

# Response Surface Methodology: Central Composite Designs

BIOE 498/598 PJ

Spring 2021

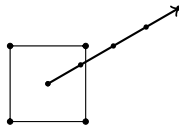
## Last time: The method of steepest ascent

- ▶ Begin with a FF+CP design.



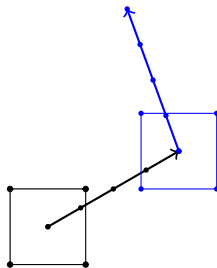
## Last time: The method of steepest ascent

- ▶ Begin with a FF+CP design.
- ▶ Follow path of steepest ascent until the model breaks.



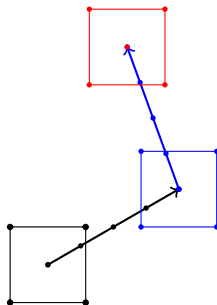
## Last time: The method of steepest ascent

- ▶ Begin with a FF+CP design.
- ▶ Follow path of steepest ascent until the model breaks.
- ▶ New FF+CP; repeat steepest ascent.



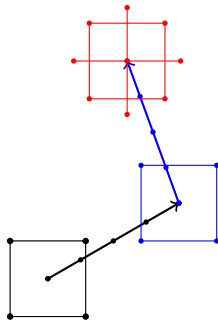
## Last time: The method of steepest ascent

- ▶ Begin with a FF+CP design.
- ▶ Follow path of steepest ascent until the model breaks.
- ▶ New FF+CP; repeat steepest ascent.
- ▶ Stop when model detects lack of fit.



## Last time: The method of steepest ascent

- ▶ Begin with a FF+CP design.
- ▶ Follow path of steepest ascent until the model breaks.
- ▶ New FF+CP; repeat steepest ascent.
- ▶ Stop when model detects lack of fit.
- ▶ **Today:** Fitting a model to a curved response surface.



# Fitting models with curvature

- ▶ We need two things to model a curved response surfaces:
  1. A model that is flexible enough to curve.
  2. Data that can detect the curvature.

# Fitting models with curvature

- ▶ We need two things to model a curved response surfaces:
  1. A model that is flexible enough to curve.
  2. Data that can detect the curvature.
- ▶ The optimal operating conditions correspond to a maximum in the response surface.
- ▶ We need models that can contain maxima.



# Fitting models with curvature

- ▶ We need two things to model a curved response surfaces:
  1. A model that is flexible enough to curve.
  2. Data that can detect the curvature.
- ▶ The optimal operating conditions correspond to a maximum in the response surface.
- ▶ We need models that can contain maxima.
- ▶ FO + TWI models are curved, but are rarely bounded.

# Fitting models with curvature

- ▶ We need two things to model a curved response surfaces:
  1. A model that is flexible enough to curve.
  2. Data that can detect the curvature.
- ▶ The optimal operating conditions correspond to a maximum in the response surface.
- ▶ We need models that can contain maxima.
- ▶ FO + TWI models are curved, but are rarely bounded.

$$y = 20 + 3.6x_1 - 1.8x_2 - 0.6x_1x_2$$

Set  $x_2 = 0$ , then  $y \rightarrow \infty$  as  $x_1 \rightarrow \infty$ .

# Nonlinear response surfaces

- ▶ The *true* model for any system is a general nonlinear function

$$y = f(x_1, x_2, \dots, x_k)$$

# Nonlinear response surfaces

- ▶ The *true* model for any system is a general nonlinear function

$$y = f(x_1, x_2, \dots, x_k)$$

- ▶ If you know  $f$  for your system, congrats! Fit its parameters with regression and use it.

# Nonlinear response surfaces

- ▶ The *true* model for any system is a general nonlinear function

$$y = f(x_1, x_2, \dots, x_k)$$

- ▶ If you know  $f$  for your system, congrats! Fit its parameters with regression and use it.
- ▶ Usually we don't know  $f$ , so we approximate it with a simpler function.

