Sheet 6 118010350

A6.1 (a) Example: Let $f(x)=e^{-x}$, $g(x)=x_1^2+x_2^2$, we can verify that $f(x)$ is convex on $\mathbb{R}$, $g(x)$ is convex on $\mathbb{R}^2$.

Then the composition is $(f\circ g)(x)=e^{-(x_1^2+x_2^2)}$,

we can compute the Hessian,

$$\nabla(f\circ g)(x)=\begin{bmatrix} e^{-x_1^2}\cdot(-2x_1) \\ e^{-x_2^2}\cdot(-2x_2)\end{bmatrix},\quad \nabla^2(f\circ g)(x)=\begin{bmatrix}(4x_1^2-2)e^{-x_1^2} & 0 \\ 0 & (4x_2^2-2)e^{-x_2^2}\end{bmatrix}$$

Since $\nabla^2(f\circ g)(x)$ is not positive semidefinite for $\forall x\in\mathbb{R}^2$, then $f\circ g:\mathbb{R}^2\to\mathbb{R}$ is not convex.

(b). Proof: Since $g:\Omega\to\mathbb{R}$ is convex, and $f:I\to\mathbb{R}$ is convex and nondecreasing, $I\supseteq g(\Omega)$, then we can have

$$\begin{aligned}(f\circ g)(\lambda x_1+(1-\lambda)x_2)&=f(g(\lambda x_1+(1-\lambda)x_2))\\ &\le f(\lambda g(x_1)+(1-\lambda)g(x_2))\\ &\le \lambda f(g(x_1))+(1-\lambda)f(g(x_2))\\ &=\lambda(f\circ g)(x_1)+(1-\lambda)(f\circ g)(x_2),\ \forall\, 0\le\lambda\le 1\end{aligned}$$

Thus, $f\circ g:\Omega\to\mathbb{R}$ is convex.

(c). Example: let $f(x)=-x^2$, $g(x)=x_1+x_2$, then we can know that $f(x):\mathbb{R}_+\to\mathbb{R}$ is concave and nonincreasing, $g(x):\mathbb{R}_+^2\to\mathbb{R}$ is convex. Then the composition is $(f\circ g)(x)=-(x_1+x_2)^2$, we can compute the Hessian,

$$\nabla(f\circ g)(x)=\begin{bmatrix}-2x_1\\-2x_2\end{bmatrix},\quad \nabla^2(f\circ g)(x)=\begin{bmatrix}-2 & 0\\ 0 & -2\end{bmatrix}$$

Since $\nabla^2(f\circ g)$ is negative semidefinite for $\forall x\in\mathbb{R}^2$, then $f\circ g:\mathbb{R}_+^2\to\mathbb{R}$ is concave.

(d) Example: let $f(x)=\begin{cases}-\frac{1}{x-1}, & x\le 0\\ -\frac{1}{(x+1)^2}+2, & x>0\end{cases}$

we can verify that $f(x)$ is increasing and nonnegative.

then we consider the function $x\to xf(x)$ on $\mathbb{R}_+$.

$xf(x)=-\frac{x}{(x+1)^2}+2x$, $x\in\mathbb{R}_+$.

we can compute the second order derivative.

$$(xf(x))'=-\frac{1}{(x+1)^2}+\frac{2x}{(x+1)^3}+2,\quad (xf(x))''=\frac{4}{(x+1)^3}-\frac{6x}{(x+1)^4}$$

Since $(xf(x))''$ can be negative on $\mathbb{R}_+$, then $xf(x)$ is not convex on $\mathbb{R}_+$.

A6.2. (a) $\Omega_1=\{(x,t)\in\mathbb{R}^n\times\mathbb{R}: x^Tx\le t^2\}$.

Suppose we have arbitrary points $(x_1,t_1),(x_2,t_2)\in\Omega_1$, then $x_1^Tx_1\le t_1^2$ and $x_2^Tx_2\le t_2^2$.

