

# Response Surface Methodology

Author: John M. Cimbala, Penn State University  
Latest revision: 22 September 2008

## Introduction

- Taguchi design arrays were introduced previously as a way to hunt intelligently for an optimum point.
- Taguchi's design arrays are not infallible, however, and may not always be practical to set up, since they work best with fairly *large changes* of the parameters.
- There are *other* methods with which one can search or hunt for an optimum. One common method is called *response surface methodology (RSM)*.
- **RSM is particularly useful for optimizing a running system.** For example, suppose a power plant is running continuously. RSM can be used to optimize the efficiency or power output without large changes in the operating parameters, *while continuing the operation without interruption*.
- The main advantage of RSM is that it can be done "**on the fly**" with *small changes* of the parameters.
- A good reference for most of this material is Design and Analysis of Experiments, 4th ed., D. C. Montgomery, John Wiley and Sons, QA279.M66 1996, Ch. 14.

## Overview of the RSM Method

- Suppose  $y = y(a, b, c, \dots)$ , where  $y$  is the *outcome* or *result* or *response* that is to be optimized, and there are  $n$  parameters,  $a, b, c, \dots$ , which can be varied.
- In these notes, it is assumed that the optimum  $y$  is the *maximum*  $y$ . A similar analysis can be performed for *minimizing*  $y$ .
- **The goal of RSM is to efficiently hunt for the optimum values of  $a, b, c, \dots$  such that  $y$  is maximized.**
- RSM works by the *method of steepest ascent*, in which the *parameters are varied in the direction of maximum increase of the response* until the response no longer increases.
- RSM is best illustrated by examples. First, some simple *qualitative* examples are given with 1 and 2 parameters ( $n = 1$  and  $n = 2$ ), followed by a *quantitative* example with three parameters ( $n = 3$ ).

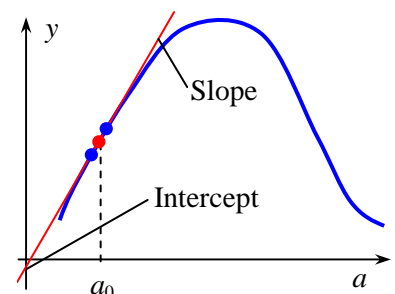
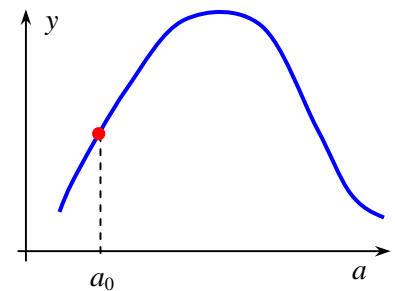
- **Example for  $n = 1$ :** [Note: This is only a *qualitative* example.]

**Given:** Response  $y = y(a)$ . Here,  $n = 1$  since there is only one parameter in the problem. The experiments begin at some arbitrary operating point, i.e., at some value of  $a$ ,  $a = a_0$ .

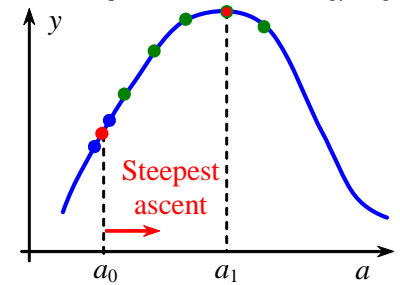
**To do:** Apply RSM to find the optimum value of  $y$ , i.e., find the value of parameter  $a$  where  $y$  is a maximum.

**Solution:**

- Imagine a plot of  $y$  as a function of  $a$ , with initial operating condition  $a_0$  indicated by the **red dot** on the plot to the right.
- Note that in an actual experiment, such a plot would not be available, but is used here for illustrative purposes only.
- With only one parameter ( $n = 1$ ) as in this simple example, all one needs to know is whether to search to the left or to the right. One of these is the *direction of steepest ascent*.
- To find this direction of steepest ascent, we measure  $y$  at some data points *in the vicinity of  $a_0$* , as illustrated by the **blue dots** in the plot to the right. Two other data points, in addition to that at  $a = a_0$  itself, are sufficient for this simple case.
- If there is a lot of scatter in the data, some *replication* may be required – **take several measurements at the same value of  $a$ , and use the average measurement – this minimizes the effect of scatter.**
- A *linear regression analysis* is performed on the data, and the  $y$  intercept and slope are found, as sketched.
  - If the slope is *positive* ( $y$  increases with  $a$ ), the direction of steepest ascent is to the *right* (towards increasing  $a$ ).
  - If the slope is *negative* ( $y$  decreases with  $a$ ), the direction of steepest ascent is to the *left* (towards decreasing  $a$ ).
- In this qualitative example, the slope is positive, so the direction of steepest ascent is to the right.
- As seen on the above plot, this slope is very accurate near  $a = a_0$ , but the inaccuracy increases when parameter  $a$  is far from  $a_0$ . (At the maximum point, of course, the slope is actually *zero*.)



