CS5321 Numerical Optimization Homework 2

Due March 25

1. (15%) Prove that $\vec{a}^T\vec{b} = \|\vec{a}\|\|\vec{b}\|\cos(\theta)$ for $\vec{a}, \vec{b} \in \mathbb{R}^n$ and θ is the angle between $\vec{a}$ and $\vec{b}$. (Hint: Let $\vec{c} = \vec{a} - \vec{b}$, and use the relation of $\|\vec{a}\|, \|\vec{b}\|, \|\vec{c}\|$ in a triangle to derive the result.)

   Let $\vec{c} = \vec{a} - \vec{b}$.

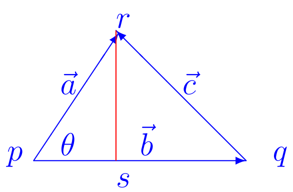
$$\|\vec{c}\|^2 = \vec{c}^T\vec{c} = (\vec{a}-\vec{b})^T(\vec{a}-\vec{b}) = \vec{a}^T\vec{a} - 2\vec{a}^T\vec{b} + \vec{b}^T\vec{b} = \|\vec{a}\|^2 - 2\vec{a}^T\vec{b} + \|\vec{b}\|^2. \tag{1}$$

   From geometry viewpoint, as shown in the figure, we have

   (a) $\|\vec{a}\| = \overline{pr}$, $\|\vec{b}\| = \overline{pq}$, $\|\vec{c}\| = \overline{rq}$.

   (b) By the Pythagorean theorem, $\overline{pr}^2 = \overline{rs}^2 + \overline{ps}^2$ and $\overline{rq}^2 = \overline{rs}^2 + \overline{qs}^2$.

   (c) By the trigonometric relations, $\overline{rs} = \|\vec{a}\|\sin\theta$ and $\overline{ps} = \|\vec{a}\|\cos\theta$. As the result, $\overline{qs} = \|\vec{b}\| - \|\vec{a}\|\cos\theta$.



   From those relations, we have

   $\|\vec{a}\|^2\sin^2\theta = \|\vec{a}\|^2 - \|\vec{a}\|^2\cos^2\theta = \|\vec{c}\|^2 - (\|\vec{b}\| - \|\vec{a}\|\cos\theta)^2$.

   Therefore, $\|\vec{c}\|^2 = \|\vec{a}\|^2 - 2\|\vec{a}\|\|\vec{b}\|\cos\theta + \|\vec{b}\|^2$.

   Comparing it to (1), we have $\vec{a}^T\vec{b} = \|\vec{a}\|\|\vec{b}\|\cos\theta$.

2. (15%) Consider a function $f(x,y) = \begin{cases} \frac{xy}{x+y} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$. Show that its partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist at $(0,0)$, but the directional derivative $D(f,[1,1]) \neq \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}$ at $(0,0)$.

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{0-0}{h} = 0.$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0}\frac{f(0,h)-f(0,0)}{h} = \lim_{h\to 0}\frac{0-0}{h} = 0.$$

$$D(f,[1,1]) = \lim_{h\to 0}\frac{f(h,h)-f(0,0)}{h} = \lim_{h\to 0}\frac{h^2/(h+h)-0}{h} = \frac{1}{2} \neq \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}$$

   This example shows a counter example of $D(f,\vec{p}) \neq \nabla f^T\vec{p}$, because f is singular at $(0,0)$.

3. (20%) Compute the LDL decomposition of the matrix

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 2 & 2 \\ 2 & 2 & 3 & 3 \\ 3 & 2 & 3 & 4 \end{pmatrix}.$$

(Hint: Use pivoting to stablize the computation, and put the pivotings into a permutation matrix P, such that $PAP^T = LDL^T$.)

First, we do the pivoting. Since the largest diagonal element is 4, we pivot (1,4). The permutation matrix is

$$P = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

And let

$$A^{(0)} = PAP^T = (PA)P^T = \begin{pmatrix} 4 & 2 & 3 & 3 \\ 2 & 2 & 2 & 1 \\ 3 & 2 & 3 & 2 \\ 3 & 1 & 2 & 0 \end{pmatrix} =. \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

We want to have $PAP^T = A^{(0)} = LDL^T$. Let

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \ell_{21} & 1 & 0 & 0 \\ \ell_{31} & \ell_{32} & 1 & 0 \\ \ell_{41} & \ell_{42} & \ell_{43} & 1 \end{pmatrix}, D = \begin{pmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{pmatrix}.$$

The Algorithm 3.4 in the lecture note page 53 gives a systematic method to perform the LDL decomposition

1. For $j = 1, 2, \ldots, n$
2. $\quad c_{jj} = a_{jj} - \sum_{s=1}^{j-1} d_s \ell_{js}^2$
3. $\quad d_j = c_{jj}$
4. $\quad$ For $i = j+1, \ldots, n$
5. $\qquad c_{ij} = a_{ij} - \sum_{s=1}^{j-1} d_s \ell_{is} \ell_{js}$
6. $\qquad \ell_{ij} = c_{ij}/d_j$