# Nonlinear response surfaces

- ▶ The *true* model for any system is a general nonlinear function

$$y = f(x_1, x_2, \dots, x_k)$$

- ▶ If you know  $f$  for your system, congrats! Fit its parameters with regression and use it.
- ▶ Usually we don't know  $f$ , so we approximate it with a simpler function.
- ▶ **We are not claiming that  $f$  is a particular shape.** Rather, we claim that an approximation is “good enough” over our domain of interest.

## Approximating $f$ with a general quadratic

Let's find the second-order Taylor series of  $f(x_1, x_2)$  centered at zero:

$$f(x_1, x_2) \approx \underbrace{f|_0}_{\beta_0} + \underbrace{\frac{\partial f}{\partial x_1} \Big|_0}_{\beta_1} x_1 + \underbrace{\frac{\partial f}{\partial x_2} \Big|_0}_{\beta_2} x_2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1^2} \Big|_0}_{\beta_{11}} x_1^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_2^2} \Big|_0}_{\beta_{22}} x_2^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1 \partial x_2} \Big|_0}_{\beta_{12}} x_1 x_2$$

$$f(x_1, x_2) \approx \underbrace{\beta_0 + \beta_1 x_1 + \beta_2 x_2}_{\text{FO}} + \underbrace{\beta_{11} x_1^2 + \beta_{22} x_2^2}_{\text{PQ}} + \underbrace{\beta_{12} x_1 x_2}_{\text{TWI}}$$

SO

- ▶ The function  $f$  and its derivatives are unknown, so we fit the parameters  $\beta$  with a linear model.
- ▶ In general we will have  $k$  factors and the quadratic approximation will be

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{j=1}^k \sum_{i=1}^{j-1} \beta_{ij} x_i x_j.$$

- ▶ This model has  $1 + 2k + k(k-1)/2$  parameters, so RSM designs must have at least this many runs.

## Approximating $f$ with a general quadratic

Let's find the second-order Taylor series of  $f(x_1, x_2)$  centered at zero:

$$f(x_1, x_2) \approx \underbrace{f|_0}_{\beta_0} + \underbrace{\frac{\partial f}{\partial x_1} \Big|_0}_{\beta_1} x_1 + \underbrace{\frac{\partial f}{\partial x_2} \Big|_0}_{\beta_2} x_2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1^2} \Big|_0}_{\beta_{11}} x_1^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_2^2} \Big|_0}_{\beta_{22}} x_2^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1 \partial x_2} \Big|_0}_{\beta_{12}} x_1 x_2$$

$$f(x_1, x_2) \approx \underbrace{\beta_0 + \beta_1 x_1 + \beta_2 x_2}_{\text{FO}} + \underbrace{\beta_{11} x_1^2 + \beta_{22} x_2^2}_{\text{PQ}} + \underbrace{\beta_{12} x_1 x_2}_{\text{TWI}}$$

SO

- ▶ The function  $f$  and its derivatives are unknown, so we fit the parameters  $\beta$  with a linear model.
- ▶ In general we will have  $k$  factors and the quadratic approximation will be

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{j=1}^k \sum_{i=1}^{j-1} \beta_{ij} x_i x_j.$$

- ▶ This model has  $1 + 2k + k(k-1)/2$  parameters, so RSM designs must have at least this many runs.



## Approximating $f$ with a general quadratic

Let's find the second-order Taylor series of  $f(x_1, x_2)$  centered at zero:

$$f(x_1, x_2) \approx \underbrace{f|_0}_{\beta_0} + \underbrace{\frac{\partial f}{\partial x_1} \Big|_0}_{\beta_1} x_1 + \underbrace{\frac{\partial f}{\partial x_2} \Big|_0}_{\beta_2} x_2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1^2} \Big|_0}_{\beta_{11}} x_1^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_2^2} \Big|_0}_{\beta_{22}} x_2^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1 \partial x_2} \Big|_0}_{\beta_{12}} x_1 x_2$$

$$f(x_1, x_2) \approx \underbrace{\beta_0 + \beta_1 x_1 + \beta_2 x_2}_{\text{FO}} + \underbrace{\beta_{11} x_1^2 + \beta_{22} x_2^2}_{\text{PQ}} + \underbrace{\beta_{12} x_1 x_2}_{\text{TWI}}$$

SO

- ▶ The function  $f$  and its derivatives are unknown, so we fit the parameters  $\beta$  with a linear model.
- ▶ In general we will have  $k$  factors and the quadratic approximation will be

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{j=1}^k \sum_{i=1}^{j-1} \beta_{ij} x_i x_j.$$

- ▶ This model has  $1 + 2k + k(k-1)/2$  parameters, so RSM designs must have at least this many runs.

