Problem 1: Fixed point iterations in practice

Suppose we want to find a solution for:

$$x^4 - x - 10 = 0.$$

Notice that the solutions for this polynomial are $x_1 \approx 1.85556$ and $x_2 = -1.6975$.

1. Plot this function (either by hand or with a software of your choice).

2. Consider the following fixed point scheme:

$$x_{k+1} = \frac{10}{x_k^3 - 1}.$$

Let $x_0 = 2$. Does this method converge? Why (/why not)? In the same plot draw the functions $y = x$, $y = \frac{10}{x^3-1}$ and add to this plot the first 5 iterations of such method.

3. Consider the following fixed point scheme:

$$x_{k+1} = (x_k + 10)^{1/4}.$$

Test this method with the following starting points: $x_0 = 1$, $x_0 = 2$, $x_0 = 4$. Does this method converge? Why (/why not)? In the same plot draw the functions $y = x$, $y = (x + 10)^{1/4}$ and add to this plot the first 5 iterations of such method for $x_0 = 2$.

4. Consider the following fixed point scheme:

$$x_{k+1} = \frac{(x_k + 10)^{1/2}}{x_k}.$$

Test this method with $x_0 = 1.8$, $x_0 = 2$. Does this method converge? Why (/why not)? In the same plot draw the functions $y = x$, $y = \frac{(x+10)^{1/2}}{x}$ and add to this plot the first 5 iterations of such method for $x_0 = 1.8$.

What can you tell about the methods that converge? How do these methods look? Can you relate this to the contraction mapping theorem? What can you tell about the speed of convergence?

Convergence of iterative methods

Consider the following theorem:

Theorem 1. *Let $f \in C^2$ on $I_\delta = [\xi - \delta, \xi + \delta], \delta > 0$. Assume that $f(\xi) = 0$ and $f''(\xi) \neq 0$. Assume that $\exists A > 0$ such that $\frac{|f''(x)|}{|f'(y)|} \leq A$ for all $x, y \in I_\delta$. If x_0 is such that $|\xi - x_0| \leq min(\delta, \frac{1}{A})$, then $x_n \to \xi$ quadratically.*

1. Consider the function $f(x) = sin(x)$ show that if $x_0 \in (\frac{-\pi}{2} + a, \frac{\pi}{2} - a)$, where $a \geq 0$ and $x_0 \neq 0$, $x_n \to 0$ quadratically.

2. What assumptions do you need to have on a?

Newton's Method

1. Suppose Newton's method is applied to find the solution of the following equation:

$$e^x - x - 2 = 0.$$

   Write down Newton's iteration. Show that if the starting value is positive, the iteration converges to the positive solution, and if the starting value is negative, it converges to the negative solution. Obtain approximate expressions for x_1 if we start in $x_0 = 100$ and if we start in $x_0 == -100$, describe the subsequent behaviour of the iteration.

2. Consider the iteration for the solution $f(x) = 0$:

$$x_{k+1} = x_k - \frac{(f(x_k))^2}{f(x_k + f(x_k)) - f(x_k)}.$$

   Explain the connection with Newton's method, and show that it converges quadratically if x_0 is sufficiently close to the solution.

Secant method

1. Show that the secant method,

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

   can be written as:

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

2. Now, denote the root of f to be ξ, so that $f(\xi) = 0$. Also assume that f is twice continuously differentiable and that $f' > 0$ and $f'' > 0$ in a neighborhood of ξ. Define the quantity φ to be:

$$\varphi(x_k, x_{k-1}) = \frac{x_{k+1} - \xi}{(x_k - \xi)(x_{k-1} - \xi)},$$

where x_{k+1} is as in 1. Compute (for fixed value of x_{k-1})

$$\psi(x_{k-1}) = \lim_{x_k \to \xi} \varphi(x_k, x_{k-1}).$$

3. Now compute

$$\lim_{x_{k-1} \to \xi} \psi(x_{k-1}),$$

and therefore show that

$$\lim_{x_k, x_{k-1} \to \xi} \varphi(x_k, x_{k-1}) = \frac{f''(\xi)}{2f'(\xi)}.$$

4. Next, assume that the secant method has convergence order q, that is to say that

$$\lim_{k \to \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^q} = A < \infty.$$

Using the above results, show that $q - 1 - 1/q = 0$, and therefore that $q = (1 + \sqrt{5})/2$.

