MA 576 Optimization for Data Science
Homework 1

Nicolas Jorquera

Problem 1

Let $L \in \mathbf{R}^{nxm}$ and $A = LL^T$

1. Prove that A is **positive semidefinite**

   a Since $A = LL^T$, we know A is symmetric because $A^T = (LL^T)^T = (L^T)^T(L)^T = LL^T = A$. By definition, A is symmetric since $A = A^T$

   b An nxn symmetric real matrix A is positive semidefinite if $x^T Ax \geq 0$ for all $x \in \mathbf{R}^n$

   c By replacing A in the equation, we get $x^T Ax = x^T LL^T x$.

   d Let $B = x^T L$, then $B^T = L^T x$

   e Therefore $x^T Ax = BB^T$. Since the product of this vector cannot be less than 0, we prove that A is indeed a positive semidefinite matrix.

2. Prove that A is positive definite iff L has full row rank.

   a An nxn symmetric real matrix A is positive definite if $x^T Ax > 0$ for all $x \in \mathbf{R}^n$ for $x \neq 0$. Since we know that L is full rank then we again know A is symmetric since $A = A^T$ and the proof still holds.

   b Recall, that the kernel of a function whose range is $\mathbf{R}^n$ consists of all values in its domain where the function is 0. In other words, ker(A) is the set of vectors x $\in \mathbf{R}^n$ for which $Ax = 0$.

   c Therefore $Ax = 0$ is the kernel of A; where A is the superset of L. $ker(L) \subset ker(A)$

   d Using the same reasoning from above, we can determine that $\|L^T x\|^2 > 0$ since it can never be 0. However because L is full rank we know that L^T, L are row independent; meaning that the $ker(L^T x) = 0$. Because $L^T x$ is $ker(L^T x) \subset ker(A)$ we can imply that A has rows that are also linearly indepenednt, making A a full rank symmetric matrix.

   e Therefore for $x \neq 0$, then $\|L^T x\|^2 > 0 (A = L^T x)$, implying that matrix A is positive definite since $x^T Ax > 0$.

Problem 2

Show that:

1. If Q is positive definite, then the diagonal elements are positive.

   a As mentioned, if Q is positive definite then $x^T Qx > 0$ for all $x \in \mathbf{R}^n$ for $x \neq 0$.

   b In order to find the diagonal elements in column i, let $x_i = [0, 0, ...1, ...0, 0]$ where it is 1 in the ith column and 0 everywhere else.

   c Therefore:

$$
\begin{aligned}
f(x) &= x^T Qx > 0 \\
f(x_i) &= x_i^T Qx_i > 0 \\
f(x_i) &= a_{ii} > 0
\end{aligned}
$$

d a_{ii} is the diagonal element in the ith row / column. Therefore we can conclude that any diagonal element is positive.

2. Let Q be a symmetric matrix. If there exist positive and negative in the diagonal, then Q is indefinite.

   a As mentioned, if Q is positive definite then $x^TQx > 0$ for all $x \in \mathbf{R}^n$ for $x \neq 0$.

   b In order to find the diagonal elements in column i, let $x_i = [1, 0, ...0, ...0, 0]$ where it is 1 in the ith column and 0 everywhere else. However this time let $x_i = [0, 1, ...0, ...0, 0]$. This time we will be calculating for the 1st and 2nd diagonal elements.

   c Using the same calculations as above, we find the diagonal element to be $a_{11} > 0$ and $a_{22} < 0$

   d Because not all the diagonal elements are positive, we can conclude that Q is indefinite, since $x^TQx > 0$ does not hold.

Problem 3

Write a program in MATLAB/Octave/Scilab or Python (any other language please con- tact me) that randomly generates a positive definite matrix and then verifies it is by using Sylvester's criterion.

Code provided as an attachment

2

Problem 4

Study the continuity of f and the existence of the directional derivatives:

$$f(x, y) = \begin{cases} xy \sin \frac{1}{x} \cos \frac{1}{y} & \text{if } (x, y) \neq (0, 0) \\ a & \text{if } (x, y) = (0, 0) \end{cases}$$

for $a \in \mathbf{R}$

To study the continuity of f we must evaluate the function at the point (0,0); which is given to be **a**. In order to test continuity we must take the limit from the right and the left and ensure they are equal:

$$\lim_{(x,y)\to(0,0)^+} xy \sin \frac{1}{x} \cos \frac{1}{y} \qquad \lim_{(x,y)\to(0,0)^-} xy \sin \frac{1}{x} \cos \frac{1}{y}$$

If we take the limit of the function we know that $\sin \frac{1}{x}$ and $\cos \frac{1}{y}$ are bounded by ; therefore we can solve that

$$\lim_{(x,y)\to(0,0)^+} xy \sin \frac{1}{x} \cos \frac{1}{y} = 0$$

Therefore in order for this function to be continuous a must be 0.

