Math Problem Set #7: Nonlinear Optimization

OSM Lab - University of Chicago
Geoffrey Kocks

Problem 1: HJ 9.3

Solution. An overview of each of the higher-dimensional methods discussed in the first section are summariezed in the table below:

Method	**Description**	**Types of Problems**	**Strengths**	**Weaknesses**
Gradient Descent	always moves in the direction of the greatest increase of $\mathbf{x}$ until convergence; each approximation given by $\mathbf{x}_{i+1} = \mathbf{x}_i - \alpha D^T(\mathbf{x}_i)$	useful for getting close to the optimum if the initial value of x is far away, and then other methods can be used	in some cases may converge quite quickly, line search can be performed to find the optimal choice of α	convergence may be very slow and zig-zag if the problem has a lot of canyons or troughs
Newton and Quasi-Newton	descent method and local approximation method that takes advantage of the Taylor polynomial with the update $\mathbf{x}_{i+1} = \mathbf{x}_i + (D^2f(\mathbf{x}_i))^{-1}Df^T(\mathbf{x}_i)$ at each iteration until convergence	typically used for problems with low dimensions and if the Hessian is positive definite and can be computed easily	quadratic convergence speed so converges very rapidly; if f is quadratic function of specific forms then will reach the optimizer from any starting point in just one iteration	difficulty converging when the starting point is far from the optimizer or when $Df^2(\mathbf{x}_i)$ is not positive definite
Conjugate Gradient	moves toward the minimizer of a function by moving along Q-conjugate directions	useful for solving large quadratic optimization problems	each step is very inexpensive compared to Newton's Method, and guaranteed to optimize a quadratic of n variables in n steps; does not have to retain much information from previous steps	may take many steps to converge

Problem 2: HJ 9.6

Solution. See the attached Jupyter notebook for the code of the steepest descent function.

Problem 3: HJ 9.7

Solution. See the attached Jupyter notebook.

Problem 4: HJ 9.10

Solution. Let $f(\mathbf{x}) = 0.5\mathbf{x}^T Q\mathbf{x} - \mathbf{b}^T\mathbf{x}$. Because the function is quadratic, we know that the unique minimizer satisfies the first order condition:

$$Df = Q\mathbf{x} - b = 0 \implies Q\mathbf{x}^* = b$$

We then apply Newton's method, also using the fact that $D^2 f(\mathbf{x_0}) = Q$. Then by Newton's method:

$$\mathbf{x}_1 = \mathbf{x}_0 - Q^{-1}Df(\mathbf{x}_0)$$

$$Q\mathbf{x}_1 = Q\mathbf{x}_0 + Df(\mathbf{x}_0) = Q\mathbf{x}_0 - Q\mathbf{x}_0 + b = b.$$

$Q\mathbf{x}^* = b = Q\mathbf{x}_1$ so one interation of Newton's method will give the unique minimizer.