37242 Introduction to Optimisation

Tutorial 8

Question 1. Solve the unconstrained optimisation problem

$$\min f(x_1,\, x_2) = x_1^2 + x_2^2 + x_1x_2 + 4x_1$$

(a) By steepest descent method
(b) By Newton's method
(c)By finding stationary points and determining their nature

Extra excercises:

Question 2. Solve the unconstrained optimisation problem with the methods above

$$\min f(x_1,\, x_2) = x_1^2 + x_2^2 + 4x_1 - 6x_2$$

Question 3. Prove that if A is an n × n positive definite matrix, then

(a) All eigenvalues of $\mathbf{A}$ are positive.

(b) $\mathbf{A}$ is invertible.

(c) All eigenvalues of $\mathbf{A}^{-1}$ are positive.

Question 4. (Winston Chapter 11, Section 3, Question $1, 2, 7, 8, 9$)

On the given set $\mathbf{S}$, determine whether each function is convex, concave, or neither.

(a) $f(x) = x^3;\;\; \mathbf{S} = [0,\, \infty).$

(b) $f(x) = x^3;\;\; \mathbf{S} = \mathbf{R}.$

(c) $f(x_1,\, x_2) = x_1^2 + x_2^2;\;\; \mathbf{S} = \mathbf{R}^2.$

(d) $f(x_1,\, x_2) = -x_1^2 - x_1x_2 - 2x_2^2;\;\; \mathbf{S} = \mathbf{R}^2.$

(e) $f(x_1,\, x_2,\, x_3) = -x_1^2 - x_2^2 - 2x_3^2 + 0.5x_1x_2;\;\; \mathbf{S} = \mathbf{R}^3.$

Question 1. Solve the unconstrained optimisation problem

$$\min f(x_1, x_2) = x_1^2 + x_2^2 + x_1 x_2 + 4x_1$$

(a) By steepest descent method
(b) By Newton's method

$\Big\} \; x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \; \varepsilon = \frac{1}{2}$

(c)By finding stationary points and determining their nature

a) $$\nabla f(x) = \begin{pmatrix} 2x_1 + x_2 + 4 \\ 2x_2 + x_1 \end{pmatrix}$$

$$\nabla^2 f(x) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

It. 0 $\quad x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

1. $\nabla f(0,0) = \begin{pmatrix} 4 \\ 0 \end{pmatrix}; \; \|\nabla f(0,0)\| = 4 > \varepsilon$
2. $d_0 = -\nabla f(0,0) = \begin{pmatrix} -4 \\ 0 \end{pmatrix}$

$$\underbrace{x^{(0)} + \alpha d_0}_{} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -4\alpha \\ 0 \end{pmatrix} = \begin{pmatrix} -4\alpha \\ 0 \end{pmatrix}$$

3. $g(\alpha) = f(-4\alpha, 0) = 16\alpha^2 - 16\alpha$

$$\boxed{\text{Find } \min_{\alpha > 0} g(\alpha)}$$

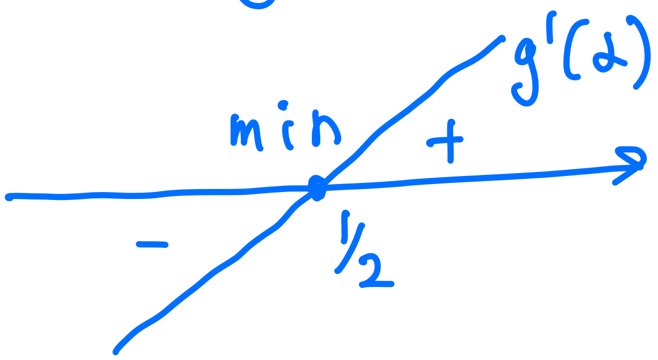
3.1. $g(\alpha)$ is quadratic f-n, graph is concaved up parabola

