01NAEX - Lecture 09
Plackett-Burman Design, 3^k Factorial and Response Surface Design

Jiri Franc

Czech Technical University
Faculty of Nuclear Sciences and Physical Engineering
Department of Mathematics

Plackett-Burman Design

Plackett-Burman Design is two-level fractional factorial design for studying $k = N - 1$ variables in N runs, where N is a multiple of 4. Sometimes it is called nongeometric design.

- Developed by Plackett & Burmann in 1946.
- These designs can not be represented as cubes.
- There are no design generators and the alias structure is complicated.
- N is a multiple of 4 (i.e. $N \in \{4, 8, 12, 16, 20, 24, 28, 32, \dots\}$).
- PB designs, where N is not a power of 2, are called non-geometric (i.e. $N \in \{12, 20, 24, \dots\}$).
- PB designs are non-regular designs. Regular design is one in witch all effects can be estimated independently of the other effects.

Plackett-Burman Design

The designs for $N = 12, 20, 24$, and 36 are obtained by writing the appropriate signs as a column (or row). A second column (or row) is then generated from the first one by moving the element of the column (or row) down (or right) one position and placing the last element in the first position. The process is continued until k columns (rows) are generated. A row of minus signs is then added.

Plus and Minus Signs for the Plackett–Burman Designs

$k = 11, N = 12$	+ + − + + + − − − + −
$k = 19, N = 20$	+ + − − + + + + − + − + − − − − + + −
$k = 23, N = 24$	+ + + + + − + − + + − − + + − − + − + − − − −
$k = 35, N = 36$	− + − + + + − − − + + + + + − + + + − − + − − − − + − + − + + − − + −

$k = 27, N = 28$

+ − + + + + − − −	− + − − − + − − +	+ + − + − + + − +
+ + − + + + − − −	− − + + − − + − −	− + + + + − + + −
− + + + + + − − −	+ − − − + − − + −	+ − + − + + − + +
− − − + − + + + +	− − + − + − − − +	+ − + + + − + − +
− − − + + − + + +	+ − − − − + + − −	+ + − − + + + + −
− − − − + + + + +	− + − + − − − + −	− + + + − + − + +
+ + + − − − + − +	− − + − − + − + −	+ − + + − + + + −
+ + + − − − + + −	+ − − + − − − − +	+ + − + + − − + +
+ + + − − − − + +	− + − − + − + − −	− + + − + + + − +

For $N = 28$, the three blocks are written as: X,Y,Z - first row, Z, X, Y - second row, Y, Z, X - third row and a rows of minus signs is added to these 27 rows.

Plackett-Burman Design

An example of the Plackett-Burman Design for $N = 12,\ k = 11$.

Run	A	B	C	D	E	F	G	H	I	J	K
1	+	−	+	−	−	−	+	+	+	−	+
2	+	+	−	+	−	−	−	+	+	+	−
3	−	+	+	−	+	−	−	−	+	+	+
4	+	−	+	+	−	+	−	−	−	+	+
5	+	+	−	+	+	−	+	−	−	−	+
6	+	+	+	−	+	+	−	+	−	−	−
7	−	+	+	+	−	+	+	−	+	−	−
8	−	−	+	+	+	−	+	+	−	+	−
9	−	−	−	+	+	+	−	+	+	−	+
10	+	−	−	−	+	+	+	−	+	+	−
11	−	+	−	−	−	+	+	+	−	+	+
12	−	−	−	−	−	−	−	−	−	−	−

For example: AB interaction is aliased with the nine main effects and each main effect is partially aliased with 45 two-factor interactions.

$$[A] = A - \frac{1}{3}BC - \frac{1}{3}BD - \frac{1}{3}BE + \frac{1}{3}BF + \cdots - \frac{1}{3}KL$$

The 3^k Factorial Design

3^k factorial design:
it is a factorial arrangement with k factors each at three levels, which are denoted by:

- 0 - low,
- 1 - intermediate,
- 2 - high.

Example:
consider the 3^2 design, quantitative factors, and let x_1 represent the factor A and x_2 represent factor B. A regression model relating the response y to x_1 and x_2 that is supported by this design is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{122} x_1 x_2^2 + \beta_{112} x_1^2 x_2 + \beta_{1122} x_1^2 x_2^2 + \epsilon$$

Notice:
The addition of third factor level allows the relationship between the response and the design factors to be modeled as a quadratic. The two factor interaction AB is then subdividing into four single degree of freedom components.

The 3^k Factorial Design

We can label the design points, similar to what we did in 2^k factorial design.