## Approximating $f$ with a general quadratic

Let's find the second-order Taylor series of  $f(x_1, x_2)$  centered at zero:

$$f(x_1, x_2) \approx \underbrace{f|_0}_{\beta_0} + \underbrace{\frac{\partial f}{\partial x_1} \Big|_0}_{\beta_1} x_1 + \underbrace{\frac{\partial f}{\partial x_2} \Big|_0}_{\beta_2} x_2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1^2} \Big|_0}_{\beta_{11}} x_1^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_2^2} \Big|_0}_{\beta_{22}} x_2^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1 \partial x_2} \Big|_0}_{\beta_{12}} x_1 x_2$$

$$f(x_1, x_2) \approx \underbrace{\beta_0 + \beta_1 x_1 + \beta_2 x_2}_{\text{FO}} + \underbrace{\beta_{11} x_1^2 + \beta_{22} x_2^2}_{\text{PQ}} + \underbrace{\beta_{12} x_1 x_2}_{\text{TWI}}$$

SO

- ▶ The function  $f$  and its derivatives are unknown, so we fit the parameters  $\beta$  with a linear model.
- ▶ In general we will have  $k$  factors and the quadratic approximation will be

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{j=1}^k \sum_{i=1}^{j-1} \beta_{ij} x_i x_j.$$

- ▶ This model has  $1 + 2k + k(k-1)/2$  parameters, so RSM designs must have at least this many runs.

## Approximating $f$ with a general quadratic

Let's find the second-order Taylor series of  $f(x_1, x_2)$  centered at zero:

$$f(x_1, x_2) \approx \underbrace{f|_0}_{\beta_0} + \underbrace{\frac{\partial f}{\partial x_1} \Big|_0}_{\beta_1} x_1 + \underbrace{\frac{\partial f}{\partial x_2} \Big|_0}_{\beta_2} x_2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1^2} \Big|_0}_{\beta_{11}} x_1^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_2^2} \Big|_0}_{\beta_{22}} x_2^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1 \partial x_2} \Big|_0}_{\beta_{12}} x_1 x_2$$

$$f(x_1, x_2) \approx \underbrace{\beta_0 + \beta_1 x_1 + \beta_2 x_2}_{\text{FO}} + \underbrace{\beta_{11} x_1^2 + \beta_{22} x_2^2}_{\text{PQ}} + \underbrace{\beta_{12} x_1 x_2}_{\text{TWI}}$$

SO

- ▶ The function  $f$  and its derivatives are unknown, so we fit the parameters  $\beta$  with a linear model.
- ▶ In general we will have  $k$  factors and the quadratic approximation will be

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{j=1}^k \sum_{i=1}^{j-1} \beta_{ij} x_i x_j.$$

- ▶ This model has  $1 + 2k + k(k-1)/2$  parameters, so RSM designs must have at least this many runs.

## Approximating $f$ with a general quadratic

Let's find the second-order Taylor series of  $f(x_1, x_2)$  centered at zero:

$$f(x_1, x_2) \approx \underbrace{f|_0}_{\beta_0} + \underbrace{\frac{\partial f}{\partial x_1}\bigg|_0}_{\beta_1} x_1 + \underbrace{\frac{\partial f}{\partial x_2}\bigg|_0}_{\beta_2} x_2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1^2}\bigg|_0}_{\beta_{11}} x_1^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_2^2}\bigg|_0}_{\beta_{22}} x_2^2 + \underbrace{\frac{1}{2} \frac{\partial^2 f}{\partial x_1 \partial x_2}\bigg|_0}_{\beta_{12}} x_1 x_2$$

$$f(x_1, x_2) \approx \underbrace{\beta_0 + \beta_1 x_1 + \beta_2 x_2}_{\text{FO}} + \underbrace{\beta_{11} x_1^2 + \beta_{22} x_2^2}_{\text{PQ}} + \underbrace{\beta_{12} x_1 x_2}_{\text{TWI}}$$

