Homework4 Report

**莱文贝格－马夸特方法**（Levenberg–Marquardt algorithm）能提供数非线性最小化（局部最小）的数值解。此算法能借由执行时修改参数达到结合高斯-牛顿算法以及梯度下降法的优点，并对两者之不足作改善（比如高斯-牛顿算法之反矩阵不存在或是初始值离局部极小值太远）。

假设 f 是一个从 \real^m \rightarrow \real^n 的非线性映射，也就是说 \mathbf{P} \in \real^m 且 \mathbf{X} \in \real^n, 那么:


f(\mathbf{P}) = \mathbf{X}


而我们的目的就是希望任意给定一个 \mathbf{x} 以及合理的初始值 \mathbf{p}_0，我们能找到一个 \mathbf{p}^{+}，使得 \mathbf{\epsilon}^T\mathbf{\epsilon} 尽量小（局部极小），其中 \mathbf{\epsilon} = f(\mathbf{p}^+)-\mathbf{x}。

像大多数最小化的方法一样，这是一个迭代的方法。首先根据泰勒展开式我们能把 f(\mathbf{p}+\mathbf{\delta_p}) 写为下面的近似，这有两个好处：第一是线性、第二是只需要一阶微分。


f(\mathbf{p}+\mathbf{\delta_p}) \approx f(\mathbf{p}) + \mathbf{J\delta_p}


对于每次的迭代我们这么作：假设这次 iteration 的点是 \mathbf{p}_k，我们要找到一个 \mathbf{\delta}_{\mathbf{p},k} 让 |\mathbf{x}-f(\mathbf{p}+\mathbf{\delta}_{\mathbf{p},k})| \approx |\mathbf{x}-f(\mathbf{p}) - \mathbf{J\mathbf{\delta}_{\mathbf{p},k}}| = |\mathbf{\epsilon}_{k}-\mathbf{J\mathbf{\delta}_{\mathbf{p},k}}| 最小。 根据投影公式我们知道当下面式子被满足的时候能有最小误差：


(\mathbf{J}^T\mathbf{J})\mathbf{\delta_p} = \mathbf{J}^T\mathbf{\epsilon}_{k}


我们将这个公式略加修改得到：


[\mu\mathbf{I}+(\mathbf{J}^T\mathbf{J})]\mathbf{\delta_p} = \mathbf{J}^T\mathbf{\epsilon}_{k}


就是莱文贝格－马夸特方法。如此一来 \mu 大的时候这种算法会接近最速下降法，小的时候会接近高斯-牛顿方法。为了确保每次 \mathbf{\epsilon} 长度的减少，我们这么作：先采用一个小的 \mu，如果 \mathbf{\epsilon} 长度变大就增加 \mu。

这个算法当以下某些条件达到时结束迭代：如果发现 \mathbf{\epsilon} 长度变化小于特定的给定值就结束。发现 \mathbf{\delta_p} 变化小于特定的给定值就结束。到达了迭代的上限设定就结束。