CS5321 Numerical Optimization Homework 1

Due Oct 28

1. (30%) For a single variable unimodal function $f \in [0, 1]$, we want to find its minimum. We have introduced the binary search algorithm in the class. But in each iteration, we need two function evaluations, $f(x_k)$ and $f(x_k+\epsilon)$. Here is another type of algorithms, called ternary search. Figure 1 illustrates the idea. The initial triplet of x values is $\{x_1, x_2, x_3\}$.

   GoldenSectionSearch.png

   Figure 1: The idea of ternary search.

   Answers are put here.
   也可以使用中文回答

   (a) (10%) For the search direction, show that to find the minimum point, if $f(x_4) = f_{4a}$, the triplet $\{x_1, x_2, x_4\}$ is chosen for the next iteration. If $f(x_4) = f_{4b}$, the triplet $\{x_2, x_4, x_3\}$ is chosen. (Hint: use the property of unimodal.)

   (b) (10%) For either case, we want these three points keep the same ratio, which means
   $$\frac{a}{b} = \frac{c}{a} = \frac{c}{b-c}.$$
   Show that under this condition, the ratio of $b/a = (\sqrt{5}+1)/2$, which is the golden ratio ϕ. (So this algorithm is called the *Golden-section search*).

   (c) (10%) If we let each iteration of the algorithm has two function evaluations, show the convergence rate of the Golden-section search is ϕ^{-2}. (This means it is faster than the binary search algorithm under the same number of function evaluations.)

   a) In this situation, if $f(x_4) = f_{4a}$, which means the lowest point is on the left of x4, so the next iteration is $\{x_1, x_2, x_4\}$; If $f(x_4) = f_{4b}$, the lowest point is on the right of x2, so the next iteration is $\{x_2, x_4, x_3\}$
   b) golden ratio means : (a + b)/a = a/b=(5 + 1)/2 (a $\succ b \succ 0).a/b = c/a = c/(b-c) : takec = a*a/b \Rightarrow a(a*a+a*b-b*b) = 0 \Rightarrow a(a+b) = b*b \Rightarrow b/a = (a+b)/b,\ then\ we\ find\ it\ is\ golden\ ratio$
   $c)we\ set\ a, b, \epsilon.$
   b - 0.618(b - a) $\Rightarrow a(1)f(a(1)) \Rightarrow f1$
   $a + 0.618(b-a) \Rightarrow a(2)f(a(2)) \Rightarrow f2$
   $iff1 < f2 : a(2) \Rightarrow b, a(1) \Rightarrow a(2), f1 \Rightarrow f2, b-0.618(b-a) \Rightarrow a(1), f(a(1)) \Rightarrow f1$

$else, a(1) \Rightarrow a, a(2) \Rightarrow a(1), f2 \Rightarrow f1, a + 0.618(b-a) \Rightarrow a(2), f(a(2)) \Rightarrow f2$
$if b - a \leq \epsilon$ end
else go if f1 less than f2...
in these steps, we have two function evaluations
so every epoch we combine two function together
so $\frac{x_n+1}{x_n} = \phi$, the rate of Golden-section search is ϕ^{-2}

2. (15%) Show that Newton's method for single variables is equivalent to build a quadratic model

$$q(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{f''(x_k)}{2}(x - x_k)^2$$

at the point x_k and use the minimum point of $q(x)$ as the next point. (Hint: to show the next point $x_{k+1} = x_k - f'(x_k)/f''(x_k)$)

in x_k, *we need to find the point where the derivative is* 0, *so we need to differentiate both sides of the above equation at the same time, which is* q'(x) = f"$(x_k)(x - x_k) + f'(x_k) = 0$, so the next point is $x(_k+1) = x_k - f'(x_k)/f''(x_k)$. just same as build a quadratic model.

3. (15%) Matrix A is an $n \times n$ symmetric matrix. Show that all A's eigenvalues are positive if and only if A is positive definite.

