LSA University of Michigan

MATH562: Continuous Optimisation

Homework 2

Name: Keegan Gyoery UM-ID: 31799451

1. **A:** min $f(\mathbf{x}) = x_1 + x_2 - \ln x_1 - \ln x_2$

   a) The first order necessary condition a point must satisfy is $\nabla f(\mathbf{x}) = \mathbf{0}$. Considering the gradient of $f(\mathbf{x})$,

   $$\nabla f(\mathbf{x}) = \left(1 - \frac{1}{x_1},\, 1 - \frac{1}{x_2}\right).$$

   Thus, the necessary condition for a point to be a minimum is $\left(1 - \frac{1}{x_1},\, 1 - \frac{1}{x_2}\right) = (0, 0)$.

   b) Considering the conditions from above,

   $$\begin{aligned} 1 - \frac{1}{x_1} &= 0 \\ \therefore x_1 &= 1, \\ 1 - \frac{1}{x_2} &= 0 \\ \therefore x_2 &= 1. \end{aligned}$$

   Thus, the point $\hat{\mathbf{x}} = (1, 1)$ satisfies the necessary condition for a local minimum.

   c) Considering the Hessian of $f(\mathbf{x})$, and $\mathbf{h} = (h_1, h_2) \neq \mathbf{0}$,

   $$\begin{aligned} Hf(\mathbf{x}) &= \begin{bmatrix} \frac{1}{x_1^2} & 0 \\ 0 & \frac{1}{x_2^2} \end{bmatrix}, \\ \mathbf{h}^T Hf(\mathbf{x})\mathbf{h} &= \mathbf{h}^T \begin{bmatrix} \frac{1}{x_1^2} & 0 \\ 0 & \frac{1}{x_2^2} \end{bmatrix} \mathbf{h} \\ &= \frac{h_1^2}{x_1^2} + \frac{h_2^2}{x_2^2} \\ &> 0 \;\forall \mathbf{h}, \mathbf{x}. \end{aligned}$$

   Thus $Hf(\mathbf{x})$ is P.S.D and P.D, and so too is $Hf(\hat{\mathbf{x}})$. Clearly, $\hat{\mathbf{x}}$ satisfies the second order necessary condition and the two second order sufficient conditions, and so $\hat{\mathbf{x}}$ is a local minimum.

   B: min $f(\mathbf{x}) = 2\ln(x_1 + x_2 + 1) - \ln x_1 - 1.5 \ln x_2$

   a) The first order necessary condition a point must satisfy is $\nabla f(\mathbf{x}) = \mathbf{0}$. Considering the gradient of $f(\mathbf{x})$,

   $$\nabla f(\mathbf{x}) = \left(\frac{2}{x_1 + x_2 + 1} - \frac{1}{x_1},\, \frac{2}{x_1 + x_2 + 1} - \frac{1.5}{x_2}\right).$$

   Thus, the necessary condition for a point to be a minimum is
   $\left(\frac{2}{x_1 + x_2 + 1} - \frac{1}{x_1},\, \frac{2}{x_1 + x_2 + 1} - \frac{1.5}{x_2}\right) = (0, 0)$.

b) Considering the conditions from above,

$$\begin{aligned}\frac{2}{x_1+x_2+1}-\frac{1}{x_1}&=0\\ 2x_1&=x_1+x_2+1\\ \therefore x_2&=x_1-1\ldots(A),\\ \frac{2}{x_1+x_2+1}-\frac{1.5}{x_2}&=0\\ 2x_2&=1.5x_1+1.5x_2+1.5\\ \therefore x_2&=3x_1+3\ldots(B).\end{aligned}$$

Combining equations (A) and (B),

$$\begin{aligned}x_1-1&=3x_1+3\\ \therefore x_1&=-2\\ \therefore x_2&=-3\end{aligned}$$

Thus, the point $\hat{\mathbf{x}}=(-2,-3)$ satisfies the necessary condition for a local minimum.