5. Finally, show that this implies that

$$\lim_{k \to \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^q} = \left(\frac{f''(\xi)}{2f'(\xi)} \right)^{q/(1+q)}.$$

Errors!

1. State the 4 causes of error in numerical schemes and give an example of each one of them.

2. Consider the following approximation to derivatives:

$$f'(x) \approx \frac{3f(x + h) - f(x) - 2f(x - h)}{5h}$$

$$f'(x) = \frac{f(x - 2h) - 8f(x - h) + 8f(x + h) - f(x + 2h)}{12h} + \frac{4h^4}{5!} f^{(5)}(\xi)$$

For the first expression show that its error is of order $\mathcal{O}(h)$ for the second expression do Taylor expansions and find the error given. We know that $\xi \in [x - 2h, x + 2h]$.

3. Consider the forward difference approximation to the derivative:

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}.$$

What happens if $f''(x) = 0$?

4. Consider the central difference approximation to the derivative:

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h},$$

what happens if $f'''(x) = 0$?

5. Consider the following polynomial:

$$p(x) = ax^2 + bx + c.$$

Show that the central difference scheme is "exact" i.e. there is no error.

6. For which type of polynomial is the forward difference approximation to the derivative exact?

7. Consider the following approximation to the integral of f on an interval $[a, b]$:

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

Suppose that $b - a = \mathcal{O}(h)$. Derive an error estimate for this approximation using a Taylor expansion. What assumption(s) do you need to take into consideration?

8. Give three examples of functions (different than constants) whose integral on the interval $[0, 2]$ can be exactly calculated using 7. What do these functions have in common?

9. Come up with a Monte-Carlo like method to approximate π. How could you make this method "better"?

Computer representation of numbers

10. If we have "a lot" of computations to do, what would be the problem if we use SP?

Numerical Linear Algebra (so far)

Practice "modelling" with Linear Algebra

1. Suppose that we have a system that can be in 3 different states. If we know that at time t we are in state i, then the probability of going to state j at time $t + 1$ is given as:

$$P(X_{t+1} = j | X_t = i) = \begin{cases} 1/2 & \text{if } i+1 = j \text{ and } i < 3 \\ 1/2 & \text{if } i = j \\ 1/2 & \text{if } i = 3 \text{ and } j = 1 \\ 0 & \text{otherwise} \end{cases}.$$

   Build a matrix such that its ij entry represents the probability $P(X_{t+1} = j | X_t = i)$.

2. Find (α_1, α_2) such that $||\vec{x} - \alpha_1 \vec{a_1} - \alpha_2 \vec{a_2}||^2$ is minimized.

Linear Algebra in a computer

1. Write in python-style pseudo-code, an algorithm for multiplying the matrices $a \in R^{m \times p}, b \in R^{p \times n}$. How many floating point adds and multiplies do you need, in terms of m,n and p?

2. Let $B \in \mathbb{R}^{m \times n}$, $A \in \mathbb{R}^{n \times n}$, $u \in \mathbb{R}^{n \times 1}$, $v^\top \in \mathbb{R}^{i \times p}$, and $z \in \mathbb{R}^{p \times 1}$, which of the following options need less floating point adds and multiplies? Why?