Consider any $\lambda\in[0,1]$, we can get a new point

$$\lambda(x_1,t_1)+(1-\lambda)(x_2,t_2)=(\lambda x_1+(1-\lambda)x_2,\ \lambda t_1+(1-\lambda)t_2)$$

Then we have to verify whether this new point is satisfied with the condition $x^Tx\le t^2$.

$$\begin{aligned}\text{Since } x^Tx&=(\lambda x_1+(1-\lambda)x_2)^T(\lambda x_1+(1-\lambda)x_2)\\ &=(\lambda x_1^T+(1-\lambda)x_2^T)(\lambda x_1+(1-\lambda)x_2)\\ &=\lambda^2x_1^Tx_1+(1-\lambda)^2x_2^Tx_2+\lambda(1-\lambda)x_1^Tx_2+\lambda(1-\lambda)\cdot x_2^Tx_1\\ &\le\lambda^2t_1^2+(1-\lambda)^2t_2^2+\lambda(1-\lambda)x_1^Tx_2+\lambda(1-\lambda)x_2^Tx_1\end{aligned}$$

$$\begin{aligned}\text{And } t^2&=(\lambda t_1+(1-\lambda)t_2)^2\\ &=\lambda^2t_1^2+(1-\lambda)^2t_2^2+2\lambda(1-\lambda)t_1t_2\end{aligned}$$

Suppose we have $x_1, x_2\in\mathbb{R}_+^n$, that is, every element is positive, and assume $x_1^Tx_1=t_1^2$, $x_2^Tx_2=t_2^2$, then if t_1, t_2 have different sign, then $t_1t_2<0$.

Then we have $\lambda(1-\lambda)x_1^Tx_2+\lambda(1-\lambda)x_2^Tx_1>2\lambda(1-\lambda)t_1t_2$.

Thus, $x^Tx>t^2$, then Ω_1 is not a convex set.

~~$\Omega_2=\{x\in\mathbb{R}^n:\|x-a\|_2\le\|x-b\|_2\}$, $a,b\in\mathbb{R}^n$, $a\ne b$.~~

~~Let $f(x)=\|x-a\|_2-\|x-b\|_2$. Suppose $x_1, x_2\in\Omega_2$, and $\forall\lambda\in[0,1]$,~~

~~Since $f(\lambda x_1+(1-\lambda)x_2)=\|\lambda x_1+(1-\lambda)x_2-a\|_2-\|\lambda x_1+(1-\lambda)x_2-b\|_2$~~

~~$\le\|\lambda x_1+(1-\lambda)x_2-a-(\lambda x_1+(1-\lambda)x_2-b)\|_2$~~

~~$=\|\lambda(x_1-a)+(1-\lambda)(x_2-a)-\lambda(x_1-b)-(1-\lambda)(x_2-b)\|_2$~~

~~$=\|\lambda(x_1-a)-\lambda(x_1-b)+(1-\lambda)(x_2-a)-(1-\lambda)(x_2-b)\|_2$~~

$\Omega_2=\{x\in\mathbb{R}^n:\|x-a\|_2\le\|x-b\|_2\}$, $a,b\in\mathbb{R}^n$, $a\ne b$.

Since the constraint $\|x-a\|_2\le\|x-b\|_2$ is equivalent to $\|x-a\|_2^2\le\|x-b\|_2^2$, then we consider $f(x)=\|x-a\|_2^2-\|x-b\|_2^2$.

