CY40014 Introduction to Computational Chemistry
Autumn 2010-2011

Module 2: Introduction to Numerical Analysis

Aim of the present module

1. Introduction to basic numerical analysis
2. Developing simple programs to implement the numerical methods

Topics of interest

1. Interpolation: least square fitting
2. Numerical integration- Trapezoidal and Simpson's one-third rule
3. Numerical solution of ordinary differential equations –Runge-Kutta method
4. Solution of non-linear equations using Newton-Raphson method
5. Solution of linear systems – LU decomposition and Gaussian elimination
6. Eigen values and Eigen vectors

Note

The methods are merely outlined in the worksheets. For a detailed description, please refer to any textbook on elementary numerical analysis. Several such books are available at the Institute's Central Library.

Reference:

1. An Introduction to Numerical Analysis, K.E. Atkinson (2nd edition), J Wiley & Sons(1978).
2. Numerical Methods using MATLAB, J.H. Mathews and K.D. Fink (4th Edition), Pearson (2004).
   [This is a good reference for those interested in using MATLAB for numerical methods]
   Both the books are available in low price editions.

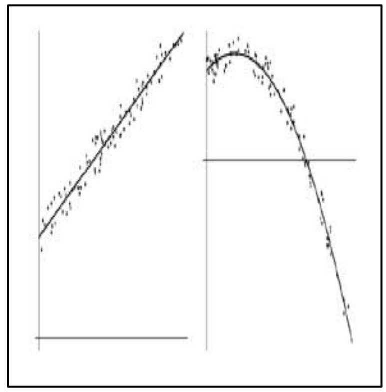
Worksheet 5

Topic 1: To fit a set of data to a straight line using the method of Least Square Fitting

Aim: To model the experimental data and to interpolate and/or extrapolate the value of a physical quantity from the measured data.

Let us consider a property *y* that depends on an independent variable *x*. In an experimental set up, several discrete values of *y* can be recorded by varying *x*. The aim of the present lesson is to show you a simple way of expressing *y* accurately as a function of *x* in terms of a simple equation.

Note: In the figure on the right, two such typical data sets are shown for which the best fit to the data are required. While the best fit is provided by a straight line in the left panel, the same is not true for the case shown in the right hand panel. Therefore, it may be necessary to use higher order polynomials for an accurate representation of the data.



Working Principle: To draw the best fit line with a minimum vertical displacement of all the data points.

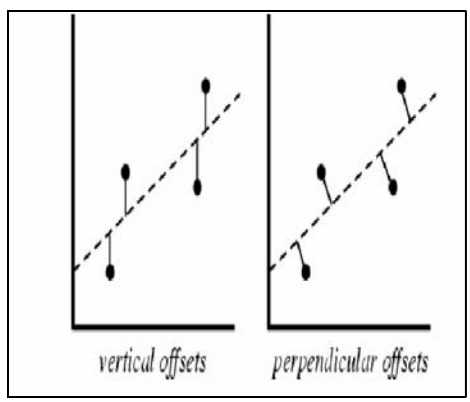
Square deviation: Let us assume that we have a data set (x_i, y_i) $(i = 1, N)$ and we would like to write $y = a + bx$. For this purpose, let us define $R^2 = \sum_{i=1}^{N} \left[y_i - (a + bx_i)\right]^2$

so that the square deviation R^2 is a function of *a* and *b*. For R^2 to be a minimum., we must have

$$\frac{\partial R^2}{\partial a} = \frac{\partial R^2}{\partial b} = 0 .$$

The above conditions lead to the following relations:

$$Na + b\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i ; \qquad a\sum_{i=1}^{N} x_i + b\sum_{i=1}^{N} x_i^2 = \sum_{i=1}^{N} x_i y_i$$



These two equations may be solved to obtain

$$a = \frac{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i y_i}{N \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i \right)^2}; \quad b = \frac{N \sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{N \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i \right)^2}$$

W5_1. Write a program to fit a given set of data to a straight line using the method of least squares.

W5_2. The heat capacity of graphite is measured at a high temperature and presented in the table below.

T (K)	300	350	400	450	500	600	700	800	900	1000
$C_{P,m}$ ($J\ K^{-1}\ mol^{-1}$)	8.581	10.241	11.817	13.289	14.623	16.844	18.57	19.827	20.824	21.610

Assume that the temperature dependence of heat capacity may be modeled as $C_{P,m}(T) = c_0 + c_1 T$. Use your program to find the heat capacity of graphite at 950 K.

W5_3. In the table given below, the results of a kinetic study on the gas-phase decomposition of N_2O_5 at 67^o C are presented. The reaction takes place as $2\,N_2O_5 \rightarrow 4\,NO_2 + O_2$.

t (min)	0.0	1.0	2.0	3.0	4.0	5.0
$[N_2O_5]$ (mol L^{-1})	1.0	0.705	0.497	0.349	0.246	0.173

Calculate the first order rate constant for this reaction at 67^o C.

