Week 5, Lecture 2

Plan:

- Finish Chapter 6
- Exercises #2. These are:
  - Theoretical 1–4
  - $\boxed{\text{Coding} \quad 5}$

  } examinable.

Convergence Rate for the Newton Method §6.3

Suppose that f is twice-differentiable and the Hessian $B(\underline{x})$ is Lipschitz in a neighbourhood of the minimizer $\underline{x}_*$ at which the sufficient conditions for optimality hold ($\nabla f(\underline{x}_*) = 0$, $B(\underline{x}_*)$ pos. definite).

Suppose that the starting-point $\underline{x}_0$ is sufficiently close to $\underline{x}_*$ and consider the iterates

$$\underline{x}_{k+1} = \underline{x}_k + \underline{p}_k^N .$$

Then:

1. The iterates converge: $\underline{x}_k \to \underline{x}_*$ as $k \to \infty$.
2. The rate of convergence is quadratic:

$$\|\underline{x}_{k+1} - \underline{x}_*\|_2 \leq C \|\underline{x}_k - \underline{x}_*\|_2^2 .$$

No proof!

Remark: By $B(\underline{x})$ Lipschitz in a neighbourhood $\mathcal{N}$ of $\underline{x}_*$, we mean, that there exists a positive constant K such that

$$\| B(\underline{y}_2) - B(\underline{y}_1) \|_2 \leq K \| \underline{y}_2 - \underline{y}_1 \|_2$$

for all $\underline{y}_1$ and $\underline{y}_2$ in $\mathcal{N}$.

Remark: As you go through the different steps in the proof, we get down to:

$$\frac{\| \underline{x}_1 - \underline{x}_0 \|_2}{\| \underline{x}_0 - \underline{x}_* \|_2} \leq \underbrace{C \| \underline{x}_0 - \underline{x}_* \|_2}$$

"Sufficiently close" means the combination on the RHS is $\leq 1/2$.

More about this in Exercises #2, Question 1.

<u>One more remark</u>: A similar result holds for quasi-Newton methods (e.g. BFGS):

$$\| \underline{x}_{k+1} - \underline{x}_* \|_2 \leq C \| \underline{x}_k - \underline{x}_* \|_2^{1+\epsilon}$$

where $\epsilon > 0$. I.e. better than SD.
For an example of how S.D. ($\epsilon = 0$) can fail, see Exercises #2, Question 5.