c) Considering the Hessian of $f(\mathbf{x})$, and $\mathbf{h}=(h_1,h_2)\neq\mathbf{0}$,

$$\begin{aligned}Hf(\mathbf{x})&=\begin{bmatrix}\frac{-2}{(x_1+x_2+1)^2}+\frac{1}{x_1^2} & \frac{-2}{(x_1+x_2+1)^2}\\ \frac{-2}{(x_1+x_2+1)^2} & \frac{-2}{(x_1+x_2+1)^2}+\frac{1}{x_1^2}\end{bmatrix},\\ \therefore Hf(-2,-3)&=\begin{bmatrix}\frac{1}{8} & -\frac{1}{8}\\ -\frac{1}{8} & \frac{1}{24}\end{bmatrix},\\ \mathbf{h}^T Hf(-2,-3)\mathbf{h}&=\mathbf{h}^T\begin{bmatrix}\frac{1}{8} & -\frac{1}{8}\\ -\frac{1}{8} & \frac{1}{24}\end{bmatrix}\mathbf{h}\\ &=\frac{h_1^2}{8}-\frac{h_1h_2}{4}+\frac{h_2^2}{24}\end{aligned}$$

Consider $\mathbf{h}=(1,1)$,

$$\begin{aligned}\mathbf{h}^T Hf(-2,-3)\mathbf{h}&=(1,1)^T Hf(-2,-3)(1,1)\\ &=\frac{1}{8}-\frac{1}{4}+\frac{1}{24}\\ &=-\frac{1}{12}\\ &<0.\end{aligned}$$

Thus $Hf(\hat{\mathbf{x}})$ is not P.D or P.S.D, so $\hat{\mathbf{x}}$ fails the second order necessary condition, and thus cannot be a local minimum.

C: min $f(\mathbf{x})=x_1x_2-\ln x_1-\ln x_2$

a) The first order necessary condition a point must satisfy is $\nabla f(\mathbf{x})=\mathbf{0}$. Considering the gradient of $f(\mathbf{x})$,

$$\nabla f(\mathbf{x})=\left(x_2-\frac{1}{x_1},\,x_1-\frac{1}{x_2}\right).$$

Thus, the necessary condition for a point to be a minimum is $\left(x_2 - \frac{1}{x_1},\, x_1 - \frac{1}{x_2}\right) = (0, 0)$.

b) Considering the conditions from above,

$$
\begin{aligned}
x_2 - \frac{1}{x_1} &= 0 \\
\therefore x_1 x_2 &= 1, \\
x_1 - \frac{1}{x_2} &= 0 \\
\therefore x_1 x_2 &= 1.
\end{aligned}
$$

Thus, the point $\hat{\mathbf{x}} = \left(t, \frac{1}{t}\right)$ satisfies the necessary condition for a local minimum.

c) Considering the Hessian of $f(\mathbf{x})$, and $\mathbf{h} = (h_1, h_2) \neq \mathbf{0}$,

$$
\begin{aligned}
Hf(\mathbf{x}) &= \begin{bmatrix} \frac{1}{x_1^2} & 1 \\ 1 & \frac{1}{x_2^2} \end{bmatrix}, \\
\therefore \mathbf{h}^T Hf(\mathbf{x}) \mathbf{h} &= \mathbf{h}^T \begin{bmatrix} \frac{1}{x_1^2} & 1 \\ 1 & \frac{1}{x_2^2} \end{bmatrix} \mathbf{h} \\
&= \frac{h_1^2}{x_1^2} + 2h_1 h_2 + \frac{h_2^2}{x_2^2} \\
&= \left(\frac{h_1}{x_1} + \frac{h_2}{x_2}\right)^2 \\
&> 0 \, \forall \mathbf{h}, \mathbf{x}.
\end{aligned}
$$

Thus $Hf(\mathbf{x})$ is P.S.D and P.D, and so too is $Hf\left(t, \frac{1}{t}\right)$. Clearly, $\hat{\mathbf{x}}$ satisfies the second order necessary condition and the two second order sufficient conditions, and so $\hat{\mathbf{x}}$ is a local minimum.