- Nevertheless, **it is assumed that the direction of steepest ascent points in the direction of the desired goal.**
- So, we then **march** in that direction (with some reasonable increment of parameter  $a$ ) **until  $y$  stops increasing** ( $y$  levels off or starts decreasing), as sketched to the right.
- After a few experimental runs, defined by the **green dots** on the plot, it is found that  $y$  reaches a maximum (the second **red dot**) at  $a = a_1$ .
- This operating point (at  $a = a_1$ ) may not be the **true optimum**, but it is certainly much closer than was the original operating point (at  $a = a_0$ ).
- We then **repeat the RSM procedure**, using  $a = a_1$  as the new starting point, to efficiently hunt for the optimum operating point.



**Discussion:** In this simple example, we do not need to proceed any further, but the *principles* apply to an optimization problem with *any number* of parameters. Next, a more complicated case is examined.

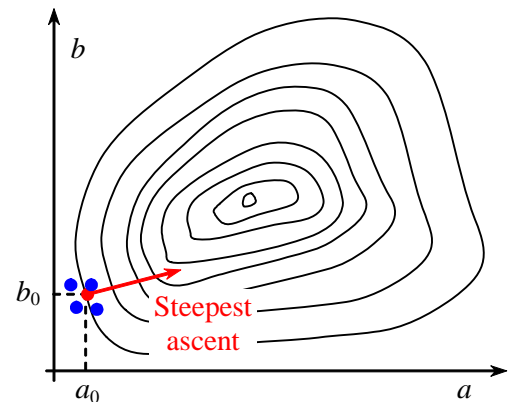
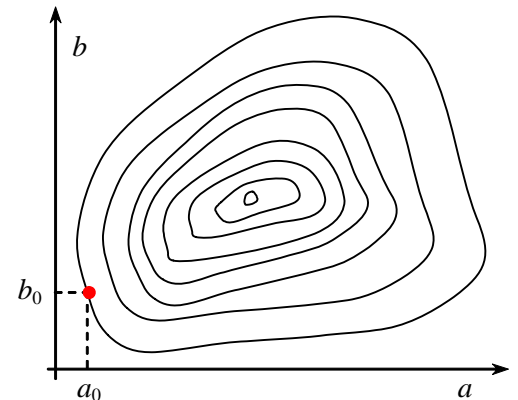
• **Example for  $n = 2$ :** [Note: This is also only a *qualitative* example.]

**Given:** Response  $y = y(a, b)$ . Here,  $n = 2$  since there are two parameters in the problem. The experiments begin at some arbitrary operating point – at  $a = a_0$  and  $b = b_0$ .

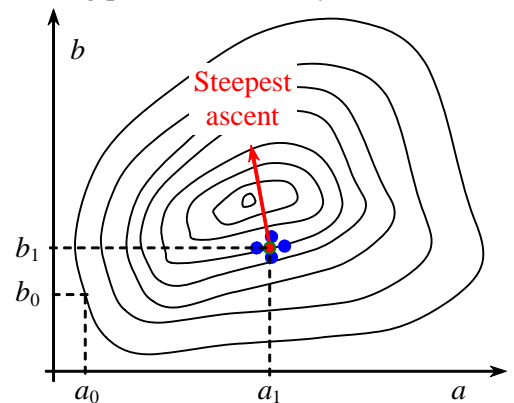
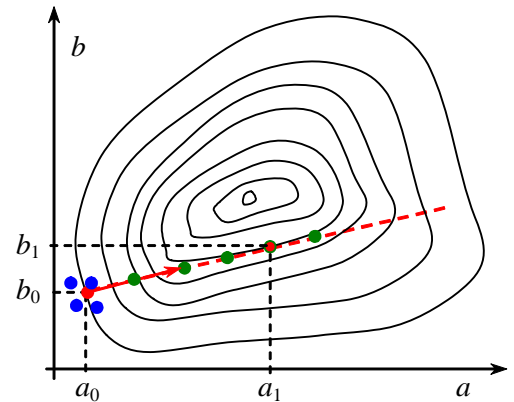
**To do:** Apply RSM to find the optimum value of  $y$ , i.e., find the value of parameters  $a$  and  $b$  where  $y$  is a maximum.

