Process Improvement: Steepest Ascent

BIOE 498/598 PJ

Spring 2021

Process Improvement

Design of Experiments is focused on **process characterization**.

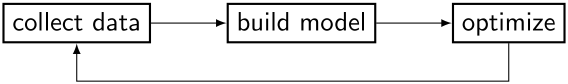
- Which factors affect the response?
- How large are the effects?

Process Improvement

Design of Experiments is focused on **process characterization**.

- Which factors affect the response?
- How large are the effects?

Process improvement asks "what factor settings yield the optimal response?"



Process improvement by steepest ascent

- Rarely are the initial factor ranges optimal. In practice we can be far away.

Process improvement by steepest ascent

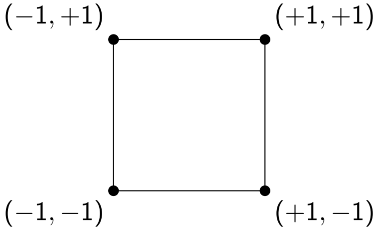
- Rarely are the initial factor ranges optimal. In practice we can be far away.
- The **method of steepest ascent** moves us quickly toward regions of better response.

Process improvement by steepest ascent

- Rarely are the initial factor ranges optimal. In practice we can be far away.
- The **method of steepest ascent** moves us quickly toward regions of better response.
- The emphasis is on moving quickly using few runs and first order models.

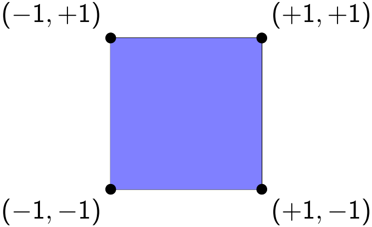
The Design Space

Runs in a factorial design sample the corners of a unit cube.



The Design Space

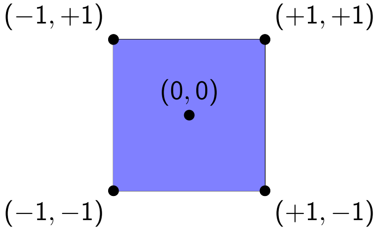
Runs in a factorial design sample the corners of a unit cube.



The region inside the factorial points is called the **design space**.

The Design Space

Runs in a factorial design sample the corners of a unit cube.



The region inside the factorial points is called the **design space**.
The origin $(0, 0)$ in *coded units* is called the **center point**.

Model Distance

- A linear model averages over runs at all corners of the design space.

Model Distance

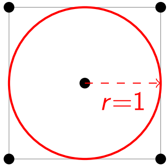
- A linear model averages over runs at all corners of the design space.
- The model's predictions are best at the center point.

Model Distance

- A linear model averages over runs at all corners of the design space.
- The model's predictions are best at the center point.
- As we move away from the center point, we switch from *interpolating* to *extrapolating*.

Model Distance

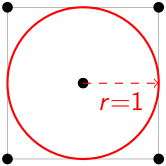
- A linear model averages over runs at all corners of the design space.
- The model's predictions are best at the center point.
- As we move away from the center point, we switch from *interpolating* to *extrapolating*.
- The **design radius** measures how far we are from the center point.



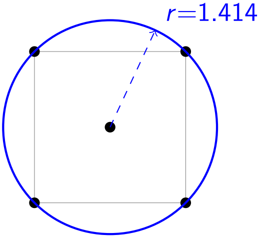
$r = 1$ touches the *faces* of the design space

Model Distance

- A linear model averages over runs at all corners of the design space.
- The model's predictions are best at the center point.
- As we move away from the center point, we switch from *interpolating* to *extrapolating*.
- The **design radius** measures how far we are from the center point.



