Assignment 6

Question a: Affine Invariance

We need to prove that $f(x_k) = g(y_k)$.

To do this, let's compute the gradient of $g(y)$ in terms of the gradient of $f(x)$.

Given that $g(y) = f(Ay + b)$, we can express x in terms of y as $x = Ay + b$.

The gradient $\nabla g(y)$ can be written using the chain rule:

$$\nabla g(y) = \nabla f(Ay + b) = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial y}.$$

Since $\frac{\partial x}{\partial y} = A$, this simplifies to:

$$\nabla g(y) = A^T \nabla f(x).$$

Now let's compute the hessian function of $g(y)$

$$H_g(y) = \frac{\partial(\nabla g(y))}{\partial y} = \frac{\partial(A^T \nabla f(x))}{\partial y} = A^T \frac{\partial \nabla f(x)}{\partial x} \cdot \frac{\partial x}{\partial y} = A^T H_f(x) A.$$

$H_f(x) = \frac{\partial(\nabla f(x))}{\partial x}$ is the Hessian function of f.

Let's show the relationship between x_{k+1} and y_{k+1}.

$$y_{k+1} = y_k - H_g(y_k)^{-1} \nabla g(y_k) = y_k - (A^T H_f(x_k) A)^{-1} (A^T \nabla f(x_k))$$
$$A(y_{k+1} - y_k) = -A(A^T H_f(x_k) A)^{-1} (A^T \nabla f(x_k)) = -H_f(x_k)^{-1} \nabla f(x_k)$$

We know that the update step of Newton's method for the function f is:

$$x_{k+1} = x_k - H_f(x_k)^{-1} \nabla f(x_k).$$

then

$$x_{k+1} = x_k + A(y_{k+1} - y_k).$$

Therefore the Newton's method is invariant under affine transformation.

Question b: Affine Invariance

Under an affine transformation, we have $x = Ay + b$, i.e., $x_k = Ay_k + b$.

The update of x_k is:

$$x_{k+1} = x_k - s^k \nabla x_k.$$

We have $\nabla x_k = A \nabla y_k$, therefore:

$$x_k - s^k \nabla x_k = Ay_k + b - s^k A \nabla y_k = A(y_k - s^k \nabla y_k) + b.$$

This satisfies the line search condition for $f(x)$:

$$f(x_k - s^k \nabla x_k) \leq f(x_k) + \alpha s^k \nabla f(x_k)^T (-\nabla x_k).$$

Since, under the affine transformation, $f(x_k - s^k \nabla x_k) = g(y_k - s^k \nabla y_k)$, it follows that $t^k = s^k$ satisfies the line search condition for $g(y)$.

Therefore, $s^k = t^k$.

Let's show that the Newton decrements of both functions are identical.

We have the decrement functions for f and g respectively:

$$\lambda(x_k) = \sqrt{\nabla f(x_k)^T H_f(x_k)^{-1} \nabla f(x_k)},$$

$$\lambda(y_k) = \sqrt{\nabla g(y_k)^T H_g(y_k)^{-1} \nabla g(y_k)}.$$

Since:

$$H_g(y_k)^{-1} = (A^T H_f(x_k) A)^{-1} = A^{-1} H_f(x_k)^{-1} (A^T)^{-1},$$

and

$$\nabla g(y_k) = A^T \nabla f(x_k),$$

we have:

$$\lambda(y_k) = \sqrt{(A^T \nabla f(x_k))^T A^{-1} H_f(x_k)^{-1} (A^T)^{-1} A^T \nabla f(x_k)}.$$

$$\lambda(y_k) = \sqrt{\nabla f(x_k)^T H_f(x_k)^{-1} \nabla f(x_k)} = \lambda(x_k).$$

Therefore the Newton decrements of both functions are identical.