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,

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

Negating this, we get

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

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

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

where we used the inequality $\ln(1/r) < 1/r - 1 = (1-r)/r$.

(c) For quadratic convergence, the bound is derived in exactly the same way as for linear convergence. To determine the number of steps, we start with the bound

$$C^N \cdot M^{2^N} \leq \varepsilon.$$

Taking logarithms and negating,

$$N \ln(1/C) - 2^N \ln(M) \geq \ln(1/\varepsilon).$$

Taking logarithms again, we get

$$N \cdot \left(\frac{\log_2(N)}{N} + \log_2(\ln(1/c) - \ln(M)) \right) \geq \log_2(\ln(1/\varepsilon)),$$

so that if

$$N > C' \cdot \ln\ln(1/\varepsilon)$$

for a constant C', we are guaranteed an error below ε.

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.

Solution (2.3) We first have to think about what it means to be a steepest descent direction with respect to a norm. If we look for a vector $\boldsymbol{p}$ with $\|\boldsymbol{p}\|_\infty = 1$ such that $\langle \boldsymbol{p}, \nabla f(\boldsymbol{x})\rangle$ is minimal. Set $\boldsymbol{v} = \nabla f(\boldsymbol{x})$, to ease notation. Since $\|\boldsymbol{p}\|_\infty = 1$ is the same as saying that $\max_{1\leq i\leq n} |p_i| = 1$, this amounts to solving the minimization problem

$$\begin{array}{ll} \text{minimize} & \langle \boldsymbol{p}, \boldsymbol{v}\rangle \\ \text{subject to} & -1 \leq p_i \leq 1, \quad 1 \leq i \leq n. \end{array}$$

Suppose that a minimizer is found and the minimum has the form

$$\sum_{i=1}^{d} p_i v_i.$$

If $p_i v_i > 0$, then we can decrease the objective function further by changing the sign of p_i, and then even further by setting $p_i = -1$ if $\operatorname{sign} v_i = 1$ and $p_i = 1$ otherwise. Therefore, the optimizer has the form

$$p_i = -\operatorname{sign} \nabla f(\boldsymbol{x})_i$$

for $1 \leq i \leq n$.