Suppose $x=(x_1,\dots,x_n)$, $a=(a_1,\dots,a_n)$, $b=(b_1,\dots,b_n)$, then the function can be rewrite as $f(x)=(x_1-a_1)^2+\dots+(x_n-a_n)^2+(x_1-b_1)^2+\dots+(x_n-b_n)^2$

We can comput the Hessian of $f(x)$.

$$\nabla f(x)=\begin{bmatrix}2(b_1-a_1)\\ \cdots\\ 2(b_n-a_n)\end{bmatrix},\quad \nabla^2f(x)=\begin{bmatrix}0&\cdots&0\\0&\cdots&0\\0&\cdots&0\end{bmatrix}$$

Since $d^T\nabla^2f(x)d\ge 0$, $\forall d\in\mathbb{R}^n$, $\forall x\in\mathbb{R}^n$, then $\nabla^2f(x)$ is positive semidefinite, by Lemma of convex constraint, the set can be expressed as $L_{\le 0}=\{x\in\mathbb{R}^n: f(x)\le 0\}$, which is a convex set.

Thus, Ω_2 is a convex set.

(b). Proof: Since the set $\{x\in\mathbb{R}_+^2: x_1x_2\ge 1\}$ is equivalent to $\{x\in\mathbb{R}_+^2:\frac{1}{x_2}-x_1\le 0\}$, then let $f(x)=\frac{1}{x_2}-x_1$, we can compute the Hessian of $f(x)$.

$$\nabla f(x)=\begin{bmatrix}-1\\-x_2^{-2}\end{bmatrix},\quad \nabla^2f(x)=\begin{bmatrix}0&0\\0&2x_2^{-3}\end{bmatrix}$$

Since $\nabla^2f(x)$ is positive semidefinite for $\forall x\in\mathbb{R}_+^2$, then by Lemma of convex constraint, the set is a convex set.

(c) Proof: Since A is positive semidefinite, then A can be diagonalized as $A=Q\Lambda Q^T$, and the diagonal entry of Λ is non-negative. Let $L=Q\sqrt{\Lambda}$, then A can be expressed as $A=Q\sqrt{\Lambda}\cdot\sqrt{\Lambda}Q^T=Q\sqrt{\Lambda}\cdot\sqrt{\Lambda}^TQ^T=(Q\sqrt{\Lambda})\cdot(Q\sqrt{\Lambda})^T=L\cdot L^T$

Thus we have $x^TAy=x^TL\cdot L^Ty=(L^Tx)^T(L^Ty)$, by Cauchy-Schwarz inequality $((L^Tx)^T(L^Ty))^2\le((L^Tx)^T(L^Tx))\cdot((L^Ty)^T(L^Ty))$,

then $(x^TAy)^2\le(x^TAx)(y^TAy)$, $x^TAy\le\sqrt{(x^TAx)(y^TAy)}$,

then $(x+y)^TA(x+y)\le x^TAx+y^TAy+2\sqrt{(x^TAx)(y^TAy)}$,

thus we have inequality $\sqrt{(x+y)^TA(x+y)}\le\sqrt{x^TAx}+\sqrt{y^TAy}$ (*).

Then we construct $z\in\mathbb{R}^{n+1}$, and $B\in\mathbb{R}^{(n+1)\times(n+1)}$, where $z=\begin{bmatrix}x\\1\end{bmatrix}$, and $B=\left[\begin{array}{c|c}A&0\\ \hline 0&1\end{array}\right]$, and it is easy to verify that B is also positive semidefinite. Then we define $g(z)=f(x)=\sqrt{x^TAx+1}=\sqrt{z^TBz}$.

Then we can use the inequality (*), for $\forall\lambda\in[0,1]$, and $\forall x_1, x_2\in\mathbb{R}^n$, $z_1=\begin{bmatrix}x_1\\1\end{bmatrix}$ and $z_2=\begin{bmatrix}x_2\\1\end{bmatrix}$, we have

$$\begin{aligned}f(\lambda x_1+(1-\lambda)x_2)&=g(\lambda z_1+(1-\lambda)z_2)\\ &=\sqrt{(\lambda z_1+(1-\lambda)z_2)^TB(\lambda z_1+(1-\lambda)z_2)}\\ &\le\sqrt{(\lambda z_1)^TB(\lambda z_1)}+\sqrt{((1-\lambda)z_2)^TB((1-\lambda)z_2)}\\ &=\lambda\sqrt{z_1^TBz_1}+(1-\lambda)\sqrt{z_2^TBz_2}\\ &=\lambda g(z_1)+(1-\lambda)g(z_2)\\ &=\lambda f(x_1)+(1-\lambda)f(x_2).\end{aligned}$$

Then we get $f(\lambda x_1+(1-\lambda)x_2)\le\lambda f(x_1)+(1-\lambda)f(x_2)$.

