Convex Optimization: Homework 2

Martín Prado

August 28, 2024
Universidad de los Andes – Bogotá Colombia

Exercise 1.

Calculate the minimum over $\mathbb{R}^2_{++}$ of the function $f(x, y) = \frac{1}{xy} + x + y$.

Solution: The function is differentiable on $\mathbb{R}^2 \backslash \{0\}$. Therefore, we have derivatives in $\mathbb{R}^2_{++}$.

$$Df = \begin{pmatrix} 1 - x^{-2}y^{-1} \\ 1 - x^{-1}y^{-2} \end{pmatrix}.$$

Critical points are found when

$$\begin{array}{rc} & x^{-2} = y \\ & y^{-2} = x \\ \hline \implies & x^4 = x \\ \iff & x^4 - x = 0 \end{array}$$

and since $x, y \in \mathbb{R}^2_{++}$, the only possible solution for this is when $x = y = 1$. Therefore, the only critical point is $(1, 1)$. Now, the Hessian matrix is the following,

$$Hf(x, y) = \begin{bmatrix} \frac{2}{x^3 y} & \frac{1}{x^2 y^2} \\ \frac{1}{x^2 y^2} & \frac{2}{x y^3} \end{bmatrix}$$

Then,

$$Hf(1, 1) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix},$$

has positive eigenvalues $\lambda_1 = 3$, $\lambda_2 = 1$. Therefore, $(1, 1)$ is a minimum on $\mathbb{R}^2_{++}$, which evaluates $f(1, 1) = 3$.

Exercise 2.

Güler: Ex 2.9

Consider the function $f(x, y) = x^2 + \alpha xy^2 + 2y^4$. Show that all the parameter values except two, the origin $(0, 0)$ is the only critical point of f.

a) Find the exceptional α's and show that f has infinitely many critical points for these α values. Determine the nature of these critical points.

b) Consider the values of α's for which the origin is the only critical point. For each α, determine the nature of the critical point. Show that in some cases, the origin is a local minimum, but in other cases, it is a saddle point.

c) Show that even when the origin is a saddle point, (0,0) is a local strict minimizer of f on every line passing through the origin. In fact, show that, except for one line, the function $g(t) := f(td)$ satisfies $g'(0) = 0$ and $g''(0) > 0$.

Solution Part (a)

The derivative of f is

$$Df(x, y) = \begin{pmatrix} 2x + \alpha y^2 \\ 2\alpha xy + 8y^3 \end{pmatrix}.$$

By solving α in $Df(x, y) = 0$ we can find where f has infinitely many critical points. Then, we get the following equations, for $x, y \neq 0$:

$$\begin{aligned} \alpha &= \frac{-2x}{y^2} \\ \alpha &= \frac{-8y^3}{2xy} \end{aligned} \iff \frac{-2x}{y^2} = \frac{-8y^3}{2xy}$$

$$\iff x^2 = 2y^4$$

$$\iff y^2 = \pm\frac{x}{\sqrt{2}}.$$

Therefore, the only parameter values where there are infinitely many solutions are:

$$\alpha = \frac{-2x}{y^2} = \frac{-2x}{\pm x \cdot \sqrt{1/2}} = \pm 2\sqrt{2}.$$

When α is different from these values, the chain of equivalences we made before tells us that $Df(x, y) \neq 0$ when $x, y \neq 0$. Since $Df(0, 0) = 0$ for any α, it follows that $(0, 0)$ is the only critical point whenever $\alpha \neq \pm 2\sqrt{2}$.

Solution Part (b)

The Hessian matrix is

$$\begin{bmatrix} 2 & 2\alpha y \\ 2\alpha y & 2\alpha x + 24y^2 \end{bmatrix}$$

Solution Part (c)

Exercise 3.

Nocedal & Wright: Ex 3.1

Program the steepest descent and Newton algorithms using the backtracking line search, Algorithm 3.1. Use them to minimize the Rosenbrock function (2.22). Set the initial step length $\alpha = 1$ and print the step length used by each method at each iteration. First try the initial point $x_0 = (1.2, 1.2)^T$ and then the more difficult starting point $x_0 = (-1.2, 1)^T$.

Exercise 4.

Boyd & Vanderberghe: Ex 9.2

Minimizing a quadratic-over-linear fractional function. Consider the problem of minimizing the function $f : \mathbb{R}^n \to \mathbb{R}$ defined as

$$f(x) = \frac{\|Ax - b\|_2^2}{c^T x + d}, \qquad \mathrm{dom}(f) = \{x \ : \ c^T x + d > 0\}.$$

We assume rank$A$

Exercise 5.

Consider the function

$$f(x) = \langle c, x \rangle - \sum_{j=1}^{m} \log(1 - \langle a_j, x \rangle) - \sum_{i=1}^{n} \log(1 - x_i^2).$$

Take $n = 5000$ and $m = 2000$, and vectors c, a_j chosen arbitrarily such that $\|a_j\| = 1$. For every method initiate the iterations in $x_0 = 0$ and the termination criterion in the form $\|\nabla f(x_k)\| \leq \varepsilon$. Graph $\log(f(x_k) - f^*)$ on each iteration (estimating the best possible f^*) and indicate the total time each method takes.

a) Use the gradient method