However, the coding using $\{0, 1, 2\}$,which is a generalization of the $\{0, 1\}$ coding that we also used in the 2^k design is more preferred.

A	B
-	-
0	-
+	-
-	0
0	0
+	0
-	+
0	+
+	+

A	B
0	0
1	0
2	0
0	1
1	1
2	1
0	2
1	2
2	2

The 3^k Factorial Design

3^k factorial design allows to model a curvature in the response function. However, two points need to be considered.

1. The 3^k design is not the most efficient way to model a quadratic relationship (the response surface designs are superior alternatives).
2. The 2^k design augmented with center points is an excellent way to obtain an indication of curvature. It allows one to keep the size and complexity of the design low and simultaneously obtain some protection against curvature. Then, if curvature is important, the 2^k design can be augmented with axial runs to obtain a central composite design.

The 3^2 Factorial Design

Let us consider 3^2 factorial design. There are 8 DF between treatment combinations, where the main effects of A and B each have 2 DF, and the AB interaction has 4 DF.

If the factor is quantitative:
Each main effect can be represented by a linear and a quadratic component, each with a single degree of freedom. The two-factor interaction AB may be partitioned in two ways.

The first method consists of subdividing AB into the four single-degree-of-freedom components corresponding to $AB_{L\times L}$, $AB_{L\times Q}$, $AB_{Q\times L}$, and $AB_{Q\times Q}$. This can be done by fitting the terms β_{12}, β_{122}, β_{112}, and β_{1122}.

And for the interaction sum of squares hold:

$$SS_{AB} = SS_{AB_{LxL}} + SS_{AB_{LxQ}} + SS_{AB_{QxL}} + SS_{AB_{QxQ}}$$

where L denoted linear and Q quadratic.

The 3^2 Factorial Design - partitioning of the AB interaction

Another way how to subdivide the interaction AB is to use Latin Squares. This method doesn't require that the factors be quantitative and it is usually associated with the case where all factors are qualitative.

Since interactions in three level designs don't have the same number of degrees of freedom as main effects we must partition the interactions into pseudo components (pseudo factors) called the AB component and the AB^2 component. These components could be called pseudo-interaction effects.

The components are defined as a linear combination as follows:

$$
\begin{aligned}
Levels_{AB} &= x_1 + x_2 (mod 3) \\
Levels_{AB^2} &= x_1 + 2x_2 (mod 3)
\end{aligned}
$$

The 3^2 Factorial Design

Pseudo components of decomposition of AB interaction in 3^2 design:

$$\begin{aligned} Levels_{AB} &= x_1 + x_2 (mod 3) \\ Levels_{AB^2} &= x_1 + 2x_2 (mod 3) \end{aligned}$$

A	B	AB	AB^2
0	0	0	0
1	0	1	1
2	0	2	2
0	1	1	2
1	1	2	0
2	1	0	1
0	2	2	1
1	2	0	2
2	2	1	0

The components AB and AB^2 each have two degrees of freedom. Note that

$$A^2B = (A^2B)^2 = A^4B^2 = AB^2.$$

The partitioning has no actual meaning and is not displayed in ANOVA results, but is useful in constructing more complex designs.

The 3^2 Factorial Design - Latin Squares construction

Let us consider 3^2 factorial design. In latin square notation we obtain:

$$\text{Latin block a } = AB \text{ component} \quad \begin{aligned} Q: &= x_1 + x_2 = 0(mod3) \\ R: &= x_1 + x_2 = 1(mod3) \\ S: &= x_1 + x_2 = 2(mod3) \end{aligned}$$

$$\text{Latin block b } = AB^2 \text{ component} \quad \begin{aligned} Q: &= x_1 + 2x_2 = 0(mod3) \\ R: &= x_1 + 2x_2 = 1(mod3) \\ S: &= x_1 + 2x_2 = 2(mod3) \end{aligned}$$

$$SS_{AB} = SS_a + SS_b$$

The AB and AB^2 components have no physical significance and can be called as **I and J components of interaction**:

$$\begin{aligned} I(AB) &= AB^2 \\ J(AB) &= AB \end{aligned}$$

The 3^2 Factorial Design - Example

The effective life of a cutting tool installed in a numerically controlled machine is thought to be affected by the cutting speed and the tool angle.

Three speeds and three angles are selected, and a 3^2 factorial experiment with two replicates is performed.