In mathematics, a symmetric matrixA with real entries is positive-definite if the real number x'Ax is positive for every nonzero real column vector. And spectral theorem is a result about when a linear operator or matrix can be diagonalized. Since the symmetric matrix A is a positive definite matrix, there is an orthogonal matrix T, so that the elements on the diagonal of T'AT are all positive values, and the elements on the diagonal are all eigenvalues of A , that is, the eigenvalues of A are all positive numbers

4. (50%) Consider a function $f(x_1, x_2) = (x_1 - x_2)^3 + 2(x_1 - 1)^2$.

   (a) Suppose $\vec{x_0} = (1, 2)$. Compute $\vec{x_1}$ using the steepest descent step with the optimal step length.

   $f(x_1, x_2) = x_1^3 + 2x_1^2 - 3x_1^2x_2 - 3x_1x_2^2 - 4x_1 + x_2^3 + 2$

   $\Delta f(x_1, x_2) = [3x_1^2 + 4x_1 - 6x_1x_2 + 3x_2^2 - 4, -3x_1^2 + 6x_1x_2 - 3x_2^2] = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$

   $H = \begin{bmatrix} 6x_1 + 4 - 6x_2 & -6x_1 + 6x_2 \\ -6x_1 + 6x_2 & -6x_2 + 6x_1 \end{bmatrix} = \begin{bmatrix} -2 & 6 \\ 6 & -6 \end{bmatrix}$

   $H^{-1} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{12} \end{bmatrix}$

   $x_0 = (1, 2)$, so $p = -g = -\Delta f(1, 2) = \begin{bmatrix} -3 \\ 3 \end{bmatrix}$

   $\alpha = \frac{-gp}{p^T Hp} = -0.1$

   $minf(x_0 - \alpha\Delta f) = (1 - 3\alpha - (2 + 3\alpha))^3 + 2(1 - 3\alpha - 1)^2$

   f'$(x_0 - \alpha\Delta f)$ *cannot be zero* *since* α should be (0,1), α should be 1 so $x_1 = x_0 + \alpha p = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$

(b) What is the Newton's direction of f at $(x_1, x_2) = (1, 2)$? Is it a descent direction?

$$g = \Delta f(1,2) = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$$

$H = \begin{bmatrix} -2 & 6 \\ 6 & -6 \end{bmatrix}$
$p = -H^{-1}g = \begin{bmatrix} 0 \\ -\frac{1}{2} \end{bmatrix}$
$-g^T H^{-1} g = \frac{3}{2}$
It is not a descent direction, because in that place, $g^T p$ is positive

(c) Compute the LDL decomposition of the Hessian of f at $(x_1, x_2) = (1, 2)$. (No pivoting)
$H = \begin{bmatrix} -2 & 6 \\ 6 & -6 \end{bmatrix} = LDL^T = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & 12 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$

(d) Compute the modified Newton step using LDL modification.
$LDL^T = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & 12 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$
$\hat{H} = L\hat{D}L^T = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 12 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ -3 & 21 \end{bmatrix}$
$p = -H\hat{^{-1}}g = \begin{bmatrix} -\frac{9}{2} \\ -\frac{1}{2} \end{bmatrix}$
$x1 = x_0 + p = \begin{bmatrix} -\frac{7}{2} \\ \frac{3}{2} \end{bmatrix}$

(e) Suppose $\vec{x}_0 = (1, 1)$ and $\vec{x}_1 = (1, 2)$, and the $B_0 = I$. Compute the quasi Newton direction p_1 using BFGS.
$g_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$p_0 = -B_0^{-1} g = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$s_0 = x_1 - x_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $g_1 = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$ $y_0 = g_1 - g_0 = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$
$B_1 = B_0 - \frac{B_0 s_0 s_0^T B_0}{s_0^T B_0 s_0} + \frac{y_0 y_0^T}{y_0^T s_0} = \begin{bmatrix} -2 & 3 \\ 3 & -3 \end{bmatrix}$
$p_1 = -B_1^{-1} g_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$