Methods in Computational Science – Iterative methods for nonlinear equations (ch.8)

Johan Hoffman

Nonlinear equations $f(x) = 0$

- Cannot be expressed as matrix equations
- To construct an iterative method we need to know how the function f(x) behaves near a solution to the equation f(x)=0.
- Is it continuous?
- Is it differentiable?

Continuous functions

A continuous function can be roughly characterized as a function for which a small change in input results in a small change in output. More formally, a function $f : R \to R$ is said to be *continuous* at $x \in R$, if

$$\lim_{h \to 0} f(x+h) = f(x),$$

uniformly continuous on the interval $I = [a, b]$, if for each $\epsilon > 0$ we can find a $\delta > 0$ such that

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon, \quad \forall x, y \in I,$$

and *Hölder continuous* on I, if

$$|f(x) - f(y)| \leq C|x - y|^{\alpha}, \quad \forall x, y \in I,$$

where C and α are real numbers such that $C > 0$ and $0 < \alpha \leq 1$. In the special case of $\alpha = 1$, we say that the function is *Lipschitz continuous*, with *Lipschitz constant* $L_f = C$.

Vector spaces of continuous functions

We denote by $C(I)$ the vector space of real valued continuous functions on the interval I, which is closed under the vector space operations of pointwise addition and scalar multiplication, that is, for each $x \in I$,

$$(f+g)(x) = f(x) + g(x),$$
$$(\alpha f)(x) = \alpha f(x),$$

for all $f, g \in C(I)$ and $\alpha \in R$. Similarly, we let $Lip(I)$ denote the vector space of Lipschitz continuous functions, with the same vector space operations. It follows from the definitions that any Lipschitz continuous function is also continuous, and hence that $Lip(I) \subset C(I)$.

Vector spaces of differentiable functions

Let $C^k(I)$ denote the vector space of continuous functions with also continuous derivatives up to the order k, where $C^0(I) = C(I)$ and $C^\infty(I)$ is the vector space of continuous functions with continuous derivatives of arbitrary order, also referred to as *smooth functions*. A natural norm on the vector spaces $C^k(I)$ is

$$\|f\|_\infty = \max_{0\leq\alpha\leq k} \sup_{x\in I} |D^\alpha f(x)|,$$

where

$$D^\alpha f = \frac{d^\alpha f}{dx^\alpha},$$

and on $Lip(I)$,

$$\|f\|_{Lip} = \sup_{x\in I} |f(x)| + \sup_{x,y\in I} \frac{|f(x) - f(y)|}{|x-y|}.$$

Equipped with these norms, $C^k(I)$ and $Lip(I)$ are Banach spaces.

Vector spaces of differentiable functions

Example 8.1. The elementary functions $\sin(x)$, $\cos(x)$ and $\exp(x)$ are all smooth functions.

Example 8.2. The Heaviside step function

$$H(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0, \end{cases}$$

is continuous away from $x = 0$, hence, an element e.g. of $C([1,2])$ but not of $C([-1,1])$.

Example 8.3. The absolute value is a function in $C([-1,1])$ but not in $C^1([-1,1])$,

$$|x| = \begin{cases} -x, & x \leq 0, \\ x, & x > 0. \end{cases}$$

Mean value theorem

Theorem 8.4 (Mean value theorem). *If $f \in C^1([a,b])$ there exists an $x \in (a,b)$, such that*

$$f'(x) = \frac{f(b) - f(a)}{b - a}.$$

Example 8.5 (Estimation of Lipschitz constant). Any function $f \in C^1([a,b])$ is Lipschitz continuous, and we can estimate its Lipschitz constant L_f by the mean value theorem, since

$$|f(b) - f(a)| = |f'(x)(b - a)| \leq \|f'\|_\infty |b - a|$$

which gives the Lipschitz constant $L_f = \|f'\|_\infty$.