Data for Tool Life Experiment

Total Angle (degrees)	Cutting Speed (in/min) 125		150		175		$y_{i..}$
15	−2 −1	−3	−3 0	−3	2 3	5	−1
20	0 2	2	1 3	4	4 6	10	16
25	−1 0	−1	5 6	11	0 −1	−1	9
$y_{.j.}$	−2		12		14		$24 = y_{...}$

The 3^2 Factorial Design - Example

Simple 3^2 Factorial Design - ANOVA with interaction.

```
summary(aov(lm(Life~Angle*Speed,data=data_55f)))
            Df Sum Sq Mean Sq F value  Pr(>F)
Angle        2  24.33  12.167   8.423 0.00868 **
Speed        2  25.33  12.667   8.769 0.00770 **
Angle:Speed  4  61.33  15.333  10.615 0.00184 **
Residuals    9  13.00   1.444
```

Regression model with interaction.

```
summary(lm(Life~Angle*Speed,data=data_55))
Call:
lm.default(formula = Life ~ Angle * Speed, data = data_55)
Coefficients:
               Estimate Std. Error t value  Pr(>|t|)
(Intercept) -34.000000  21.704028  -1.567    0.140
Angle         1.366667   1.063276   1.285    0.220
Speed         0.213333   0.143372   1.488    0.159
Angle:Speed  -0.008000   0.007024  -1.139    0.274
Residual standard error: 2.483 on 14 degrees of freedom
Multiple R-squared:  0.3038, Adjusted R-squared:  0.1546
F-statistic: 2.036 on 3 and 14 DF,  p-value: 0.1551
```

The 3^2 Factorial Design - Example

Regression with quadratic terms and 4 interaction terms.

```
summary(lm(Life~Angle+Speed+Angle:Speed+I(Angle^2)+I(Speed^2)
    +I(Angle^2):Speed+I(Speed^2):Angle+I(Angle^2):I(Speed^2))

Coefficients:
                       Estim.Std. Error    t value Pr(>|t|)
(Intercept)          -1.068e+03  7.022e+02 -1.521 0.1626
Angle                 1.363e+02  7.261e+01  1.877 0.0932 .
Speed                 1.448e+01  9.503e+00  1.524 0.1619
I(Angle^2)           -4.080e+00  1.810e+00 -2.254 0.0507 .
I(Speed^2)           -4.960e-02  3.164e-02 -1.568 0.1514
Angle:Speed          -1.864e+00  9.827e-01 -1.897 0.0903 .
Speed:I(Angle^2)      5.600e-02  2.450e-02  2.285 0.0481 *
Angle:I(Speed^2)      6.400e-03  3.272e-03  1.956 0.0822 .
I(Angle^2):I(Speed^2) -1.920e-04  8.158e-05 -2.353 0.0431 *
Residual standard error: 1.202 on 9 degrees of freedom
Multiple R-squared:  0.8952, Adjusted R-squared:  0.802
F-statistic: 9.606 on 8 and 9 DF,  p-value: 0.001337
```

The 3^2 Factorial Design - Example

ANOVA with quadratic terms and 4 interaction terms.

```
summary(aov(lm(Life~Angle+Speed+Angle:Speed
+I(Angle^2)+I(Speed^2)+I(Angle^2):Speed+I(Speed^2):Angle
+I(Angle^2):I(Speed^2))))
                         Df  Sum Sq Mean Sq  F value   Pr(>F)
Angle                     1    8.33    8.33   5.769 0.039772 *
Speed                     1   21.33   21.33  14.769 0.003948 **
I(Angle^2)                1   16.00   16.00  11.077 0.008824 **
I(Speed^2)                1    4.00    4.00   2.769 0.130451
Angle:Speed               1    8.00    8.00   5.538 0.043065 *
Speed:I(Angle^2)          1    2.67    2.67   1.846 0.207306
Angle:I(Speed^2)          1   42.67   42.67  29.538 0.000414 ***
I(Angle^2):I(Speed^2)     1    8.00    8.00   5.538 0.043065 *
Residuals                 9   13.00    1.44
```

$$SS_{AB} = SS_{LxL} + SS_{LxQ} + SS_{QxL} + SS_{QxQ} = 8 + 42.67 + 2.67 + 8.00 = 61.34.$$

The 3^2 Factorial Design - Example

Interaction decomposition by Latin Squares

Factor A \ Factor B	0	1	2
0	Q −3	R −3	S 5
1	R 2	S 4	Q 10
2	S −1	Q 11	R −1

(*a*)

