CV Assignment 3 - Math Part(1)

1.

To prove that $L(\mathbf{h})$ is strictly convex, we need to show that its Hessian matrix $\mathbf{H}(\mathbf{x})$ is positive definite. The Hessian matrix of $L(\mathbf{h})$ is derived from the second derivative of $L(\mathbf{h})$ with respect to $\mathbf{h}$.

Given the expression

$$\begin{aligned} L(\mathbf{h}) &= \frac{1}{2}(f(\mathbf{x}+\mathbf{h}))^T f(\mathbf{x}+\mathbf{h}) + \frac{1}{2}\mu\mathbf{h}^T\mathbf{h} && (1) \\ &= \frac{1}{2}(f(\mathbf{x}))^T f(\mathbf{x}) + \mathbf{h}^T(\mathbf{J}(\mathbf{x}))^T f(\mathbf{x}) + \frac{1}{2}\mathbf{h}^T(\mathbf{J}(\mathbf{x}))^T\mathbf{J}(\mathbf{x})\mathbf{h} + \frac{1}{2}\mu\mathbf{h}^T\mathbf{h} && (2) \end{aligned}$$

$\mathbf{h} \in \mathbb{R}^n$, $L(\mathbf{h})$ is scalar, $f(\mathbf{x}) \in \mathbb{R}^m$, $\mathbf{J}(\mathbf{x}) \in \mathbb{R}^{m\times n}$.

We can get its Hessian:

$$\nabla L(\mathbf{h}) = \mathbf{J}(\mathbf{x})^T f(\mathbf{x}) + (\mathbf{J}(\mathbf{x}))^T\mathbf{J}(\mathbf{x})\mathbf{h} + \mu\mathbf{h} \tag{3}$$

$\nabla L(\mathbf{h}) \in \mathbb{R}^n$.

$$\mathbf{H}(\mathbf{x}) \equiv \nabla^2 L(\mathbf{h}) = (\mathbf{J}(\mathbf{x}))^T\mathbf{J}(\mathbf{x}) + \mu I \tag{4}$$

$\nabla^2 L(\mathbf{h}) \in \mathbb{R}^{n\times n}$, where I is the identity matrix.

To prove that $\mathbf{H}(\mathbf{x})$ is positive definite, we need to prove that for all non-zero vectors $\mathbf{z}$, $\mathbf{z}^T\mathbf{H}(\mathbf{x})\mathbf{z} > 0$.

$$\begin{aligned} \mathbf{z}^T\mathbf{H}(\mathbf{x})\mathbf{z} &= \mathbf{z}^T((\mathbf{J}(\mathbf{x}))^T\mathbf{J}(\mathbf{x}) + \mu I)\mathbf{z} && (5) \\ &= \mathbf{z}^T(\mathbf{J}(\mathbf{x}))^T\mathbf{J}(\mathbf{x})\mathbf{z} + \mathbf{z}^T\mu I\mathbf{z} && (6) \\ &= (\mathbf{J}(\mathbf{x})\mathbf{z})^T\mathbf{J}(\mathbf{x})\mathbf{z} + \mu\mathbf{z}^T\mathbf{z} && (7) \\ &= ||\mathbf{J}(\mathbf{x})\mathbf{z}||_2^2 + \mu||\mathbf{z}||_2^2 && (8) \end{aligned}$$

Obviously, $||\mathbf{J}(\mathbf{x})\mathbf{z}||_2^2 \geq 0$, $\mu||\mathbf{z}||_2^2 > 0$(because $\mu > 0$ and $\mathbf{z}$ is non-zero vector), thus $\mathbf{z}^T\mathbf{H}(\mathbf{x})\mathbf{z} > 0$.

So $\mathbf{H}(\mathbf{x})$ is positive definite, $L(\mathbf{h})$ is strictly convex.