Thus, $f(x)=\sqrt{x^TAx+1}$ is convex on $\mathbb{R}^n$.

(d) ① check the objective function.

Since $-x_1-x_2$ is linear, then it is convex; By Lemma of maximum, x_3, x_4 are linear, then $\max\{x_3,x_4\}$ is convex. By Lemma of sum rule, thus $-x_1-x_2+\max\{x_3,x_4\}$ is convex.

② Chek the constraints.

Let $g_1(x)=(x_1-x_2)^2+(x_3+2x_4)^4$, $g_2(x)=x_1+2x_2+x_3+2x_4$

Since $g_2(x)$ is linear of x_1,x_2,x_3,x_4, then $g_2(x)$ is convex, by lemma of constraint, the set $\{x\in\mathbb{R}^4: g_2(x)\le 6\}$ is convex.

let $A=\begin{bmatrix}1&-1&0&0\\0&0&1&2\end{bmatrix}$, $x=\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}$, then $Ax=\begin{bmatrix}x_1-x_2\\x_3+2x_4\end{bmatrix}$

Let $f(x)=x_1^2+x_2^4$, we can easily verify $f(x)$ is convex, then by lemma of composition, $g_1(x)=f(Ax)$, $x\in\mathbb{R}^4$ is convex. By Lemma of constraint, the following set $\{x\in\mathbb{R}^4: g_1(x)\le 5\}$ is convex.

③ Therfore, the optimization problem is convex.

Use MATLAB to solve the optimization problem, we can get the optimal solution is $x_1=3.4907$, $x_2=1.2546$, $x_3=0$, $x_4=0$. and the optimal value is -4.7454.

A6.3 (a). Use MATLAB to solve the optimization problem, we can find the solution is $x^*=2.18870459$. with accuracy less than 10^{-5}.

(b). ① Bisection: 18 iterations, $x^*=5.88535309\times10^{-1}$

② Golden Section: 24 iterations, $x^*=5.88530599\times10^{-1}$.

We find that bisection method use less iterations than golden section method to find the solution with accuracy less than 10^{-5}

A6.4. (a) Since $f(x)=e^{1-x_1-x_2}+e^{x_1+x_2-1}+x_1^2+x_1x_2+x_2^2+2x_1-3x_2$

then we can compute the gradient,

$$\nabla f(x)=\begin{bmatrix}(-1)\cdot e^{1-x_1-x_2}+e^{x_1+x_2-1}+2x_1+x_2+2\\ (-1)\cdot e^{1-x_1-x_2}+e^{x_1+x_2-1}+x_1+2x_2-3\end{bmatrix}$$

then we set $d^k=-\nabla f(x^k)$, ① use gloden section method to find α^k which minimize $f(x^k+\alpha^k d^k)$ ② or use the constraint $f(x^k+\alpha d^k)\le f(x^k)+r\alpha\cdot\nabla f(x^k)^Td^k$ to find α^k.

① Exact Line Search: 47 iterations, optimal value is -4.142309.

$x_1^*=-2.1418$, $x_2^*=2.8582$.