**Solution:**

- Imagine a **contour plot** of  $y$  as a function of  $a$  and  $b$ , as sketched to the right. The maximum value of  $y$  is in the center of the contours. The initial operating condition at  $(a_0, b_0)$  is indicated by the **red dot**:
- Note that in an actual experiment, such a plot would not be available, but is used here for illustrative purposes only.
- With two parameters ( $n = 2$ ), we need to find the **direction of steepest ascent**, which is now a **vector**, in the  $a$ - $b$  plane.
- To find this direction of steepest ascent, we measure  $y$  at several data points **in the vicinity of  $(a_0, b_0)$** , as illustrated by the **blue dots** on the contour plot to the right.
- **At least three data points**, in addition to that at  $(a_0, b_0)$  itself, are required for a reasonable linear regression when  $n = 2$ . Four are shown here, but we recommend more than that if time and cost permit.
- If there is scatter in the data, replication may be required.
- A **multiple variable linear regression analysis** is performed on the data. (Multiple variable regression analysis is discussed in a previous learning module.)
- The  $y$  intercept and **two slopes** are found (with respect to parameter  $a$  and with respect to parameter  $b$ ). The **vector produced by these slopes** is sketched above.
  - If the slope with respect to parameter  $a$  is *positive* ( $y$  increases with  $a$ ), the direction of steepest ascent is to the *right* (towards increasing  $a$ ).
  - If the slope with respect to parameter  $a$  is *negative* ( $y$  decreases with  $a$ ), the direction of steepest ascent is to the *left* (towards decreasing  $a$ ).
  - If the slope with respect to parameter  $b$  is *positive* ( $y$  increases with  $b$ ), the direction of steepest ascent is to the *top* (towards increasing  $b$ ).
  - If the slope with respect to parameter  $b$  is *negative* ( $y$  decreases with  $b$ ), the direction of steepest ascent is to the *bottom* (towards decreasing  $b$ ).
- With  $n = 2$  parameters, however, we know more than simply the *direction* (right or left, up or down). We actually know the **vector direction of steepest ascent**, as sketched above.
- Here the direction of steepest ascent is to the right and up, at some small angle above the horizontal.
- As seen on the plot, this direction is very accurate near  $(a_0, b_0)$ , but becomes more inaccurate when parameter  $a$  is far from  $a_0$  or parameter  $b$  is far from  $b_0$ . (At the maximum point, of course, both slopes are zero.)
- Nevertheless, **it is assumed that the direction of steepest ascent points toward the desired goal.**



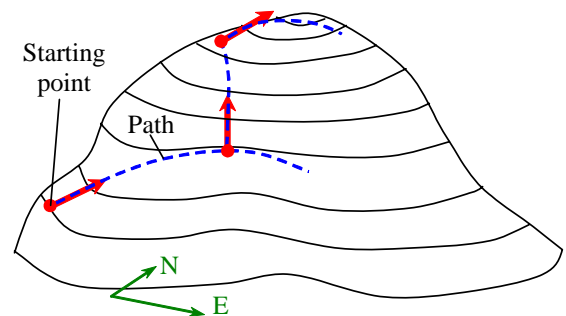
- We must now choose some reasonable increments for both parameters. **In practice, the increment for one of the parameters is chosen or picked, and that of the other parameter(s) is calculated based on the direction of steepest ascent.**
- Typically, **the parameter with the largest magnitude of slope is the one for which the increment is chosen**, but this is not a hard and fast rule. For the sake of argument, let us assume this is parameter  $a$ .
- Increment  $\Delta a$  is chosen – not too small or too large, but a reasonable value, based on the experiment.
  - If the slope of  $y$  with respect to parameter  $a$  is positive, a positive value of  $\Delta a$  is chosen.
  - If the slope of  $y$  with respect to parameter  $a$  is negative, a negative value of  $\Delta a$  is chosen.
- Once the increment in parameter  $a$  is chosen, the increment in parameter  $b$  is *calculated*, such that the vector  $(\Delta a, \Delta b)$  points in the *direction of steepest ascent*.
- We now **march in the direction of steepest ascent** (the **green dots** in the sketch to the right) **until  $y$  stops increasing** ( $y$  levels off and then starts decreasing).
- After a few experimental runs  $y$  reaches a maximum (the second **red dot**) at  $(a_1, b_1)$ . (The next **green dot** has a lower value of  $y$ , as can be seen from the contour lines.)
- This operating point  $(a_1, b_1)$  is not the **true optimum**, but it is much closer to the optimum than was the original operating point  $(a_0, b_0)$ .
- We then **repeat the RSM procedure**, using  $(a_1, b_1)$  as the new starting point, to efficiently hunt for the optimum operating point.
- On the second pass, a *new* direction of steepest ascent is calculated from measurements at data points in the vicinity of  $(a_1, b_1)$ , as indicated by the blue dots in the figure to the right. Here, four measurements are taken surrounding the red dot at  $(a_1, b_1)$ .
- In the sketch, the *new* **direction of steepest ascent** is indicated by the **red arrow**.
- We now march in this new direction until the response,  $y$ , starts to decrease, as previously.
- *Note:* as the operating point gets closer and closer to the optimum point, we typically choose a **smaller increment** in the parameters.
- As seen on the sketch, the true optimum point is still not found on this second pass, but we are closer!



