Challenges in Chemistry: Calculus challenge

Ove Christiansen, ove@chem.au.dk

Contents

1 Introduction

Solving equations and finding extremal points are key problems in the application of mathematics. Often one can not solve the problem at hand analytically with pen and paper but must resort to numerical and computational methods. These methods are of course still rooted in mathematics.

Here we shall consider the Newton method (also sometimes called Newton-Raphson) for functions of one variable and functions of many variables. We shall illustrate its use by simple functions that can be given analytical forms. Note, however, that one of the advantages of such methods for use in chemistry is that they do actually not require an analytical form, only on the availability to obtain function values and derivatives at specific points. The general idea is that one can iteratively improve upon a guess until a sufficient accurate solution to the equation at hand is obtained.

2 Practicalities

The challenges in Chemistry are made by staff at the Department of Chemistry and is in this case a FOLLOW UP on the Calculus course but is not part of the calculus course. Your lecturers, teachers, instructors, etc. in Calculus know nothing about this challenge, and please DO NOT USE their time on this project.

You may use English or Danish as you like in your answer, but mathematics is mandatory. I refer also to apps on the www such as

http://www.math.umn.edu/~garrett/qy/Newton.html
or
http://www.personal.psu.edu/dpl14/java/calculus/newtonsmethod.html
to get a feeling for the method. A few parts requires, however, a hand calculation. If you want, you may choose to program the algorithm yourself if you have such experience (Matlab, C, C++, Python, Fortran, or other) and attach the computer code and the succesful runs instead. This is, however, not required!

3 The Newton method for functions of one variable

3.1 The Newton method for functions of one variable

Read relevant material on the Newton method in your Calculus material, including exercises!, and if necessary other sources. Explain the Newton method concisely with your own words.

3.2 Example

Consider the function

$$f(x) = \cos(x) * \sin(x) + x * (x - \pi) \tag{1}$$

Use Newtons method to find x for which $f(x) = 0$ using respectively 1 and 2 as start points. A Table should be made with iteration number, current prediction for the root, $f(x)$, $f'(x)$. Try 6 iterations. How many digits accuracy is obtained?

Compare to results obtained by http://www.personal.psu.edu/dpl14/java/calculus/newtonsmethod.html, and report a "dump" of the result.

3.3 The Newton method for finding extremal points for functions of one variable

The Newton method can also be used for finding extremal points such as minima or maxima if an expression for both the first and second derivative can be obtained. Explain in your own words how.

3.4 Example

Find the first derivative of $f(x)$ from the previous example. Use the Newton method to find the minimum of $f(x)$. First try using 2 as a starting points. Thereafter try using 1. Discuss the dependence on start guess.

It is here ok to use http://www.personal.psu.edu/dpl14/java/calculus/newtonsmethod.html reporting a "dump" of the result.

3.5 Convergence properties of the Newton method

We have in the last example seen interesting cases of the convergence of the Newton method. It is here important to distinguish "local" and "global" convergence. We shall now analyse the "local" convergence. That is, we assume we have a reasonable starting guess.

We index iteration by n, so the current guess is denoted x_n, and the current error is noted e_n. Denoting the exact root as r we have

$$e_n = x_n - r \tag{2}$$

First argue that the Newton sequence imply that

$$e_{n+1} = x_{n+1} - r = e_n - \frac{f(x_n)}{f'(x_n)} \tag{3}$$

We now need to use a property related to Taylor expansions (covered in Calculus beta) namely that if we expand $f(x)$ around the current x_n we can write $f(r)$ as

$$f(r) = f(x_n - e_n) = f(x_n) - e_n f'(x_n) + \frac{f^{(2}(\zeta)}{(2)!} e_n^2 \tag{4}$$

where ζ is a number between x_n and r. NB! The remainder (the last term) can be given in many ways. One is the Lagrange form used here. This may not have been used much in Calculus beta, but is needed here. It can be taken as a fact.

Recall that $f(r) = 0$. Use the equations to show first that

$$e_n f'(x_n) - f(x_n) = \frac{1}{2} f''(\zeta) e_n^2 \tag{5}$$

and thereafter

$$e_{n+1} = \frac{1}{2} \frac{f''(\zeta)}{f'(x_n)} e_n^2 \tag{6}$$

Assume now that the above derivaties are "wellbehaved" (continous, $f'(x_n) \neq 0$, etc.). Argue that this shows the celebrated "quadratic convergence" of the Newton method when a guess is provided that have a small enough starting error.

3.6 Diatomic molecule

The equilibrium structure of a diatomic molecule is defined as the minimum energy structure. Thus, we consider the energy of a molecule as function of distance between the nuclei. E.g. for H_2 the relevant variable is the $H - H$ distance. The energy is actually the electronic energy (more about that in second year physical chemistry). Generally such energies can be evaluated approximately but accurately by theoretical methods. For our analysis here, we shall, however, use a model potential that is very widely used to describe binding, vibration and bondbreaking of diatomic molecules, namely the Morse potential. The Morse potential is a convenient interatomic interaction model for the potential energy (denoted V) of a diatomic molecule of the form

$$V(R) = D_e(1 - \exp(-\alpha(R - R_e)))^2 \tag{7}$$

Here R is the internuclear distance, R_e is the equilibrium bond distance, D_e is the well depth (relative to the dissociated atoms). Fianlly α is related to the width of the potential.