② Backtracking: 40 iterations, optimal value is -4.142309

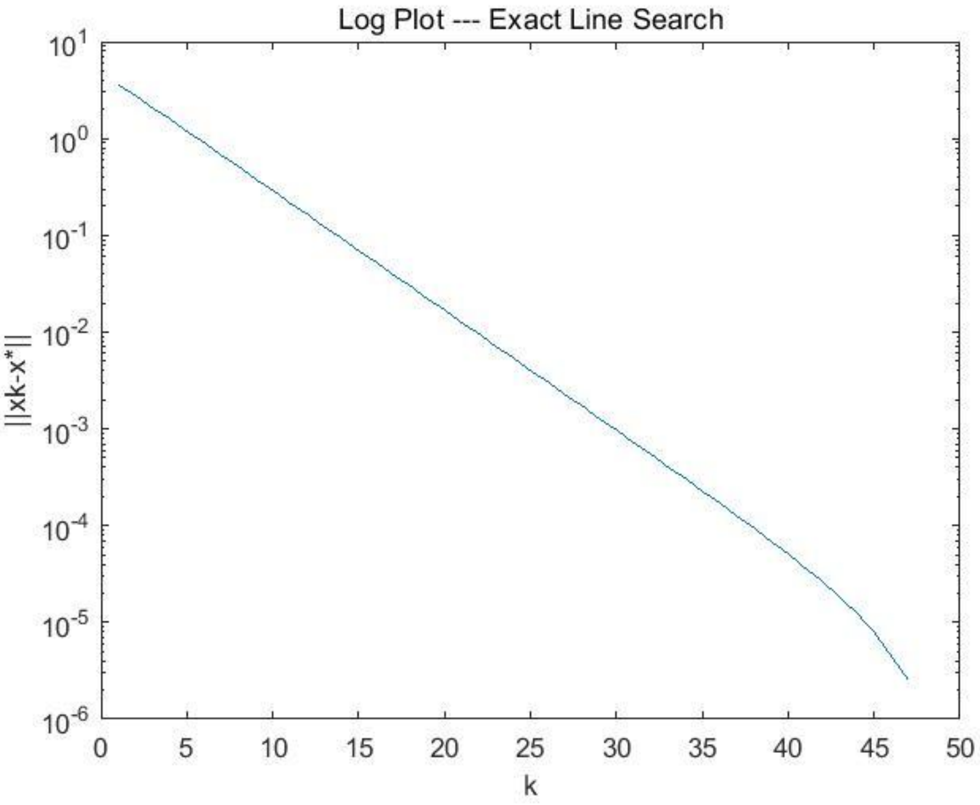
$x_1^*=-2.1418$, $x_2^*=2.8582$

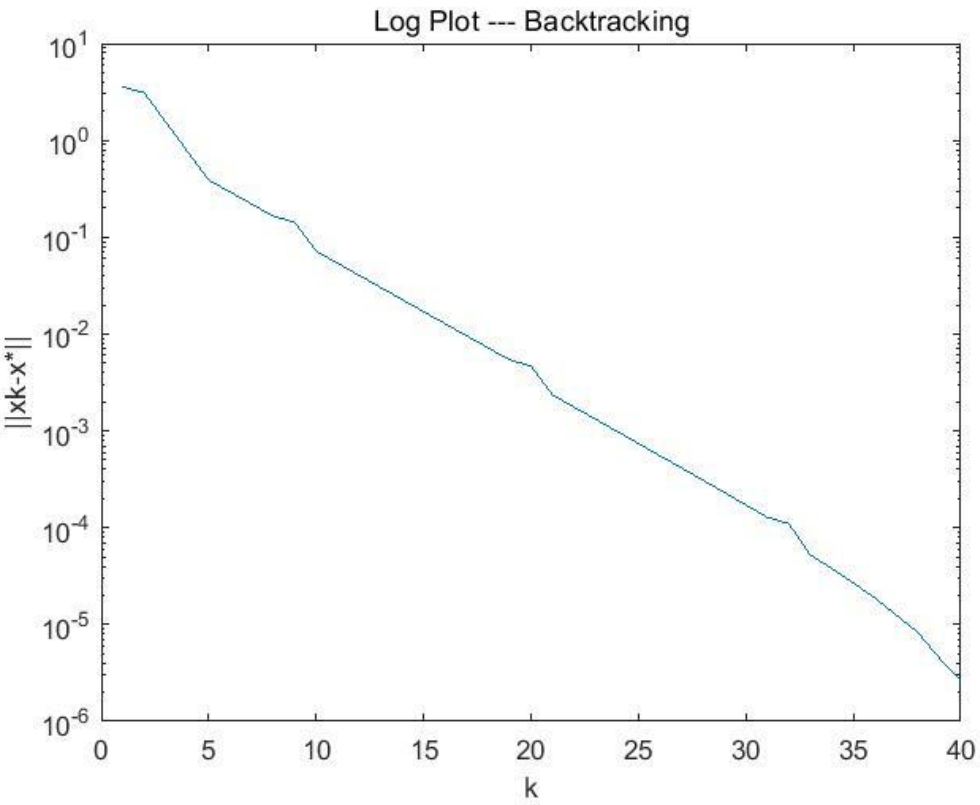
(b) Use logarithmic scale, we can observe that both the sequences $(\|x^k-x^*\|)_k$ and $(\|\nabla f(x^k)\|)_k$ decrease nearly in a straight line.