$\rightarrow$ min is at vertex:

$$\alpha_{verx} = \frac{16}{2 \times 16} = \frac{1}{2}$$

or

3.2 $g'(\alpha) = 32\alpha - 16 \rightarrow g'(\alpha) = 0$ when $\alpha = 1/2$



or

$g''(\alpha) = 32 > 0$ for all $\alpha \rightarrow$ $\alpha = 1/2$ is mi

4. $x^{(1)} = \begin{pmatrix} -4\alpha \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$

It. 1 $\quad x^{(1)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$

1. $\nabla f(-2,0) = \begin{pmatrix} -4+0+4 \\ 0-2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$

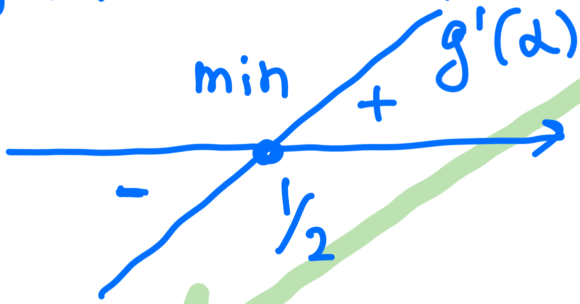
$\|\nabla f(-2,0)\| = 2 > \varepsilon$

2. $d_1 = -\nabla f(-2,0) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$

$$x^{(1)} + \alpha d_1 = \begin{pmatrix}-2\\0\end{pmatrix} + \begin{pmatrix}0\\2\alpha\end{pmatrix} = \begin{pmatrix}-2\\2\alpha\end{pmatrix}.$$

3. $g(\alpha) = f(-2, 2\alpha) = 4 + 4\alpha^2 - 4\alpha - 8$

$g'(\alpha) = 8\alpha - 4 = 0 \rightarrow \alpha = \frac{1}{2}$



4. $x^{(2)} = \begin{pmatrix}-2\\1\end{pmatrix}$

It. 2 $\quad x^{(2)} = \begin{pmatrix}-2\\1\end{pmatrix}$

1. $\nabla f(-2, 1) = \begin{pmatrix}-4+1+4\\2-2\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$

$\|\nabla f(-2,1)\| = 1 > \varepsilon$

2. $d_2 = -\nabla f(-2,1) = \begin{pmatrix}-1\\0\end{pmatrix}$

$x^{(2)} + \alpha d_2 = \begin{pmatrix}-2\\1\end{pmatrix} + \begin{pmatrix}-\alpha\\0\end{pmatrix} = \begin{pmatrix}-2-\alpha\\1\end{pmatrix}$

3. $g(\alpha) = f(-2-\alpha, 1) = (2+\alpha)^2 + 1 \underbrace{-(2+\alpha) - 4(2+\alpha)}_{-5(2+\alpha)}$

Find $\min_{\alpha>0} g(\alpha)$:

$g'(\alpha) = 2(2+\alpha) - 5 = 0 \rightarrow$

$2\alpha - 1 = 0 \quad \alpha = 1/2$

$g''(\alpha) = 2 > 0 \rightarrow \alpha = 1/2$ is min

4. $x^{(3)} = \begin{pmatrix} -5/2 \\ 1 \end{pmatrix}$

It. 3 $\quad x^{(3)} = \begin{pmatrix} -5/2 \\ 1 \end{pmatrix}$

1. $\nabla f(-5/2, 1) = \begin{pmatrix} -5+1+4 \\ 2 - 5/2 \end{pmatrix} = \begin{pmatrix} 0 \\ -1/2 \end{pmatrix}$

$\|\nabla f(-5/2, 1)\| = \frac{1}{2} = \varepsilon \rightarrow$ STOP

$x = \begin{pmatrix} -5/2 \\ 1 \end{pmatrix}$ is suffiently close approximation of local minimiser

(b) By Newton's method

(c)By finding stationary points and determining their nature

$$\nabla f(x) = \begin{pmatrix} 2x_1 + x_2 + 4 \\ 2x_2 + x_1 \end{pmatrix}$$

$$\nabla^2 f(x) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

It. 0 $\quad x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

1. $\nabla f(0,0) = \begin{pmatrix} 4 \\ 0 \end{pmatrix}; \; \|\nabla f(0,0)\| = 4 > \varepsilon$

2. $x^{(1)} = x^{(0)} - \left[\nabla^2 f(x^{(0)})\right]^{-1} \nabla f(x^{(0)})$

$$\left[\nabla^2 f(x)\right]^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$x^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ 0 \end{pmatrix} = -\frac{1}{3}\begin{pmatrix} 8 \\ -4 \end{pmatrix} = \begin{pmatrix} -8/3 \\ 4/3 \end{pmatrix}$$

It. 1.

