4opt

Alain, Jonas, Vincent

Exercise 6

Affine Invariance

a) Show that Newton's method is invariant under affine transformations. Suppose $f: \mathbb{R}^n \to \mathbb{R}$ is a twice differentiable function. Consider the function $g(y) = f(Ay+b)$, where A is a non-singular constant matrix and $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}$. Prove that the sequence of point $\{x^k\}$ of f and the sequence of points $\{y^k\}$ of g generated by Newton's Method, starting from x^0, y^0 respectively with $x^0 = Ay^0+b$, have one-to-one correspondence under the transformation, ie. $f(x^k) = g(y^k)$. Here we assume that the step lengths of the search directions of both functions are equal.

We need to prove that

$$f(x^k) = g(y^k) \Rightarrow f(x^k) = f(Ay^k+b) \Rightarrow x^k = Ay^k+b$$

by induction:

Note: $\nabla g(y) = \nabla(f(Ay+b))$
$= A^T \nabla f(Ay+b)$ and
$\nabla^2 g(y) = A^T \nabla^2 f(Ay+b) A$

- Step 0:

$x^0 = Ay^0+b$ directly by hypothesis

- Induction step:

Suppose we have $x^k = Ay^k+b$

$$y^{k+1} = y^k - t^k \nabla^2 g(y^k)^{-1} \nabla g(y^k)$$ Newton's formula step

$$\begin{aligned}
\Rightarrow Ay^{k+1}+b &= Ay^k - t^k A \nabla^2 g(y^k)^{-1} \nabla g(y^k) + b \\
&= Ay^k+b - t^k A \left(A^T \nabla^2 f(Ay^k+b) A\right)^{-1} \left(A^T \nabla f(Ay^k+b)\right) \\
&= \underbrace{Ay^k+b}_{x^k} - t^k \underbrace{AA^{-1}}_{Id} \nabla^2 f(Ay^k+b)^{-1} \underbrace{A^{T^{-1}} A^T}_{Id} \nabla f(Ay^k+b) \\
&= x^k - t^k \nabla^2 f(Ay^k+b)^{-1} \nabla f(Ay^k+b) \\
&= x^k - t^k \nabla^2 f(x^k)^{-1} \nabla f(x^k) \\
&= x^{k+1}
\end{aligned}$$

t^k = step length equal

$$\Rightarrow x^k = Ay^k+b \quad \forall k \in \mathbb{N}$$

$$\Rightarrow f(x^k) = g(y^k) \quad \forall k \in \mathbb{N}$$ #

b) Under affine transformations, show that the Newton's method with backtracking line search generates the same length step, ie. the step length s^k of the function $f(x)$ equals to t^k of the function $g(y)$

In addition, show that the newton decrements of both functions are identical

---

- Newton decrements:

$$\lambda(x) = \left(\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)\right)^{1/2}$$

$$= \left(\nabla f(x)^T A A^{-1} \nabla^2 f(x)^{-1} A^{T^{-1}} A^T \nabla f(x)\right)^{1/2}$$

$$= \left(\left(\nabla f(Ay+b)^T A\right)\left(A^{-1} \nabla^2 f(x)^{-1} A^{T^{-1}}\right)\left(A^T \nabla f(x)\right)\right)^{1/2}$$

$$= \left(\left(A^T \nabla f(Ay+b)\right)^T \left(A^T \nabla^2 f(x) A\right)^{-1} A^T \nabla f(x)\right)^{1/2}$$

$$= \left(\nabla g(y)^T \nabla^2 g(y)^{-1} \nabla g(y)\right)^{1/2} = \lambda(y)$$

$$\Rightarrow \lambda(x) = \lambda(y) \quad \#$$

- Step size:

we need for the calculation:

$$\begin{cases} \Delta x = -\nabla^2 f(x)^{-1} \nabla f(x) \\ \Delta y = -\nabla^2 g(y)^{-1} \nabla g(y) \\ \quad = -A^{-1} \nabla^2 f(x) A^{T^{-1}} A^T \nabla f(x) \\ \quad = -A^{-1} \nabla^2 f(x) \nabla f(x) \\ \quad \Rightarrow A \Delta y = \Delta x \end{cases}$$

(I) $f(x^k + t^k \Delta x^k) < f(x^k) + \alpha t^k \nabla f(x^k)^T \Delta x^k$

$f(x^k + t^k \Delta x^k) < f(x^k) - \alpha t^k \lambda(x^k)^2$

(II) $g(y^k + s^k \Delta y^k) < g(y^k) + \alpha s^k \nabla g(y^k)^T \Delta y^k$

$f(A(y^k + s^k \Delta y^k) + b) < f(Ay^k + b) + \alpha s^k \nabla f(x^k)^T A A^{-1} \Delta x^k$

$f(Ay^k + s^k A \Delta y^k + b) < f(x^k) + \alpha s^k \nabla f(x^k)^T \Delta x^k$

$f(x^k + s^k \Delta x^k) < f(x^k) - \alpha s^k \lambda(x^k)^2$

(I) and (II) $\Rightarrow$ The backtracking line search will always stop at the same t^k, s^k for $f(x^k), g(y^k)$

$\Rightarrow s^k = t^k \quad \forall k \in \mathbb{N} \quad \#$