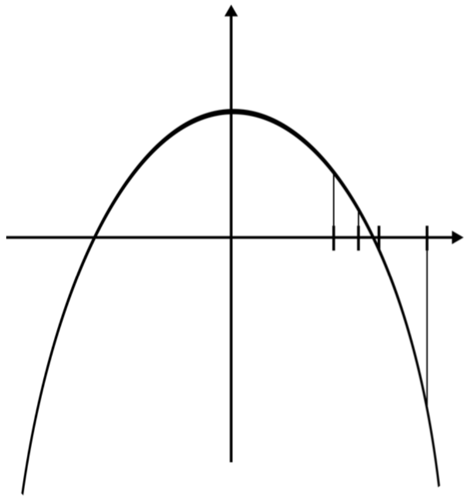
Intermediate value theorem

Theorem 8.6 (Intermediate value theorem). *If $f \in C([a,b])$, then for every $y \in [f(a), f(b)]$ there exits an $x \in (a,b)$ such that $f(x) = y$.*

Example 8.7 (Bisection method). The intermediate value theorem suggests a simple bisection method for solving the equation $f(x) = 0$ on an interval $[a,b]$, in the special case of $y = 0$.

ALGORITHM 8.1. x = bisection(f, a, b).
Input: a function **f**, and an interval **[a,b]**.
Output: an approximate root **x**.

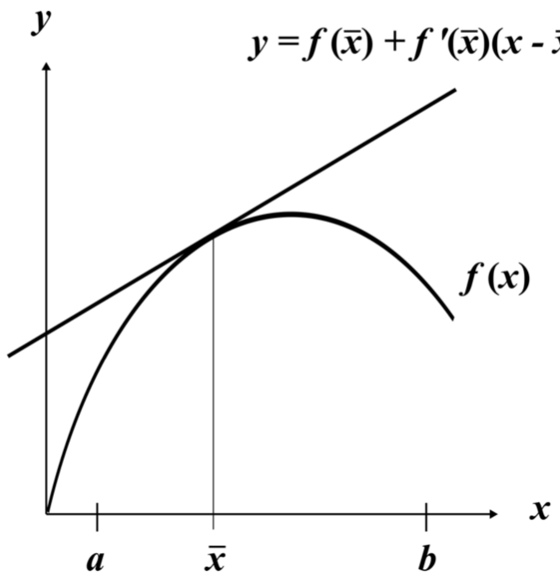
```
while (b-a)/2 > TOL do
    x = (a+b)/2
    if f(a)*f(x) > 0 then
        a = x
    else
        b = x
    end if
end while
return x
```



Taylor’s theorem

Theorem 8.8 (Tayor’s theorem for linear approximation). *For* $f \in C^2(I)$ *and* $y \in I$,

$$f(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{1}{2} f''(\xi)(x - \bar{x})^2, \quad \xi \in [x, \bar{x}].$$



Fixed point iteration

For a nonlinear function $f : I \rightarrow R$, we seek a solution $x \in R$ to the equation

$$f(x) = 0,$$

for which we formulate a fixed point iteration

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

where $\alpha \in R$ is a parameter of the method to be chosen.

Fixed point iteration

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

so that for all $m > n$,

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

by the triangle inequality. If $L_g < 1$, then $\{x^{(n)}\}_{n=1}^{\infty}$ is a Cauchy sequence,

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

which implies that there exists an $x \in R$, such that

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

because R is a Banach space.

Fixed point iteration

Uniqueness of *x* follows from assuming that there exists another solution *y* such that *y=g(y)*. But then

$$|x - y| = |g(x) - g(y)| \leq L_g|x - y| \Leftrightarrow (1 - L_g)|x - y| \leq 0 \Rightarrow |x - y| = 0,$$

and hence $x = y$ is the unique solution to the equation $x = g(x)$.

Fixed point iteration

Example 8.9. Consider the nonlinear equation

$$\cos(x) = 0$$

for which we seek a root in the interval $[\pi/4, 3\pi/4]$. The fixed point iteration takes the form

$$x^{(k+1)} = x^{(k)} + \alpha\cos(x^{(k)}),$$

with the fixed point function $g(x) = x + \alpha\cos(x)$. By the mean value theorem, with $\alpha = 1$,

$$L_g = \|g'\|_\infty = \|1 - \sin(x)\|_\infty < 1.$$

Hence, the fixed point iteration converges to the unique solution to the equation.

