Optimization

Lusine Poghosyan

AUA

February 25, 2019

Newton's Method

Let $f \in \mathbb{C}^2(\mathbb{R}^n)$ and our aim is to find the minimizer of f.

Let $x^{(0)} \in \mathbb{R}^n$ be the starting point. Then we construct a quadratic function that matches its value, first and second derivatives at $x^{(0)}$ with that of the function f. This quadratic function has the form

$$q(x) = f\left(x^{(0)}\right) + \nabla f\left(x^{(0)}\right)^T \left(x - x^{(0)}\right) + \frac{1}{2}\left(x - x^{(0)}\right)^T \nabla^2 f\left(x^{(0)}\right)\left(x - x^0\right).$$

Then, instead of minimizing f, we minimize its approximation q.

The FONC for q yields

$$\nabla q(x) = \nabla f\left(x^{(0)}\right) + \nabla^2 f\left(x^{(0)}\right)\left(x - x^{(0)}\right) = 0.$$

The solution of this system

$$x^{(1)} = x^{(0)} - \left[\nabla^2 f\left(x^{(0)}\right)\right]^{-1} \nabla f\left(x^{(0)}\right)$$

will be our next approximation. Reapplying this procedure we get the sequence defined by Newton's Method

$$x^{(k+1)} = x^{(k)} - \left[\nabla^2 f\left(x^{(k)}\right)\right]^{-1} \nabla f\left(x^{(k)}\right), \quad k = 0, 1, \ldots.$$

The k-th iteration can be written in two steps:

1. Solve $\nabla^2 f\left(x^{(k)}\right) d^{(k)} = -\nabla f\left(x^{(k)}\right)$.
2. Set $x^{(k+1)} = x^{(k)} + d^{(k)}$.

- The convergence is local.
- Suppose that $f \in C^3$ and $x^* \in \mathbb{R}^n$ is a point such that $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is invertible. Then, for all $x^{(0)}$ sufficiently close to x^*, Newton's method is well-defined for all k and converges to x^* with an order of convergence at least 2.
- The direction of search is

$$d^{(k)} = -\left[\nabla^2 f\left(x^{(k)}\right)\right]^{-1} \nabla f\left(x^{(k)}\right).$$

If $\nabla^2 f\left(x^{(k)}\right)$ is positive definite, then $d^{(k)}$ is a descent direction.

Modification of Newton's method

The step size is usually $\alpha_k = 1$ but sometimes one takes other step size and gets

$$x^{(k+1)} = x^{(k)} - \alpha_k \left[\nabla^2 f\left(x^{(k)}\right)\right]^{-1} \nabla f\left(x^{(k)}\right), \quad k = 0, 1, \ldots .$$

For example we can take

$$\alpha_k = \text{arg min}_{\alpha \geq 0} f\left(x^{(k)} - \alpha \left[\nabla^2 f\left(x^{(k)}\right)\right]^{-1} \nabla f\left(x^{(k)}\right)\right)$$

to ensure that $f\left(x^{(k+1)}\right) < f\left(x^{(k)}\right)$.

Stopping conditions

- $||\nabla f\left(x^{(k)}\right)|| < \varepsilon$
- $||x^{(k+1)} - x^{(k)}|| < \varepsilon$ or $\frac{||x^{(k+1)} - x^{(k)}||}{||x^{(k)}||} < \varepsilon$ if $||x^{(k)}|| \neq 0$
- $\left|f\left(x^{(k+1)}\right) - f\left(x^{(k)}\right)\right| < \varepsilon$ or $\frac{\left|f\left(x^{(k+1)}\right) - f\left(x^{(k)}\right)\right|}{\left|f\left(x^{(k)}\right)\right|} < \varepsilon$ if $f\left(x^{(k)}\right) \neq 0$.

Example

Assume we want to use the Newton's Method to minimize

$$f(x_1, x_2) = 2x_1^2 + x_2^2 - 2x_1x_2.$$

We start with $x^{(0)} = (1, 1)^T$. Calculate $x^{(2)}$ by using the Newton's Method. Explain why after one iteration we have that $\nabla f\left(x^{(1)}\right) = 0$.