Theorem 2.20 Assume that $f(x)$ is twice differentiable and $\nabla^2 f(x)$ is Lipschitz continuous in a neighborhood of the solution x^*, which satisfies the sufficient second order conditions. Then, by applying Algorithm 2.1 and with x^0 sufficiently close to x^*, there exists a constant $\hat{L} > 0$ such that

- $\|x^{k+1} - x^*\| \leq \hat{L}\|x^k - x^*\|^2$, i.e., the convergence rate for $\{x^k\}$ is quadratic;
- the convergence rate for $\{\nabla f(x^k)\}$ is also quadratic.

Claim 1. The convergence rate for $\{x^k\}$ is quadratic.

Step 1. What does the continuity of the second derivative tell us?

Step 2. Consider the Newton step $p^k = -\left(\nabla^2 f(x^k)\right)^{-1} \nabla f(x^k)$. Add $x^k - x^*$ to both sides.

Step 3. Work through algebra on RHS to obtain…

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

Step 4. Recall $\nabla f(x+p) - \nabla f(x) = \int_0^1 \nabla^2 f(x^k + tp)dt$ and apply to RHS.

Step 5. Take norms of both sides. Recall the Cauchy–Schwarz inequality $|u^T v| \leq ||u|| \cdot ||v||$.

Step 6. Invoke continuity properties (see Step 1).

Claim 2. The convergence rate for $\{\nabla f(x^k)\}$ is quadratic.

Start with $0 = \nabla f(x^k) + \nabla^2 f(x^k) p^k$.

Step 1. Add $\nabla f(x^{k+1})$ to both sides.

Step 2. Take norm of both sides.

Step 3. Recall $\nabla f(x^{k+1}) - \nabla f(x^k) = \int_0^1 \nabla^2 f(x^k + tp^k) p^k dt$.

Step 4. Invoke Lipschitz continuity.

Step 5. Substitute $p^k = -\nabla^2 f(x^k)^{-1} \nabla f(x^k)$

ALGORITHM 3.1.
Choose a starting point x^0 and tolerances $\epsilon_1, \epsilon_2 > 0$.

For $k \geq 0$ while $\|p^k\| > \epsilon_1$ and $\|\nabla f(x^k)\| > \epsilon_2$:

1. At x^k, evaluate $\nabla f(x^k)$ and the matrix B^k, which is positive definite and bounded in condition number.
2. Solve the linear system $B^k p^k = -\nabla f(x^k)$.
3. Set $x^{k+1} = x^k + p^k$ and $k = k + 1$.