Factor A \ Factor B	0	1	2
0	Q −3	R −3	S 5
1	S 2	Q 4	R 10
2	R −1	S 11	Q −1

(*b*)

$$SS_{block\ a} = \frac{18^2 + (-2)^2 + 8^2}{(3)(2)} - \frac{24^2}{(9)(2)} = 33.34$$

$$SS_{block\ b} = \frac{0^2 + 6^2 + 18^2}{(3)(2)} - \frac{24^2}{(9)(2)} = 28.00$$

$$SS_{AB} = SS_{block\ a} + SS_{block\ b} = 33.34 + 28.00 = 61.34$$

The 3^k Factorial Design

In 3^3 design the two-factor interaction can be decomposed into eight single degree of freedom components (linear and quadratic combinations) or into four orthogonal two degree of freedom components, which are usually called W, X, Y, and Z.

$$\begin{aligned}
Levels_{AB^2C^2} = W(ABC): &= x_1 + 2x_2 + 2x_3 = 0(mod3) \\
Levels_{AB^2C} = X(ABC): &= x_1 + 2x_2 + \ \ x_3 = 0(mod3) \\
Levels_{ABC^2} = Y(ABC): &= x_1 + \ \ x_2 + 2x_3 = 0(mod3) \\
Levels_{ABC} = Z(ABC): &= x_1 + \ \ x_2 + \ \ x_3 = 0(mod3)
\end{aligned}$$

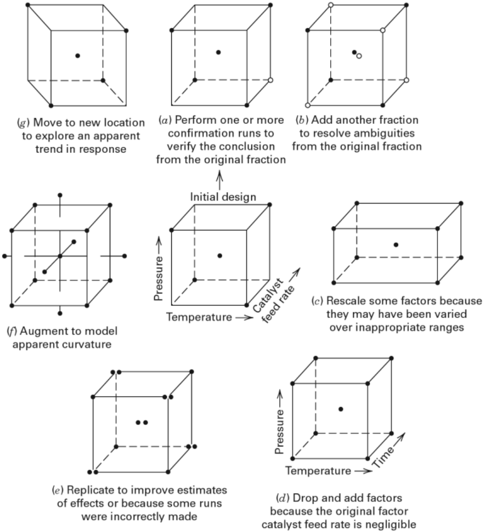
In general 3^k design, the only exponent allowed on the first letter is 1.

Like the I and J components, the W, X, Y, and Z components have no practical interpretation.

This decomposition is very useful, when we want to block 3^k design into 3^p blocks. We take one of the component of the interaction and confound it with blocks.

How to follow up experimentation

Iterative experimentation with alternatives for a subsequent set of runs depending on results from a previous set.



Response Surface Designs (RSD)

Response Surface Designs, sometime called Response Surface Methods (RSM), is a collection of techniques useful for modeling and analysis of problems in which a response of interest y is influenced by several variables $x_1, x_2, \ldots, x_k$ and the goal is to optimize this response,

$$y = f(x_1, x_2, \ldots, x_k) + \varepsilon,$$

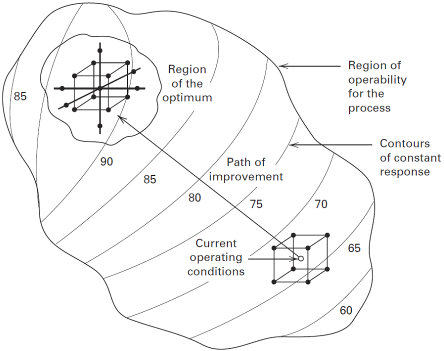
where ε represents the noise or error observed in the response y and

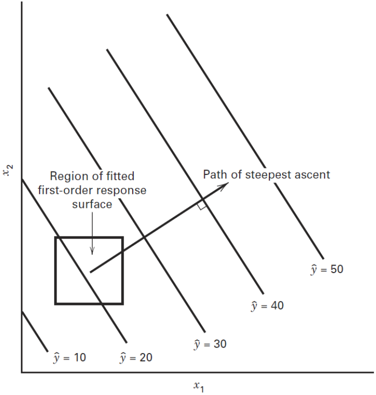
$$f(x_1, x_2, \ldots, x_k) = E[y]$$

is called a response surface.

Response Surface Designs (RSD)

RSD is a sequential procedure and frequently, the initial estimate of the optimum operating conditions is far from the actual optimum. The method of steepest ascent is a procedure for moving sequentially in the direction of the maximum increase in the response. We usually assume that in a small region of the *x*'s is the response fitted by first-order model and the steepest ascent is a gradient procedure.





An Example of Steepest Ascent

The yield y of a chemical process depends on reaction time A and temperature B. Current conditions are 35 min. and 155 F.