SO

- ▶ The function  $f$  and its derivatives are unknown, so we fit the parameters  $\beta$  with a linear model.
- ▶ In general we will have  $k$  factors and the quadratic approximation will be

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{j=1}^k \sum_{i=1}^{j-1} \beta_{ij} x_i x_j.$$

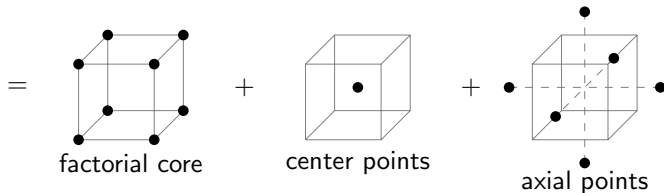
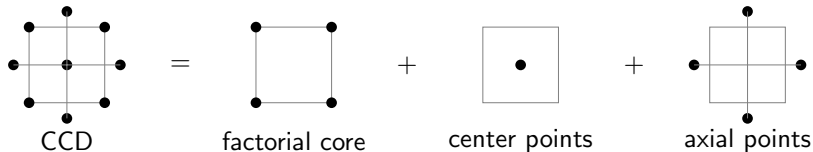
- ▶ This model has  $1 + 2k + k(k-1)/2$  parameters, so RSM designs must have at least this many runs.

# The Central Composite Design (CCD)

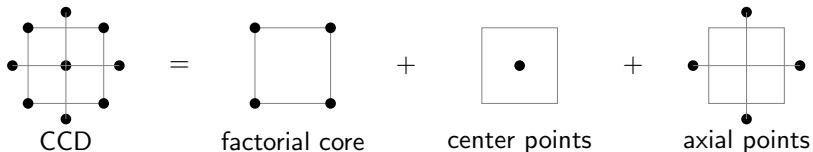
- ▶ A factorial or FF design can estimate FO and TWI terms.
- ▶ Estimating curvature requires points beyond the factorial corners.
- ▶ One popular option is the **Central Composite Design**.

# The Central Composite Design (CCD)

- ▶ A factorial or FF design can estimate FO and TWI terms.
- ▶ Estimating curvature requires points beyond the factorial corners.
- ▶ One popular option is the **Central Composite Design**.

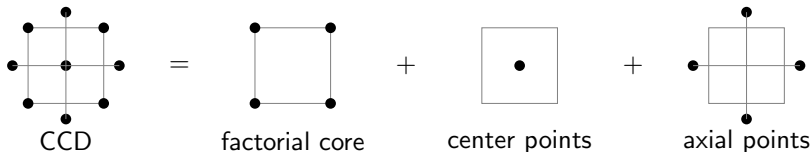


## Parts of the CCD



- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.

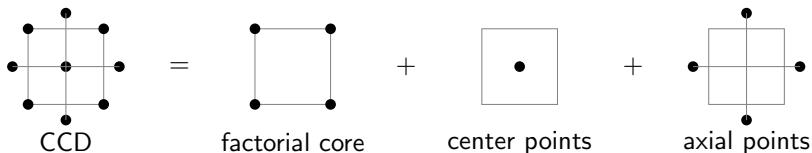
## Parts of the CCD



- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.
- ▶ Axial points allow estimation of the PQ terms. Without axial points we could only estimate the sum of all PQ terms.