**Discussion:** This example is still only qualitative, but it helps us visualize the RSM process.

### The Mountain Analogy

- Here is an analogy to help us better understand the philosophy of RSM.
- Imagine a person walking up a mountain. Furthermore, suppose that it is very *foggy* on the mountain. The climber cannot see the mountain peak, and has only an **altimeter** and a **compass** to guide him. (An altimeter measures altitude or elevation.) His goal is to reach the top of the mountain.
- From some arbitrary starting point, he can determine the **direction of steepest ascent** by measuring the elevation at several points close by. He notes the direction in which the compass points when he moves in the direction of steepest ascent, which is **the direction of most rapid increase in elevation**.
- Then he walks in that direction (**red arrow** at the starting point), checking his altimeter each step of the way. The altimeter reading increases as he climbs – he is making progress.
- He continues to climb in that same direction, with the help of his compass, until his altimeter shows *decreasing* altitude. Obviously, his first direction of steepest ascent is



no longer the correct direction of steepest ascent! (This is because the direction was determined only *locally*, around the original starting point, and he has now moved away from that point.)

- At this point, he stops, determines a *new* direction of steepest ascent (second **red arrow**), and moves along that direction instead. The process is repeated.
- Eventually, he will reach the top of the mountain (in a fairly efficient manner).
- His path up the mountain is illustrated in the sketch by the **dashed blue lines** (compass directions of north and east are shown to provide a perspective view). The isocontour lines are contours of constant elevation.
- The path to the top may not be the shortest possible, but it eventually gets him there – certainly faster than by trial and error!

### Coded Variables

- In general, the units of parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , etc. differ from each other. Even if some of the parameters have the same units, not all of these parameters will be tested over the *same range*. For example, parameter  $a$  may be a voltage, with a range of 4 V, parameter  $b$  may be a length, with a range of 0.10 m, parameter  $c$  may be a length, with a range of 2.0 m, etc.
- Since parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , etc. have different *units* and/or different *ranges* in the experimental data set, **regression analysis should not be performed on the raw, physical (dimensional) parameters themselves**. Instead, one must first **normalize** the parameters  $a$ ,  $b$ ,  $c$ ,  $d$ , etc. *before* performing a regression analysis.
- The normalized variables are called **coded variables** – **dimensionless variables that range from  $-1$  to  $1$** .
- Instead of using values of  $a$ ,  $b$ ,  $c$ , etc. directly in the regression analysis, coded variables,  $x_1$ ,  $x_2$ ,  $x_3$ , etc. are used as the independent variables in the regression analysis. A coded variable must be defined for *each* of the physical variables:
  - $x_1$  is defined for parameter  $a$ ,
  - $x_2$  is defined for parameter  $b$ ,
  - $x_3$  is defined for parameter  $c$ , etc.

- **Each of the coded variables is forced to range from  $-1$  to  $1$ , so that they all affect response  $y$  more evenly, and so the units and ranges of parameters  $a$ ,  $b$ ,  $c$ , etc. are irrelevant.**
- To convert parameter  $a$  to its coded variable  $x_1$ , the following formula is applied to each value of  $a$  in the

data set:  $x_1 = 2 \left( \frac{a - a_{\text{mid value}}}{a_{\text{range}}} \right)$ , where  $a_{\text{mid value}}$  is the **middle value of  $a$**  in the data set,  $a_{\text{mid value}} = \frac{a_{\text{min}} + a_{\text{max}}}{2}$

and  $a_{\text{range}}$  is the **range of parameter  $a$**  from its minimum to its maximum,  $a_{\text{range}} = a_{\text{max}} - a_{\text{min}}$ .

- You can verify that  $a_{\text{min}}$  maps into coded variable  $x_1 = -1$ ,  $a_{\text{mid value}}$  maps into coded variable  $x_1 = 0$ , and  $a_{\text{max}}$  maps into coded variable  $x_1 = 1$ , for *any* set of values of parameter  $a$ , regardless of its range or units.
- A similar equation is applied to the other parameters ( $b$ ,  $c$ , ...) in the problem. For example, coded variables

$x_2$  and  $x_3$  are defined for each value of parameter  $b$ ,  $x_2 = 2 \left( \frac{b - b_{\text{mid value}}}{b_{\text{range}}} \right)$  and  $c$ ,  $x_3 = 2 \left( \frac{c - c_{\text{mid value}}}{c_{\text{range}}} \right)$ .

