Aopr

Schaller Alain, Fontana Jonas, Carrel Vincent

Exercise 5: line Search Method

1) Exact line search:

$$f(x) = \frac{1}{2} x^T Q x + q^T x + c$$

where $x \in \mathbb{R}^n$, Q symmetric positive definite matrix $\in \mathbb{R}^{n \times n}$, $q \in \mathbb{R}^n$, $c \in \mathbb{R}$.

Given a search direction Δx, compute t for an arbitrary $x \in \operatorname{dom} f$

$$t := \underset{s \geq 0}{\operatorname{argmin}}\ f(x + s\Delta x)$$

Be x arbitrary, then

$$g(s) = f(x + s\Delta x) = \frac{1}{2}(x + s\Delta x)^T Q (x + s\Delta x) + q^T(x + s\Delta x) + c$$

$$= \frac{1}{2} x^T Q x + \frac{1}{2} x^T Q s \Delta x + \frac{1}{2} s \Delta x^T Q x + \frac{1}{2} s^2 \Delta x^T Q \Delta x$$

$$+ q^T x + s q^T \Delta x + c$$

$$\nabla f = \frac{\partial f(x)}{\partial x} = Qx + q.$$

To minimize, we need $\frac{\partial g}{\partial s} = 0$

$$\Rightarrow \frac{\partial g}{\partial s} = \frac{1}{2} x^T Q \Delta x + \frac{1}{2} \Delta x^T Q x + s \Delta x^T Q \Delta x + q^T \Delta x = 0$$

$$\Rightarrow -s(\Delta x^T Q \Delta x) = x^T Q \Delta x + q^T \Delta x$$

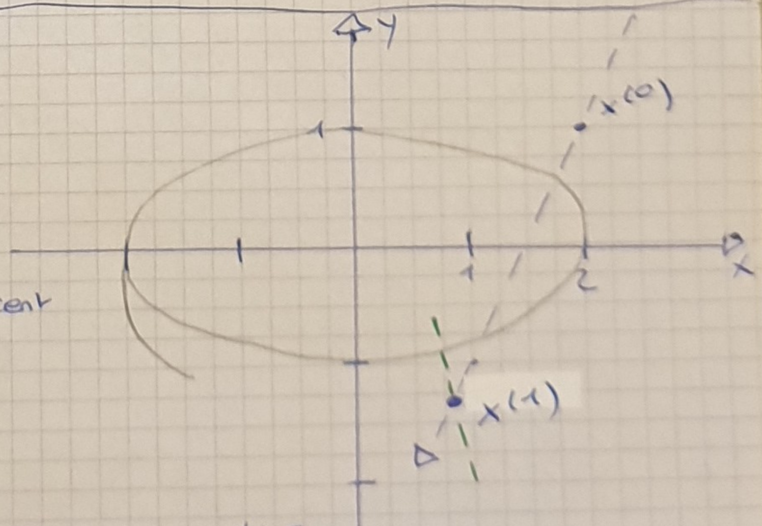
$$\Rightarrow -s(\Delta x^T Q \Delta x) = (x^T Q + q^T)^T \Delta x = \nabla f(x)^T \Delta x$$

$$\Delta x^T Q x = (x^T Q^T \Delta x)^T = (x^T Q \Delta x)^T = x^T Q \Delta x$$

$$\boxed{t = \frac{-\nabla f(x)^T \Delta x}{\Delta x^T Q \Delta x}}$$

---

2) Gradient descent:



minimize $\frac{1}{4} x^2 + y^2$

starting from $x^{(0)} = (2, 1)$

perform one iteration of gradient descent

find the value of $\|\nabla f(x)\|^2$

before and after the update

$$x^{(1)} = x^{(0)} + t^{(0)} \Delta x^{(0)}$$

we can guess the next $\Delta x^{(1)}$ approx in green

- find the search direction:

$$\Delta x = -\nabla f(x) = \begin{pmatrix} -x/2 \\ -2y \end{pmatrix}, \text{ with } x^{(0)} = (2,1) \Rightarrow \Delta x^{(0)} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

$$f(x,y) = (x \;\; y)\begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \text{ so we have } Q = \begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix}$$

We can use the formula from ex 1 to find $t^{(0)}$:

$$t^{(0)} = \frac{-\nabla f(x)^T \Delta x}{\Delta x^T Q \Delta x} = \frac{-(1 \;\; 2)\begin{pmatrix} -1 \\ -2 \end{pmatrix}}{(-1 \;\; -2)\begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ -2 \end{pmatrix}} = \frac{5}{(-1/4 \;\; -2)\begin{pmatrix} -1 \\ -2 \end{pmatrix}} = \frac{5}{\frac{17}{4}} = \frac{20}{17}$$

$$x^{(1)} = x^{(0)} + t^{(0)} \Delta x^{(0)}$$

$$= \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \frac{20}{17}\begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 14/17 \\ -23/17 \end{pmatrix} \approx \begin{pmatrix} 0{,}82 \\ -1{,}35 \end{pmatrix}$$

$$\Rightarrow x^{(1)} = \left(\frac{14}{17}, \frac{-23}{17}\right)$$

Now calculate:

$$\|\nabla f(x^{(0)})\|^2 = \left\|\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\|^2 = 1 + 4 = 5$$

$$\|\nabla f(x^{(1)})\|^2 = \left\|\begin{pmatrix} 14/34 \\ -46/17 \end{pmatrix}\right\|^2 \approx 0{,}17 + 7{,}32 \approx 7{,}5$$