Université Paris Saclay	**MSc AI & MSc Data Science.**
TC2-Optimization for Machine Learning	Faïcel Chamroukhi

TP - December 05, 2024.

Exercice 1. Gradient Descent and variants: Given $f(x) = x^2$, $x \in \mathbb{R}$. Write a Python/Jupyter code to

1. Minimize f by gradient descent with a fixed step size (eg. $\alpha = 0.1$)
2. Minimize f by classical momentum (eg. use $\mu = 0.6$).
3. Consider now a Nesterov accelerated gradient descent.
4. Compare and comment your results by considering the same starting values (for example $x^{(0)} = 0.6, \alpha = 0.1$ for $f(x) = x^2$). eg. observe across algorithm's iterations, the solution's paths, the gradient paths, the function values along the paths, the norm of the gradients , etc
5. Use it for $f(x) = (x-3)^2 + 2x$, $x \in \mathbb{R}$.

Exercice 2. Newton Method Now write a Python/Jupyter code to

1. Minimize f with the Newton method
2. What do you observe ? How do you explain this ?
3. Use the Newton method for $f(x) = (x-3)^2 + 2x$, $x \in \mathbb{R}$.
4. Comment.

Exercice 3. Gradient Descent and variants, and Newton method for multivariate problms Given $f(x) = x^t A x$, with $A = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2, \end{pmatrix}$ $x \in \mathbb{R}^2$.
Write a Python/Jupyter code to

1. Minimize f by gradient descent with a fixed step size (eg. $\alpha = 0.1$)
2. Minimize f by gradient descent with the step size chozen by Armijo's method (eg. set $\sigma = 0.1$, and $\beta^{(0)} = 1$ and then reduce it by some factor).
3. Minimize f by classical momentum (eg. use $\mu = 0.8$).
4. Consider Nesterov momentum (use the same momentum eg. $\mu = 0.8$).
5. Consider the Newton method. What do you observe ? Why this ?
6. Compare and comment your results by considering the same starting values (eg. $x^{(0)} = (2.5, 2.5)^T, \alpha = 0.1$)
   eg. observe across algorithm's iterations, the solution's paths, the gradient paths, the function values along the paths, the norm of the gradients similarly, etc