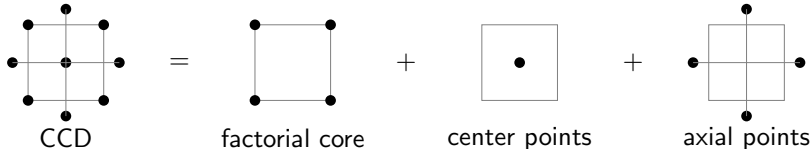


## Parts of the CCD



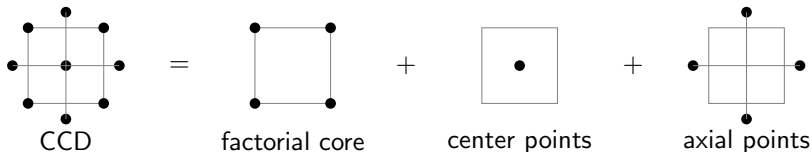
- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.
- ▶ Axial points allow estimation of the PQ terms. Without axial points we could only estimate the sum of all PQ terms.
- ▶ CCDs have a pair of axial runs for each factor:

## Parts of the CCD



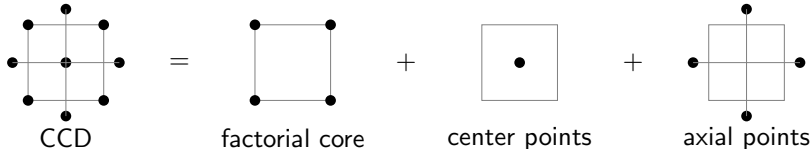
- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.
- ▶ Axial points allow estimation of the PQ terms. Without axial points we could only estimate the sum of all PQ terms.
- ▶ CCDs have a pair of axial runs for each factor:
  - ▶ One factor is set to  $\pm\alpha$

## Parts of the CCD



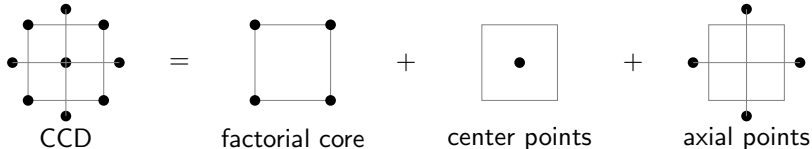
- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.
- ▶ Axial points allow estimation of the PQ terms. Without axial points we could only estimate the sum of all PQ terms.
- ▶ CCDs have a pair of axial runs for each factor:
  - ▶ One factor is set to  $\pm\alpha$
  - ▶ All other factors are set to 0.

## Parts of the CCD



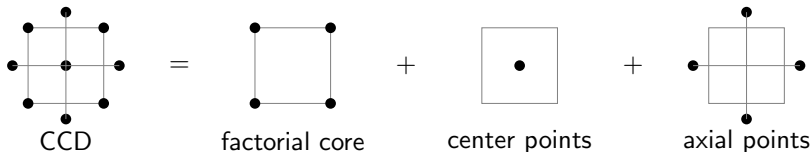
- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.
- ▶ Axial points allow estimation of the PQ terms. Without axial points we could only estimate the sum of all PQ terms.
- ▶ CCDs have a pair of axial runs for each factor:
  - ▶ One factor is set to  $\pm\alpha$
  - ▶ All other factors are set to 0.
- ▶ Center points estimate pure error and help (some) with PQ terms.

## Parts of the CCD



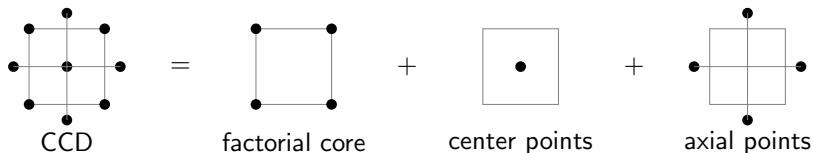
- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.
- ▶ Axial points allow estimation of the PQ terms. Without axial points we could only estimate the sum of all PQ terms.
- ▶ CCDs have a pair of axial runs for each factor:
  - ▶ One factor is set to  $\pm\alpha$
  - ▶ All other factors are set to 0.
- ▶ Center points estimate pure error and help (some) with PQ terms.
- ▶ To build a CCD you need to decide:

