CONVERGENCE THEORY

* FIXED-POINT ITERATIONS $x_{n+1} = g(x_n)$

Using the true solution R, we write an expression for the error in the n-th iteration

$$R - x_{n+1} = R - g(x_n) = g(R) - g(x_n)$$

Multiplying and dividing by $(R - x_n)$

$$R - x_{n+1} = \frac{g(R) - g(x_n)}{R - x_n} (R - x_n)$$

Using the mean-value theorem

$$R - x_{n+1} = g'(\xi_n) \cdot (R - x_n), \text{ where } \xi_n \in [x_n, R]$$

Thus, denoting $e_n = R - x_n$, we have

$$|e_{n+1}| = |g'(\xi_n)| \cdot |e_n|$$

Remarks

* if $|g'(\xi_n)| < 1$ on some interval, the fixed-point iterations will converge for initial values in that interval (this explains why certain rearrangements of $f(x) = 0$ may result in divergent iterations)

* ~~the~~ the error at a given iteration is a fraction of the error at the previous iteration $\Rightarrow$ ~~linear~~ linear convergence

* to ensure convergence, the function $g(x)$ must be contractive, i.e., Lipschitz-continuous with the constant $L < 1$

* Newton's METHOD

Newton's method uses iterations similar to the fixed-point approach

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = g(x_n)$$

Iterates converge if $|g'(x)| < 1$

Thus

$$|g'(x)| = \left| \frac{f(x) \cdot f''(x)}{[f'(x)]^2} \right| < 1$$ // condition for convergence

Error relation: $R - x_{n+1} = g(R) - g(x_n)$

Taylor-expand $g(x_n)$ about $x = R$ up to the second order

$$g(x_n) = g(R) + g'(R)(x_n - R) + \frac{g''(\xi)}{2}(x_n - R)^2, \quad \xi \in [x_n, R]$$

$$f(R) = 0 \Rightarrow g'(R) = \frac{f(R) \cdot f''(R)}{[f'(R)]^2} = 0$$

$$\to g(x_n) = g(R) + \frac{g''(\xi)}{2}(x_n - R)^2$$

Putting this into the error relation

$$|e_{n+1}| = |g(R) - g(x_n)| = \left|\frac{g''(\xi)}{2}\right| |e_n|^2$$

Remarks

* smallness of the second derivative required for convergence
* the error at a given iteration is the square of the error at the previous iteration — quadratic convergence

SECANT AND FALSE POSITION METHODS

In both cases the consecutive iterates can be written as

$$x_{n+1} = x_n - \frac{f(x_n)}{f(x_n) - f(x_{n-1})} (x_n - x_{n-1}) = \underline{g(x_n, x_{n-1})}$$

Following a similar approach as above, but with more tedious details, we obtain

$$e_{n+1} = \frac{g(\xi_1, \xi_2)}{2} e_n e_{n-1}, \quad \xi_1, \xi_2 \in [x_n, R]$$

Error is proportional to the two previous errors.

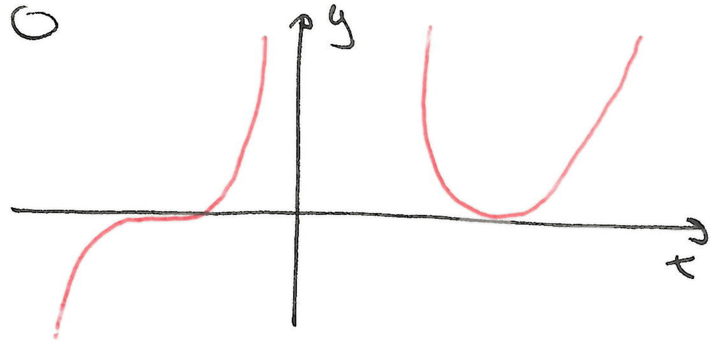
Remark

* It can be shown that the rate of convergence is faster than linear, but slower than quadratic
$\Rightarrow$ <u>Superlinear convergence</u>

General Remarks

* Higher-order methods (Newton's method) are faster, but come with more stringent requirements on the function $f(x)$ (boundedness of higher derivatives)
* One can combine different methods:
  - start with a slower, but more robust method
  - switch to a faster method close to the root.

* problems occur when multiple roots are present, i.e., when $f(R) = f'(R) = \ldots = 0$



In Newton's method

$$g'(x) = \frac{f(x)\cdot f''(x)}{[f'(x)]^2} \sim \frac{0}{0}$$

doesn't vanish and second-order convergence is lost.

1.3 Numerical Linear Algebra – A Short Review

(MATH 2T03, Grasselli & Pelinovsky Sections 2.3–2.6)

We are interested in the solution of problems of the type

$$Ax = b, \quad x, b \in \mathbb{R}^N, \; A \in \mathbb{R}^{N\times N}$$

N-large

General remarks about existence of solutions:

* if $\mathrm{Det}(A) \neq 0$ a unique solution x exists
  if $b = 0$, $x = 0$

* if $\mathrm{Det}(A) = 0$, let $x = x_0 + x_1$
  - $Ax_0 = 0$ has m solutions defined up to multiplicative constants, m – the number of distinct eigenvalues
  - $Ax_1 = b$ – additional solutions exist if $b \perp N(A^T)$ – Fredholm alternative

Cramer's Rule – sol'n involving determinants (elegant, but useless for large N)

* Elimination Methods

Start with the system below; the goal is to convert it to an upper-triangular form

$$
\begin{array}{rl}
4x_1 - 2x_2 + x_3 = 15 & / \cdot 3 \quad / \cdot (-1) \\
-3x_1 - x_2 + 4x_3 = 8 & / \cdot 4 \rightarrow \oplus \\
x_1 - x_2 + 3x_3 = 13 & / \cdot 4 \rightarrow \oplus
\end{array}
\Rightarrow
$$

$$
\Rightarrow
\begin{array}{rl}
4x_1 - 2x_2 - x_3 = 15 & \\
-10x_2 + 19x_3 = 77 & / \cdot 2 \\
-2x_2 + 11x_3 = 37 & / \cdot (-10) \rightarrow \oplus
\end{array}
\Rightarrow
\begin{cases}
4x_1 - 2x_2 + x_3 = 15 \\
-10x_2 + 19x_3 = 77 \\
-72x_3 = -216
\end{cases}
$$

Given the triangular form, the solution can be found trivially via back-substitution: $x_3 = 3, x_2 = -2, x_1 = 2$

The same goal can be achieved by performing elementary row operations on the corresponding system matrix:

- multiplication of a row by a constant
- adding a multiple of a row to another row
- interchanging the orders of two rows

Problem

- the magnitude of the coefficients grows; this may become an issue for large systems (round-off errors)