Newton’s method

The analysis suggests a linear order of convergence of the fixed point iteration

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

so that $|e^{(k+1)}| \leq L_g |e^{(k)}|$ for the error $e^{(k)} = x - x^{(k)}$. But if we choose $\alpha = -f'(x^{(k)})^{-1}$, which corresponds to *Newton’s method*, the fixed point iteration exhibits a quadratic order of convergence. A geometric interpretation is that $x^{(k+1)}$ is determined from the tangent line of the function $f(x)$ at $x^{(k)}$, and the quadratic order of convergence follows from Taylor’s theorem.

Newton's method

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

with $\xi \in [x, x^{(k)}]$. Divide by $f'(x^{(k)})$ to get

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

so that by the fixed point update formula $x^{(k+1)} = x^{(k)} - f'(x^{(k)})^{-1} f(x^{(k)})$,

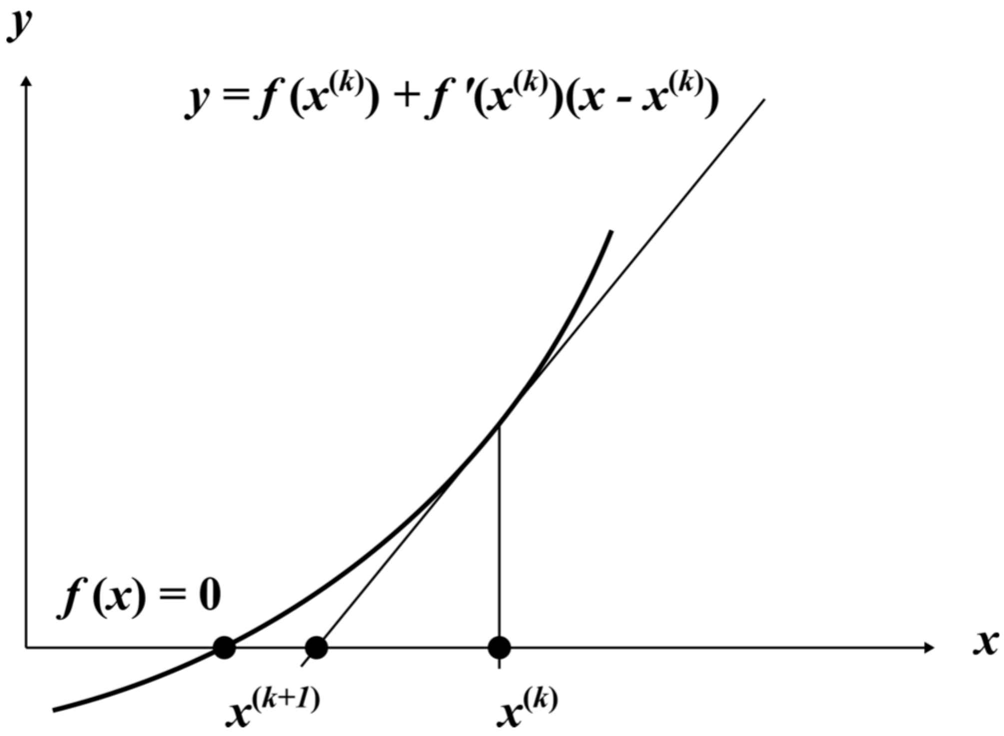
$$|e^{(k+1)}| = \frac{1}{2} |f'(x^{(k)})^{-1} f''(\xi)| \, |e^{(k)}|^2,$$

which displays the quadratic convergence of the sequence $x^{(k)}$ close to x, assuming $f \in C^2(I)$.