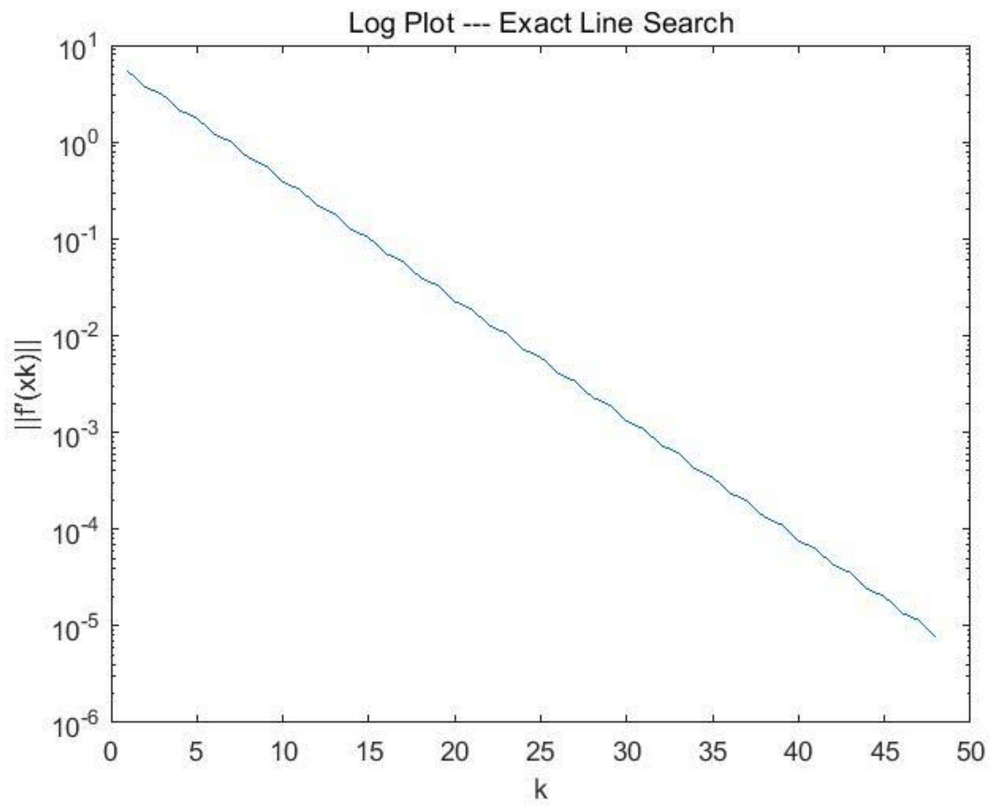
$\Rightarrow$ Graphs are attached at the end.

