ETH zürich

MAD exercise session 3

Newtons Method

Pascal Auf der Maur

Recap

- Conditionnumber of LSQ
- QR/SVD decomposition

Last Exercise

Problem 1

- Generate own data with noise
- Observe behaviour of LSQ with noise and outlier
  - LSQ stable in the presence of noise
  - LSQ unstable with outliers

Last Exercise

Problem 2 / P & P 1

- Least squares in 2D
- Least squares is robust to noise
- Accuracy is decreased if less points are sampled

$$A = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{bmatrix} \quad w = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \quad z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{bmatrix}$$

Last Exercise

P & P 2

- Very difficult exercise
- Results shows that the error of least squares scales with $\frac{\sigma}{\sqrt{N}}$ where σ is the standard deviation of the noise and N is the number of samples.

Non-linear equations

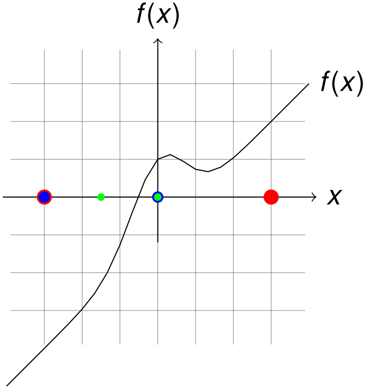
- Usually nature follows complex rules which are not always linear in nature. For example exponetials or logarithms (ex. Diodes).
- These problems can be very complex and might be impossible to solve by hand.
- We try to approximate the solution numerically

Every equation can be converted into a rootfinding problem:

$$y = f(x) \quad \rightarrow \quad y - f(x) = 0$$

Bisection

- Based on intermediate value theorem.
- We find the root by closing in on the interval where it is located.



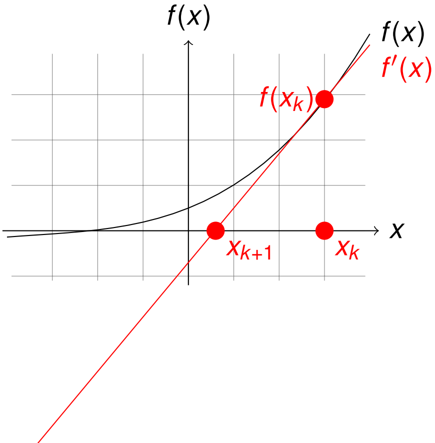
Bisection

Properties

- Linear convergence
- Stable
- Only the function has to be evaluated
- Finding an interval may prove difficult

Newtons method

- Faster method than bisection
- Uses derivatives to estimate the roots



Newtons method

$$f'(x_k) = \frac{f(x_k) - 0}{x_k - x_{k+1}} \qquad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

- Converges quadratically
- Is not guaranteed to converge
- Evaluates the derivative aswell
- Can get “stuck“

Algorithm 1 Newton's Method

Input

E: maximum error

$f(x)$: function to be evaluated

$df(x)$: derivative function of f

x_0: initial value of x, higher than the root

N: maximum iteration steps

Output

x: approximated root

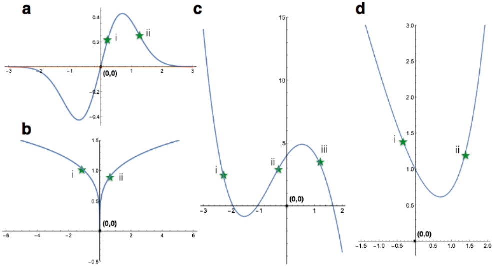
```
procedure NEWTON
    x ← x_0
    for i ← 1, N do
        if abs(f(x)) ≤ E then
            Break
        end if
        x ← x − f(x)/df(x)
    end for
end procedure
```

Old exam question

Question 8: Newton's method (8 points)

You are given the function graphs below. For every function graph state whether it is possible to evaluate one or more roots of the function with Newton's method, starting from the marked points (stars) as initial guesses for the method. Provide a short explanation for your answers.

Note: The x-axis in graph (a) is an asymptote.



Secant Method

- Same idea as Newtons method
- Derivative is numerically approximated
- Needs two initial guesses
- Convergence between first and second order

$$x_{k+1} = x_k - f(x_k)\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}$$

Systems of equations

- Newtons method can be expanded for more dimensions
- $f(x) \rightarrow F(x) \quad f'(x) \rightarrow J(F(x))$

$$x_{k+1} = x_k - J^{-1}(x_k)F(x_k)$$

- In praxis we solve the linear system of equations

Systems of equations

- If $J(x)$ is not invertible we can calculate $J^{\dagger}(x)$
- The error can be calculated with $\|F(x_{k+1})\|$
- We can also formulate an update formula for J

Optimisation

- We can adopt solving nonlinear equations into multivariable optimisation
- $F(x) = \nabla E(x) = 0$
- $J(F(x)) = \nabla^2 E(x)$
- $\nabla^2 E(x_k) z = -\nabla E(x_k)$
- $x_{k+1} = x_k + z$
- With Newton we can perform non linear least squares by optimising:

$$E(w) = \sum_{i=1}^{N} (y_i - f(x_i, w))^2$$

Gradient descent

- $x_{k+1} = -\nabla E(x_k)$
- $x_{k+1} = x_k + \eta z$
- Cheaper to calculate than the Jacobian
- Finds local minimas
- Used for neural networks

Old exam question

Question 11: Newton's method (7 points)

a) Derive the formula used for one iteration of Newton's method.
Hint: Taylor series expansion around the starting point x_0.

b) How many iterations does it take for Newton's method to find the root of the function $f(x) = ax + b$ for $a, b \in \mathbb{R}$ and $a \neq 0$? Justify your answer.

c) We are interested in solving the nonlinear equation $\cos(x) = e^x - 1$. Write down the formula for the estimated solution x_1 after one iteration of Newton's method starting from a given x_0.