Chapter 8

Nonlinear algebraic equations

8.1 Nonlinear scalar equation

Fixed point iteration

We seek the solution $x \in I = [a, b] \subset \mathbb{R}$ of the equation

$$f(x) = 0, \tag{8.1}$$

with $f : I \to \mathbb{R}$ a nonlinear function, for which we can formulate the following fixed point iteration:

$$x^{(k+1)} = g(x^{(k)}) = x^{(k)} + \alpha f(x^{(k)}) \tag{8.2}$$

The fixed point iteration (8.2) converges to a unique solution $x = g(x)$, corresponding to $f(x) = 0$, if the function $g : I \to I$ is a *contraction mapping*, meaning that there exits a constant $L_g < 1$, such that

$$|g(x) - g(y)| \leq L_g |x - y|, \tag{8.3}$$

for all $x, y \in I$, where L_g is the *Lipschitz constant* of $g(x)$, a *Lipschitz continuous* function for $x \in I$.

Convergence of the fixed point iteration (8.2) is proven similar to the case of a linear system of equations (7.12). For any $k > 1$, we have that

$$|x^{(k+1)} - x^{(k)}| = |g(x^{(k)}) - g(x^{(k-1)})| \leq L_g |x^{(k)} - x^{(k-1)}| \leq L_g^k |x^{(1)} - x^{(0)}|,$$

and for $m > n$,

$$\begin{aligned} |x^{(m)} - x^{(n)}| &= |x^{(m)} - x^{(m-1)}| + ... + |x^{(n+1)} - x^{(n)}| \\ &\leq (L_g^{m-1} + ... + L_g^n)\, |x^{(1)} - x^{(0)}|, \end{aligned}$$

so that for $L_g < 1$, we have that

$$\lim_{n\to\infty} |x^{(m)} - x^{(n)}| = 0, \tag{8.4}$$

which implies that there exists an $x \in \mathbb{R}$ such that

$$x = \lim_{n\to\infty} x^{(n)}, \tag{8.5}$$

since $\{x^{(n)}\}_{n=1}^{\infty}$ is a Cauchy sequence in the complete space $\mathbb{R}$.

Uniqueness follows from assuming that there exists another solution $y \in \mathbb{R}$ such that $y = g(y)$, which leads to a contradiction, since

$$|x - y| = |g(x) - g(y)| \leq L_g|x - y| < |x - y|. \tag{8.6}$$

Thus x is the unique solution to the equation $x = g(x)$.

Rate of convergence

We are not only interested in *if* an iterative method converges, but also *how fast*, that is the *rate of convergence*. We say that a sequence of approximate solutions $\{x^{(k)}\}_{k=1}^{\infty}$ converges with *order* p to the exact solution x, if

$$\lim_{k\to\infty} \frac{|x - x^{k+1}|}{|x - x^{k+1}|^p} = C, \quad C > 0, \tag{8.7}$$

where $p = 1$ corresponds to a linear order of convergence, and $p = 2$ a quadratic order of convergence. We can approximate the rate of convergence by extrapolation, so that

$$p \approx \frac{\log \dfrac{|x^{k+1} - x^k|}{|x^k - x^{k-1}|}}{\log \dfrac{|x^k - x^{k-1}|}{|x^{k-1} - x^{k-2}|}} \tag{8.8}$$

which is useful in practice when the exact solution is not available.

Newton's method

The analysis above suggests that (8.2) converges linearly with the Lipschitz constant L_g, since

$$|x - x^{(k+1)}| = |g(x) - g(x^{(k)})| < L_g|x - x^{(k)}| \tag{8.9}$$

Algorithm 11: Newton's method

```
Given initial approximation x^(0)
while |f(x^(k))| >= TOL do
  | x^(k+1) = x^(k) - f'(x^(k))^(-1) f(x^(k))
end
```

so that the error $e^{(k)} = x - x^{(k)}$ decreases linearly as $|e^{(k+1)}| < L_g|e^{(k)}|$.

Although, by the choice $\alpha = -f'(x^{(k)})^{-1}$, the fixed point iteration (8.2) exhibits quadratic converge for $x^{(k)}$ close to x, the exact solution to the equation $f(x) = 0$. This is Newton's method.

The quadratic convergence of Newton's method follows from *Taylor's formula*, with $f(x)$ expanded around $x^{(k)}$, with $\xi \in I$,

$$0 = f(x) = f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + \frac{1}{2}f''(\xi)(x - x^{(k)})^2. \tag{8.10}$$

We divide by $f'(x^{(k)})$ to get

$$x - (x^{(k)} - f'(x^{(k)})^{-1}f(x^{(k)})) = -\frac{1}{2}f'(x^{(k)})^{-1}f''(\xi)(x - x^{(k)})^2, \tag{8.11}$$

so that for the error $e^{(k)} = x - x^{(k)}$,

$$|e^{(k+1)}| = \frac{1}{2}|f'(x^{(k)})^{-1}f''(\xi)|\,|e^{(k)}|^2, \tag{8.12}$$

which displays the quadratic convergence for $x^{(k)}$ close to x.

