Solutions to Part A of Problem Sheet 2

Solution (2.1)

(a) We apply the bound inductively,

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\| \leq r \cdot \|\boldsymbol{x}_{k-1} - \boldsymbol{x}^*\| \leq r \cdot (r \cdot \|\boldsymbol{x}_{k-2} - \boldsymbol{x}^*\|) \leq \cdots \leq r^k \cdot \|\boldsymbol{x}_0 - \boldsymbol{x}^*\|.$$

(b) Let $\varepsilon > 0$. We are guaranteed to have an error bounded by ε if $r^N \cdot M < \varepsilon$, by Part (a). Taking logarithms of this inequality,

$$N \ln(r) + \ln(M) < \ln(\varepsilon).$$

Negating this, we get (where we use the identity $-\ln(x) = \ln(1/x)$)

$$-N \ln(r) - \ln(M) = N \ln(1/r) - \ln(M) > \ln(1/\varepsilon).$$

Dividing by $\ln(1/r)$ gives

$$N > \frac{1}{\ln(1/r)} \left(\ln(1/\varepsilon) + \ln(M)\right).$$

The important part of this bound is the dependence on ε.

(c) For quadratic convergence, the bound is derived in exactly the same way as for linear convergence. Specifically, one has

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\| \leq C\|\boldsymbol{x}_{k-1} - \boldsymbol{x}^*\|^2 \leq C(C\|\boldsymbol{x}_{k-2} - \boldsymbol{x}^*\|^2)^2 = C^3\|\boldsymbol{x}_{k-2} - \boldsymbol{x}^*\|^{2^2},$$

and so on, until one reaches the term involving $\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| = M$.

To determine the number of steps, we start with the bound

$$C^{2^N-1} \cdot M^{2^N} < \varepsilon \Rightarrow (CM)^{2^N} < C\varepsilon.$$

Taking logarithms and negating,

$$2^N \ln(1/CM) > \ln(1/C\varepsilon)$$

Taking logarithms again and rearranging for N, we get

$$N > \frac{1}{\ln(2)} \left(\ln\ln(1/C\varepsilon) - \ln\ln(1/MC)\right). \qquad (*)$$

Summarising, we know that if (*) is satisfied, then the error will be bounded by ε. We can reformulate the expression (*) a bit as follows

$$N > \frac{1}{\ln(2)} \left(\ln(\ln(1/\varepsilon) - \ln(C)) - \ln\ln(1/MC)\right).$$

With the above problems, the important bit is the dependence on ε. In double precision arithmetic, the *machine epsilon*, the smallest number that can be represented on a computer, is 2^{-53}. In words, from a computer's point of view, anything smaller than $\varepsilon = 2^{-53}$ is equal to zero. Since

$$\ln\ln(1/\varepsilon) = \ln\ln(2^{53}) \approx 3.6,$$

this shows that very few iterations of a quadratically convergent algorithm are needed to solve a problem "excatly", which means to the amount of precision allowed by the computer.

Solution (2.2) We first compute the derivatives,

$$\begin{aligned} f(x) &= \sqrt{x^2+1} \\ f'(x) &= \frac{x}{\sqrt{x^2+1}} \\ f''(x) &= \frac{1}{(x^2+1)^{3/2}}. \end{aligned}$$

Note that the second derivative is always positive. Newton's method then has the following form

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} = x_k - x_k(1+x_k^2) = -x_k^3.$$

For $|x_0| < 1$ this clearly converges to 0, while for $|x_0| > 1$ this diverges. For $|x_0| = 1$ the sequence alternates between 1 and -1.