Exercise class 8

Learning Objectives:

- Convexity of functions
- Line search: Bisection and Newton's methods

Demo 1: Bisection method

a) Minimise $x^4 - 3x^3 + 2x$ using the bisection method with input interval [-4,4] and tolerance $\varepsilon = 0.01$.

b) Can you be sure that the solution of the algorithm is indeed an optimal solution?

c) Identify all stationary points. Start by using [-2,4] as new starting interval for the bisection method.

d) Find the global minimum.

Algorithm 1 Bisection method

1: **initialise.** tolerance $l > 0$, $[a_0, b_0] = [a, b]$, $k = 0$.
2: **while** $b_k - a_k > l$ **do**
3: $\quad\lambda_k = \frac{(b_k + a_k)}{2}$ and evaluate $f'(\lambda_k)$.
4: $\quad$**if** $f'(\lambda_k) = 0$ **then** return λ_k.
5: $\quad$**else if** $f'(\lambda_k) > 0$ **then**
6: $\quad\quad a_{k+1} = a_k,\ b_{k+1} = \lambda_k$.
7: $\quad$**else**
8: $\quad\quad a_{k+1} = \lambda_k,\ b_{k+1} = b_k$.
9: $\quad$**end if**
10: $\quad k = k + 1$.
11: **end while**
12: **return** $\overline{\lambda} = \frac{a_k + b_k}{2}$.

Solution

a) First, one iteration by hand: $f'(x) = 4x^3 - 9x^2 + 2$
$|b_0 - a_0| = 4 + 4 > 0.01$
$\lambda_0 = \frac{4-4}{2} = 0$
$f'(\lambda_0) = 2 > 0$
$a_1 = -4,\ b_1 = \lambda_0 = 0$
This iterative process continues until $|b_k - a_k| < 0.01$.
Applying the Bisection method in Julia gives the following iterations:

```
a_0 = -4 b_0 = 0.0
a_1 = -2.0 b_1 = 0.0
a_2 = -1.0 b_2 = 0.0
a_3 = -0.5 b_3 = 0.0
a_4 = -0.5 b_4 = -0.25
a_5 = -0.5 b_5 = -0.375
a_6 = -0.4375 b_6 = -0.375
a_7 = -0.4375 b_7 = -0.40625
a_8 = -0.4375 b_8 = -0.421875
a_9 = -0.4375 b_9 = -0.4296875
λ = -0.4296875
```

b) Can we trust that the final iteration is the global minimum? No, we cannot fully trust the algorithm since we have not checked if the function is convex. We can only trust the Bisection method if it is applied to convex function.

Let's check it: x^4 is convex, $-3x^3$ is not convex since we are not restricted just by positive real values, and $2x$ is convex (actually, linear so both convex and concave). So, the function $x^4 - 3x^3 + 2x$ is not convex. Therefore, the point we have found $\bar{x_1} = -0.4297$ is a stationary point (inflection), but not necessarily a global optimum.

c) Let's investigate a new starting interval $[-2, 4]$. First iteration by hand:

$|b_0 - a_0| = 4 + 2 > 0.01$
$\lambda_0 = \frac{4-2}{2} = 1$
$f'(\lambda_0) = 4(1^3) - 9(1^2) + 2 = -3 < 0$
$a_1 = \lambda_0 = 1,\ b_1 = 4$

Applying the Bisection method in Julia gives the following iterations:

```
a_0 = 1.0 b_0 = 4
a_1 = 1.0 b_1 = 2.5
a_2 = 1.75 b_2 = 2.5
a_3 = 2.125 b_3 = 2.5
a_4 = 2.125 b_4 = 2.3125
a_5 = 2.125 b_5 = 2.21875
a_6 = 2.125 b_6 = 2.171875
a_7 = 2.125 b_7 = 2.1484375
a_8 = 2.13671875 b_8 = 2.1484375
a_9 = 2.13671875 b_9 = 2.142578125
λ = 2.142578125
```

We have identified another stationary point $\bar{x_2} = 2.1426$. We have now identified 2 stationary points, but, in fact, we can have upto 3 stationary points in this function as $f'(x) = 4x^3 - 9x^2 + 2$.

Solving for $f'(x) = 0$ gives $\bar{x_1}$, $\bar{x_2}$, and $\bar{x_3} = 0.54088$.

d) Let's check optimality conditions to investigate the type of stationary points.

Both $\bar{x_1} = -0.4297$ and $\bar{x_2} = 2.1426$ satisfy necessary conditions (see slide 16 for definitions) as:

$$f'(\bar{x_1}) = f'(\bar{x_2}) = 0$$

Do the points also satisfy sufficient conditions (slide 17)?

