSM for highly variable responses
    - Requires replication of all treatments, sometimes more than 2 replicates
- Screen variables to focus on ones with large effects
- Choose range of values for each variable $\Rightarrow$ low, high levels
  - mid is average of low, high
- Check whether optimum is likely to be within the range
  - If not, change the range(s) and recheck

- Estimate parameters of the 2nd order model
- Estimate optimum and a confidence region for it

In pictures, from Box and Draper 2006, *Response surfaces, mixtures, and ridge analyses*, 2nd ed, Wiley, p 484:



Parameterization of RSM models

- Standard coding of levels
  - -1: "low" level: often − in design descriptions
  - 0: "mid" level
  - +1: "high" level: often + in design descriptions
- 'Sum to zero" constraints (from factorial ANOVA)
- coded levels and actual values related by linear equations
- Use regression to estimate effects

Screening variables

- Goal: identify variables with largest influence on response
- Use a factorial design with 2 levels to see whether X has an effect
  - Choose practically relevant low (-1) and high (+1) levels
  - Often evaluate 7 or 8 variables
  - Use a fractional factorial design to reduce number of treatments
  - Fit 1st order model (main effects only)
  - Focus on estimates of effects
  - May be unreplicated
- Use subset of variables for further experimentation

Checking whether ranges are appropriate

- Want to estimate curvature for each variable
  - Need 3 levels

  - Add 2 "center" points (0 for all coded variables)
- Could reuse some of the screening runs
- Better not to
  - especially if new data collected on a new day, with different raw materials or machine calibrations
  - I.e., if expect unwanted variability between experimental runs
  - Consider screening runs and checking runs as two blocks
- For each variable:
  - Mean at center should be larger than average of low and high responses
- If not, expand the range for the offending variable

If uncertain whether range is appropriate

- Add the other half of the fractional factorial + more center points
- Result is a 2^k complete factorial plus center points
  - Where k is the number of variables under investigation
  - More precise estimates of 1st order effects
  - and the curvature

Central Composite Design, for k variables

- 3 components:
- Center point: 0 for all coded variables
  - Almost always replicated, 3-6 replicates are common
- Cube points: $+1$ / -1 for all coded variables
  - Complete factorial, 2^k runs
  - Not replicated in low error variance situations
- Axial ("star") points:
  - For each variable:
  - α / $-\alpha$ for that variable, 0's for other variables
  - Greatly increases precision of 2nd order coefficients
  - And provides extra df to assess lack of fit
  - Choosing α: two approaches
    * Spherical designs
    * distance from center to axial same as distance from center to cube
    * $\alpha = \sqrt{k}$
    * e.g. $\alpha = \sqrt{2} = 1.414$ for 2 variables, $\sqrt{3} = 1.732$ for 3 variables
    * Rotatable designs

    - * Prediction variance same for all points the same distance from the center
    - * $\alpha = \sqrt[4]{2^k}$
    - * gives $\alpha = 1.414$ for 2 variables, 1.68 for 3 variables
  - My sense is that spherical is more common
  - I don't see much difference between them

Thoughts about the CCD

- How much does it reduce # runs?
  - Depends on # variables
  - Without replication in either design:

	# variables			
Design	$k = 2$	$k = 3$	$k = 4$	$k = 5$
Factorial (3^k)	9	27	81	243
CCD	9	15	25	42
Cube (2^k)	4	8	16	32
Axial ($2k$)	4	6	8	10
Center, w/o repl.	1	1	1	1

- Why replicate only one point?
  - replication provides an estimate of the error variance
  - keep # runs small by not replicating all points
- Why replicate the center point?
  - Expect optimum "in the middle" of the design space
    - * more replicates there ⇒ more precise estimate there
  - If variance linearly related to levels of a variable
    - * Variance at the center point is the average variance across the design space
- What can you do when there is a boundary to the design space?
  - e.g. concentration of a reactant can't be negative
  - want to model response at concentrations close to 0
  - Put axial point for that variable at 0
  - More common:
    - * Put low edge of cube for that variable at 0
    - * "Pull in" that axial point to 0
    - * No longer statistical optimal, but practically useful
  - Or, use a design without axial points

Box-Behnken designs

- General idea: avoid "extreme" points in the design space
  - axial points
  - corners of the cube
- Concept, BB compared to spherical CCD:
  - spherical CCD: design points are on a sphere, extends beyond the cube
  - BB: design points are on a sphere that is completely "inside" the cube
- Points in a BB are closer to the center
  - Common to use fewer center reps than in a CCD
  - Don't explore as large as design space

Choosing a design

- Both CCD and BB are effective
- BB requires slightly fewer runs
- CCD is more commonly used
  - CCD gives better information about where quadratic approx. "works"
  - And can be done in 2 phases:
    - Phase 1: the factorial "cube": explore whether a factor has small effects
    - Could be done in two subphases using fractional factorial designs
    - Phase 2: the axial and center points - allow fitting quadratics and blocks = phases
- But, if extreme points are a concern (or not typical), use BB

Why don't other folks use response surface designs?