- **Objective:** Determine the operating conditions that maximize yield.
- **Variables:** Reaction Time & Temperature.
- **Parameters:** Time (30, 40) min, Temperature (150, 160) F.

$$x_1 = \frac{time - 35}{5}$$

$$x_2 = \frac{temperature - 155}{5}$$

First order model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

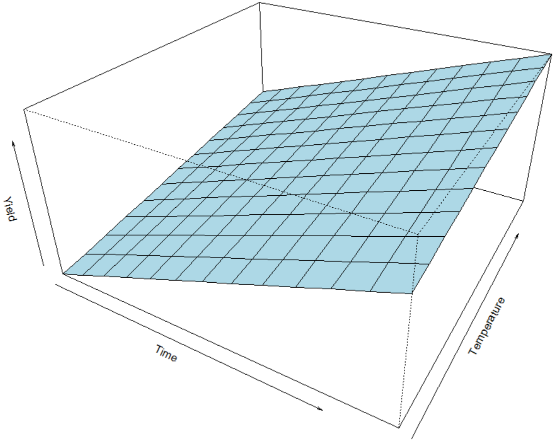
An Example of Steepest Ascent

```
> time  = c(30, 40, 30, 40)
> temp  = c(150, 150, 160,160)
> yield = c(39.3, 40.9, 40.0, 41.5)
> x1 = (time - 35)/5
> x2 = (temp - 155)/5
> data11_1 = data.frame(time,temp,x1,x2,yield)
> summary(lm( yield ~ x1 + x2,data = data11_1))

Call:
lm.default(formula = yield ~ x1 + x2, data = data11_1)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   40.425      0.025    1617 0.000394 ***
x1             0.775      0.025      31 0.020529 *
x2             0.325      0.025      13 0.048875 *
Residual standard error: 0.05 on 1 degrees of freedom
Multiple R-squared: 0.9991, Adjusted R-squared: 0.9973
F-statistic:   565 on 2 and 1 DF,  p-value: 0.02974
```

An Example of Steepest Ascent

First order model:

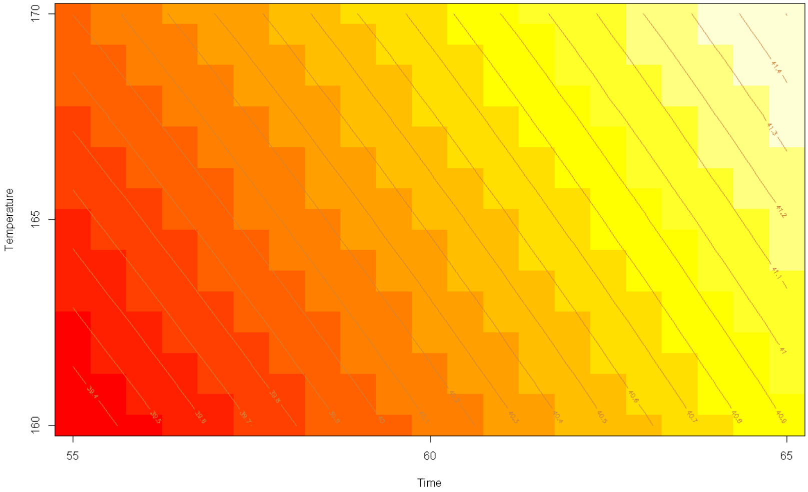
$$y = 40.425 + 0.775x_1 + 0.325x_2$$



An Example of Steepest Ascent

Contour plot of Yield process example for first order model:

$$y = 40.425 + 0.775x_1 + 0.325x_2$$



An Example of Steepest Ascent

To estimate parameters in interaction model we have to add some center points, which

- allows us to estimate the experimental error σ,
- allows us to test the interaction effect,
- allows us to check for curvature.

Natural Variables		Coded Variables		Response
ξ_1	ξ_2	x_1	x_2	y
30	150	−1	−1	39.3
30	160	−1	1	40.0
40	150	1	−1	40.9
40	160	1	1	41.5
35	155	0	0	40.3
35	155	0	0	40.5
35	155	0	0	40.7
35	155	0	0	40.2
35	155	0	0	40.6

We added four observations at the center to get an independent estimation of σ and check the fit of the first-order model or a curvature.