2. Consider the problem

$$
\min f(\mathbf{x}) = 5(x_1^2 - x_2)^2 + x_1^2 + (x_2 - 5)^2
$$

a) The gradient and Hessian for the function $f(\mathbf{x})$ are

$$
\begin{aligned}
\nabla f(\mathbf{x}) &= \left(20x_1^3 - 20x_1 x_2 + 2x_1, -10x_1^2 + 12x_2 - 10\right), \\
Hf(\mathbf{x}) &= \begin{bmatrix} 60x_1^2 - 20x_2 + 2 & -20x_1 \\ -20x_1 & 12 \end{bmatrix}.
\end{aligned}
$$

b) Considering the neccessary and sufficient conditions for a P.S.D. matrix, we get the restrictions,

$$
\begin{aligned}
60x_1^2 - 20x_2 + 2 &\geq 0 \ldots (1), \\
12 &\geq 0, \text{ which is true,} \\
12(60x_1^2 - 20x_2 + 2) - 400x_1^2 &\geq 0 \\
\therefore 40x_1^2 - 30x_2 + 3 &\geq 0 \ldots (2).
\end{aligned}
$$

c) To check that $\hat{\mathbf{x}} = \left(0, \frac{5}{6}\right)$ is a local minimum, it must first satisfy the first and second order necessary conditions. Examining the gradient at $\hat{\mathbf{x}}$, and the restrictions above,

$$\nabla f\left(0, \frac{5}{6}\right) = (0, 0)$$
$$(2) \rightarrow -22 \ngeq 0.$$

Thus, $Hf\left(0, \frac{5}{6}\right)$ is not P.S.D., and so $\hat{\mathbf{x}}$ cannot be a local minimum.

3. Using the MATLAB code,

```
load('hw3/m5.mat');
disp('5x5 Matrix');
disp(A);
d = eig(A);
disp('Eigenvalues for 5x5 Matrix');
disp(d);
load('hw3/m10.mat');
disp('10x10 Matrix');
disp(A);
d = eig(A);
disp('Eigenvalues for 10x10 Matrix');
disp(d);
```

we get the output for the 5x5 matrix, .

```
5x5 Matrix
    92    57    94    34    17
    29     8    13    17    61
    76     6    57    80    27
    76    54    47    32    66
    39    78     2    53    69

Eigenvalues for 5x5 Matrix
   1.0e+02 *

   2.4278 + 0.0000i
   0.8473 + 0.0000i
  -0.1484 + 0.1350i
  -0.1484 - 0.1350i
  -0.3982 + 0.0000i
```

Clearly, this matrix is not even symmetric and and thus cannot be P.D. or P.S.D.

From the output for the 10x10 matrix, .

```
10x10 Matrix
  Columns 1 through 6

       33013       12697       30768       33018       21593       30411
       12697       15105       13077       14789        9262       14112
       30768       13077       37332       30123       22348       34225
       33018       14789       30123       38842       21910       33404
       21593        9262       22348       21910       21567       18235
       30411       14112       34225       33404       18235       45860
       24658       12789       29603       27045       17257       34599
       30946       14456       32304       32365       18993       36148
       26746        9934       28436       29322       17190       29540
       21768       10257       25177       22599       16307       28274

  Columns 7 through 10

       24658       30946       26746       21768
       12789       14456        9934       10257
       29603       32304       28436       25177
       27045       32365       29322       22599
       17257       18993       17190       16307
       34599       36148       29540       28274
       32713       28418       20952       20917
       28418       39818       31948       29437
       20952       31948       31540       24930
       20917       29437       24930       24586

Eigenvalues for 10x10 Matrix
   1.0e+05 *

    0.0001
    0.0053
    0.0281
    0.0348
    0.0441
    0.0884
    0.1000
    0.1480
    0.1720
    2.5829
```

We can clearly see that all the eigenvalues are positive, and thus, the matrix is P.D.