1. $\nabla f\left(-8/3, 4/3\right) = \begin{pmatrix} -\frac{16}{3} + \frac{4}{3} + 4 \\ \frac{8}{3} - \frac{8}{3} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow$ Hence $x = \begin{pmatrix} -8/3 \\ 4/3 \end{pmatrix}$ is Local min

(c)By finding stationary points and determining their nature

$$\nabla f(x) = \begin{pmatrix} 2x_1 + x_2 + 4 \\ 2x_2 + x_1 \end{pmatrix}$$

$$\nabla^2 f(x) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

1. find where $\nabla f(x) = 0$

$$\begin{cases} 2x_1 + x_2 + 4 = 0 & ① \times 2 \\ 2x_2 + x_1 = 0 & ② \end{cases} \quad -$$

$$4x_1 + 2x_2 + 8 - 2x_2 - x_1 = 0$$

$$3x_1 = -8$$

$$x_1 = -8/3 \rightarrow x_2 = \frac{4}{3}$$

The only st. point is $\begin{pmatrix} -8/3 \\ 4/3 \end{pmatrix}$

2. Nature of the st. point:

$$\nabla^2 f(x) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \rightarrow \text{for all } x \in R^n$$

Use Th 4: find eigenvalues by solving $\det(\nabla^2 f(x) - \lambda I) = 0$.

$$\det\begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = 0$$

$\downarrow$

$$(2-\lambda)^2 = 1$$

$2-\lambda = 1$ $\qquad\qquad$ $2-\lambda = -1$

$\lambda = 1 > 0$ $\qquad\qquad$ $\lambda = 3 > 0$

$\downarrow$

by Th. 4 $\nabla^2 f(x)$ is pos.-def. at every $x \in R^n$

- at $x^* = \begin{pmatrix} -8/3 \\ 4/3 \end{pmatrix}$ $\nabla f = 0$ and $\nabla^2 f$ is pos.-def

$\downarrow$

by Th. 6 x^* is Local minimiser

- as $\nabla^2 f(x)$ is pos.-def. at every $x \in R^n$, by Th 7, $f(x)$ is strictly convex

- by Th. 3 x^* is unique global minimiser

Q3 A is pos.-def. Prove that

(a) All eigenvalues of **A** are positive.

A pos.-def $\rightarrow$ for any $x \neq 0$

$$x^T A x > 0$$

Let $x^{(i)}$: $A x^{(i)} = \lambda_i x^{(i)}$

$$0 < x^{(i)T} A x^{(i)} = x^{(i)T} \lambda_i x^{(i)} = \lambda_i \underbrace{\|x^{(i)}\|^2}_{> 0}$$

$$\downarrow$$

$$\lambda_i > 0$$

Challenge

- if A is negative-definite,
  show that all eigenvalues are negative

Question 2. Solve the unconstrained optimisation problem with the methods above

$$\min f(x_1, x_2) = x_1^2 + x_2^2 + 4x_1 - 6x_2$$

a) Steepest descent

$\nabla f(x_1, x_2)^T = \langle 2x_1 + 4, \; 2x_2 - 6 \rangle$

$x^{(0)} = (0, 0) \quad ; \; \varepsilon = 1/2$

It. 0.