- Big debate ca 40 years ago
- A big practical difference
  - Missing values are bad news for RSM designs
  - Especially for BB designs without a Cube (a 2^k factorial)
  - Because RSM deliberately minimizes the number of runs
  - Missing cells are bad for complete factorial designs
  - But can still estimate lots of relevant quantities
- My resolution is based on the nature of the experiment
- If experimentation is sequential
  - I.e., plan next runs after seeing data from early runs
  - missing data can be fixed - rerun the missing treatments
- If set up experiment, then wait 4 months for results

  - Care a lot more about robustness to some missing observations

Fractional factorials

- Another design approach to reduce the number of runs while exploring many factors
- Most common designs have 2 levels of each factor. Principles apply to 3 levels per factor
  - Complete factorial with k factors, each with 2 levels, 2^k treatments
  - Lots of runs, even without any replication: 5 factors: 32 treatments, 8 factors: 264 treatments
- Only use 1/2 of the 2^k treatments.
- Example with 3 factors:

A	B	C	BC
+	+	+	+
+	-	-	+
-	+	-	-
-	-	+	-
+	+	-	-
+	-	+	-
-	+	+	+
-	-	-	+

- 2^{3-1} design uses 1st four treatment combinations
- Fit a model: $Y_{ij} = \mu + \beta_A A_i + \beta_B B_i + \beta_C C_i + \varepsilon_i$
- Notice the problem: BC confounded with A
- β_A estimates main effect of A and the interaction of B and C
- Fewer runs comes with a cost: estimates assume no 2 way interactions

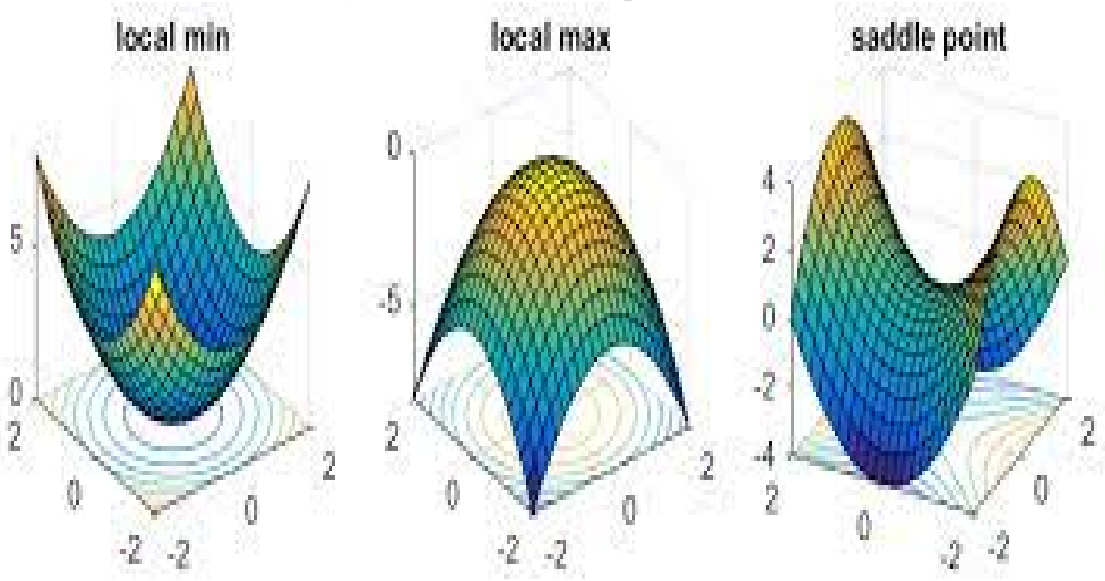
Fractional factorials: details

- Lots of tables of designs, with or without various choices of blocking
- With 4 or more factors, can choose what is confounded
- Resolution: summarizes the confounding
  - Resolution III: estimate main effects, may be confounded with 2 way interactions
  - Resolution IV: main effects not confounded with 2 way interactions;
    2 way interactions may be confounded with other 2 way
  - Resolution V: main effects not confounded with 3 way interactions;
    2 way interactions not confounded with other 2 ways
    3 way interactions may be confounded with 2 ways
- Deriving designs is an exercise in combinatorics.
- Extensively covered in engineering / industrial design of experiments texts.

Analyzing RSM data

- Fit the 2nd order model, estimate parameters for coded variables
- Now what?
- Is the quadratic model reasonable?

  - Look at residuals, especially standardized residuals
  - Are any large (large + or large −)?
  - Or do a lack-of-fit test
- Some folks like to simplify the model
  - Remove terms that have large p-values
  - But respect hierarchy:
    don't remember a linear term if used in a 2nd order term
  - My perspective: ok only if subject matter knowledge supports the simplification
- Estimate the location of the maximum
  - 1 variable: $Y = \beta_0 + \beta_1 X + \beta_2 X^2$
  - Maximum / minimum is $X_{opt} = -\hat{\beta}_1/(2\hat{\beta}_2)$
  - k variables: solve system of k equations
- Check whether it is maximum, minimum or saddlepoint



  - Can be done from the eigenvalues of the Hessian matrix
  - reported by most response-surface fitting algorithms
- There are methods to compute a confidence region for the optimum
- Specialized software for fitting response surfaces:
  - SAS: RSREG
  - R: rsm() in rsm package
  - JMP: use the Response surface macro in the fit model box
- Software provides
  - Whether max, min, or saddle point
  - Estimate of the optimum
  - Can get the confidence region for the optimum