Global exercise - GUE12

Tuan Vo

19th January, 2023

Summary of content covered: Numerics

- ✓ Optimization: Line search method + Step-by-step standard Newton method
- ✓ Optimization: Demo the 4th programming exercise PRU03

1 Line search method: Main points

1. Line search method is an iteration scheme with the form

$$x^{(k+1)} = x^{(k)} + \alpha_k \, d^{(k)} \tag{1}$$

2. Typically, for the sake of minimum-searching problem we take

$$\nabla f\left(x^{(k)}\right) d^{(k)} < 0. \tag{2}$$

3. Searching direction $d^{(k)}$ takes the form

$$d^{(k)} = -B_k^{-1} \nabla f\left(x^{(k)}\right) \tag{3}$$

   (i) When $B_k = I$: **Gradient descent** or **Steepest descent** method.

$$\begin{cases} x^{(k+1)} = x^{(k)} + \alpha_k \, d^{(k)} \\ d^{(k)} = -\nabla f\left(x^{(k)}\right) \end{cases} \tag{4}$$

   (ii) When $B_k = \nabla^2 f\left(x^{(k)}\right)$: **Newton** method.

$$\begin{cases} x^{(k+1)} = x^{(k)} + \alpha_k \, d^{(k)} \\ d^{(k)} = -\left(\nabla^2 f\left(x^{(k)}\right)\right)^{-1} \nabla f\left(x^{(k)}\right) \end{cases} \tag{5}$$

   (iii) When $B_k \approx \nabla^2 f\left(x^{(k)}\right)$: Quasi-**Newton** method.

4. Controlling the coefficient a_k leads to examining the two **Wolfe** conditions:

   (i) Upper bound of a_k: **Armijo** condition.

   (ii) Lower bound of a_k: Curvature condition.

2 Step-by-step: standard Newton method

Example 1. *Given the following function f as follows*

$$f : \begin{cases} \mathbb{R} \to \mathbb{R}, \\ x \mapsto f(x) := \sqrt{1+x^2}, \end{cases} \tag{6}$$

and the standard Newton method is taken into consideration

$$\begin{cases} x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)} \\ d^{(k)} = -\left(\nabla^2 f\left(x^{(k)}\right)\right)^{-1} \nabla f\left(x^{(k)}\right) \end{cases} \tag{7}$$

(i) Compute 3 steps of the Newton method using $\alpha_k = 1/2$ and $x^{(0)} = -3/2$.

(ii) Given $\alpha_k = 1$. Check all possibilities of the starting points, and examine whether the standard Newton method is converged, diverged, or oscillating.

Approach:

(i) First, we shall need to compute the gradient $\nabla f(x)$ as follows

$$\nabla f(x) = \frac{d}{dx} f(x) = \frac{\left(1+x^2\right)'}{2\sqrt{1+x^2}} = \frac{x}{\sqrt{1+x^2}} \tag{8}$$

and the second gradient $\nabla^2 f$ is computed as follows

$$\nabla^2 f(x) = \frac{d^2}{dx^2} f(x) = \frac{d}{dx}\left(\frac{x}{\sqrt{1+x^2}}\right) = \left(1+x^2\right)^{-3/2} \tag{9}$$

Next, the direction $d^{(k)}$ for the standard Newton method is ready computed

$$\begin{aligned} d^{(k)} &= -\left(\nabla^2 f\left(x^{(k)}\right)\right)^{-1} \nabla f\left(x^{(k)}\right) && (10) \\ &= -\left(1+\left(x^{(k)}\right)^2\right)^{3/2} \frac{x^{(k)}}{\left(1+\left(x^{(k)}\right)^2\right)^{1/2}} && (11) \\ &= -x^{(k)}\left(1+\left(x^{(k)}\right)^2\right) && (12) \end{aligned}$$

Then, taking into consideration $\alpha_k = 1/2$ the next step of the standard Newton method takes the form

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \alpha_k\, d^{(k)} && (13) \\ &= x^{(k)} - \frac{1}{2} x^{(k)}\left(1+\left(x^{(k)}\right)^2\right) && (14) \end{aligned}$$

With the starting point $x^{(0)} = -3/2$ we gradually obtain the next steps

(a) First step:

$$x^{(1)} = x^{(0)} - \frac{1}{2}x^{(0)}\left(1 + \left(x^{(0)}\right)^2\right) = \dots \tag{15}$$

(b) Second step:

$$x^{(2)} = x^{(1)} - \frac{1}{2}x^{(1)}\left(1 + \left(x^{(1)}\right)^2\right) = \dots \tag{16}$$

(c) Third step:

$$x^{(3)} = x^{(2)} - \frac{1}{2}x^{(2)}\left(1 + \left(x^{(2)}\right)^2\right) = \dots \tag{17}$$

(ii) Examine the convergence

Observation 1. *When studying the convergence of the line search in general or the Newton method in particular, we shall need to know first the minimum x^* of the function $f(x)$, so that we may examine the following expression*

$$\lim_{k\to\infty} \left|x^{(k+1)} - x^*\right|. \tag{18}$$

Should the scheme converges, the limit results in 0.