- **Regression analysis must be performed on  $y$  as a function of  $x_1$ ,  $x_2$ ,  $x_3$ , etc., not on the original variables, and the slopes with respect to these coded variables are used to determine the direction of steepest ascent.**
- When using coded variables, the vector of steepest ascent must then be converted back to the original (uncoded) **physical variables**, using the inverse of the above equations. In other words, the steepest ascent increment in  $x_1$  is converted to the corresponding increment in  $a$ , the steepest ascent increment in  $x_2$  is converted to the corresponding increment in  $b$ , etc., so that the marching process can be performed in physical variables.
- For example, once the increment  $\Delta x_1$  in coded variable  $x_1$  is chosen, we must convert back to increment  $\Delta a$

in the physical (uncoded) parameter  $a$ , using the following equation:  $\Delta a = \frac{a_{\text{range}}}{2} \Delta x_1$ , with similar equations

for the increments in  $b$ ,  $c$ , ... etc.

- Response surface methodology (RSM) and the use of coded variables is best illustrated by example.

#### **Example:**

**Given:** A factory is continuously producing a product on an assembly line. Three parameters,  $a$ ,  $b$ , and  $c$ , which may improve the quality of the product, are identified. *Note:* To simulate the experiment, a function  $y(a, b, c)$  has been created in an Excel spreadsheet. The function has nonlinear cross-talk

between the three parameters, and has random noise to simulate experimental error. The maximum value of the function is 300, which occurs at  $(a, b, c) = (15, 10, -5)$ . For this example, *pretend that this information is not known*.

- The system is currently operating at  $(a_0, b_0, c_0) = (20, 6, -2)$ . At this operating point, the goodness or quality  $y$  of the product is about 260.0 (in units which represent some measure of the quality).
- The manager wants to improve the quality of the product (increase  $y$ ) *without shutting down the operation*, and with only *small changes at a time* to the parameters, so as not to disrupt the production.

**To do:** Use RSM to optimize (maximize) quality  $y$ .

**Solution:**

- Here,  $n = 3$  (3 parameters), since  $y = y(a, b, c)$ .
- Following the RSM procedure, a linear regression analysis is to be performed on a cluster of data points in the vicinity of the initial operating point.
- A 3-level, 3-parameter experimental design array is constructed, *centered around the initial operating point*. Note: A 3-level, 3-parameter Taguchi design array with each level of each parameter repeated twice is chosen here to generate the data cluster. It requires 6 data points (runs) in the vicinity of the initial operating point. The initial operating point value is also included in the regression analysis below, since that data point is also readily available.
- It should be pointed out that in practice, a formal Taguchi design array may not be necessary. Instead, some random variations of each variable around the operating condition can produce an effective regression analysis. In cases where it is expensive and/or time consuming to change parameters, however, a Taguchi design array is advised in order to minimize the number of necessary runs.
- Note also that for cases with large scatter in the data, several runs at a given operating condition should be recorded, so that the effect of random noise can be somewhat cancelled out.
- For this simple example,
  - An increment of 2 units is chosen between the low and medium levels and between the medium and high levels for parameter  $a$  – the levels for parameter  $a$  are chosen to be 18, 20, and 22.
  - An increment of 0.5 units is chosen between the low and medium levels and between the medium and high levels for parameter  $b$  – the levels for parameter  $b$  are chosen to be 5.5, 6, and 6.5.
  - An increment of 1 unit is chosen between the low and medium levels and between the medium and high levels for parameter  $c$  – the levels for parameter  $c$  are chosen to be  $-3$ ,  $-2$ , and  $-1$ .

	original parameters			response	coded variables		
	$a$	$b$	$c$	$y$	$x_1$	$x_2$	$x_3$
<b>operating point</b>	20	6	-2	260.412	0	0	0
	18	5.5	-3	274.883	-1	-1	-1
	18	6	-2	274.376	-1	0	0
	20	6.5	-1	258.338	0	1	1
	20	5.5	-2	257.051	0	-1	0
	22	6	-1	234.401	1	0	1
	22	6.5	-3	247.363	1	1	-1
<b>mimimum</b>	18	5.5	-3		-1	-1	-1
<b>maximum</b>	22	6.5	-1		1	1	1
<b>range</b>	4	1	2		2	2	2
<b>mid value</b>	20	6	-2		0	0	0

- The data appear to the right (the image is taken from an Excel spreadsheet), along with coded variables  $x_1$ ,  $x_2$ , and  $x_3$  which have also been calculated in the spreadsheet.
- A regression analysis is performed with the seven  $y$  values above (highlighted in turquoise) selected for Excel's *Input Y Range*, and with the values of coded variables  $x_1$ ,  $x_2$ , and  $x_3$  (highlighted in yellow) selected for Excel's *Input X Range*. (A total of 21  $X$  cells are highlighted – all seven rows of the coded variables.)

Intercept	258.1154076
X Variable 1	-17.4966747
X Variable 2	3.150786002
X Variable 3	-3.80031225

- A portion of Excel's regression analysis output is shown to the right, showing the calculated intercept and slopes, which Excel calls "X Variable 1", "X Variable 2", and "X Variable 3".
- The results of interest are the slopes of  $y$  with respect to coded variables  $x_1$ ,  $x_2$ , and  $x_3$ , which are approximately  $-17.5$ ,  $3.15$ , and  $-3.80$  respectively, to three significant digits.