An Example of Steepest Ascent

The replicates at the center can be used to calculate an estimate of experimental error σ^2:

$$\hat{\sigma}^2 = MSE = \frac{\sum_{centerpoints}(y_i - \bar{y_C})^2}{n_c - 1} = \frac{0.172}{4} = 0.043$$

Second order model - checking for curvature

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon$$

Estimation of curvature is an estimate for $\beta_{11} + \beta_{22}$

$$\bar{y}_F - \bar{y}_C = 40.425 - 40.46 = -0.035$$

Test of curvature

$$SS_{curv} = \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C} = 0.0027$$

No evidence of Curvature (Pure quadratic effect), $F = \frac{SS_{curv}}{\hat{\sigma}^2} = 0.063$

```
> pf(0.063,1,4,lower.tail=F)
[1] 0.8141821
```

An Example of Steepest Ascent

Apply the method of steepest ascent on the first order model:

$$y = 40.425 + 0.775x_1 + 0.325x_2$$

Direction of steepest ascent - slope: $\frac{0.325}{0.775} = 0.42$
Step size in minutes: $\Delta x_1 = 5$

	Coded Variables		Natural Variables		Response
Steps	x_1	x_2	ξ_1	ξ_2	y
Origin	0	0	35	155	
Δ	1.00	0.42	5	2	
Origin + Δ	1.00	0.42	40	157	41.0
Origin + 2Δ	2.00	0.84	45	159	42.9
Origin + 3Δ	3.00	1.26	50	161	47.1
Origin + 4Δ	4.00	1.68	55	163	49.7
Origin + 5Δ	5.00	2.10	60	165	53.8
Origin + 6Δ	6.00	2.52	65	167	59.9
Origin + 7Δ	7.00	2.94	70	169	65.0
Origin + 8Δ	8.00	3.36	75	171	70.4
Origin + 9Δ	9.00	3.78	80	173	77.6
Origin + 10Δ	10.00	4.20	85	175	80.3
Origin + 11Δ	11.00	4.62	90	179	76.2
Origin + 12Δ	12.00	5.04	95	181	75.1

Local minimum find in step 10 and we should run another first order model around this new conditions.

An Example of Steepest Ascent

```
> x1 = (time - 85)/5
> x2 = (temp - 175)/5
> data11_3b = data.frame(time,temp,x1,x2,y1)
> data11_3b
   time   temp     x1     x2   y1
1  80.00 170.00 -1.000 -1.000 76.5
2  80.00 180.00 -1.000  1.000 77.0
3  90.00 170.00  1.000 -1.000 78.0
4  90.00 180.00  1.000  1.000 79.5
5  85.00 175.00  0.000  0.000 79.9
6  85.00 175.00  0.000  0.000 80.3
7  85.00 175.00  0.000  0.000 80.0
8  85.00 175.00  0.000  0.000 79.7
9  85.00 175.00  0.000  0.000 79.8
10 92.07 175.00  1.414  0.000 78.4
11 77.93 175.00 -1.414  0.000 75.6
12 85.00 182.07  0.000  1.414 78.5
13 85.00 167.93  0.000 -1.414 77.0
```

An Example of Steepest Ascent

Linear regression model with interaction:

```
> summary(model3b)
Call:
lm.default(formula = y1 ~ x1 + x2 + I(x1^2) + I(x2^2) + x1:x2,
            data = data11_3b)
Residuals:
Min        1Q   Median       3Q      Max
-0.23995 -0.18089 -0.03995  0.17758  0.36005
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 79.93995    0.11909 671.264  < 2e-16 ***
x1           0.99505    0.09415  10.568 1.48e-05 ***
x2           0.51520    0.09415   5.472 0.000934 ***
I(x1^2)     -1.37645    0.10098 -13.630 2.69e-06 ***
I(x2^2)     -1.00134    0.10098  -9.916 2.26e-05 ***
x1:x2        0.25000    0.13315   1.878 0.102519
---
Signif. codes:  0 Â‘***Â’ 0.001 Â‘**Â’ 0.01 Â‘*Â’ 0.05 Â‘.Â’ 0.1 Â‘ Â’
Residual standard error: 0.2663 on 7 degrees of freedom
Multiple R-squared:  0.9827, Adjusted R-squared:  0.9704
F-statistic: 79.67 on 5 and 7 DF,  p-value: 5.147e-06
```

An Example of Steepest Ascent

ANOVA of Linear regression model with interaction:

```
> summary(aov(model3b))
Df Sum Sq Mean Sq F value   Pr(>F)
x1           1  7.920   7.920 111.687 1.48e-05 ***
x2           1  2.123   2.123  29.941 0.000934 ***
I(x1^2)      1 10.982  10.982 154.866 4.98e-06 ***
I(x2^2)      1  6.972   6.972  98.323 2.26e-05 ***
x1:x2        1  0.250   0.250   3.526 0.102519
Residuals    7  0.496   0.071
---
Signif. codes:  0 Â‘***Â’ 0.001 Â‘**Â’ 0.01 Â‘*Â’ 0.05 Â‘.Â’ 0.1 Â‘ Â
```

