Optimization 1 — Tutorial 8

December 10, 2020

Definition (Stationary Point)
Consider the following optimization problem:

$$\begin{array}{lll} (P) & \min & f(\mathbf{x}) \\ & \text{s.t.} & \mathbf{x} \in C, \end{array}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function and C is a closed and convex set. Then $\mathbf{x}^* \in C$ is called a stationary point of (P) if $\nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \geq 0$ for all $\mathbf{x} \in C$.

Theorem (Optimality Condition)
Let $\mathbf{x}^*$ be a local minimum of (P). Then $\mathbf{x}^*$ is a stationary point of (P).

Problem 1
Consider the problem

$$\begin{array}{lll} (P) & \min & f(\mathbf{x}) \\ & \text{s.t.} & \sum\limits_{i=1}^{n} \mathbf{x}_i = 1, \end{array}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. Denote $C = \left\{\mathbf{x} \in \mathbb{R}^n \colon \sum\limits_{i=1}^{n} \mathbf{x}_i = 1\right\}$. Show that $\mathbf{x}^* \in C$ is a stationary point of (P) if and only if

$$\frac{\partial f}{\partial \mathbf{x}_1}(\mathbf{x}^*) = \frac{\partial f}{\partial \mathbf{x}_2}(\mathbf{x}^*) = \ldots = \frac{\partial f}{\partial \mathbf{x}_n}(\mathbf{x}^*).$$

Problem 2
Consider the problem

$$\begin{array}{lll} (P) & \min & f(\mathbf{x}) \\ & \text{s.t.} & \|\mathbf{x}\|_2 \leq 1, \end{array}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. Show that $\mathbf{x}^* \in B[\mathbf{0}_n, 1]$ is a stationary point of (P) if and only if

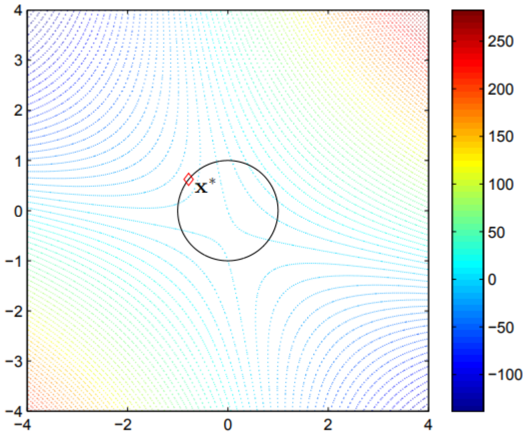
- $\nabla f(\mathbf{x}^*) = \mathbf{0}_n$, or
- $\|\mathbf{x}^*\| = 1$ and there exists $\lambda \leq 0$ such that $\nabla f(\mathbf{x}^*) = \lambda \mathbf{x}^*$.

Trust Region Sub-Problem

Consider the following optimization problem

$$\begin{aligned} \min \quad & 2x^2 + 2y^2 + 12xy + 3x + y \\ \text{s.t.} \quad & x^2 + y^2 \leq 1. \end{aligned}$$

Plotting this TRS we have



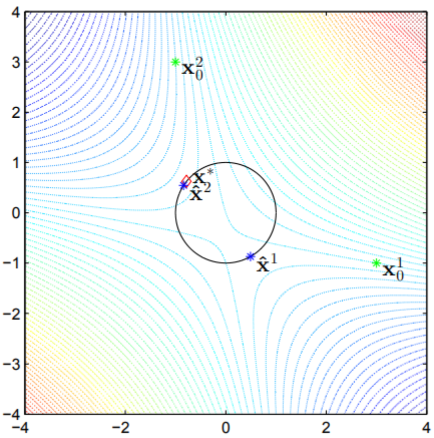
We saw in class that this problem can be rewritten as a convex optimization problem of the form

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \sum_{i=1}^{n} d_i \mathbf{z}_i - 2 \sum_{i=1}^{n} |\mathbf{f}_i| \sqrt{\mathbf{z}_i} + c \\ \text{s.t.} \quad & \sum_{i=1}^{n} \mathbf{z}_i \leq 1, \\ & \mathbf{z} \geq 0, \end{aligned}$$

where $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^T$ by the spectral decomposition theorem and $\mathbf{f} = \mathbf{U}^T\mathbf{b}$. Notice that $\mathbf{z}$ is a variable (obtained by the non-linear transformation $\left(\mathbf{U}^T\mathbf{x}\right)_i = -\text{sign}\left(\mathbf{f}_i\right)\sqrt{\mathbf{z}_i}$). Solving this problem with CVX obtains the solution $\mathbf{x}^* = (-0.781, 0.624)$ with function value -7.292.

Clearly, the original problem is non-convex and non-concave. Thus, convergence to the global optimal solution cannot be guaranteed and it depends on the starting point. For example, the Projected Gradient method with backtracking applied on TRS with initialization $\mathbf{x}^0 = (3, -1)$ yields the locally optimal point $(0.491, -0.871)$ with function value -1.929. The initialization $(-1, 3)$ results in convergence to the optimal point.

The following figure illustrates the initial points and the solutions obtained by the Projected Gradient method:



In conclusion – solving the problem using PG is less computationally demanding, but there are no guarantees to convergence to the optimum. Solving the problem using the non-linear transformation converges to the optimum, but with a great computational effort.

Problem 3

Show that the following function is convex and find its domain:

$$f(x,y,z) = \sqrt{2x^2 + 2y^2 + 5z^2 + 2xy + 2xz + 4yz - 4y + 4},$$

What are the minimum points?

Solution

We have

$$\begin{aligned} f(x,y,z) &= \sqrt{(x+y+2z)^2 + (x-z)^2 + (y-2)^2} \\ &\equiv \sqrt{k(x,y,z)} \\ &= \left\| \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} \right\|, \end{aligned}$$

So it is convex and $\operatorname{dom}(f) = \mathbb{R}^3$. Moreover, notice that finding the minimum pint of k is a least squares problem with solution $k\left(-\frac{2}{3}, 2, -\frac{2}{3}\right) = 0$. Meaning that f has no stationary points, $f \geq 0$, and its umique minimum point is $\left(-\frac{2}{3}, 2, -\frac{2}{3}\right)$ which is a non-differentiable point.

Some other techniques that do not involve completing the sqaures exist.