Additional certificate:
$f''(\bar{x_1}) = 12\bar{x_1}^2 - 18\bar{x_1} = 9.9503 > 0$,thus $\bar{x_1}$ is a local minimum.
$f''(\bar{x_2}) = 12\bar{x_2}^2 - 18\bar{x_2} = 16.4652 > 0$ thus $\bar{x_2}$ is a local minimum.

So both points satisfy necessary and sufficient optimality conditions.

It only remains to check optimality conditions for $\bar{x_3}$.

$f'(\bar{x_3}) = 4\bar{x_3}^3 - 9\bar{x_3}^2 + 2 \approx 0$ (within our tolerance).
$f''(\bar{x_3}) = 12\bar{x_3}^2 - 18\bar{x_3} = -6.2252 < 0$, thus $\bar{x_3}$ is a local maximum.

Finally, as $-4.1483 = f(\bar{x_2}) < f(\bar{x_1}) = -0.5872$, $\bar{x_2}$ is the global minimum for this function.

Demo 2: Newton's method vs. approx. Newton's method

Solve $f(x) = \frac{2}{3}x^3 - \frac{8}{3}x$ with Newton's method and the approximated Newton's method (also known as the secant method), which approximate the second-order derivative by $\frac{f'(b)-f'(a)}{b-a}$.

Start with initial value $x_0 = -3$ and (for the approximation) $x_1 = -2.9$. Set the tolerance to 0.001.

Algorithm 2 Newton's method

1: **initialise.** tolerance $\epsilon > 0$, initial step size x_0, iteration count $k = 0$.
2: **while** $|f'(x_k)| > \epsilon$ **do**
3: $\quad x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$.
4: $\quad k \leftarrow k + 1$.
5: **end while**
6: **return** $\overline{x} = x_k$.

Solution

$f'(x) = 2x^2 - 8/3$
$f''(x) = 4x$

First iteration by hand:

$|f'(x_0)| = 2(-3)^2 - 8/3 = 15.33 > 0.001$
$x_1 = -3 + \frac{15.33}{12} = -1.722$
$|f'(x_1)| = 2(-1.722)^2 - 8/3 = 3.26 > 0.001$

Applying Newton's line search in Julia gives the following iterations:

```
Iteration 1: -1.7222222222222222
Iteration 2: -1.2482078853046594
Iteration 3: -1.158203009414713
Iteration 4: -1.1547058342139431
x = -1.1547058342139431
```

Now, for the approximate Newton's method.

This is the same, except we approximate the second-order derivative by:

$$d = \frac{f'(b) - f'(a)}{b - a}$$

As $x_0 = -3$ and $x_1 = -2.9$, $d_1 = -11.799$.

Doing the first iteration by hand:

$|f'(x_0)| = 2(-3)^2 - 8/3 = 15.33 > 0.001$
$x_1 = -3 + \frac{15.33}{11.799} = -1.70056$
$|f'(x_1)| = 2(-1.70056)^2 - 8/3 = 3.26543 > 0.001$

Applying approximate Newton's method in Julia gives:

```
Iteration 1: -1.7005649717514109
Iteration 2: -1.3617831266118132
Iteration 3: -1.191613069566423
Iteration 4: -1.1576941758366486
Iteration 5: -1.1547475746859013
Iteration 6: -1.1547005992714445
x = -1.1547475746859013
```

Newton's method gave $x = -1.15471$ and approximate Newton's gave $x = -1.15475$.

So why use approx. method? Because you don't need the second-order derivative, which, depending on the function, can be extremely hard to calculate.

Problem 1: Convexity

Show that, if f is convex and differentiable, then $f(y) \geq f(x) + f'(x)(y - x)$ for any y (i.e., the function is 'above' the first-order approximation at every point).

Solution

f is convex, therefore we have $\forall \lambda : 0 < \lambda < 1$ and $\forall x, y \in \mathbb{R}^n$:

$$f(\lambda y + (1 - \lambda)x) \leq \lambda f(y) + (1 - \lambda) f(x)$$

Reorder the left hand side:

$$f(x + \lambda(y - x)) \leq \lambda f(y) + (1 - \lambda) f(x)$$

Divide both sides by λ and reorder:

$$f(y) \geq \frac{f(x + \lambda(y - x))}{\lambda} - \frac{f(x)}{\lambda} + f(x)$$

Combine the fractions from the right hand side and divided and multiplied it by (y-x):

$$f(y) \geq \frac{f(x + \lambda(y - x)) - f(x)}{\lambda(y - x)}(y - x) + f(x)$$

Take the limit $\lambda \to 0$, making the fraction the definition of the derivative:

$$f(y) \geq f'(x)(y - x) + f(x)$$

Problem 2: Convex or concave?