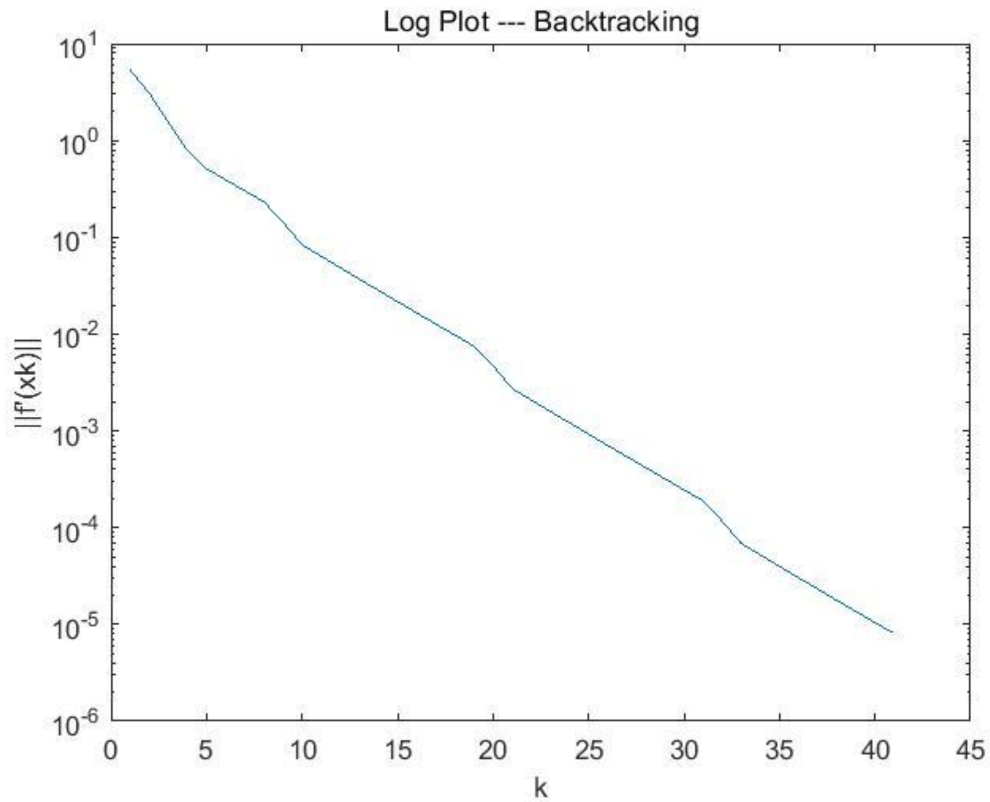
(c) $\Rightarrow$ Graphs are attached at the end

A 6.4(b)

