METHODS OF NONLINEAR OPTIMIZATION: HW#2

(1) Show that the Rosenbrock (see the text) is not a convex function.

At what point (x_1, x_2) is the Hessian of the function positive definite?

At what point is the Hessian not positive definite?

Recall the basic (and easy to prove) fact that the Newton direction $-[\nabla^2 f(x)]^{-1}\nabla f(x)$ is a descent direction when $\nabla^2 f(x)$ is positive definite. This, however, does not say what happens to the Newton direction when $\nabla^2 f(x)$ is not positive definite.

For the Rosenbrock function, check that the Newton direction at $x = (2, 5)$ is **not** a descent direction.

(Extra credits.) Is it possible that the Hessian is not positive definite at a point but yet the Newton direction is a descent direction?

(2) Consider again the function $f(x) = x^4$. What happens if we apply Newton's method to find its minimizer? For what initial guesses would the method converge? What is the rate of convergence?
Is the rate of convergence quadratic? If not, which condition in Theorem 5.2 is violated?

(3) In HW#1 we see that the gradient descent methods are invariant under **orthogonal** – *but not arbitrary affine* – changes of coordinates. In fact, their performances can be severely affected by an non-orthogonal change of coordinates.

Prove that the pure Newton's method is invariant under arbitrary affine changes of coordinates, i.e. if $f : \mathbb{R}^n \to \mathbb{R}$ and $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ is defined by $\bar{f}(\bar{x}) = f(U\bar{x} + v)$, for any invertible matrix U and a vector v, then the pure Newton's method applied to f and $\bar{f}$ with initial guesses x_0 and $\bar{x}_0 = U^{-1}(x_0 - v)$ (respectively) results in iterates x_k and $\bar{x}_k$ that are related by $\bar{x}_k = U^{-1}(x_k - v)$ for any $k \geq 0$.

In a sense, Theorem 5.2 does not quite respect the affine invariance of Newton's method: the constants m, L, and the ratio L/m in the theorem all change under an non-orthogonal change of coordinates. Yurii Nesterov and Arkadi Nemirovski, among their many contributions to optimization theory, found a different way to analyze Newton's method that addresses this drawback.

Thomas Yu, Department of Mathematics, Drexel University, 3141 Chestnut Street, 206 Korman Center, Philadelphia, PA 19104, U.S.A.

E-mail address: `yut@drexel.edu`

Date: April 24, 2020.