4. Suppose M is P.D.

   a) If M is P.D., then M has all positive eigenvalues. Thus, no eigenvalue of M is 0, and hence M is invertible, and thus M^{-1} exists. By definition, $M^{-1}M = I$, and noting M is symmetric, so taking the transpose gives,

   $$\begin{aligned}\left(M^{-1}M\right)^T &= I^T \\ M^T\left(M^{-1}\right)^T &= I \\ M\left(M^{-1}\right)^T &= I \\ \therefore \left(M^{-1}\right)^T &= M^{-1}.\end{aligned}$$

   Thus, M^{-1} is real and symmetric. Finally, condsider the eigenvalues λ_i of M, given by $M\mathbf{x} = \lambda\mathbf{x}$. Manipulating for M^{-1},

   $$\begin{aligned}M\mathbf{x} &= \lambda\mathbf{x} \\ \therefore \mathbf{x} &= M^{-1}\lambda\mathbf{x} \\ &= \lambda M^{-1}\mathbf{x} \\ \therefore \frac{1}{\lambda}\mathbf{x} &= M^{-1}\mathbf{x} \\ \therefore M^{-1}\mathbf{x} &= \frac{1}{\lambda}\mathbf{x}.\end{aligned}$$

   Clearly, the eigenvalues of M^{-1} are $\dfrac{1}{\lambda_i}$, which are all positive. Thus, M^{-1} is P.D.

   b) Refer to the proof above, showing that the eigenvalues of M^{-1} are the inverse of the eignevalues of M.

   c) Let $\mathbf{u}$ be an eigenvector of M, where λ is the associated eigenvalue. Thus,

   $$\begin{aligned}M\mathbf{u} &= \lambda\mathbf{u} \\ \mathbf{u} &= M^{-1}\lambda\mathbf{u} \\ \therefore M^{-1}\mathbf{u} &= \frac{1}{\lambda}\mathbf{u}.\end{aligned}$$

   Clearly, $\mathbf{u}$ is an eigenvector of M^{-1} with eigenvalue $\dfrac{1}{\lambda}$. Now let $\mathbf{u}$ be an eigenvector of M^{-1}, where $\dfrac{1}{\lambda}$ is the associated eigenvalue. Thus,

   $$\begin{aligned}M^{-1}\mathbf{u} &= \frac{1}{\lambda}\mathbf{u} \\ \mathbf{u} &= M\frac{1}{\lambda}\mathbf{u} \\ \therefore M\mathbf{u} &= \lambda\mathbf{u}.\end{aligned}$$

   Clearly, $\mathbf{u}$ is an eigenvector of M with eigenvalue λ.

5. Suppose that M is P.S.D. Therefore, M is real and symetric, and so can be rewritten $M = QDQ^T$, where Q is orthonormal, and D is diagonal, with the eigenvalues of M along the diagonal. All eigenvalues of M are non-negative, and so let N be given by QLQ^T, where Q is the same as above, and L is another diagonal matrix with the square root of the eigenvalues of M along the diagonal. That is, $D = \text{diag}[\lambda_1\ \lambda_2\ \ldots\ \lambda_n]$, and $L = \text{diag}[\sqrt{\lambda_1}\ \sqrt{\lambda_2}\ \ldots\ \sqrt{\lambda_n}]$. Noting that

because Q is orthonormal, $Q^T = Q^{-1}$, and so,

$$\begin{aligned}
N^2 &= NN \\
&= \left(QLQ^T\right)\left(QLQ^T\right) \\
&= (QL)\left(QQ^T\right)\left(LQ^T\right) \\
&= (QL)(I)\left(LQ^T\right) \\
&= (QL)\left(LQ^T\right) \\
&= \left(QLLQ^T\right) \\
&= \left(QL^2Q^T\right) \\
&= \left(QDQ^T\right) \\
\therefore N^2 &= M.
\end{aligned}$$

As $N = QLQ^T$, the eigenvalues of N are all non-negative, and as N is written in the above form, where L is diagonal and Q is orthonormal, then N is also P.S.D. Thus, there exists an N P.S.D. such that $N^2 = M$.