## Parts of the CCD



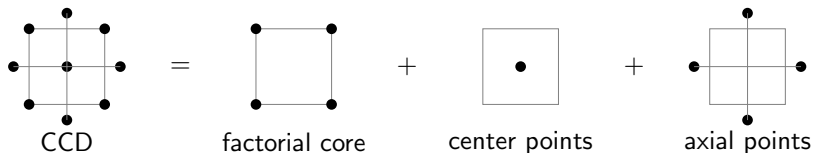
- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.
- ▶ Axial points allow estimation of the PQ terms. Without axial points we could only estimate the sum of all PQ terms.
- ▶ CCDs have a pair of axial runs for each factor:
  - ▶ One factor is set to  $\pm\alpha$
  - ▶ All other factors are set to 0.
- ▶ Center points estimate pure error and help (some) with PQ terms.
- ▶ To build a CCD you need to decide:
  1. The size of the FF core

## Parts of the CCD



- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.
- ▶ Axial points allow estimation of the PQ terms. Without axial points we could only estimate the sum of all PQ terms.
- ▶ CCDs have a pair of axial runs for each factor:
  - ▶ One factor is set to  $\pm\alpha$
  - ▶ All other factors are set to 0.
- ▶ Center points estimate pure error and help (some) with PQ terms.
- ▶ To build a CCD you need to decide:
  1. The size of the FF core
  2. The number of center runs

# Parts of the CCD



- ▶ Factorial points alone estimates the FO and TWI terms. The core must be Resolution V or higher.
- ▶ Axial points allow estimation of the PQ terms. Without axial points we could only estimate the sum of all PQ terms.
- ▶ CCDs have a pair of axial runs for each factor:
  - ▶ One factor is set to  $\pm\alpha$
  - ▶ All other factors are set to 0.
- ▶ Center points estimate pure error and help (some) with PQ terms.
- ▶ To build a CCD you need to decide:
  1. The size of the FF core
  2. The number of center runs
  3. The value of  $\alpha$

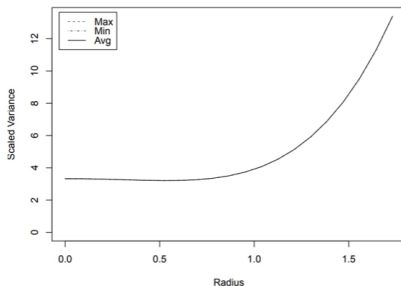


# Uniform precision

- ▶ A model has **uniform precision** if the variance at design radius 1 is the same as at the center.

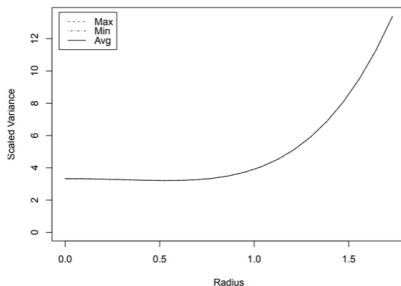
# Uniform precision

- ▶ A model has **uniform precision** if the variance at design radius 1 is the same as at the center.



# Uniform precision

- ▶ A model has **uniform precision** if the variance at design radius 1 is the same as at the center.



- ▶ Choosing the correct number of center points in a CCD ensures uniform precision.

# Rotatable designs

- ▶ Models are most precise at the center of the design.

## Rotatable designs

- ▶ Models are most precise at the center of the design.
- ▶ Ideally, the change in precision should be independent of the *direction* we move away from the center.

# Rotatable designs

- ▶ Models are most precise at the center of the design.
- ▶ Ideally, the change in precision should be independent of the *direction* we move away from the center.

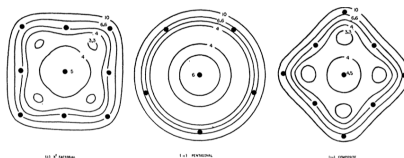


FIG. 2. Variance contours for some 2 dimensional designs

Image from Box and Hunter 1957.

# Rotatable designs

- ▶ Models are most precise at the center of the design.
- ▶ Ideally, the change in precision should be independent of the *direction* we move away from the center.

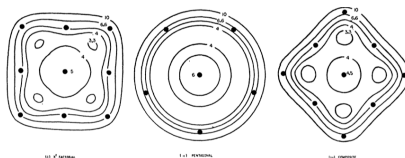


FIG. 2. Variance contours for some 2 dimensional designs

Image from Box and Hunter 1957.