1. $\nabla f(0,0)^T = \langle 4, -6 \rangle \; ; \; \|\nabla f(0,0)\| = \sqrt{52} > \varepsilon$
2. $d_0 = -\nabla f(x^{(0)}) = \begin{pmatrix} -4 \\ 6 \end{pmatrix} \qquad x^{(0)} + \alpha d_0 = \begin{pmatrix} -4\alpha \\ 6\alpha \end{pmatrix}$
3. $\alpha_0 : \; g(\alpha) = f(-4\alpha, 6\alpha) = 16\alpha^2 + 36\alpha^2 - 16\alpha - 36\alpha$

   $g'(\alpha) = 104\alpha - 52 \; ; \; g'(\alpha) = 0$ when $\alpha = 1/2$

   $g''(\alpha) = 104 > 0$ for all $\alpha \Rightarrow g''(1/2) > 0 \Rightarrow \alpha = 1/2$ is min
4. $x^{(1)} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$

It. 1

1. $\nabla f(x^{(1)})^T = \langle -4+4, \; 6-6 \rangle = \langle 0, 0 \rangle \rightarrow$

   $x^* = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$ is local minimiser

b) Newton's method:

$\nabla f(x_1, x_2)^T = \langle 2x_1 + 4,\ 2x_2 - 6 \rangle$

$x^{(0)} = (0, 0)$; $\varepsilon = 1/2$

$A = \nabla^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ $A^{-1} = \frac{1}{4}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

It. 0.

1. $\nabla f(x^{(0)})^T = \langle 4, -6 \rangle$; $\|\nabla f(x^{(0)})\| = \sqrt{52} > \varepsilon$

2. $x^{(1)} = x^{(0)} - [\nabla^2 f(0,0)]^{-1} \nabla f(0,0) =$

$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{4}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ -6 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$

It. 1.

1. $\nabla f(x^{(1)})^T = \langle 0, 0 \rangle \to \begin{pmatrix} -2 \\ 3 \end{pmatrix}$ is local minimiser

Q2 part c Solve the unconstrained optimisation problem

$$\min f(x_1,\, x_2) = x_1^2 + x_2^2 + 4x_1 - 6x_2$$

(a) Solve $\nabla f(x_1,\, x_2) = \mathbf{0}$ to find stationary point(s) of $f(x_1, x_2)$.

$$\nabla f(x_1,\, x_2) = \begin{pmatrix} 2x_1 + 4 \\ 2x_2 - 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$$

is the unique stationary point of $f(x_1, x_2)$.

(b) Find the nature of the stationary point(s) in part (a).

For any point $(x_1,\, x_2)^T$ we have $\nabla^2 f(x_1,\, x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2\,\mathbf{I}$. Since $\mathbf{I}$ is positive definite, so is $2\,\mathbf{I}$. Hence, $f(x_1, x_2)$ is strictly convex on $\mathbf{R}^2$. So, the stationary point $(x_1^*,\, x_2^*)$ is its global minimizer.

Note. The fact that $\nabla^2 f(x_1,\, x_2)$ is positive definite could also be checked by

- directly considering

$$(x_1,\, x_2)\nabla^2 f(x_1,\, x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2(x_1^2 + x_2^2 > 0 \quad \text{for any} \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \mathbf{0}, \quad \text{or}$$

- finding the eigenvalues of $\nabla^2 f(x_1,\, x_2)$, which is $\lambda_1 = \lambda_2 = 2 > 0$.

Question 3. Prove that if $\mathbf{A}$ is an $n \times n$ positive definite matrix, then

(a) All eigenvalues of $\mathbf{A}$ are positive.

If $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, corresponding to the eigenvalue λ, then $\mathbf{x} \neq \mathbf{0}$, and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Therefore $\mathbf{x}^T\mathbf{A}\mathbf{x} = \lambda\|\mathbf{x}\|^2 > 0 \quad \Rightarrow \quad \lambda > 0$.

(b) $\mathbf{A}$ is invertible.

If it was not, then there must be a non-zero vector $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{0}$. Therefore $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$, which contradicts our assumption about $\mathbf{A}$ being positive definite.

(c) All eigenvalues of $\mathbf{A}^{-1}$ are positive.