Newton's method

ALGORITHM 8.2. x = newtons_method(f, x0).
Input: a function **f**, an initial guess **x0**.
Output: approximate root **x**.

```
x = x0
while |f(x)| > TOL do
    df = derivative(f, x)
    x = x - f(x)/df
end while
return x
```



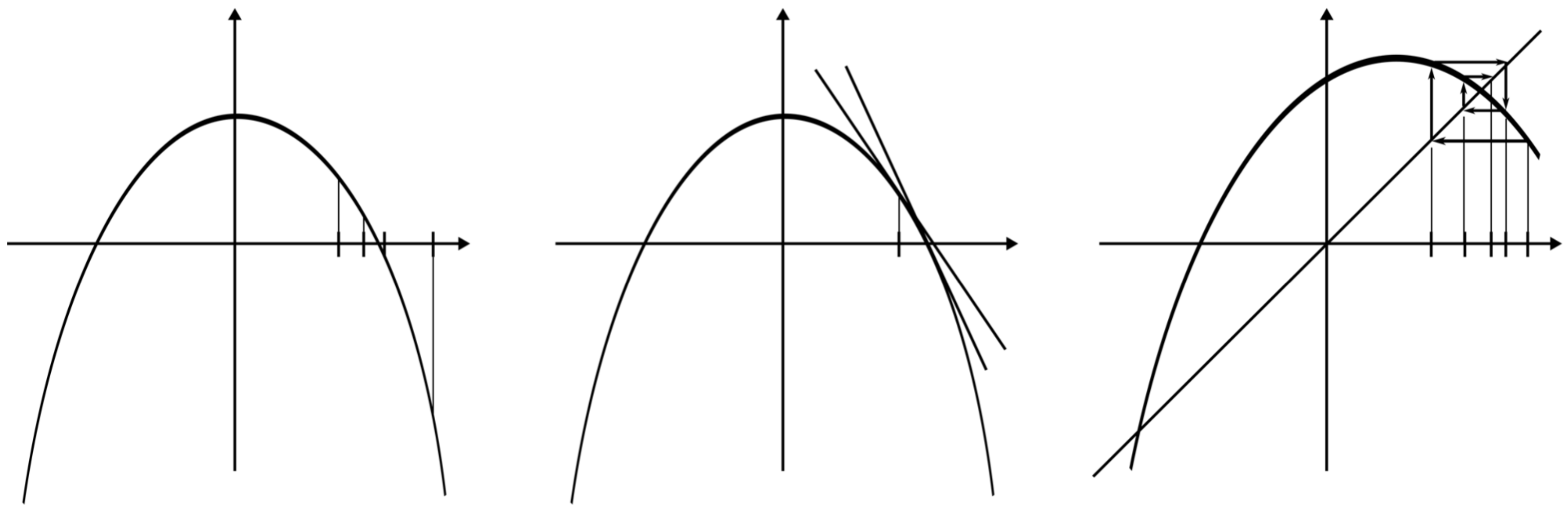
Newton's method

Example 8.10. Newton's method for the nonlinear equation

$$\cos(x) = 0$$

takes the form $x^{(k+1)} = x^{(k)} + \cos(x^{(k)}) / \sin(x^{(k)})$.