- ▶ Designs where the variance only depends on the radius are called **rotatable designs**.

# Rotatable designs

- ▶ Models are most precise at the center of the design.
- ▶ Ideally, the change in precision should be independent of the *direction* we move away from the center.

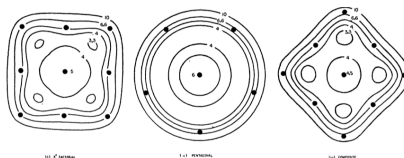


FIG. 2. Variance contours for some 2 dimensional designs

Image from Box and Hunter 1957.

- ▶ Designs where the variance only depends on the radius are called **rotatable designs**.
- ▶ A CCD with  $F$  factorial points is rotatable when  $\alpha = \sqrt[4]{F}$ .



## Rotatable, uniform precision CCDs

factors ( $k$ )	2	3	4	5	5 - 1	6
factorial points	4	8	16	32	16	64
axial points	4	6	8	10	10	12
center points	5	6	7	10	6	15
axial distance ( $\alpha$ )	1.414	1.682	2.000	2.378	2.000	2.828

factors ( $k$ )	6 - 1	7	7 - 1	8	8 - 1	8 - 2
factorial points	32	128	64	256	128	64
axial points	12	14	14	16	16	16
center points	9	21	14	28	20	13
axial distance ( $\alpha$ )	2.378	3.364	2.828	4.000	3.364	2.828

## Factor levels in a CCD

Each factor in the CCD will be set at five levels:

$$-\alpha \quad -1 \quad 0 \quad 1 \quad \alpha$$

## Factor levels in a CCD

Each factor in the CCD will be set at five levels:

$$-\alpha \quad -1 \quad 0 \quad 1 \quad \alpha$$

Unlike a 2-level design, the coded units in a CCD have meaning!

## Coding the CCD

Let's say we're designing a combination screening of three drugs. The absolute widest concentration range we can use for drug A is  $[-3.2, 1.0]$  on a  $\log_{10}$ - $\mu\text{M}$  scale. What are the five levels assuming a full-factorial CCD?

## Coding the CCD

Let's say we're designing a combination screening of three drugs. The absolute widest concentration range we can use for drug A is  $[-3.2, 1.0]$  on a  $\log_{10}\text{-}\mu\text{M}$  scale. What are the five levels assuming a full-factorial CCD?

$$F = 2^3 = 8 \Rightarrow \alpha = \sqrt[4]{8} = 1.68$$

## Coding the CCD

Let's say we're designing a combination screening of three drugs. The absolute widest concentration range we can use for drug A is  $[-3.2, 1.0]$  on a  $\log_{10}\text{-}\mu\text{M}$  scale. What are the five levels assuming a full-factorial CCD?

$$F = 2^3 = 8 \Rightarrow \alpha = \sqrt[4]{8} = 1.68$$

$$\begin{aligned} A &= \text{center}(A) + \frac{\text{range}(A)}{2\alpha}[\text{code}] \\ &= -1.1 + \frac{1 - (-3.2)}{2(1.68)}[\text{code}] \end{aligned}$$

## Coding the CCD

Let's say we're designing a combination screening of three drugs. The absolute widest concentration range we can use for drug A is  $[-3.2, 1.0]$  on a  $\log_{10}$ - $\mu$ M scale. What are the five levels assuming a full-factorial CCD?

$$F = 2^3 = 8 \Rightarrow \alpha = \sqrt[4]{8} = 1.68$$

$$\begin{aligned} A &= \text{center}(A) + \frac{\text{range}(A)}{2\alpha}[\text{code}] \\ &= -1.1 + \frac{1 - (-3.2)}{2(1.68)}[\text{code}] \end{aligned}$$

code:	$-\alpha$	$-1$	$0$	$1$	$\alpha$
$\log_{10}$ - $\mu$ M:	$-3.2$	$-2.4$	$-1.1$	$0.2$	$1.0$