So, follow the steps. For $j=1$, $d_1 = c_{11} = a_{11}$ (step 2 and step 3), and

$$\ell_{21} = a_{21}/d_1 = .5, \ell_{31} = a_{31}/d_1 = .75, \ell_{41} = a_{41}/d_1 = .75, \text{(step 4,5,6)}.$$

For $j=2$, $d_2 = c_{22} = a_{22} - d_1\ell_{21}^2 = 2 - 4\times .5^2 = 1$ (step 2,3). Step 4,5,6 give

$$\begin{array}{rclclcl} c_{32} & = & a_{32} - d_1\ell_{31}\ell_{21} & = & 2 - 4\times .75\times .5 & = & .5 \quad \to \quad \ell_{32} = c_{32}/d_2 = .5 \\ c_{42} & = & a_{42} - d_1\ell_{41}\ell_{21} & = & 1 - 4\times .75\times .5 & = & -.5 \quad \to \quad \ell_{42} = c_{42}/d_2 = -.5 \end{array}$$

For $j=3$, $d_3 = c_{33} = a_{33} - d_1\ell_{31}^2 - d_2\ell_{32}^2 = 3 - 4\times .75^2 - 1\times .5^2 = .5$ (step 2,3). Step 4,5,6 give

$$c_{43} = a_{43} - d_1\ell_{41}\ell_{31} - d_2\ell_{42}\ell_{32} = 2 - 4\times .75\times .75 - 1\times .5\times(-.5) = 0,$$

Therefore,

$$\ell_{43} = c_{43}/d_3 = 0.$$

For $j=4$, step 2 and step 3 give

$$\begin{aligned} d_4 &= c_{44} = a_{44} - d_1\ell_{41}^2 - d_2\ell_{42}^2 - d_3\ell_{43}^2 \\ &= 0 - 4\times .75^2 - 1\times(-.5)^2 - .5\times 0^2 = -2.5 \end{aligned}$$

The final answer is

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ .5 & 1 & 0 & 0 \\ .75 & .5 & 1 & 0 \\ .75 & -.5 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & .5 & 0 \\ 0 & 0 & 0 & -2.5 \end{pmatrix}, P = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

4. (50%) Let $f(x,y) = \frac{1}{2}x^2 + \frac{9}{2}y^2$. This is a positive definite quadratic with minimizer at $(x^*, y^*) = (0,0)$.
   (a) Derive the gradient g and the Hessian H of f.
   (b) Write Matlab codes to implement the steepest descent method and Newton's method with $\vec{x}_0 = (9,1)$, and compare their convergent results. The formula of the steepest descent method is

   $$\vec{x}_{k+1} = \vec{x}_k - \frac{\vec{g}_k^T\vec{g}_k}{\vec{g}_k^T H_k \vec{g}_k}\vec{g}_k,$$

   and the formula of Newton's method is

   $$\vec{x}_{k+1} = \vec{x}_k - H_k^{-1}\vec{g}_k,$$

   where $\vec{g}_k = g(\vec{x}_k)$ and $H_k = H(\vec{x}_k)$.
   (c) Draw the trace of $\{\vec{x}_k\}$ for the steepest descent method and Newton's method. Figure 1 gives an example code for trace drawing.

```
function draw_trace()

% draw the contour of the function z = (x*x+9*y*y)/2;
step = 0.1;
X = 0:step:9;
Y = -1:step:1;

n = size(X,2);
m = size(Y,2);

Z = zeros(m,n);
for i = 1:n
    for j = 1:m
        Z(j,i) = f(X(i),Y(j));
    end
end

contour(X,Y,Z,100)

% plot the trace
%   You can record the trace of your results and use the following
%   code to plot the trace.
xk = [9 8 8 7 7 6 6 5 5 4 4 3 3 2 2];
yk = [.5 .5 -.5 -.5 .5 .5 -.5 -.5 .5 .5 -.5 -.5 .5 .5 -.5];
hold on;   % this is important!! This will overlap your plots.
plot(xk,yk,'-','LineWidth',3);
hold off;

% function definition
    function z = f(x,y)
        z = (x*x+9*y*y)/2;
    end
end
```

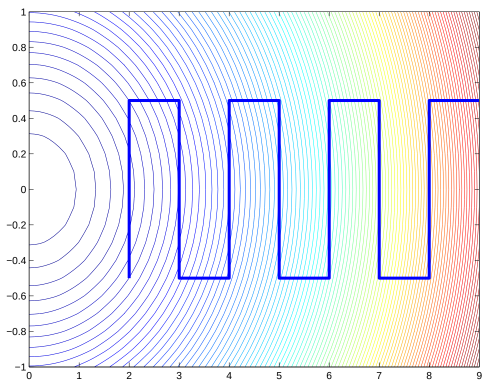


Figure 1: Function contour and a trace of (xk,yk).