- Mathematically, these are **partial derivatives** of  $y$  with respect to  $x_1$ ,  $x_2$ , and  $x_3$ ,  

$$\frac{\partial y}{\partial x_1} = -17.5, \quad \frac{\partial y}{\partial x_2} = 3.15, \quad \text{and} \quad \frac{\partial y}{\partial x_3} = -3.80.$$
- These three values represent the **direction of steepest ascent** in terms of coded variables – **the direction of steepest ascent is in the direction of vector**  $\left( \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \frac{\partial y}{\partial x_3} \right)$ .
- For this sample set of data, the direction of steepest ascent is in the direction of the vector  $(-17.5, 3.15, -3.80)$ .
- The **magnitude** of the vector is unimportant – it is the **direction** that really matters. Some authors like to convert the vector into a **unit vector** by dividing each component by the magnitude of the vector, but this is not really necessary. The unit vector for the above example is  $(-0.962, 0.173, -0.209)$ , to three significant digits.
- Since coded variable  $x_1$  (corresponding to uncoded physical parameter  $a$ ) has the largest magnitude of slope in this example, coded variable  $x_1$  is chosen as the parameter for which to pick an increment.
- An increment of  $-1.0$  unit for coded variable  $x_1$  is chosen here,  $\Delta x_1 = -1$ .
- Note:* Since the sign of the slope of  $y$  with respect to coded variable  $x_1$  is **negative**, a **negative** increment in coded variable  $x_1$  is chosen. If the slope had been positive, the increment in coded variable  $x_1$  would have been positive instead.
- Based on this selected increment in coded variable  $x_1$ , the increment in the other two coded variables  $x_2$  and  $x_3$  must be calculated. This must be done carefully such that the direction of marching is the **same** as the direction of steepest ascent.
- In other words, the increments in coded variables  $x_2$  and  $x_3$  are selected such that vector  $(\Delta x_1, \Delta x_2, \Delta x_3)$  points in the **same direction** as the vector  $\left( \frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \frac{\partial y}{\partial x_3} \right)$ .
- Simple formulas can be constructed to ensure proportionality of these vectors, namely,  

$$\Delta x_2 = \Delta x_1 \frac{\frac{\partial y}{\partial x_2}}{\frac{\partial y}{\partial x_1}} \quad \text{and} \quad \Delta x_3 = \Delta x_1 \frac{\frac{\partial y}{\partial x_3}}{\frac{\partial y}{\partial x_1}},$$
from which we can calculate the increment in coded variables  $x_2$  and  $x_3$  for a chosen value of the increment in coded variable  $x_1$ .
- Using the above equations, the increments in coded variables  $x_2$  and  $x_3$  are calculated for our example to be  $\Delta x_2 = \Delta x_1 \frac{\frac{\partial y}{\partial x_2}}{\frac{\partial y}{\partial x_1}} = (-1) \frac{3.15}{-17.5} = 0.180$  and  $\Delta x_3 = \Delta x_1 \frac{\frac{\partial y}{\partial x_3}}{\frac{\partial y}{\partial x_1}} = (-1) \frac{-3.80}{-17.5} = -0.217$ , to three significant digits.
- Notice that the **signs** (+ or -) of the increments are automatically taken care of, provided that the signs of the derivatives and the increments are included properly in the calculations, as above.
- We must recognize that these are increments in the **coded** variables. To march in the direction of steepest ascent, we must first convert back to increments in the **physical** (dimensional) variables,  $a$ ,  $b$ , and  $c$ .
- Using the conversions defined above, the increments in parameters  $a$ ,  $b$ , and  $c$  are calculated to be  

$$\Delta a = \frac{a_{\text{range}}}{2} \Delta x_1 = \frac{4}{2}(-1) = -2, \quad \Delta b = \frac{b_{\text{range}}}{2} \Delta x_2 = \frac{1}{2}(0.180) = 0.090, \quad \Delta c = \frac{c_{\text{range}}}{2} \Delta x_3 = \frac{2}{2}(-0.217) = -0.217,$$
which are rounded off to  $-2.0$ ,  $0.1$ , and  $-0.2$  respectively for convenience and simplicity in the next step of the RSM calculations.
- The **direction of steepest ascent** in the original (dimensional, uncoded, physical) variables has now been found – the direction in which parameters  $a$ ,  $b$ , and  $c$  are incremented by approximately  $-2.0$ ,  $0.1$ , and  $-0.2$ , respectively, that is,  $\Delta a = -2$ ,  $\Delta b = 0.1$ , and  $\Delta c = -0.2$ .
- Now all three parameters are incremented **simultaneously**, and  $y$  is measured. This **marching** process is continued until  $y$  begins to decrease. The Excel cells that show this are indicated to the right.