$r = 1$ touches the *faces* of the design space



$r = \sqrt{2} \approx 1.414$ touches the factorial points

First-order response surfaces

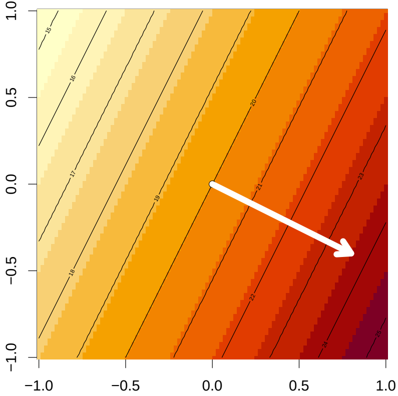
Consider the first-order linear model (without interactions)

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

First-order response surfaces

Consider the first-order linear model (without interactions)

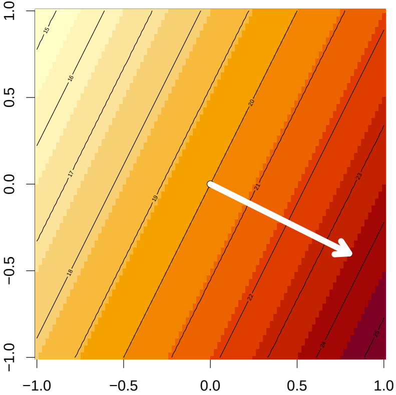
$$y = 20 + 3.6x_1 - 1.8x_2$$



First-order response surfaces

Consider the first-order linear model (without interactions)

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



We want to move "uphill" to improve the response using the **method of steepest ascent**.

If our goal was to minimize the response, we use *steepest descent* by

1. Moving opposite of the uphill direction, or
2. Multiplying the response by -1.

Finding the ascent direction for first-order models

Let's compute the partial derivatives along each factor's dimension.

$$\frac{\partial y}{\partial x_1} = \frac{\partial}{\partial x_1}\left(20 + 3.6x_1 - 1.8x_2\right) = 3.6$$
$$\frac{\partial y}{\partial x_2} = \frac{\partial}{\partial x_2}\left(20 + 3.6x_1 - 1.8x_2\right) = -1.8$$

Finding the ascent direction for first-order models

Let's compute the partial derivatives along each factor's dimension.

$$\frac{\partial y}{\partial x_1} = \frac{\partial}{\partial x_1}(20 + 3.6x_1 - 1.8x_2) = 3.6$$
$$\frac{\partial y}{\partial x_2} = \frac{\partial}{\partial x_2}(20 + 3.6x_1 - 1.8x_2) = -1.8$$

Two things to note:

1. The rate of ascent along each direction is simply the effect size β_i.
2. The rate of change is different for the two dimensions. For every step of unit length along x_1 we must move $3.6/(-1.8) = -1/2$ units along x_2.

Standardized step sizes for steepest ascent

Consider the general first-order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

1. Find the effect size with the largest **magnitude**. We'll call this β_j and the associated factor x_j.
2. Choose a step size (in coded units) along this dimension, called Δx_j.
3. For all other dimensions $i \neq j$, the step size is

$$\Delta x_i = \frac{\beta_i}{\beta_j} \Delta x_j$$

Standardized step sizes for steepest ascent

Consider the general first-order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

1. Find the effect size with the largest **magnitude**. We'll call this β_j and the associated factor x_j.
2. Choose a step size (in coded units) along this dimension, called Δx_j.
3. For all other dimensions $i \neq j$, the step size is

$$\Delta x_i = \frac{\beta_i}{\beta_j} \Delta x_j$$

Example: $y = 20 + 3.6x_1 - 1.8x_2$.

Standardized step sizes for steepest ascent

Consider the general first-order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

1. Find the effect size with the largest **magnitude**. We'll call this β_j and the associated factor x_j.
2. Choose a step size (in coded units) along this dimension, called Δx_j.
3. For all other dimensions $i \neq j$, the step size is

$$\Delta x_i = \frac{\beta_i}{\beta_j} \Delta x_j$$

Example: $y = 20 + 3.6x_1 - 1.8x_2$.

1. $|3.6| > |-1.8|$, so we standardize using x_1 ($j \equiv 1$).