An Example of Steepest Ascent

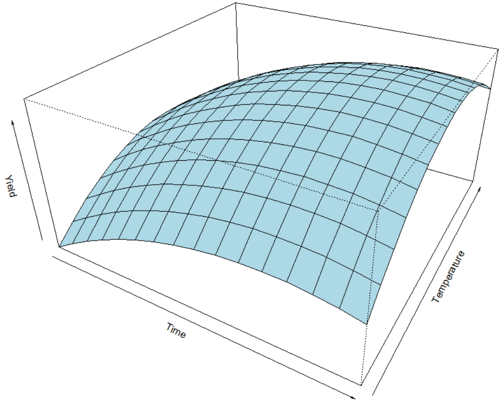
Linear regression model without interaction:

```
> summary(model3bb)
Call:
lm.default(formula = y1 ~ x1 + x2 + I(x1^2) + I(x2^2), data = data11_3
Residuals:
Min         1Q   Median        3Q       Max
-0.23995 -0.18089 -0.08232  0.06005  0.44808
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  79.9400     0.1366 585.215  < 2e-16 ***
x1            0.9951     0.1080   9.213 1.56e-05 ***
x2            0.5152     0.1080   4.770  0.00141 **
I(x1^2)      -1.3764     0.1158 -11.883 2.31e-06 ***
I(x2^2)      -1.0013     0.1158  -8.645 2.49e-05 ***
---
Signif. codes:  0 Â‘***Â’ 0.001 Â‘**Â’ 0.01 Â‘*Â’ 0.05 Â‘.Â’ 0.1 Â‘ Â’
Residual standard error: 0.3054 on 8 degrees of freedom
Multiple R-squared:  0.974, Adjusted R-squared:  0.961
F-statistic: 75.02 on 4 and 8 DF,  p-value: 2.226e-06
```

Model without interaction has some disadvantages. We can not use standard second order model and determine stationary points easily.

An Example of Steepest Ascent

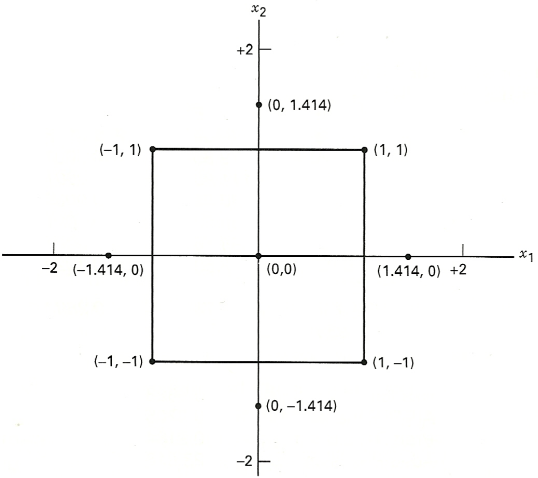
Location of stationary point - **Response surface plot** of Yield process example



The Second-Order Response Surface Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon$$

Central composite design (CCD): optimization and fitting the model is easy.



The Second-Order Response Surface Model

$$\hat{y} = \hat{\beta}_0 + \mathbf{x}^T\mathbf{b} + \mathbf{x}^T\mathbf{B}\mathbf{x},$$

where $\mathbf{x}^T = (x_1, x_2, \ldots, x_k)$, $\mathbf{b}^T = (\hat{\beta}_1, \ldots, \hat{\beta}_k)$ and

$$\mathbf{B} = \begin{pmatrix} \hat{\beta}_{11} & \frac{\hat{\beta}_{12}}{2} & \cdots & \frac{\hat{\beta}_{1k}}{2} \\ \frac{\hat{\beta}_{12}}{2} & \hat{\beta}_{22} & \cdots & \frac{\hat{\beta}_{2k}}{2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\hat{\beta}_{1k}}{2} & \frac{\hat{\beta}_{1k}}{2} & \cdots & \hat{\beta}_{kk} \end{pmatrix}$$

The general mathematical solution for the location of the stationary point for fitted second order model is:

$$\frac{\partial \hat{y}}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} = 0$$

$$\mathbf{x}_{stationary} = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b}$$

and the predicted response at the stationary point is:

