1 Gold prospects

Weight	x-coordinate	y-coordinate
0.54	1.7	2.8
0.77	2	1.4
0.92	3.4	3.7
0.41	4.5	5
0.16	5.1	4.9

You are a gold prospector, searching in a gold-rich area. It is well-known that there is a vein of gold in the area which is a straight line, but it is not known where it is. You have talked to some other prospectors in the local drinking holes, and they have all claimed to have struck gold — but their co-ordinates might be slightly inaccurate, as they were all in varying stages of drunkenness at the time!

To find the gold line, you have decided to fit a linear regression line to the co-ordinates given by your fellow prospectors. This is done by fitting a line of form $y = ax + b$ to the data. Naturally, this creates 'predictions' for the y-values of each of the data points that we already have, and these predictions will have errors. We find the coefficients of the regression line by minimising the sum of the squares of all the errors.

However, to take into account the inaccuracy of the information, you have decided to weight the co-ordinates according to how drunk the prospector was at the time. You have assigned the weights as in the table above, and wish to find the regression line which minimises the weighted sum of squares.

To do this, apply three different numerical methods to the above problem, and analyse the performance of the methods, discussing in particular the advantages and disadvantages of each method in relation to this particular problem. At least one of the methods should be not covered in the lecture notes.

For the second part of this project, you have decided that you might have been pretty drunk as well when you were talking to your fellow prospectors, and so your judgement of their drunkenness may have been fairly inaccurate. Instead, you now want to assign a weight to each point so the the weighted sum of squared errors of the regression line is minimised. In addition, the sum of the weights must be 1, and each weight must be at least $\frac{1}{10}$. Apply two different methods to this problem, and discuss the advantages and disadvantages of them.

2 Satellite location

x-coordinate	y-coordinate	z-coordinate
1.88	0.24	-0.19
-1.21	1.77	0.98
-0.46	-1.69	-0.52
0.33	-0.17	2.03

Australia wishes to send up a communications satellite, which may or may not be a spy satellite with onboard missile. This satellite must be positioned so as to communicate effectively and quickly with other satellites from other countries, which are positioned according to the table above. The scale is set assuming that the Earth is a sphere centred at the origin of radius 1. It is estimated that it costs $1 million per year to send data along a length of 1 Earth radius to another satellite, and this cost scales linearly with the length that data is sent. In addition, it costs $\ln(d+1)$ million dollars to send up the satellite, where d is the distance (in Earth-radii) from the centre of the Earth to the satellite. We wish to minimise the cost of the first year of operation (ignoring the fact that the location of the satellite may be inside the Earth!).

To do this, apply three different numerical methods to the above problem, and analyse the performance of the methods, discussing in particular the advantages and disadvantages of each method in relation to this particular problem. At least one of the methods should be not covered in the lecture notes.

For the second part of this project, we also add the constraints that the satellite must be outside the sphere of the Earth, and in addition, to prevent hostile countries from shooting it down, the satellite must be positioned above Antarctica (it must have a latitude which is lower than 45 degrees south). Apply two different methods to this problem, and discuss the advantages and disadvantages of them.

3 Newton methods

The original Newton method was probably derived (by Newton) from a method by French mathematician Francois Viéte. Newton used it as an algebraic method to retrieve the zeroes of polynomials. The method was simplified, interpreted, and developed over time by many mathematicians.

Consider the problem

$$\min f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T H\mathbf{x}$$

where

$$\mathbf{c} = (5.04, -59.4, 146.4, -96.6)^T$$

and

$$H = \begin{pmatrix} 0.16 & -1.2 & 2.4 & -1.4 \\ -1.2 & 12.0 & -27.0 & 16.8 \\ 2.4 & -27.0 & 64.8 & -42.0 \\ -1.4 & 16.8 & -42.0 & 28.0 \end{pmatrix}$$

Consider three Newton methods (including quasi-Newton methods) when finding the solution to this problem. Discuss in detail the advantages and disadvantages of using these methods over others for this particular problem and in the general case. How do these methods compare to other available methods for global convergence to a local minimum? Consider also the constrained case where $\mathbf{x}^T\mathbf{x} = 1$.

4 The Rosenbrock function

The Rosenbrock function,

$$f(x) = \sum_{i=1}^{N-1} [(1 - x_i)^2 + 100(x_{i+1} - x_i^2)^2] \qquad \forall\ x \in \Re^N,$$

is commonly used to test unconstrained numerical minimisation algorithms. It is designed specifically to cause algorithms to converge slowly or possibly fail. Apply every n-dimensional algorithm discussed in the course to this function. Create an in-depth comparison of these methods, drawing on examples of other functions if necessary. In order to perform a comparison, first consider a list of criteria with which to compare the algorithms. Also discuss the suitability of the Rosenbrock function as a test function for unconstrained n-dimensional algorithms.

5 Penalty methods

Provide a survey of penalty methods. Consider the following examples while comparing the advantages and disadvantages of these methods.

Problem 1:

$$\min x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4$$

subject to

$$\begin{aligned} -x_1^2 - x_2^2 - x_3^2 - x_4^2 - x_1 + x_2 - x_3 + x_4 + 8 &\geq 0 \\ -x_1^2 - 2x_2^2 - x_3^2 - 2x_4^2 + x_1 + x_4 + 10 &\geq 0 \\ -2x_1^2 - x_2^2 - x_3^2 - 2x_1 + x_2 + x_4 + 5 &\geq 0 \end{aligned}$$

Problem 2:

$$\min[1.5 - x_1(1 - x_2)]^2 + [2.25 - x_1(1 - x_2^2)]^2 + [2.625 - x_1(1 - x_2^3)]^2$$

subject to

$$\mathbf{x}^T\mathbf{x} = 1$$

where $\mathbf{x} = (x_1, x_2, x_3)$.

Provide solutions for these examples using penalty method algorithms you have implemented in a programming language of your choice. Also discuss when penalty methods are appropriate or inappropriate for a constrained nonlinear optimisation problem, and in which real-life situations these methods would be useful.

6 Make-your-own project!

Find an optimisation situation which involves nonlinearity. Model the situation and solve it using two different methods. Pay particular attention to the suitability and effectiveness (or lack of it) of the methods.