Problem1

1. Problem Description

Please prove that $L(\mathbf{h})$ is a strictly convex function.

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

where $\mathbf{J}(\mathbf{x})$ is $f(\mathbf{x})$'s Jacobian matrix, and $\mu > 0$ is the damped coefficient.

2. Problem Solution

According to the prompts given by the title, If a function $L(\mathbf{h})$ is differentiable up to at least second order, L is strictly convex if its Hessian matrix is positive definite.)

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

so we can calculate derivative of $L(\mathbf{h})$ to the first order (Jacobian matrix):

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

And then, we can calculate derivative of $L(\mathbf{h})$ to the second order (Hessian matrix):

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

$\nabla L(\mathbf{h}) \in R^n$ and $\nabla^2 L(\mathbf{h}) \in R^{n\times n}$. I is the indentity matrix. Next, we should prove that $\mathbf{H}(\mathbf{h})$ is positive definite.

For all $\mathbf{x} \in R^n$ and $\mathbf{x} \neq 0$, we have:

$$\mathbf{x}^T\mathbf{H}(\mathbf{h})\mathbf{x} = \mathbf{x}^T(\mathbf{J}(\mathbf{x}))^T\mathbf{J}(\mathbf{x})\mathbf{x} + \mathbf{x}^T\mu I\mathbf{x}$$
$$= (\mathbf{J}(\mathbf{x})\mathbf{x})^T\mathbf{J}(\mathbf{x})\mathbf{x} + \mu\mathbf{x}^T\mathbf{x}$$
$$= ||\mathbf{J}(\mathbf{x})\mathbf{x}||_2^2 + \mu||\mathbf{x}||_2^2$$

Obviously, $||\mathbf{J}(\mathbf{x})\mathbf{x}||_2^2 \geq 0$. Beacause $\mu > 0$ is the damped coefficient, so $\mu||\mathbf{x}||_2^2 > 0$. Therefore, $\mathbf{x}^T\mathbf{H}(\mathbf{h})\mathbf{x} > 0$. So $\mathbf{H}(\mathbf{x})$ is positive definite, and $L(\mathbf{h})$ is a strictly convex function.