Consider the unconstrained nonlinear program:

$$\min_{x \in \mathbb{R}^2} f(x) = \frac{4}{3} x_1^3 - x_1x_2^2 - 8x_2 + 3x_2^2.$$  

(a) Show that $x = (1, 2)^T$ is a stationary point of $f$.

First order condition $\nabla f(x^*) = 0$:

$$\nabla f(x) = \begin{pmatrix} 4x_1^2 - x_2^2 \\ -2x_1^2 - 8 + 6x_2 \end{pmatrix}, \quad \text{so:} \quad \nabla f(1, 2) = \begin{pmatrix} 4 - 4 \\ -4 - 8 + 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Therefore $(1, 2)$ is a stationary point of $f$.

(b) Does the second-order sufficiency condition hold for the stationary point $x = (1, 2)^T$? Is $x = (1, 2)^T$ a local minimum of $f$? Briefly justify your answer.

Hessian:

$$\nabla^2 f(x) = \begin{pmatrix} 8x_1 & -2x_2 \\ -2x_2 & -2x_1 + 6 \end{pmatrix}, \quad \text{so:} \quad \nabla^2 f(1, 2) = \begin{pmatrix} 8 & -4 \\ -4 & 4 \end{pmatrix}$$

which is positive definite: $\det(\nabla^2 f(1, 2)) = 32 - (-16) = 48 > 0$. Hence by Lemma 3, $(1, 2)$ is a local minimum. Alternatively, solve for the eigenvalues:

$$(8 - \lambda)(4 - \lambda) - 16 = 0 \Rightarrow \lambda^2 - 12\lambda + 16 = 0 \Rightarrow \lambda = \frac{12 \pm \sqrt{80}}{2} > 0$$

for both roots. As all eigenvalues are positive, the point $(1, 2)^T$ is a local minimum.
(c) Find the direction of steepest descent for $f$ at the point $x^0 = (2, 0)^T$. The steepest descent algorithm uses a stepsize $t^*$ to calculate $x^1$ of the form $t^* = \arg\min_{t \geq 0} q(t)$. Write down the function $q(t)$ for this problem (note that you only have to write down the function $q(t)$, you are not required to find its minimum, or to simplify your expression for $q(t)$).

**The steepest descent algorithm**
To minimise a unimodal function $f : \mathbb{R}^N \rightarrow \mathbb{R}$ to within tolerance $\epsilon$.

1. Select $x^0 \in \mathbb{R}^N$.
   Set $k = 0$.
2. Calculate $d_k = -\nabla f(x^k)$.
   If $||d_k|| < \epsilon$ then stop.
3. Select step length $t_k$ by solving the single-variable minimisation problem:
   \[ \min_{q(t)} q(t) = f(x^k + td_k). \]
4. Set $k = k + 1$.
   Set $x^{k+1} = x^k + t_k d_k$.
   Return to step 2.

Steepst descent direction is negative of gradient, so it is $-\nabla f(2, 0) = (-16, 8)^T$. To obtain $x^1$ the method uses $x^1 = (2, 0)^T + t^*(-16, 8)^T$, where $t^* = \arg\min_{t \geq 0} q(t)$, and
\[
q(t) = f[(2, 0) + t(-16, 8)^T] = f(2 - 16t, 8t) = \frac{4}{3}(2 - 16t)^3 - (8t)^2(2 - 16t) + 64t + 3(-8t)^3 = \frac{4}{3}(2 - 16t)^3 + 1024t^3 + 64t^2 - 64t.
\]

(d) Find the Newton direction for $f$ at $x^0 = (2, 0)^T$. Is the Newton direction a descent direction? Justify your answer.

**Newton’s Method**
To minimise a unimodal function $f : \mathbb{R}^N \rightarrow \mathbb{R}$ to within tolerance $\epsilon$.

1. Select $x^0 \in \mathbb{R}^N$.
   Set $k = 0$.
2. If $||\nabla f(x^k)|| < \epsilon$ then stop.
   If $\nabla^2 f(x^k)$ is positive definite, then
   Set $d^k = -\nabla^2 f(x^k)^{-1} \nabla f(x^k)$.
   Else, set $d^k = -\nabla f(x^k)$
3. Select step length $t_k$ either
   • by solving the single-variable minimisation problem: $\min q(t) = f(x^k + td^k)$.
   • by using our procedure for finding a step length that satisfies the Armijo-Goldstein and Wolff conditions.

4. Set $k = k + 1$.
   Set $x^{k+1} = x^k + t_k d^k$.
   Return to step 2.

Newton direction is $-\left[ \nabla^2 f(2, 0) \right]^{-1} \nabla f(2, 0)$, that is:

$$
\begin{pmatrix}
16 & 0 \\
0 & 2
\end{pmatrix}^{-1} 
\begin{pmatrix}
-16 \\
8
\end{pmatrix} = 
\begin{pmatrix}
1/16 & 0 \\
0 & 1/2
\end{pmatrix} 
\begin{pmatrix}
-16 \\
8
\end{pmatrix} = 
\begin{pmatrix}
-1 \\
4
\end{pmatrix}
$$

Consider Lemma 6:

**Lemma 6**

If $\nabla f(x^k) \neq 0$ and $\nabla^2 f(x^k)$ is positive definite, then $\nabla^2 f(x^k)$ is invertible and the Newton direction is a descent direction for $f$ at $x^k$.

We can see that $\nabla^2 f(x^k)$ is positive definite as it has only positive eigenvalues.

**(e)** Show that a step of length $t_0 = 1$ along the Newton direction $d$ at $x^0$ satisfies the Armijo-Goldstein condition, with the linesearch parameter $\sigma = \frac{1}{9}$.

Use $t = 1$ and $d = (-1, 4)^T$. Armijo-Goldstein is satisfied if:

$$f(x^0 + td) \leq f(x^0) + t\sigma \nabla f(x^0)^Td.$$  

Verification: $x^0 + td = (2, 0)^T + (1) \times (-1, 4)^T = (1, 4)^T$. $f(1, 4) = 4/3$, and

$$f(x^0) + t\sigma \nabla f(x^0)^Td = \frac{32}{3} + 1 \left( \frac{1}{9} \right) (-48) = \frac{48}{9} > \frac{4}{3}.$$