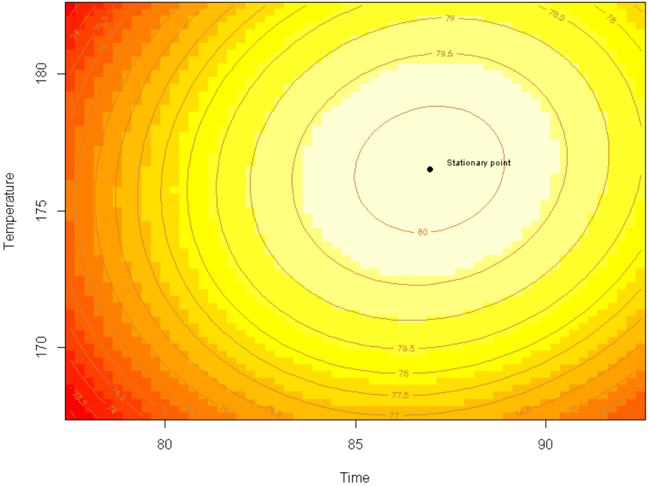
$$\hat{y}_{stationary} = \hat{\beta}_0 + \frac{1}{2}\mathbf{x}_{stationary}^T\mathbf{b}$$

The Second-Order Response Surface Model

```
> b = matrix(c(model3b$coeff[2],model3b$coeff[3]),2,1)
> B = matrix(c(model3b$coeff[4], model3b$coeff[6]/2,
+                  model3b$coeff[6]/2,model3b$coeff[5]),2,2)
> cbind(b,B)
[,1]      [,2]      [,3]
[1,] 0.9950503 -1.376449  0.125000
[2,] 0.5152028  0.125000 -1.001336
> x_stat       = -1/2 * solve(B) %*% b
> x_stat_natur = c((5*x_stat[1]+85), (5*x_stat[2] +175))
> y_stat_natur = predict(model3b,
                       data.frame(x1=x_stat[1],x2=x_stat[2]))
> x_stat
[,1]
[1,] 0.3892304
[2,] 0.3058466
> x_stat_natur
[1]  86.94615 176.52923
> y_stat_natur
1
80.21239
```

An Example of Steepest Ascent

Location of stationary point - **Contour plot** of Yield process example



Multiple Responses

Designs with multiple responses are common in practice. Typically, we want to simultaneously optimize all responses, or find a set of conditions where certain product properties are achieved.

A simple approach is to model all responses and overlay the contour plots.

Today Exercises

If necessary, continue in measurement to finish HW 02 from the last lesson.

Solve problems 11.8 from D. C. Montogomery DAOE.

The data were collected in an experiment to optimize crystal growth as a function of three variables x_1, x_2, and x_3. Large values of y (yield in grams) are desirable. Fit a second-order model and analyze the fitted surface. Under what set of conditions is maximum growth achieved?

x_1	x_2	x_3	y
−1	−1	−1	66
−1	−1	1	70
−1	1	−1	78
−1	1	1	60
1	−1	−1	80
1	−1	1	70
1	1	−1	100
1	1	1	75
−1.682	0	0	100
1.682	0	0	80
0	−1.682	0	68
0	1.682	0	63
0	0	−1.682	65
0	0	1.682	82
0	0	0	113
0	0	0	100
0	0	0	118
0	0	0	88
0	0	0	100
0	0	0	85

Today Exercises

Solve problems 11.12 from D. C. Montogomery DAOE.

Consider the three-variable central composite design. Analyze the data and draw conclusions, assuming that we wish to maximize conversion (y_1) with activity (y_2) between 55 and 60 achieved?

Run	Time (min)	Temperature (°C)	Catalyst (%)	Conversion (%) y_1	Activity y_2
1	−1.000	−1.000	−1.000	74.00	53.20
2	1.000	−1.000	−1.000	51.00	62.90
3	−1.000	1.000	−1.000	88.00	53.40
4	1.000	1.000	−1.000	70.00	62.60
5	−1.000	−1.000	1.000	71.00	57.30
6	1.000	−1.000	1.000	90.00	67.90
7	−1.000	1.000	1.000	66.00	59.80
8	1.000	1.000	1.000	97.00	67.80
9	0.000	0.000	0.000	81.00	59.20
10	0.000	0.000	0.000	75.00	60.40
11	0.000	0.000	0.000	76.00	59.10
12	0.000	0.000	0.000	83.00	60.60
13	−1.682	0.000	0.000	76.00	59.10
14	1.682	0.000	0.000	79.00	65.90
15	0.000	−1.682	0.000	85.00	60.00
16	0.000	1.682	0.000	97.00	60.70
17	0.000	0.000	−1.682	55.00	57.40
18	0.000	0.000	1.682	81.00	63.20
19	0.000	0.000	0.000	80.00	60.80
20	0.000	0.000	0.000	91.00	58.90