8.2 Systems of nonlinear equations

Fixed point iteration for nonlinear systems

Now consider systems of nonlinear equations: find $x = (x_1, ..., x_n)^T \in \mathbb{R}^n$, such that

$$f(x) = 0, \tag{8.13}$$

with $f = (f_1, ..., f_n)^T : \mathbb{R}^n \to \mathbb{R}^n$. We can formulate a fixed point iteration $x^{(k+1)} = g(x^{(k)})$, with $g : \mathbb{R}^n \to \mathbb{R}^n$, just as in the case of the scalar problem.

Existence of a unique solution to the algorithm follows by the Banach fixed point theorem.

Theorem 13 (Banach fixed point theorem)**.** *The fixed point iteration of Aglorithm 12 converges to a unique solution if $\|1 + \alpha L_f\| < 1$, with L_f the Lipschitz constant of the function $f(x)$.*

Algorithm 12: Newton's method for systems of nonlinear equations

Given initial approximation $x^{(0)}$
while $\|f(x^{(k)})\| \geq TOL$ **do**
  $x^{(k+1)} = x^{(k)} + \alpha f(x^{(k)})$
end

Proof. For $k > 1$ we have that

$$\begin{aligned} \|x^{(k+1)} - x^{(k)}\| &= \|x^{(k)} - x^{(k-1)} + \alpha(f(x^{(k)}) - f(x^{(k-1)}))\| \\ &\leq (1+\alpha L_f)\|x^{(k)} - x^{(k-1)}\| \leq (1+\alpha L_f)^k \|x^{(1)} - x^{(0)}\|, \end{aligned}$$

and for $m > n$,

$$\begin{aligned} \|x^{(m)} - x^{(n)}\| &= \|x^{(m)} - x^{(m-1)}\| + ... + \|x^{(n+1)} - x^{(n)}\| \\ &\leq ((1+\alpha L_f)^{m-1} + ... + (1+\alpha L_f)^n)\|x^{(1)} - x^{(0)}\|. \end{aligned}$$

For $(1+\alpha L_f) < 1$, $\{x^{(n)}\}_{n=1}^{\infty}$ is a Cauchy sequence since $\|x^{(m)} - x^{(n)}\| \to 0$, for $n \to \infty$, which implies that there exists an $x \in \mathbb{R}^n$ such that

$$x = \lim_{n\to\infty} x^{(n)}, \tag{8.14}$$

since $\mathbb{R}^n$ is a complete vector space. Uniqueness follows from assuming that there exists another solution $y \in \mathbb{R}^n$ such that $f(y) = 0$, which leads to a contradiction, since

$$\|x-y\| = \|x-y+\alpha(f(x)-f(y))\| \leq (1+\alpha L_f)\|x-y\| < \|x-y\|. \tag{8.15}$$

Thus x is the unique solution to the equation $f(x) = 0$. □

Newton's method for nonlinear systems

The *Jacobian matrix* $f'(x)$ is defined as

$$f'(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}, \tag{8.16}$$

which we use to formulate Newton's method for systems of equations.

The quadratic convergence of Newton's method follows from *Taylor's formula* in $\mathbb{R}^n$, with $f(x)$ expanded around $x^{(k)}$,

$$0 = f(x) = f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + \frac{1}{2}(x - x^{(k)})^T f''(\xi)(x - x^{(k)}),$$

Algorithm 13: Newton's method for systems of nonlinear equations

Given initial approximation $x^{(0)}$
while $\|f(x^{(k)})\| \geq TOL$ **do**
  Solve $f'(x^{(k)})\Delta x^{(k+1)} = -f(x^{(k)})$ ▷ `solve for` $\Delta x^{(k+1)}$
  $x^{(k+1)} = x^{(k)} + \Delta x^{(k+1)}$ ▷ `update by` $\Delta x^{(k+1)}$
end

where $f''(x^{(k)}$ is the *Hessian*. We have that

$$x - (x^{(k)} - f'(x^{(k)})^{-1} f(x^{(k)})) = -\frac{1}{2} f'(x^{(k)})^{-1} (x - x^{(k)})^T f''(\xi)(x - x^{(k)}),$$

with $f'(x^{(k)})^{-1}$ the inverse of the Jacobian, so that,

$$\|e^{(k+1)}\| = \frac{1}{2} \|f'(x^{(k)})^{-1} f''(\xi)\| \, \|e^{(k)}\|^2, \tag{8.17}$$

for the error $e^{(k)} = x - x^{(k)}$, which shows quadratic convergence for $x^{(k)}$ close to x.