Optimization of a Multivariate Chemical Reaction by a Simplex Process

This experiment was developed by Dr. Edward Navarre at Southern Illinois University Edwardsville.
https://acpo.github.io/SimplexOptimization/

Objectives:

to optimize the experimental conditions of a simulated chemical reaction
to use the Simplex method to perform reaction optimization

References

Shavers, C. L.; Parsons, M. L.; Deming, S. N. Simplex Optimization of Chemical Systems. *J. Chem. Educ.* **1979**, *56* (5), 307.
Leggett, D. J. Instrumental Simplex Optimization: Experimental Illustrations for an Undergraduate Laboratory Course. *J. Chem. Educ.* **1983**, *60* (9), 707.

Background Readings

Analytical Chemistry 2.1. David Harvey. Chapter 14 section A (pages 912 to 915) available at
https://sites.google.com/depauw.edu/dth/analytical-curriculum-projects/ac2-1

Apparatus and Chemicals

This is an at-home experiment without apparatus.

Safety

No safety precautions are required.

Theory

One aspect of experimental design that a chemist will encounter is deciding the correct amount (as concentration, mass, *etc.*) of reagents necessary to optimize a chemical process. Optimization can focus on getting the greatest yield, using the least time, or having the best economic outcomes.

If there were only two *factors* (variables) in the chemical process, for example A + B + dye ⇌ products, then it would be relatively simple to perform a full factorial design (or fully crossed design) experiment in which all possible values of [A] and [B] were explored. With two factors at six concentrations each, only $6^2 = 36$ measurements would be required. Searching over a greater number of concentrations would expand the experimental time requirement by the square of the number of concentrations. However, if there were more factors, let's say [A], [B], pH, and temperature, the experiment becomes much more complex. Again let's assume six concentrations of [A] and [B], six pH values, and six temperatures so that a full factorial design would require $6^4 = 1296$ measurements. We can readily and correctly guess that most of the 1296 values are far from the optimum condition and that a more efficient approach is needed.

Simplex optimization is one approach to efficiently search within the factor space (the matrix of all possible combinations of factor values) to locate a maximum or minimum response. There are five rules to the Simplex method and you will need to read about them from the literature (Leggett, 1983). You have some choice in terms of what Simplex approach to use. The simplest is the *fixed size simplex* and somewhat more advanced is the *variable size simplex*. How you do the work also has a choice: it is easiest to draw the Simplex diagram on graph paper (*e.g.*, in your notebook) and identify experimental conditions from the plot; or you can do the simplex calculations algebraically if the math in the Leggett article makes sense to you. In both cases, we trade efficiency of the method for simplicity of the method, so all of the approaches require a similar amount of effort.

The chemical system that you will use is affected by pH and temperature. Because the chemical system is an aqueous solution, the pH can range from 0 to 14 and the temperature from 278 to 365 K. Although the specific chemical system is not named for this experiment, it is a good example of many aqueous reactions that have kinetic and thermodynamic dependence on pH and temperature. A prominent example of this type is the reaction of glycoside hydrolases with polysaccharides, especially in the production of ethanol by fermentation and refining of cellulosic materials.

The goal of this experiment is to maximize the yield of the reaction. When ranking the steps in the Simplex optimization you will use % yield as the measure of effectiveness.

Procedure

1. **Find you data**. You will be sent an Excel spreadsheet file in .xlsx format. The file contains calculations that allow you to do this experiment. It is strongly recommended that you work with the file in Excel (either an installed version or the on-line Excel365). The spreadsheet has two spaces marked in yellow for entering the experimental conditions. The space marked in green will return the yield for the conditions that you enter in the yellow spaces.
2. **Set up the factor space.** Start with a factor space that covers pH 0 to 14 and temperature 278 to 365 K. The instructor will give you a starting point pair of values within that space for the first set of conditions (see the graph below for a rough idea of how the factor space relates to your solutions).
3. **Measure the first reaction condition.** Enter the initial conditions that you are given and record the yield reported by the spreadsheet. This is the first vertex in the Simplex process.
4. **Pick the next two conditions.** What you pick next is up to you. There are some guides that can help: 1) don't move along a single axis in the factor space, doing so only changes one factor so it doesn't use the method effectively; 2) don't make giant or tiny steps, think about what the overall Simplex process looks like and pick steps sizes that will allow you to get near a maximum in about 12 steps; 3) don't pick three points in a straight line, you need triangular shapes in the factor space. Consult the reference readings for some help. Record the conditions and yield that you selected.
5. **Start measuring.** The objective of this experiment is that you come to understand how to work with the Simplex method. Since you are not required to make each combination of experimental conditions in a laboratory, the "experimental" time is very short. You have plenty of time to think carefully about how to make the Simplex process work. After finding the yield of the first 3 reaction conditions, you will add the yield values to the Simplex plot at the first three vertices. The Simplex process has you rank the three vertices and use the outcome to select the next set of conditions.
6. **When am I done?** You need to find the global maximum – so more than four sets of conditions; probably more than ten.