Standardized step sizes for steepest ascent

Consider the general first-order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

1. Find the effect size with the largest **magnitude**. We'll call this β_j and the associated factor x_j.
2. Choose a step size (in coded units) along this dimension, called Δx_j.
3. For all other dimensions $i \neq j$, the step size is

$$\Delta x_i = \frac{\beta_i}{\beta_j} \Delta x_j$$

Example: $y = 20 + 3.6x_1 - 1.8x_2$.

1. $|3.6| > |-1.8|$, so we standardize using x_1 ($j \equiv 1$).
2. Let $\Delta x_1 = 1$.

Standardized step sizes for steepest ascent

Consider the general first-order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

1. Find the effect size with the largest **magnitude**. We'll call this β_j and the associated factor x_j.
2. Choose a step size (in coded units) along this dimension, called Δx_j.
3. For all other dimensions $i \neq j$, the step size is

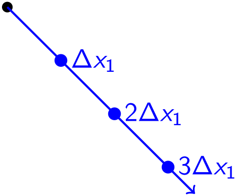
$$\Delta x_i = \frac{\beta_i}{\beta_j} \Delta x_j$$

Example: $y = 20 + 3.6x_1 - 1.8x_2$.

1. $|3.6| > |-1.8|$, so we standardize using x_1 ($j \equiv 1$).
2. Let $\Delta x_1 = 1$.
3. $\Delta x_2 = \frac{\beta_2}{\beta_1} \Delta x_1 = \frac{-1.8}{3.6}(1) = -0.5$

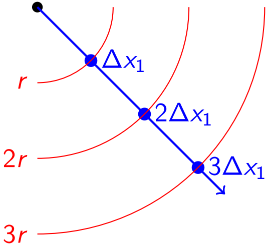
Why standardize step sizes?

Uniform steps give uniform differences in design radii.



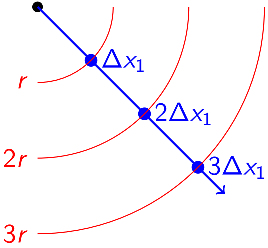
Why standardize step sizes?

Uniform steps give uniform differences in design radii.

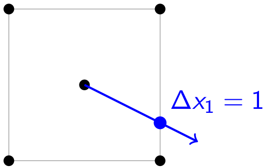


Why standardize step sizes?

Uniform steps give uniform differences in design radii.

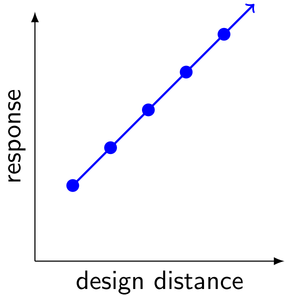


A standardized step of 1 always defines a point on the design space boundary.



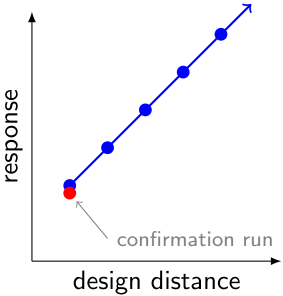
How far do we go?

- A first order model predicts the response will increase *forever*.
- We perform additional runs at every step along the ascent path.



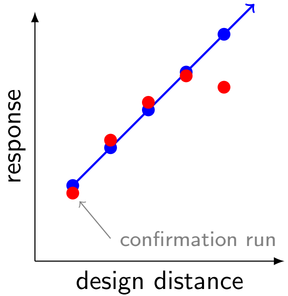
How far do we go?

- A first order model predicts the response will increase *forever*.
- We perform additional runs at every step along the ascent path.
- The first run is close to the center to confirm the system behaves as expected.



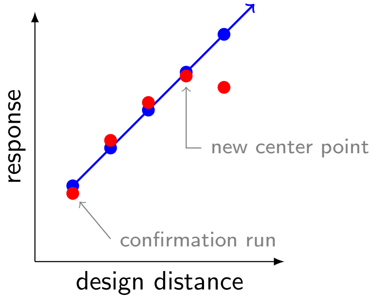
How far do we go?

