1 Second Order Methods

For unconstrained problem $\min_{x\in\mathbb{R}^n} f(x)$, the standard Newton scheme updates according to

$$x^{k=1} = x^k - \left[\nabla^2 f(x^k)\right]^{-1} \nabla f(x^k) \tag{1}$$

where $f \in C_L^{2,2}(\mathbb{R}^n)$. The method has quadratic local convergence rate when initial iterate is close to the optimum of f [1]. Cubic regularized Newton's Method converges globally to second order stationary points ($\nabla f(x) = 0$ and $\nabla^2 f(x) \succeq 0$) assuming $f \in C_L^{2,2}$, i.e. twice continuously differentiable with lipschitz continuous hessian [2, 3]. The idea is to iteratively minimize a global upper bound of the objective,

$$x^{k+1} = \arg\min_{y\in\mathbb{R}^n} \tilde{f}_{x^k} y \tag{2}$$

where $\tilde{f}(y)$ is a cubic regularized quadratic model of the objective,

$$\tilde{f}_x(y) = f(x) + \langle \nabla f(x), y - x\rangle + \frac{1}{2}\left\langle \nabla^2 f(x)(y-x), y-x\right\rangle + \frac{L}{6}\|y-x\|^3 \tag{3}$$

This modified Newton step ensures that function values of iterates are monotonic non-increasing. Cubic regularized Newton's Method converges globally to second order stationary points ($\nabla f(x) = 0$ and $\nabla^2 f(x) \succeq 0$) assuming $f \in C_L^{2,2}$, i.e. twice continuously differentiable with lipschitz continuous hessian [2, 3]. The method has quadratic global convergence rate when initial iterate is close to the optimum of f [1]. Under weak non-degeneracy assumption of the Hessian matrix, the local convergence rate is super-linear of the order $\frac{4}{3}$ or $\frac{3}{2}$.