Data Analysis

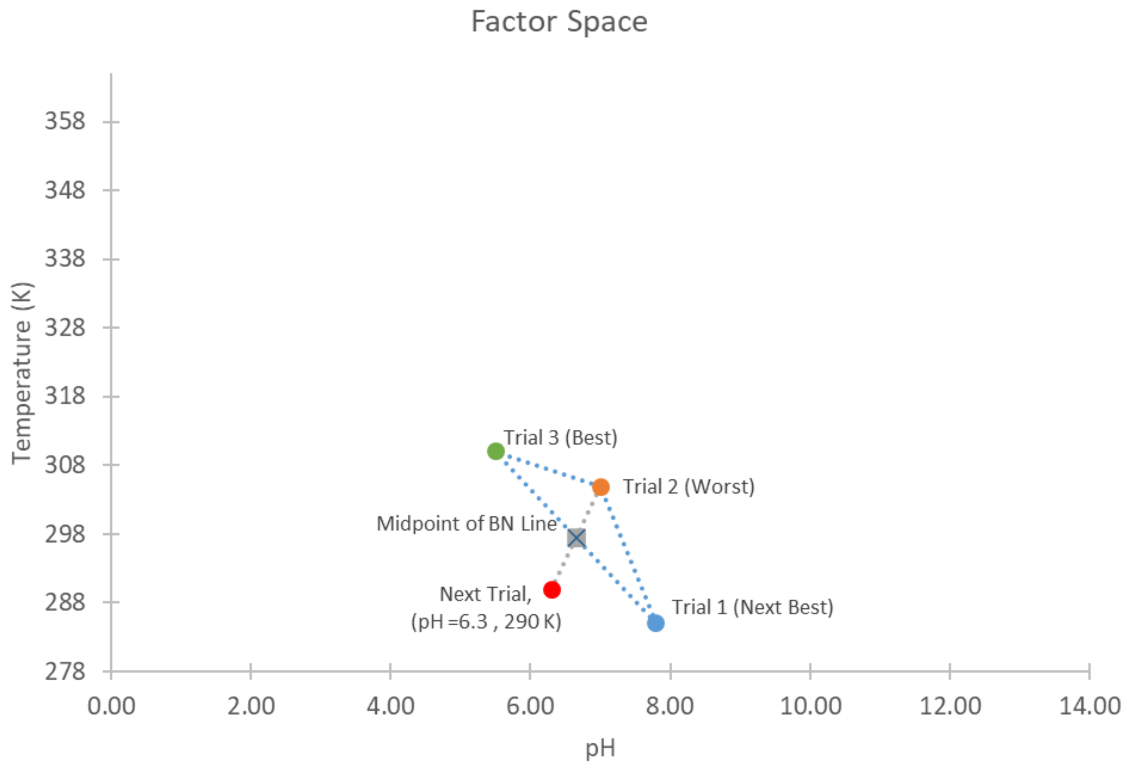
1) Report the pH, temperature, and % yield for each condition that you test.

2) Plot the conditions that you used on the factor space of the optimization. It will be a figure similar to what is shown in the reference readings. Label the points so that the order of each vertex is clear.

3) Calculate an estimated uncertainty for pH and temperature (how well you know the optimized result). For example, perhaps you know the pH = 2.3 ± 0.15. Think about how you would make the estimate of uncertainty.

4) The report needs to show an example of how you selected or calculated successive vertices in the factor space. The selection needs to be systematic and use the Simplex method. Arbitrary selection of reaction conditions will be graded unfavorably.

5) The data that you collect needs to demonstrate that you arrived at the optimized condition. Consult the reference readings for an example of what that looks like.



*Above is a depiction of what a Simplex method might look like for the first three measurements and the prediction of the next measurement. What is shown is **not** a complete experiment, this is only the start of the process to demonstrate how the next set of conditions would be selected.*

Questionnaire

(*Please copy the numbered questions to your submission. Answers must be in the form of complete sentences or data figures and tables.*)

1) Present a well-formatted table showing the solution conditions (pH and temperature) for each measurement (vertex) and the corresponding % yield.
2) Present a plot of the conditions that you tried on the factor space; the plot will look a lot like the example above. The plot must be large enough that conditions can be estimated visually (should occupy about half or more of a sheet of paper). Label the measurements in the order that you made them and lightly draw in the lines connecting the points to show the simplex process. Labelling points and drawing lines in Excel is <u>not easy</u> and could quickly become a way to lose 4 hours of your time. Carefully labelling by hand and drawing lines with the aid of a ruler is likely much faster.
3) Plot the % yield versus the vertex number (the order in which you performed the experiment). Use a scatterplot with straight lines connecting the points. This plot helps to show how the Simplex progressed toward the maximum value.

4) Show an example of how you calculated or selected the successive points in the factor space. This description needs to show how you systematically used the Simplex method to pick each vertex. Make reference to the Simplex plot that you made above to facilitate the writing.
5) State the optimized reaction conditions and calculate an uncertainty estimate for both pH and temperature (how well you know the location of the optimized result).