By taking into consideration (8) we may compute the extremum

$$\nabla f(x) = \frac{x}{\sqrt{1+x^2}} = 0 \Rightarrow x^* = 0. \tag{19}$$

Then, the extremum $x^* = 0$, which is a local minimum by considering the (9), can be inserted into the (18) to examine the converged ability

$$\lim_{k\to\infty} \left|x^{(k+1)} - x^*\right| \overset{!}{=} 0 \Leftrightarrow \lim_{k\to\infty} \left|x^{(k+1)} - 0\right| \overset{!}{=} 0 \tag{20}$$

$$\Leftrightarrow \lim_{k\to\infty} \left|x^{(k+1)} - 0\right| \overset{!}{=} 0 \tag{21}$$

$$\Leftrightarrow \lim_{k\to\infty} \left|-\left(x^{(k)}\right)^3\right| \overset{!}{=} 0 \overset{(\star)}{\Rightarrow} \left|x^{(0)}\right| < 1 \tag{22}$$

where $(\star)$ is enabled only if the radius of the starting point $x^{(0)}$ is less than 1.

Therefore, it leads us to the following conclusion

(i) $\left|x^{(0)}\right| < 1$: The method will be converged.

(ii) $\left|x^{(0)}\right| = 1$: The method will be oscillating, i.e. by flipping around.

(iii) $\left|x^{(0)}\right| > 1$: The method will be diverged.

Note in passing that we have computed the Newton step in the context of $\alpha_k = 1$ as follows

$$x^{(k+1)} = x^{(k)} + \alpha_k \, d^{(k)} \tag{23}$$

$$= x^{(k)} - 1 \cdot x^{(k)} \left(1 + \left(x^{(k)} \right)^2 \right) \tag{24}$$

$$= - \left(x^{(k)} \right)^3 \tag{25}$$

3 Demo programming exercise PRU03

(White board + Projector)

Remark 1. *Rosenbrock function*

$$f : \begin{cases} \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \\ (x, y) \mapsto f(x, y) := a\,(y - x^2)^2 + (1 - x)^2, \end{cases} \tag{26}$$

where $a = 100$, *is well-known in Optimization. It is often used as a test problem for evaluating algorithms in Optimization. This function has a* ***global minimum*** $f(x^*, y^*) = 0$ *at the point* $(x^*, y^*) = (1, 1)$.

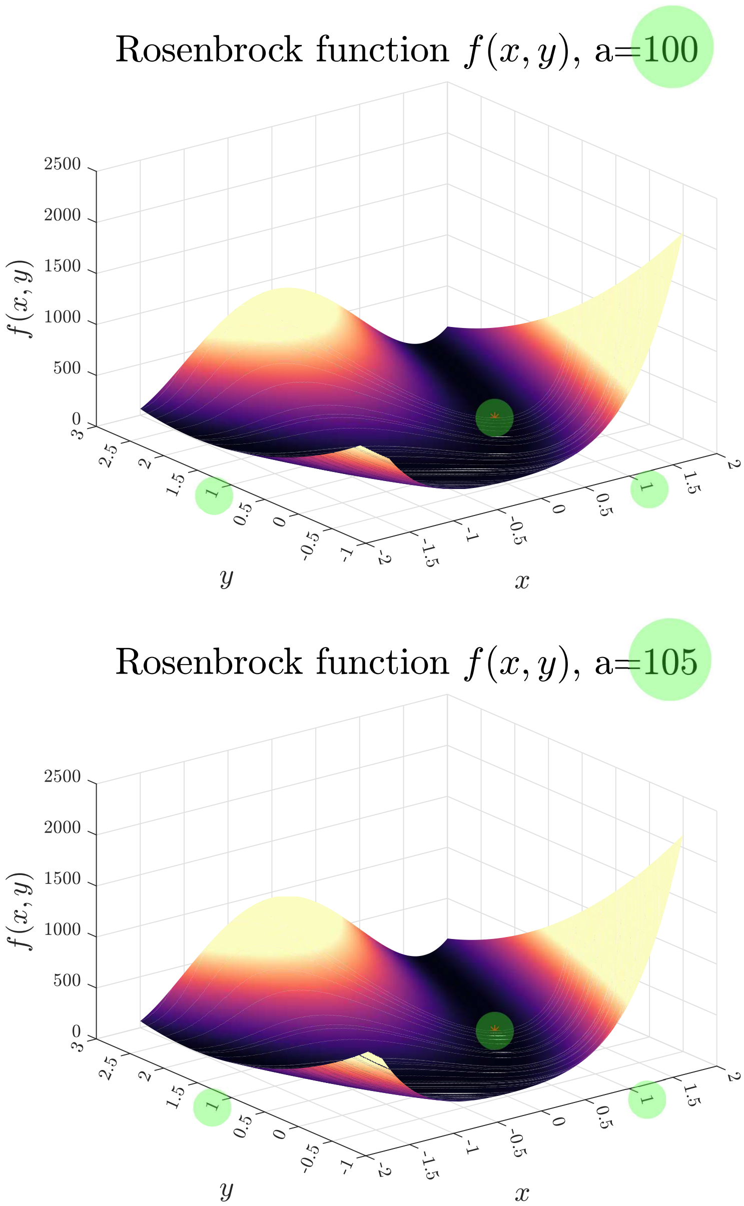

Figure 1: Rosenbrock function and its only global minimum.