Using $D_e = 1$, $\alpha = 1$, and $R_e = 1$, use the Newton algorithm with start guess 1.5 for finding a minimum. Check that the correct equilibrium distance is found in this case (Why is the result no surprise in this case in this case?).

4 The Newton algorithm for more variables and identifying structures for molecules

4.1 Background

The preceeding diatomic molecule was only a warm up to the case of a more general molecule. In a general molecule the energy is a function of many variables. For example, for water there are 3 internal coordinates. We may think of 2 OH bonds and an angle. Generally, if there are N_{atoms}, there will be $3N_{atoms} - 6$ internal variables ($3N_{atoms} - 5$ for linear molecules). We will below use general notation but limit ourself to two-dimensional case for easy of calculation.

We may still conceive of identifying an equilibrium structure as the structure having minimal energy. For this purpose we have to consider the Newton algorihm for more variables.

Consider now a set of variables collected in the vector $\mathbf{x}$ with elements x_1, x_2 etc. As is standard in much of the physics and chemistry literature we use bold face to denote vectors and matrices. Thus $\mathbf{x}_1$ is vector, while x_1 is a number (here an element of a vector). The potential energy function, $V(\mathbf{x})$, is now a function of several variables collected in the vector $\mathbf{x}$. We assume there exist a partial derivative vector denoted the gradient with elements defined as

$$G_i(\mathbf{x}) = \frac{\partial V}{\partial x_i}(\mathbf{x}) \tag{8}$$

When all elements of the gradient are zero simultaneously we are at an extremal point (and we may check if it is a minima).

Let us now in analogy with the one-dimensional case consider an expansion of the gradient around a point given by the vector $\mathbf{x}_1$

$$\mathbf{G}(\mathbf{x}) = \mathbf{G}(\mathbf{x}_1) + \mathbf{J}(\mathbf{x}_1)(\mathbf{x} - \mathbf{x}_1) + \cdots \tag{9}$$

where the second derivative matrix, often denoted the Jacobian, is defined by the elements

$$J_{ij}(\mathbf{x}_1) = \frac{\partial G_i}{\partial x_j}(\mathbf{x}_1) \tag{10}$$

Take care with the notation here: The vectors and matrices (meaning their elements) are functions of the evaluation point, as indicated by $(\mathbf{x}_1)$. In the term $\mathbf{J}(\mathbf{x}_1)(\mathbf{x}-\mathbf{x}_1)$ the first $(\mathbf{x}_1)$ is for the functional dependence of the matrix while $(\mathbf{x} - \mathbf{x}_1)$ is the vector which the matrix $\mathbf{J}(\mathbf{x}_1)$ is multiplied on. Thus elementwise we can write out Eq.(9) as

$$G_i(\mathbf{x}) = G_i(\mathbf{x}_1) + \sum_j J_{ij}(\mathbf{x}_1)(x_j - (\mathbf{x}_1)_j) + \cdots \tag{11}$$

If we now ignore the higher order terms (ignore the $\cdots$ in Eq.(9)) we can construct a multi-variable Newton algorithm by determining the step to the next point from

$$0 = \mathbf{G}(\mathbf{x}_1) + \mathbf{J}(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1) \tag{12}$$

This of course generalises to

$$0 = \mathbf{G}(\mathbf{x}_n) + \mathbf{J}(\mathbf{x}_n)(\mathbf{x}_{n+1} - \mathbf{x}_n) \tag{13}$$

in the n'th step.

In each step in the Newton algorithm we need to solve the linear set of equations in Eq.(13) to obtain $\mathbf{x}_{n+1} - \mathbf{x}_n$ which provides the new estimate when added to $\mathbf{x}_1$.

In the following we consider this theory for the function

$$V(x_1, x_2) = 2x_1^2 + 3x_2^2 + x_1x_2 + \frac{1}{2}x_2 \tag{14}$$

4.2 The gradient vector as a function

Derive an expression of the gradient vector $\mathbf{G}$ with elements that are functions of x_1 and x_2.

4.3 The Jacobian matrix as a function

Derive an expression of the Jacobian matrix **J**. Generally, the elements would be functions of x_1 and x_2. In this case the Jacobian is constant. Why?

4.4 The first Newton step

Starting at $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and find a Newton step to arrive at an improved guess for the position of the minimum.

For this you need to solve the two equations for the two unknowns. Use "Determinant-metoden" to do this (See here for what I mean, http://www.webmatematik.dk/lektioner/matematik-c/ligninger/to-ligninger-med-to-ubekendte. Notice that in our particular case the Determinant is constant and therefore the same determinant could be use in subsequent steps).

4.5 Optimal value

What is the Gradient at the updated point?

Take care here if you have done numerical calculations with limited precision. Is you answer significantly different from zero?

A zero gradient reflects that a minimum has been found. Why was a minimum found after only one iteration/step in this case? If we add the term $5x_1^2x_2^2$ would Newtons method still deliver the true minimum in one iteration?

In general many steps may be required in the Newton algorithm, but the above examples illustrate how one can, given a potential and a starting point, use the Newton algorithm to find equilibrium structures of molecules. Actually, the potential is rarely know as a function. However, what is available is methods for calculating the energy and its first and perhaps second derivative for a given set of coordinates for the atomic nuclei. This is sufficient to use the Newton algorithm for finding the equilibrium structure.