Bisection, Newton, fixed point iteration

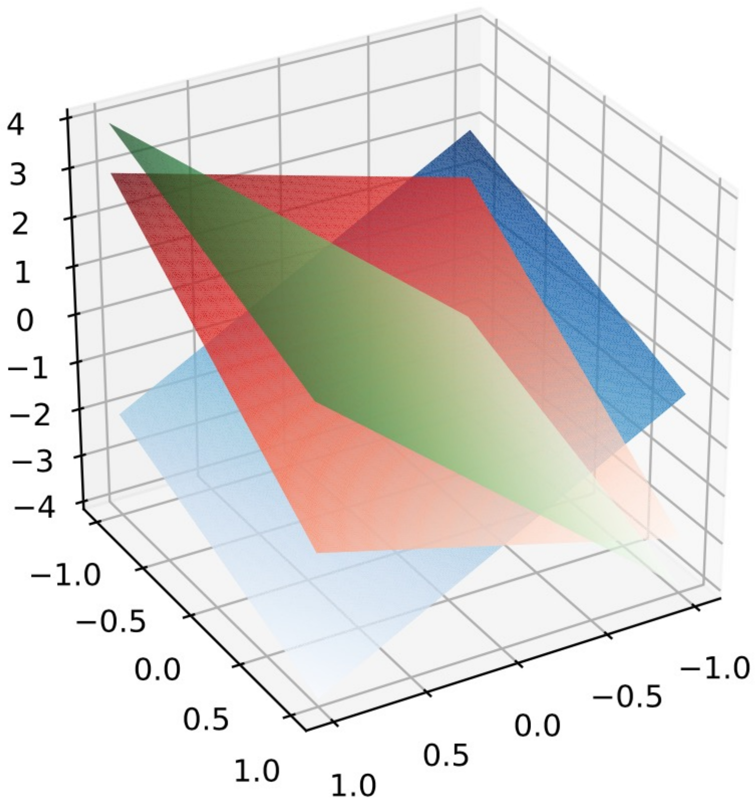


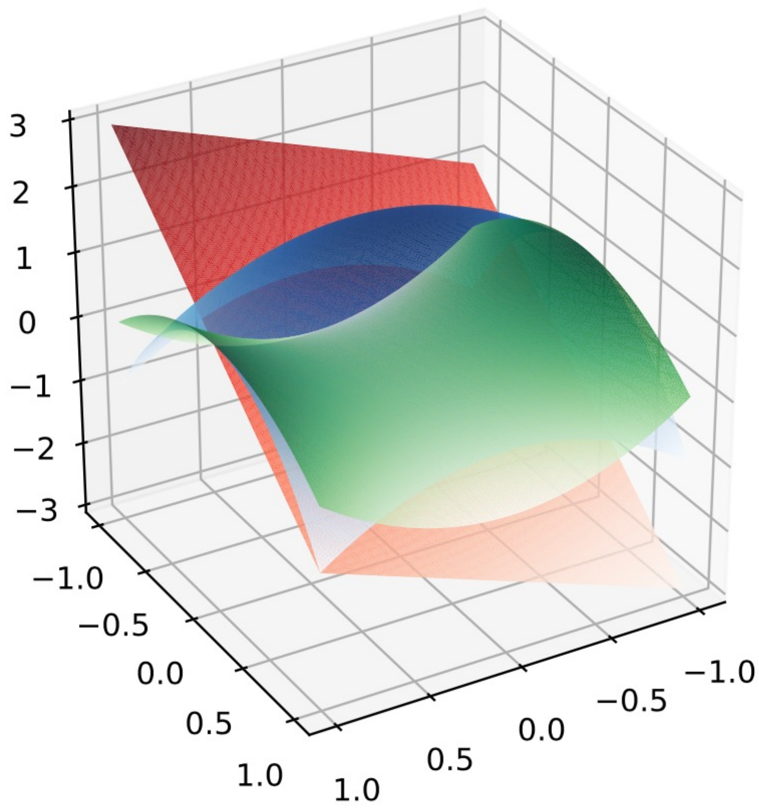
Systems of nonlinear equations

$$\left\{\begin{array}{rcr} x_1 - 2x_2 - x_3 & = & 0, \\ 3x_1 - x_2 - x_3 & = & 0, \\ 2x_1 + x_2 + x_3 & = & -1. \end{array}\right.$$