- A first order model predicts the response will increase *forever*.
- We perform additional runs at every step along the ascent path.
- The first run is close to the center to confirm the system behaves as expected.
- Eventually the actual response will stop increasing.



How far do we go?

- A first order model predicts the response will increase *forever*.
- We perform additional runs at every step along the ascent path.
- The first run is close to the center to confirm the system behaves as expected.
- Eventually the actual response will stop increasing.
- When the response drifts, we use the best response location as the center for a new set of experiments.



What about interactions?

Models with interactions have **curved** paths of steepest ascent since the gradient changes with $\boldsymbol{x}$.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

$$\nabla y = \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \beta_1 + \beta_{12} x_2 \\ \beta_2 + \beta_{12} x_1 \end{pmatrix}$$

What about interactions?

Models with interactions have **curved** paths of steepest ascent since the gradient changes with $\boldsymbol{x}$.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

$$\nabla y = \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \beta_1 + \beta_{12} x_2 \\ \beta_2 + \beta_{12} x_1 \end{pmatrix}$$

We can follow this path by integrating: $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + (\nabla y)\Delta x$.

What about interactions?

Models with interactions have **curved** paths of steepest ascent since the gradient changes with $\boldsymbol{x}$.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

$$\nabla y = \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \beta_1 + \beta_{12} x_2 \\ \beta_2 + \beta_{12} x_1 \end{pmatrix}$$

We can follow this path by integrating: $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + (\nabla y)\Delta x$.

However, in practice we usually ignore the interactions.

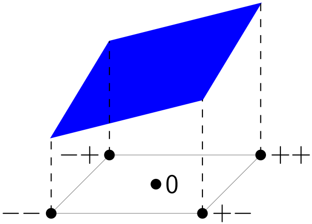
- The model will often break down before the curvature becomes significant.
- We rarely have enough runs in the initial design to identify interactions.

When is a first order model not good enough?

- The FF designs used for process improvement are usually augmented by **center points** — repeated runs at the design center $(0, 0)$.
- Center points serve two purposes:
  1. Estimate the *pure error* via the standard deviation of the repeated runs.
  2. Test for *lack of fit* to detect curvature.

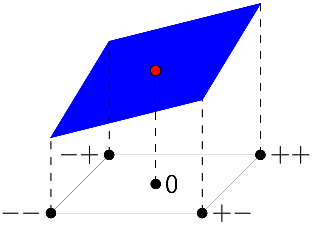
When is a first order model not good enough?

- The FF designs used for process improvement are usually augmented by **center points** — repeated runs at the design center $(0, 0)$.
- Center points serve two purposes:
  1. Estimate the *pure error* via the standard deviation of the repeated runs.
  2. Test for *lack of fit* to detect curvature.



When is a first order model not good enough?

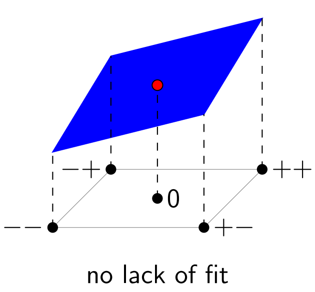
- The FF designs used for process improvement are usually augmented by **center points** — repeated runs at the design center $(0, 0)$.
- Center points serve two purposes:
  1. Estimate the *pure error* via the standard deviation of the repeated runs.
  2. Test for *lack of fit* to detect curvature.



no lack of fit

When is a first order model not good enough?

- The FF designs used for process improvement are usually augmented by **center points** — repeated runs at the design center $(0, 0)$.
- Center points serve two purposes:
  1. Estimate the *pure error* via the standard deviation of the repeated runs.
  2. Test for *lack of fit* to detect curvature.



Testing for lack of fit due to curvature

We want to compare the degree of curvature to the uncertainty (pure error) in our center points. We compare using a sum-of-squares approach.

Testing for lack of fit due to curvature

We want to compare the degree of curvature to the uncertainty (pure error) in our center points. We compare using a sum-of-squares approach.