W5_4. Note the following three definitions on standard deviation.

(a) the standard deviation about **regression**: $s_r = \sqrt{\frac{S_{yy} - m^2 S_{xx}}{N-2}}$,

(b) the standard deviation of the **slope**: $s_m = \sqrt{\frac{s_r^2}{S_{xx}}}$

(c) the standard deviation of the **intercept**: $s_b = s_r \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i \right)^2}}$

where $N\,S_{xx} = N \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i \right)^2$ and $N\,S_{yy} = N \sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i \right)^2$. Using these definitions, calculate all the three deviations in the data presented in the two examples cites above.

Topic 2: Numerical integration

Aim

Let us consider the following integral $I = \int_a^b f(x)\,dx$. The aim of this exercise is to approximate the integral I [that is, the area under the curve *f(x)*] by evaluating the integrand *f(x)* at a finite number of sample points in the interval *[a,b]*.

Method

The idea behind numerical integration is to interpolate the integrand *f(x)* and then to integrate the interpolating polynomial *p* analytically.

Trapezoidal Rule

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2}\left[f(a)+f(b)\right]$$

Simpson's Rule

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6}\left[f(a)+4f\left(\frac{a+b}{2}\right)+f(b)\right]$$

where the integral has been approximated by summing the integrand at 2 and 3 equidistant points, respectively. The former method replaces the curve *f(x)* by a straight line while the latter uses a quadratic approximation.

To decrease the error caused by such severe approximations, several such rules are linked together to give the composite rules which are as follows:

Composite Trapezoidal Rule

In this method, the interval *[a,b]* is sub-divided into *n* intervals and in each sub-interval, *f(x)* is replaced by a straight line.

$$\int_a^b f(x)\,dx \approx \frac{h}{2}\left[f(a)+f(b)+2\sum_{i=1}^{n-1} f(x_i)\right], \quad where\ \ h=\frac{b-a}{n},\ x_i = a+ih,\ x_n = b$$

Composite Simpson's One-third Rule

This is an approximate method valid only for an **even** number of intervals.

$$\int_a^b f(x)\,dx \approx \frac{h}{3}\left[f(a) + f(b) + \sum_{i=1}^{n-1} a_i f(x_i) \right], \; where \;\; h = \frac{b-a}{n}, \; x_i = a + ih, \; x_n = b$$

$$Note: \quad a_i = 2 \quad for \;\; i \;\; even$$
$$\quad = 4 \quad for \; i \;\; odd$$

$$\therefore \int_a^b f(x)\,dx \approx \frac{h}{3}\left[f(a) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + ... + 2f(x_{n-2}) + 4f(x_{n-1}) + f(b) \right]$$

W5_5. Write a program to evaluate the integral $I = \int_0^1 dx \; x$ using

(i) the trapezoidal rule and
(ii) Simpson's one-third rule.

W5_6. The luminous efficiency (ratio of the energy in the visible spectrum to the total energy) of a black body radiator may be expressed as a percentage by the following expression

$$E = \frac{64.77}{T^4} \int_{4X10^{-5}}^{7X10^{-5}} x^{-5} \left[\exp\left(\frac{1.432}{Tx} \right) - 1 \right]^{-1} dx$$

where *T* is the absolute temperature in degrees Kelvin, *x* is the wavelength in cm and the range of integration is over the visible spectrum. Taking *T=3500 K*, write a program that uses Simpson's one-third rule to compute *E*.

W5_7. It is well known experimentally that at high temperatures, irrespective of the nature of the solid, its heat capacity of is a constant. This is known as the Dulong-Petit law. On the other hand, at low temperatures, the heat capacity becomes a function of temperature and $C_v \sim T^3$ at very low temperatures. Under these conditions, only a small amount of thermal energy is available to the system for the atoms/molecules to execute vibrational motion. Therefore, the low temperature vibrational pattern of the solid is dominated by low-energy (i.e. long wavelength) vibrational modes. It has been shown by Debye that these modes can be approximated by the long wavelength vibrational modes of a continuous elastic body and the following expression is obtained for C_v

$$\frac{C_v}{3Nk_B} = \frac{3\pi^4}{5}\left(\frac{T}{T_D}\right)^3 \int_0^{T_D/T} dx \frac{x^4 \exp(x)}{[\exp(x)-1]^2}$$

<u>Eq. 1</u>

Since the integral in Eq.1 cannot be evaluated analytically, therefore, the methods of numerical integration are applied to estimate C_v at a given temperature T. T_D is a characteristic constant of the solid known as the Debye temperature.

Write a program that

(i) reads in the temperature in units of T_D ,

(ii) calculates $\frac{C_v}{3Nk_B}$ by evaluating the above integral numerically using Simpson's one-third rule and

(iii) gives as output a table containing the dimensionless temperature T/T_D in the first column and $\frac{C_v}{3Nk_B}$ in the second column.