$a$	$b$	$c$	$y$
20	6	-2	260.407
18	6.1	-2.2	275.811
16	6.2	-2.4	284.519
14	6.3	-2.6	286.611
12	6.4	-2.8	282.024
10	6.5	-3	270.766
8	6.6	-3.2	252.892

- Notice that the starting point is at (20, 6, -2), the original operating conditions, at which the result (y) is approximately 260. Parameter *a* is incremented by -2.0, parameter *b* by 0.1, and parameter *c* by -0.2. The marching is done with *all three variables simultaneously, in the direction of steepest ascent*.
- The response, *y*, increases with each increment until the **highlighted row**, after which it begins to decrease. The response surface methodology (RSM) procedure has successfully increased response *y* from an initial value of around 260 to a new value of around 287. *We have improved the response*.
- The new suggested operating condition is at (*a, b, c*) = (14.0, 6.3, -2.6), where *y* is around 287. This can be compared with the *actual* optimum (which we are pretending that we don't know), at (*a, b, c*) = (15, 10, -5), where *y* is 300. While not at the exact optimum, the operation has moved much closer than previously.

	original parameters			response	coded variables		
	<i>a</i>	<i>b</i>	<i>c</i>	<i>y</i>	<i>x</i> <sub>1</sub>	<i>x</i> <sub>2</sub>	<i>x</i> <sub>3</sub>
<b>operating point</b>	<b>14</b>	<b>6.3</b>	<b>-2.6</b>	<b>286.595</b>	<b>0</b>	<b>0</b>	<b>-0</b>
	13	6	-2.9	283.785	-1	-1	-1
	13	6.3	-2.6	284.673	-1	0	-0
	14	6.6	-2.3	287.187	0	1	1
	14	6	-2.6	285.035	0	-1	-0
	15	6.3	-2.3	286.063	1	0	1
	15	6.6	-2.9	289.267	1	1	-1
<b>mimimum</b>	<b>13</b>	<b>6</b>	<b>-2.9</b>		<b>-1</b>	<b>-1</b>	<b>-1</b>
<b>maximum</b>	<b>15</b>	<b>6.6</b>	<b>-2.3</b>		<b>1</b>	<b>1</b>	<b>1</b>
<b>range</b>	<b>2</b>	<b>0.6</b>	<b>0.6</b>		<b>2</b>	<b>2</b>	<b>2</b>
<b>mid value</b>	<b>14</b>	<b>6.3</b>	<b>-2.6</b>		<b>0</b>	<b>0</b>	<b>-0</b>

- At this point, *further optimization is performed by starting RSM over at this new operating condition*.

- A new Taguchi design array is constructed locally, centered around the new operating point (14.0, 6.3, -2.6) as shown to the right.

- Notice that a *smaller* increment in each variable was chosen for this second round of tests (e.g., 1.0 instead of 2.0 for

parameter *a*), since we are closer to the optimum point, and will need to calculate the slopes more precisely this time.

- Note also that coded variables *x*<sub>1</sub>, *x*<sub>2</sub>, and *x*<sub>3</sub> are once again calculated for each *a, b*, and *c* in the new experimental data set.
- From these data, a second regression analysis is performed.
- The results of interest are the slopes of *y* with respect to coded variables *x*<sub>1</sub>, *x*<sub>2</sub>, and *x*<sub>3</sub>, which are now 1.11, 1.51, and -0.601

respectively,  $\frac{\partial y}{\partial x_1} = 1.11$ ,  $\frac{\partial y}{\partial x_2} = 1.51$ , and  $\frac{\partial y}{\partial x_3} = -0.601$ , as shown to the right.

	<i>yy</i>
Intercept	286.0844914
X Variable 1	1.10810072
X Variable 2	1.505755348
X Variable 3	-0.601358

- This time, coded variable *x*<sub>2</sub> has the largest magnitude of slope. Coded variable *x*<sub>2</sub> (or its corresponding physical variable *b*) is thus the variable for which we choose the increment.

- We select an increment of  $\Delta b = 0.1$ , for which the increment for coded variable *x*<sub>2</sub> is calculated as

$$\Delta x_2 = 2 \frac{\Delta b}{b_{\text{range}}} = 2 \frac{0.1}{0.6} = 0.333.$$

- Based on this selected increment in *x*<sub>2</sub>, the increment in the other two coded variables must be calculated in the same manner as above,  $\Delta x_1 = \Delta x_2 \frac{\partial y}{\partial x_1} / \frac{\partial y}{\partial x_2} = (0.333) \frac{1.11}{1.51} = 0.245$  and

$$\Delta x_3 = \Delta x_2 \frac{\partial y}{\partial x_3} / \frac{\partial y}{\partial x_2} = (0.333) \frac{-0.601}{1.51} = -0.133, \text{ to three significant digits.}$$

