Assignment 1 Answers

1. **Find the gradient and Hessian of the function $||Ax - b||_2^2$. Is the Hessian positive definite? Justify your answer.**

   Let $f(x) = ||Ax - b||_2^2 = (Ax - b)^T(Ax - b)$.

   Gradient:

   $$\begin{aligned}\nabla f(x) &= \nabla \left[(Ax - b)^T(Ax - b)\right] \\ &= 2A^T(Ax - b).\end{aligned}$$

   Hessian:

   $$H_f(x) = \nabla^2 f(x) = 2A^TA.$$

   Is the Hessian positive definite?

   A^TA is not always positive definite.

2. **Find the gradient and Hessian of $\Phi(x) = \frac{1}{2}x^TAx - b^Tx$, where $A \in \mathbb{R}^n$.**

   Gradient:

   $$\begin{aligned}\nabla\Phi(x) &= \nabla\left(\frac{1}{2}x^TAx - b^Tx\right) \\ &= \frac{(A + A^T)x}{2} - b.\end{aligned}$$

   Hessian:

   $$H_\Phi(x) = \nabla^2\Phi(x) = \frac{(A + A^T)}{2}.$$

   (a) **Suppose A is positive definite. Let $x = x_0 - \alpha p$, where x_0 and $p \in \mathbb{R}^n$ are fixed. And $\alpha \geq 0$Now, Φ is a univariate function of α. Find the minimizer α^*.**

   Let $\Phi(\alpha) = \Phi(x_0 - \alpha p) = \frac{1}{2}(x_0 - \alpha p)^TA(x_0 - \alpha p) - b^T(x_0 - \alpha p)$.

   The derivative with respect to α is:

   $$\begin{aligned}\frac{d\Phi(\alpha)}{d\alpha} &= \frac{d}{d\alpha}\left[\frac{1}{2}(x_0 - \alpha p)^TA(x_0 - \alpha p) - b^T(x_0 - \alpha p)\right] \\ &= -p^TA(x_0 - \alpha p) + b^Tp.\end{aligned}$$

   Setting this derivative to zero to find the minimizer:

   $$\begin{aligned}-p^TA(x_0 - \alpha^* p) + b^Tp = 0 \\ -p^TAx_0 + \alpha^* p^TAp + b^Tp = 0 \\ \alpha^* = \frac{p^TAx_0 - b^Tp}{p^TAp}.\end{aligned}$$

(b) **In part (a), let p be the gradient of $\Phi(x)$ at x_0 and repeat the previous step.**
Let $p = \nabla\Phi(x_0) = Ax_0 - b$.
Then, the minimizer α^* is:

$$\alpha^* = \frac{(Ax_0 - b)^T Ax_0 - b^T(Ax_0 - b)}{(Ax_0 - b)^T A(Ax_0 - b)}.$$

(c) **Let p be the negative gradient of $\Phi(x)$ at x_0 and repeat the previous step.**
Let $p = -\nabla\Phi(x_0) = -(Ax_0 - b)$.
Then, the minimizer α^* is:

$$\begin{aligned}
\alpha^* &= \frac{-(Ax_0 - b)^T Ax_0 + b^T(Ax_0 - b)}{(Ax_0 - b)^T A(Ax_0 - b)} \\
&= \frac{b^T(Ax_0 - b) - (Ax_0 - b)^T Ax_0}{(Ax_0 - b)^T A(Ax_0 - b)} \\
&= \frac{b^T Ax_0 - b^T b - x_0^T A^T Ax_0 + x_0^T Ab}{(Ax_0 - b)^T A(Ax_0 - b)} \\
&= \frac{b^T Ax_0 - b^T b - (Ax_0)^T(Ax_0) + x_0^T Ab}{(Ax_0 - b)^T A(Ax_0 - b)}. \\
&= -\frac{p^T p}{p^T Ap} \leq 0
\end{aligned}$$

So α^*=0