1. $\bar{y}_{\text{center}}$ = mean response of the n_{center} center points
   $\bar{y}_{\text{fact}}$ = mean response of the n_{fact} factorial points

Testing for lack of fit due to curvature

We want to compare the degree of curvature to the uncertainty (pure error) in our center points. We compare using a sum-of-squares approach.

1. $\bar{y}_{\text{center}}$ = mean response of the n_{center} center points
   $\bar{y}_{\text{fact}}$ = mean response of the n_{fact} factorial points
2. $$SS_{\text{curve}} = \frac{n_{\text{fact}}\, n_{\text{center}} (\bar{y}_{\text{fact}} - \bar{y}_{\text{center}})^2}{n_{\text{fact}} + n_{\text{center}}}, \quad \text{DF}(SS_{\text{curve}}) = 1$$

Testing for lack of fit due to curvature

We want to compare the degree of curvature to the uncertainty (pure error) in our center points. We compare using a sum-of-squares approach.

1. $\bar{y}_{\text{center}}$ = mean response of the n_{center} center points
   $\bar{y}_{\text{fact}}$ = mean response of the n_{fact} factorial points
2. $$SS_{\text{curve}} = \frac{n_{\text{fact}} n_{\text{center}} (\bar{y}_{\text{fact}} - \bar{y}_{\text{center}})^2}{n_{\text{fact}} + n_{\text{center}}}, \quad \text{DF}(SS_{\text{curve}}) = 1$$
3. $$SS_{\text{error}} = \sum_{\substack{\text{center}\\\text{points}}} (y_i - \bar{y}_{\text{center}})^2, \quad \text{DF}(SS_{\text{error}}) = n_{\text{center}} - 1$$

Testing for lack of fit due to curvature

We want to compare the degree of curvature to the uncertainty (pure error) in our center points. We compare using a sum-of-squares approach.

1. $\bar{y}_{\text{center}}$ = mean response of the n_{center} center points
   $\bar{y}_{\text{fact}}$ = mean response of the n_{fact} factorial points
2. $$SS_{\text{curve}} = \frac{n_{\text{fact}} n_{\text{center}} (\bar{y}_{\text{fact}} - \bar{y}_{\text{center}})^2}{n_{\text{fact}} + n_{\text{center}}}, \quad \text{DF}(SS_{\text{curve}}) = 1$$
3. $$SS_{\text{error}} = \sum_{\substack{\text{center}\\\text{points}}} (y_i - \bar{y}_{\text{center}})^2, \quad \text{DF}(SS_{\text{error}}) = n_{\text{center}} - 1$$
4. $$F_{\text{curve}} = \frac{SS_{\text{curve}}/\text{DF}(SS_{\text{curve}})}{SS_{\text{error}}/\text{DF}(SS_{\text{error}})}$$

Example: Testing for curvature (Myers 2009)

temp	time	yield
$-$	$-$	39.3
$-$	$+$	40.0
$+$	$-$	40.9
$+$	$+$	41.5
0	0	40.3
0	0	40.5
0	0	40.7
0	0	40.2
0	0	40.6

Example: Testing for curvature (Myers 2009)

1. $\bar{y}_{\text{fact}} = (39.3 + 40.0 + 40.9 + 41.5)/4 = 40.425$
   $\bar{y}_{\text{center}} = (40.3{+}40.5{+}40.7{+}40.2{+}40.6)/5 = 40.46$

temp	time	yield
$-$	$-$	39.3
$-$	$+$	40.0
$+$	$-$	40.9
$+$	$+$	41.5
0	0	40.3
0	0	40.5
0	0	40.7
0	0	40.2
0	0	40.6

Example: Testing for curvature (Myers 2009)