- A new *direction of steepest ascent* from this new operating point has now been found, in terms of coded variables. We must now convert back to physical variables:  $\Delta a = \frac{a_{\text{range}}}{2} \Delta x_1 = \frac{2}{2} (0.245) = 0.245$ , and

$$\Delta c = \frac{c_{\text{range}}}{2} \Delta x_3 = \frac{0.6}{2} (-0.133) = -0.040. \text{ [Recall that } \Delta b \text{ has already been chosen as } \Delta b = 0.1.]$$

- These increments in the physical variables are rounded off to  $\Delta a = 0.25$ ,  $\Delta b = 0.1$ , and  $\Delta c = -0.04$  respectively for convenience and simplicity in the next step of the RSM calculations.
  - Thus, the direction of steepest ascent in terms of physical parameters  $a$ ,  $b$ , and  $c$  is the direction in which parameters  $a$ ,  $b$ , and  $c$  are incremented by approximately  $\Delta a = 0.25$ ,  $\Delta b = 0.1$ , and  $\Delta c = -0.04$ .
  - Now all three parameters are incremented *simultaneously*, and  $y$  is measured, as previously. This process is continued until  $y$  begins to decrease. The Excel cells which show this are indicated to the right.
  - Notice that the starting point is at (14.0, 6.3, -2.6), the second set of operating conditions. Parameter  $a$  is incremented by 0.25 units, parameter  $b$  by 0.1 units, and parameter  $c$  by -0.04 units. The marching is done by **incrementing all three variables simultaneously, in the direction of steepest ascent**.
  - The response,  $y$ , increases with each increment until the **highlighted row**, after which it begins to decrease. The response surface methodology (RSM) procedure has successfully increased response  $y$  from an initial value of around 260 to a value of around 287, and now to a value of around 290.
  - Further optimization can be performed by starting RSM over at this new maximum operating condition.
  - At this point, since the optimum value is probably not too far away, we should use even smaller increments in the parameters (take data *closer* to the current operating point) in order to determine the direction of steepest ascent more accurately.
  - Note that one soon encounters **diminishing returns**. The true optimum may never be found precisely using RSM, due to randomness in the response.
  - When one is close to the optimum, it is wise to either use a larger design array, or to **replicate – repeat the measurements several times before performing the regression analysis**. This reduces the effect of random noise errors.
  - In this example, the RSM procedure was performed one more time (not shown, but available on the Excel spreadsheet that goes along with this learning module. The final optimum value of result  $y$  is around 298.5 at operating conditions (13.9, 9.2, -4.6).
  - This result is compared to the (known) analytical result,  $y_{\max} = 300.0$  at (15, 10, -5). [Recall that an analytical function, with random noise, was created for the Excel spreadsheet in order to simulate this experiment. The RSM was performed, however, without any knowledge of the analytical result.]
- Discussion:** The RSM result is not perfect, but the new operating conditions are *very close* to the actual optimum values. **RSM has been used successfully to optimize result  $y$  in this example.**

$a$	$b$	$c$	$y$
<b>14</b>	<b>6.3</b>	<b>-2.6</b>	<b>286.61</b>
14.25	6.4	-2.64	287.43
14.5	6.5	-2.68	288.20
14.75	6.6	-2.72	288.85
15	6.7	-2.76	289.36
15.25	6.8	-2.8	289.78
15.5	6.9	-2.84	290.13
15.75	7	-2.88	290.33
<b>16</b>	<b>7.1</b>	<b>-2.92</b>	<b>290.44</b>
16.25	7.2	-2.96	290.42
16.5	7.3	-3	290.29
16.75	7.4	-3.04	290.07

### Final Comments

- In the above example, the **point of diminishing returns** of the RSM technique may have been reached.
- Further improvement is possible only by the following techniques:
  - Take data around a **smaller region** in the vicinity of the current operating point. This increases the accuracy of using linear regression analysis to calculate the direction of steepest ascent.
  - **Replicate** the data. This helps to cancel the effect of random noise (experimental error).
  - Use a **higher-order regression scheme**. Note that here only a *linear* (first-order) regression analysis has been used. We can instead use a second-order or higher-order regression analysis (as discussed in a previous learning module). Some RSM schemes have been devised which can even take into account cross-talk between variables. These are beyond the scope of this learning module.
- A caution about response surface methodology must be given here:
  - RSM will always find a **local** maximum response. If there is more than one peak in the function, one of the *other* peaks may have a larger value of  $y$ . In other words, **the local maximum response determined by RSM may not necessarily be the optimum response!**
  - Returning to the mountain analogy above, the mountain climber may reach a local peak in the mountain, but when the fog clears, he sees that there is an even higher peak somewhere else!
  - If this happens, a *different starting point* must be selected, and the RSM process must be repeated.
- Overall, however, RSM is a very powerful technique for optimizing a response.