Assume that λ is an eigenvalue of $\mathbf{A}^{-1}$. The there is a vector $\mathbf{x} \neq \mathbf{0}$ such that

$$\begin{aligned}
\mathbf{A}^{-1}\mathbf{x} &= \lambda\mathbf{x} \\
\mathbf{A}(\mathbf{A}^{-1}\mathbf{x}) &= \mathbf{A}(\lambda\mathbf{x}) \\
\mathbf{x} &= \lambda\mathbf{A}\mathbf{x}. \ \text{ Since } \mathbf{x} \neq \mathbf{0} \ \Rightarrow \lambda \neq 0, \text{ so we have} \\
\mathbf{A}\mathbf{x} &= \tfrac{1}{\lambda}\mathbf{x}.
\end{aligned}$$

So, $\frac{1}{\lambda}$ is an eigenvalue of $\mathbf{A}$. By part (a) we have $\frac{1}{\lambda} > 0 \quad \Rightarrow \quad \lambda > 0$.

Question 4. (Winston Chapter 11, Section 3, Question 1, 2, 7, 8, 9)

On the given set **S**, determine whether each function is convex, concave, or neither.

(a) $f(x) = x^3$; $\mathbf{S} = [0, \infty)$.
$f(x)$ is convex on **S**.

$f'(x) = 3x^2$

$f''(x) = 6x$ ← Hessian

1st principal minor of $f''(x) = 6x$ is $6x$

$6x \geq 0$ for any $x \geq 0 \rightarrow f(x)$ convex (Th. 1)

(b) $f(x) = x^3$; $\mathbf{S} = \mathbf{R}$.
$f(x)$ is neither convex nor concave on **S**.

(c) $f(x_1, x_2) = x_1^2 + x_2^2$; $\mathbf{S} = \mathbf{R}^2$.
The Hessian of the function $f(\mathbf{x})$ at any point $\mathbf{x} \in \mathbf{R}^2$ is

$$H(\mathbf{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

The eigenvalues of this matrix are $\lambda_1 = \lambda_2 = 2 > 0$. Hence, $f(x_1, x_2)$ is strictly convex on $\mathbf{R}^2$.

(d) $f(x_1, x_2) = -x_1^2 - x_1x_2 - 2x_2^2$; $\mathbf{S} = \mathbf{R}^2$.
The Hessian of the function $f(\mathbf{x})$ at any point $\mathbf{x} \in \mathbf{R}^2$ is

$$H(\mathbf{x}) = \begin{pmatrix} -2 & -1 \\ -1 & -4 \end{pmatrix}$$

The firs principal minors of this matrix are $-2 < 0$ and $-4 < 0$. The second principal is $(-2).(-4) - (-1).(-1) = 7 > 0$. Hence, by Theorem [redacted]2, $f(\mathbf{x})$ is a concave function on $\mathbf{R}^2$.

(e) $f(x_1, x_2, x_3) = -x_1^2 - x_2^2 - 2x_3^2 + 0.5x_1x_2$; $\mathbf{S} = \mathbf{R}^3$.
The Hessian of the function $f(\mathbf{x})$ at any point $\mathbf{x} = (x_1, x_2, x_3) \in \mathbf{R}^3$ is

$$H(\mathbf{x}) = \begin{pmatrix} -2 & 0.5 & 0 \\ 0.5 & -2 & 0 \\ 0 & 0 & -4 \end{pmatrix}$$

- The first principal minors of this matrix are $-2 < 0$, $-2 < 0$, and $-4 < 0$.
- The second principal minors of this matrix are

$$\begin{vmatrix} -2 & 0.5 \\ 0.5 & -2 \end{vmatrix} = 3.75 > 0, \quad \begin{vmatrix} -2 & 0 \\ 0 & -4 \end{vmatrix} = 8 > 0, \quad \begin{vmatrix} -2 & 0 \\ 0 & -4 \end{vmatrix} = 8 \geq 0.$$

- The third principal minors of this matrix is

$$\begin{vmatrix} -2 & 0.5 & 0 \\ 0.5 & -2 & 0 \\ 0 & 0 & -4 \end{vmatrix} = -15 < 0.$$

Hence, by Theorem [redacted]2, $f(\mathbf{x})$ is a concave function on $\mathbf{R}^3$.