Are these functions convex, concave, or both?

1. $f_1(x) = 4x - x^2$
2. $f_2(x) = ax + b$ for $a, b \in \mathbb{R}$
3. $f_3(x) = \frac{8}{log(x)^2} : x > 1$
4. $f_4(x) = e^x - 3x$

Solution

1. Concave. x^2 is convex, thus, $-x^2$ is concave. $4x$ is linear, thus, both concave and convex. In our case we will treat is as concave. The linear combination of the concave functions with positive coefficients is concave.

2. Both. Affine function is always simultaneously concave and convex

3. Convex. First, reformulate to $f_3(x) = 2 \cdot (\frac{2}{log(x)})^2$. Now, $log(x)$ is concave and > 0 when $x > 1$, Thus. $h(x) = 2/log(x) : x > 1$ is convex. $g(y) = 2y^2$ is convex and non decreasing for $y = h(x) > 0$. Therefore, the combination $g(h(x))$ of these functions is convex for $x > 1$.

4. Convex. e^x is convex. $-3x$ is both, but we will consider it as convex here. The linear combination of the convex functions with positive coefficients is convex.

Problem 3: Bisection method

Apply the bisection method for solving $\max. -4x^2 + 2x - 1$. Use [-1,1] and tolerance 0.01. Solve by hand, then check with Julia. Does the method find a global optimum?

Solution

f is concave as (concave + convex/concave + constant). Thus, Bisection method returns a global optima for maximisation.

Iterations in Julia are:

```
a_0 = 0.0 b_0 = 1
a_1 = 0.0 b_1 = 0.5
λ = 0.25
```

Problem 4: Newton's method

Apply Newton's method for $\max. - 4x^2 + 2x - 1$ from $x_0 = 2$. Why does it only need one step? Does this depend on x_0?

Solution

Solution: $x = 0.25$.

If we consider quadratic approximation:
$q = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}f''(x_k)(x - x_k)^2$

In our case:
$q = -4{x_k}^2 + 2x_k - 1 + (-8x_k + 2)(x - x_k) + \frac{1}{2}(-8)(x - x_k)^2$
$q = -4{x_k}^2 + 2x_k - 1 + -8x_kx + 8{x_k}^2 + 2x - 2x_k - 4x^2 + 8xx_k - 4{x_k}^2$

Thus, $q = -1 + 2x - 4x^2$ which is exact and we can find the roots of $q'(x_{k+1}) = -1 + 2x_{k+1} - 4{x_{k+1}}^2 = 0$ which do not depend on x_k or on the x_0 in our case.

The quadratic approximation is exact thus the method converged in one step and does not depend on x_0.

Problem 5: Newton's method

Solve $x^4 - 3x^3 + 2x$ using Newton's method with $\varepsilon = 0.01$, once with starting point 1 and once with 2. Why are the solutions different?

Solution

$f'(x) = 4x^3 - 9x^2 + 2$
$f''(x) = 12x^2 - 18x$

Applying Newton's method in Julia for each starting point gives:

```
starting value 1:
Iteration 1: 0.5
Iteration 2: 0.5416666666666666
x = 0.5416666666666666

 starting value 2:
Iteration 1: 2.1666666666666665
Iteration 2: 2.1415598290598292
Iteration 3: 2.1409137121182757
x = 2.1409137121182757
```

Solutions are different because the function is not convex (f'' can be both > 0 and < 0). Then, $f''(0.5417) = -6.2293$ and $f''(2.1409) = 16.4652$. Thus, 0.54 is a max., 2.14 is a min.

Home Exercise 8: Convexity

Find optimum/ optima for the following functions, justifying whether they are local or global minimum/ maximum.

1. $e^{(x^3-3x)} : x \geq 0$
2. $-2ln(x) + (x-1)^2$
3. $x^4 - 2x^3 + 2x$ (quasi convex - has a flat bit)

Hint. Use your preferred optimisation method and check the sufficiency of optimality conditions (i.e., check for convexity to assert whether the solution can be classified as global optimum).

Solution

1. What is the stationary point of function 1? Solution: 1
2. The stationary point for function 1 is a global minimum.
3. What is the stationary point of x for function 2? (to 2 decimal places) Solution: 1.62
4. The stationary point for function 2 is a global minimum.
5. What is an stationary point for function 3? (to 2 decimal places) Solution: -0.5
6. What is an stationary point for function 3? (to 2 decimal places) Solution: 1

7. For the stationary point that is a candidate for being a minimum or maximum, is, in fact, a global minimum.