   - $(BA)(uv^\top)z$
   - $B(Au)(v\top z)$

3. In numerical analysis inverting a matrix is seen as "terrible". But why? For an invertible $n \times n$ matrix, how many FLOPs would we need to invert it? If you're not quite sure try inverting this matrix by hand and count how many operations you needed.

$$A = \begin{bmatrix} 4 & -2 & 1 \\ 5 & 0 & 3 \\ -1 & 2 & 6 \end{bmatrix}.$$

LU factorization

1. Count FLOPs for forward or backward substition.

2. Prove that the product of two lower unit triangular matrices is unit lower triangular.

3. For the LU factorization, if we label the kth Gauss transform M_k where $M_K = I - \tau^k e_k^T$, show that $M_k^{-1} = I + \tau^k e_k^T$.

4. Prove that $M_1^{-1}...M_{n-1}^{-1}$ is a lower triangular matrix L.

5. Describe how you can create a recursive algorithm for computing $A = LDU$, where L is unit lower triangular and U is unit upper triangular and D is diagonal (actually diagonal). To make it easy, assume n (number: $A \in \mathbb{R}^{n \times n}$) is a power of two. Hint: what is the recursive base case (n=1)? Think about block matrices.

Norms

1. Let $v \in \mathbb{R}^n$ prove the following inequalities:

$$\|v\|_\infty \le \|v\|_2$$
$$\|v\|_2^2 \le \|v\|_1 \|v\|_\infty.$$

2. For each of the previous inequalities, give an example of a non-zero vector in $\mathbb{R}^n$ for which equality is obtained.

3. Show that:

$$\|v\|_\infty \le \|v\|_2 \le \|v\|_1$$
$$\|v\|_2 \le \sqrt{n}\|v\|_\infty$$

4. Let $A \in \mathbb{R}^{m \times n}$ be any real-values matrix. Show that:

$$\|A\|_\infty \le \sqrt{n}\|A\|_2, \|A\|_2 \le \sqrt{m}\|A\|_\infty.$$

5. The Frobenius norm for a matrix $A \in \mathbb{R}^{m \times n}$ as:

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}.$$

Prove that $\|A\|_F = \sqrt{\text{trace}(A^\top A)}$.

6. Prove that $\|A\|_F = \sqrt{\sum_j \sigma_j}$, where σ_j are the singular values of A.

7. Show that $\|A\|_2 \le \|A\|_F$.

Exercises for mid term

Cholesky and QR

1. Show that if A is positive definite, then A^k is also positive definite for all $k \geq 1$.

2. Find the QR factorization of the following matrix:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

3. Let A be an $m \times n$ matrix with independent columns. If $A = QR$ and $A = \hat{Q}\hat{R}$ are QR-factorization of A, show that $Q = \hat{Q}$ and $R = \hat{R}$.

4. Let Q be an orthogonal matrix, prove that for any other $n \times n$ matrix A we have that $\|AQ\|_2 = \|A\|_2$.

5. Let Q be an orthogonal matrix, prove that for any other orthogonal matrix U, QU is also an orthogonal matrix.

6. Prove that for any real matrix $A \in \mathbb{R}^{m \times n}$ with $m \geq n$ $A^\top A$ is symmetric and if $A^\top A$ is nonsingular, then it is positive definite.

7. Consider the following lemma:

   Lemma 1. *Let $A \in \mathbb{R}^{m \times m}$ be spd (symmetric positive definite matrix). Let $X \in \mathbb{R}^{m \times n}$ where $m \geq n$ be full rank (column rank because $m \geq n$). Then $X^T A X$ is spd.*

   Now consider the following A spd and blocking like:

   $$\begin{aligned} A &= \begin{bmatrix} a_{11} & w_i^T \\ w_1 & k_1 \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1 & 0 \\ \frac{w_1}{\alpha_1} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & k_1 - w_1 \end{bmatrix} \begin{bmatrix} \alpha_1 & \frac{w_1^T}{\alpha_1} \\ 0 & I \end{bmatrix} \\ &= L_1 A_1 L_1^\top. \end{aligned}$$

   Use the last lemma to prove that A_1 spd implies $k_1 - \frac{w_1 w_1^T}{d}$ is spd