1. $\bar{y}_{\text{fact}} = (39.3 + 40.0 + 40.9 + 41.5)/4 = 40.425$
   $\bar{y}_{\text{center}} = (40.3+40.5+40.7+40.2+40.6)/5 = 40.46$
2. $$SS_{\text{curve}} = \frac{4 \times 5 \times (40.425 - 40.46)^2}{4 + 5} = 0.0026$$

temp	time	yield
$-$	$-$	39.3
$-$	$+$	40.0
$+$	$-$	40.9
$+$	$+$	41.5
0	0	40.3
0	0	40.5
0	0	40.7
0	0	40.2
0	0	40.6

Example: Testing for curvature (Myers 2009)

1. $\bar{y}_{\text{fact}} = (39.3 + 40.0 + 40.9 + 41.5)/4 = 40.425$
   $\bar{y}_{\text{center}} = (40.3+40.5+40.7+40.2+40.6)/5 = 40.46$
2. $$SS_{\text{curve}} = \frac{4 \times 5 \times (40.425 - 40.46)^2}{4+5} = 0.0026$$
3. $$\begin{aligned} SS_{\text{error}} &= (40.3 - 40.46)^2 + \cdots + (40.6 - 40.46)^2 \\ &= 0.172 \end{aligned}$$

temp	time	yield
−	−	39.3
−	+	40.0
+	−	40.9
+	+	41.5
0	0	40.3
0	0	40.5
0	0	40.7
0	0	40.2
0	0	40.6

Example: Testing for curvature (Myers 2009)

1. $\bar{y}_{\text{fact}} = (39.3 + 40.0 + 40.9 + 41.5)/4 = 40.425$
   $\bar{y}_{\text{center}} = (40.3+40.5+40.7+40.2+40.6)/5 = 40.46$
2. $$SS_{\text{curve}} = \frac{4 \times 5 \times (40.425 - 40.46)^2}{4 + 5} = 0.0026$$
3. $$SS_{\text{error}} = (40.3 - 40.46)^2 + \cdots + (40.6 - 40.46)^2$$
   $$= 0.172$$
4. $$F_{\text{curve}} = \frac{0.0026/1}{0.172/(5-1)} = 0.0605$$

temp	time	yield
$-$	$-$	39.3
$-$	$+$	40.0
$+$	$-$	40.9
$+$	$+$	41.5
0	0	40.3
0	0	40.5
0	0	40.7
0	0	40.2
0	0	40.6

Example: Testing for curvature (Myers 2009)

1. $\bar{y}_{\text{fact}} = (39.3 + 40.0 + 40.9 + 41.5)/4 = 40.425$
   $\bar{y}_{\text{center}} = (40.3+40.5+40.7+40.2+40.6)/5 = 40.46$
2. $$SS_{\text{curve}} = \frac{4 \times 5 \times (40.425 - 40.46)^2}{4 + 5} = 0.0026$$
3. $$SS_{\text{error}} = (40.3 - 40.46)^2 + \cdots + (40.6 - 40.46)^2 = 0.172$$
4. $$F_{\text{curve}} = \frac{0.0026/1}{0.172/(5-1)} = 0.0605$$

temp	time	yield
$-$	$-$	39.3
$-$	$+$	40.0
$+$	$-$	40.9
$+$	$+$	41.5
0	0	40.3
0	0	40.5
0	0	40.7
0	0	40.2
0	0	40.6

`pf(0.0605, 1, 4, lower.tail=FALSE)` $\rightarrow p < 0.818$.

The steepest ascent method

1. Run a FF design augmented with replicated center points.

The steepest ascent method

1. Run a FF design augmented with replicated center points.
2. Fit a first order model and check for lack of fit.
   - If significant lack of fit, switch to Response Surface Methodology.

The steepest ascent method

1. Run a FF design augmented with replicated center points.
2. Fit a first order model and check for lack of fit.
   - If significant lack of fit, switch to Response Surface Methodology.
3. Perform runs along the steepest ascent path until the response diminishes.

The steepest ascent method

1. Run a FF design augmented with replicated center points.
2. Fit a first order model and check for lack of fit.
   - If significant lack of fit, switch to Response Surface Methodology.
3. Perform runs along the steepest ascent path until the response diminishes.
4. Go to (1) and repeat using the location of maximum response as the new center point.