In order to test the directional derivative of f at (x_0, y_0) in the direction of the vector $u =< b, c >$ is:

$$D_u f(x_0, y_0) = \frac{f(x_0 + hb, y_0 + hc) - f(x_0, y_0)}{h}$$

We are trying to find the directional derivative at point (0,0), and we know the value of our function at that point is a; therefore:

$$\begin{aligned} D_u f(x_0, y_0) &= \frac{f(hb, hc) - a}{h} \\ &= h^2 bc \sin \frac{1}{hb} \cos \frac{1}{hc} \end{aligned}$$

Because of small angle approximation we can rewrite this function to be:

$$D_u f(0, 0) = h^2 bc \frac{1}{hb} \frac{1}{hc}$$

therefore if this function is continuous we know that a = 0; and the directional derivative at point (0,0).

Problem 5

Compute the gradient and the Hessian of:

1. $f(x) = e^{\|x\|_2^2}$, with $x \in \mathbf{R}$ and $\|x\|_2^2 = \sum_j x_j^2$

   a We are given:
   $$f(x) = e^{\|x\|_2^2} = e^{\sum_j x_j^2} = e^{(x_1^2+x_2^2+...+x_n^2)}$$

   b We know that the gradient of $f(x)$ is $\nabla f(x) = [\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}]$

   c To find the partial derivatives of x_1 and x_2:
   $$\frac{\partial f}{\partial x_1} = 2x_1 e^{(x_1^2+x_2^2+...+x_n^2)} \qquad \frac{\partial f}{\partial x_2} = 2x_2 e^{(x_1^2+x_2^2+...+x_n^2)}$$

   d We can generalize; and conclude that the gradient is calculated by $\nabla f(x) = 2xe^{\sum_j x_j^2}$ where $x = [x_1, x_2, ...x_n]$

   e To calculate the Hessian: $H(f(x))_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. Therefore we need to calculate the second order partial derivatives for each element in the matrix H.
   $$\frac{\partial f}{\partial x_1 \partial x_1} = 4x_1^2 e^{\sum_j x_j^2} + 2e^{\sum_j x_j^2} \qquad \frac{\partial f}{\partial x_1 \partial x_2} = 4x_1 x_2 e^{\sum_j x_j^2}$$

   f Therefore we can generalize the Hessian to:
   $$H(f(x))_{ij} = \begin{cases} 4x_i x_j e^{\sum_j x_j^2} & \text{if } i \neq j \\ 4x_i^2 e^{\sum_j x_j^2} + 2e^{\sum_j x_j^2} & \text{if } i = j \end{cases}$$

2. $g(x) = \prod_{j=1}^n x_j$.

   a We are given:
   $$g(x) = \prod_{j=1}^n x_j = x_1 x_2 x_3 ... x_n$$

   b We find the partial derivatives to find the gradient:
   $$\frac{\partial g}{\partial x_1} = x_2 x_3 ... x_n \qquad \frac{\partial g}{\partial x_2} = x_1 x_3 ... x_n$$

   d We can generalize; and conclude that the gradient is calculated by $\nabla g(x) = \frac{g(x)}{x_i}$ where $x = [x_1, x_2, ...x_n]$ and x_i the gradient for element i

   e To calculate the Hessian we will calculate the second order partial derivatives for each element in the matrix H.
   $$\frac{\partial g}{\partial x_1 \partial x_1} = 0 \qquad \frac{\partial g}{\partial x_1 \partial x_2} = x_3 ... x_n.$$

   f Therefore we can generalize the Hessian to:
   $$H(g(x))_{ij} = \begin{cases} \frac{g(x)}{x_i x_j} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}$$

Problem 6

Compute

$$\lim_{(x,y)\to(0,0)} \frac{xy - \sin(x)\sin(y)}{x^2+y^2}$$

using Taylor's Theorem.

When trying to solve this problem using Taylor Theorem; I was trying to solve the 1st-order Taylor expansion (Linear Approximation) of the equation above, but the solution proved to be indeterminate. Therefore; I will first expand the equation to:

$$\frac{xy - \sin(x)\sin(y)}{x^2+y^2} = \frac{xy}{x^2+y^2}(1 - \frac{\sin(x)}{x}\frac{\sin(y)}{y})$$

However because $\frac{xy}{x^2+y^2}$ is bounded, we can focus on the Taylor expansion around $1 - \frac{\sin(x)}{x}\frac{\sin(y)}{y}$. In order to do this we calculate the partial derivatives and evaluate them at 0; however these calculations also lead to an indeterminate solution. Lastly, we can evaluate the limit (similar to problem above) and see that:

$$\lim_{(x,y)\to(0,0)} 1 - \frac{\sin(x)}{x}\frac{\sin(y)}{y} \to 0$$