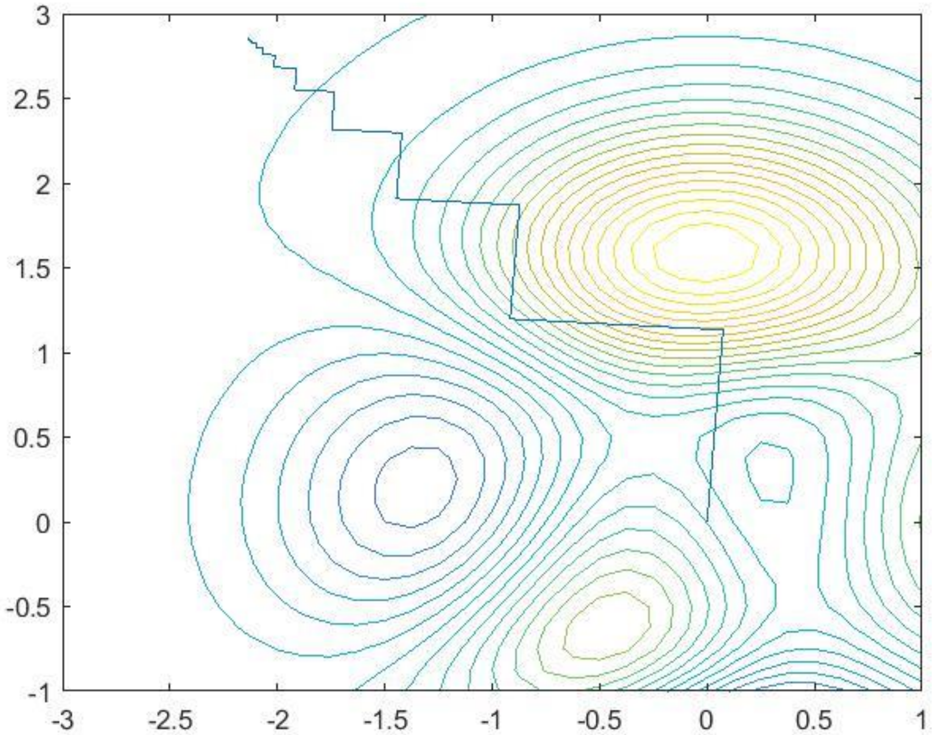




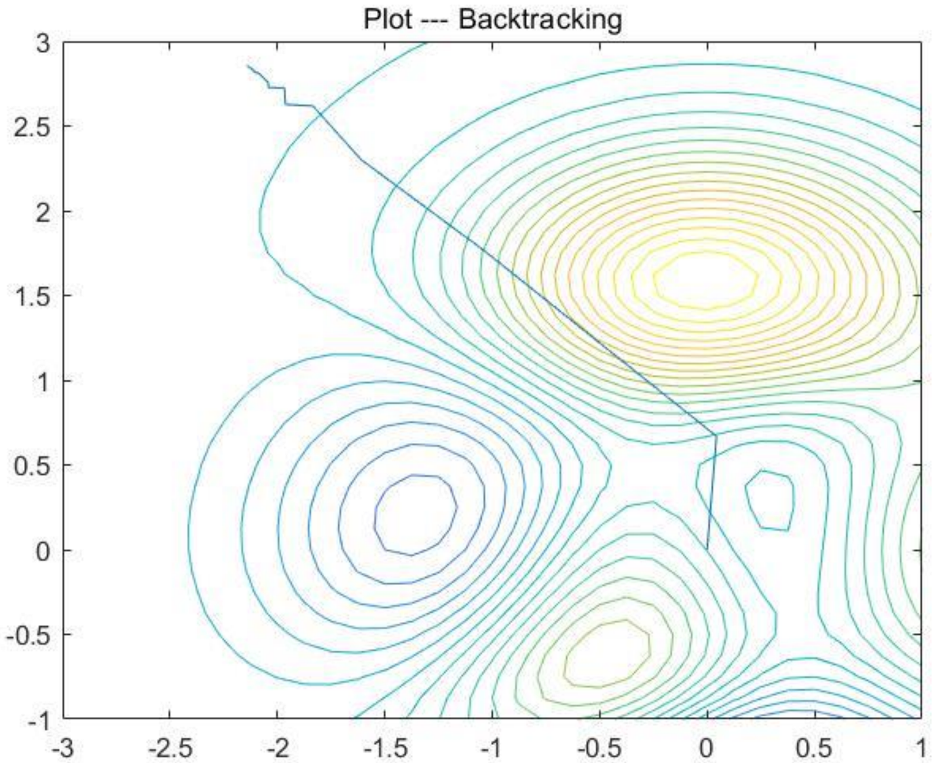

Log Plot --- Exact Line Search
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
||f'(xk)||
0 5 10 15 20 25 30 35 40 45 50
k


Log Plot --- Backtracking
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
||f'(xk)||
0 5 10 15 20 25 30 35 40 45
k

A 6.4(c)



3
2.5
2
1.5
1
0.5
0
-0.5
-1
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1



Plot --- Backtracking
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1