Review 2 (2nd Exam Study Questions) Solutions

Problem 1. Non-linear minimization:

A. Derive Newton's method for finding the local extremum (minimum or maximum) for a function f(x) starting with some initial guess x_0

This is in the notes

B. By using Taylor series analysis, prove that this method converges quadratically when close to the actual extremum x^*, provided that the second derivative at this point is non-zero.

This is in the notes

C. Apply this method to find the minimum of the function $f(x) = x^3 - x$ with the initial guess $x_0 = 1/2$. Do three iterations.

Solution:

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)} \qquad f'(x_i) = 3x_i^2 - 1 \qquad f''(x_i) = 6x_i$$

Thus:

$$x_{i+1} = x_i - \frac{3x_i^2 - 1}{6x_i}$$

and hence: $x_0 = 0.5$, $x_1 = 0.0.583333$, $x_2 = 0.57738$, $x_3 = 0.57735027$

Problem 2. Root Finding:

A.

1. Under what conditions is the bisection method guaranteed to converge? (be brief but complete)

2. What is the **rate** of convergence of each of the following methods?

a. Bisection _____1_______

b. Newton's Method ______2________

c. Secant Method ______1.618________

3. If a function locally behaves as $f(x) \sim (x - c)^3$ which of the above techniques will converge the fastest?

Bisection - all methods are linear because of the triple root, but bisection has a smaller constant in this case.

B. Derive the secant rule for finding the root to some function f(x). Don't try to compute the rate of convergence, however -- simply show where the rule comes from.

This is in the notes

Problem 3. Regression Error Propagation:

Suppose you are asked to fit a set of data $b_i(t_i)$ to a model of the form:

$$b(t_i) = x_1 \phi_1(t_i) + x_2 \phi_2(t_i) + x_3 \phi_3(t_i)$$

where the $\phi_j(t_i)$ are the modelling functions and the x_j are the modelling parameters. You are to determine the best fit values of the parameters x_j (set up the regression problem in matrix form) and then to compute the function $y = f(\underset{\sim}{x})$. If we have n data points, show how we may compute the random error in y. I want equations in matrix form here: be as specific as possible, and state all of your assumptions.

Solution:

$$\underset{\sim}{r} = \underset{\approx}{A}\,\underset{\sim}{x} - \underset{\sim}{b} \qquad \underset{\approx}{A} = \begin{bmatrix} \phi_1(t_1) & \cdots & \phi_3(t_1) \\ \vdots & \ddots & \vdots \\ \phi_1(t_n) & \cdots & \phi_3(t_n) \end{bmatrix} \qquad \underset{\sim}{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \qquad \underset{\sim}{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_3 \end{bmatrix}$$

$$\underset{\approx}{K} = \left[\underset{\approx}{A}\,\underset{\approx}{A}^T\right]^{-1} \underset{\approx}{A}^T \qquad \underset{\sim}{x} = \underset{\approx}{K}\,\underset{\sim}{b} \qquad \underset{\approx}{\Sigma}_x^2 = \underset{\approx}{K}\,\underset{\approx}{\Sigma}_b^2\,\underset{\approx}{K}^T$$

Now if 1) All the b_i are independent, and 2) they all have the same variance, then:

$$\underset{\approx}{\Sigma}_b^2 = \sigma_b^2\,\underset{\approx}{I} \qquad \text{where:} \qquad \sigma_b^2 = \frac{\underset{\sim}{r}^T\underset{\sim}{r}}{n-3}$$

Finally, we have:

$$\sigma_y^2 \approx \underset{\sim}{\nabla} f\,\underset{\approx}{\Sigma}_x^2 \left(\underset{\sim}{\nabla} f\right)^T \qquad \text{provided that: } \underset{\approx}{\Sigma}_x^2 : \underset{\sim}{\nabla}\underset{\sim}{\nabla} f << f$$

Problem 4. Statistics:

An automatic food bagging and boxing machine is loading boxes with crackers. The boxes are each supposed to hold 12oz of crackers, but analysis shows that the average weight of crackers in each box is actually 12.36oz with a population standard deviation of ±0.18oz. You may assume that they are normally distributed.

A. What is the probability that any given box has less than 12oz of crackers in it? Use the table for a normal distribution given below.

This is equal to the probability that a box is more than 2σ below the mean, which from the table is 1 - 0.97725 = 0.02275, or about 2.3%.

B. The boxes of crackers are packed 16 to a case. If all of the cracker boxes can be considered independent, what is the probability that the average of the weight of the boxes in a case is less than 12oz?

The mean of 16 boxes has a standard deviation which, if the boxes are independent, is 1/4 that of a single box. Thus the probability is that of the mean being 8σ below the expectation value, or 6.22×10^{-16}. Note that this answer is unreasonable - odds are that the boxes are either not governed exactly by the normal distribution (this is often the case near the tails of a distribution) or the boxes in a case are not all independent.

C. What is the probability that none of the boxes in a case have a weight less than 12oz?

This is equivalent to the probability that all of the boxes are above 12oz, which is equal to the probability that an individual box is above this weight raised to the 16th power (again assuming independence). Thus the answer is $(0.97725)^{16} = 0.692$ or about 69%.

You may find the following table helpful:

x	P(x)
0	0.5
1	0.84134
2	0.97725
3	0.99865
4	$1 - 3.17 \times 10^{-5}$
5	$1 - 2.87 \times 10^{-7}$
6	$1 - 9.87 \times 10^{-10}$
7	$1 - 1.28 \times 10^{-12}$
8	$1 - 6.22 \times 10^{-16}$
9	$1 - 1.13 \times 10^{-19}$
10	$1 - 7.62 \times 10^{-24}$

where $P(x) = \int_{-\infty}^{x} Z(t)\,dt$

and $Z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

P(x) is known as the cumulative probability distribution. Z(x) is the Gaussian distribution.