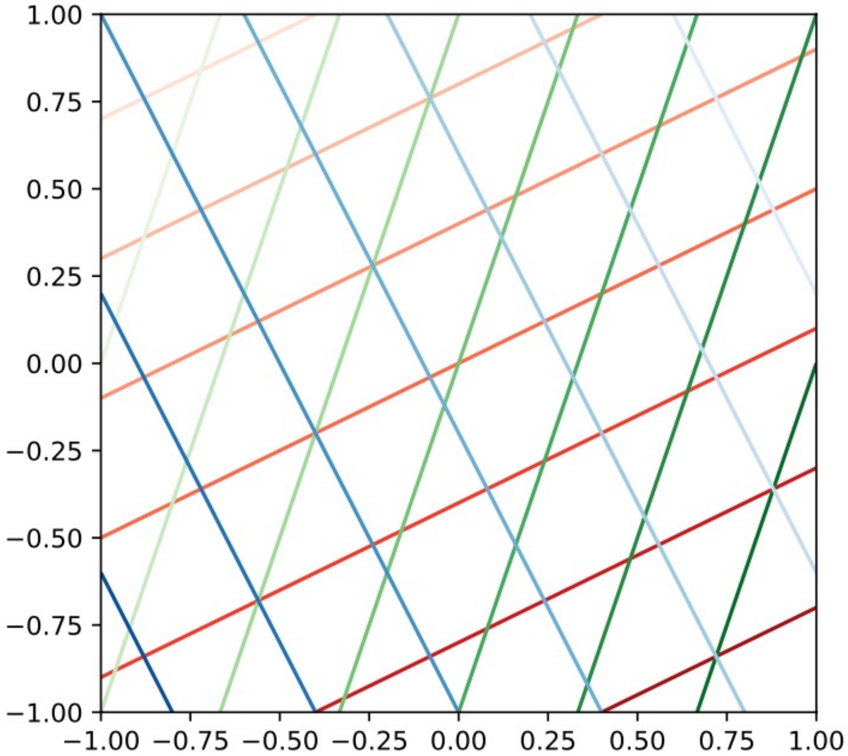
$$\left\{\begin{array}{ccr} x_1 - 2x_2 - x_3 & = & 0, \\ x_1^2 - x_2^2 - x_3 & = & 0, \\ -x_1^2 - 2x_2^2 - x_3 & = & -2. \end{array}\right.$$

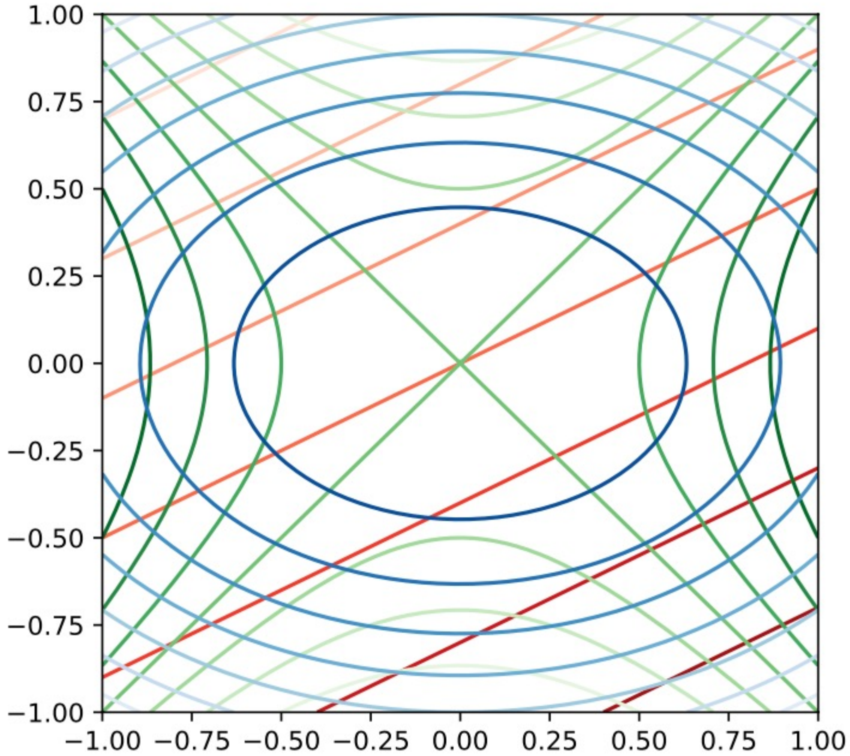
Systems of nonlinear equations





Systems of nonlinear equations





Continuous functions in R^n

We say that a vector function $f : R^n \to R^m$ is continuous at $x \in R^n$, if for all $\epsilon > 0$ there is a $\delta > 0$, such that for $y \in R^n$,

$$\|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \epsilon.$$

If the function is continuous at all $x \in R^n$ then $f \in C(R^n; R^m)$, with $C(R^n; R^m)$ the vector space of continuous functions $f : R^n \to R^m$, where vector addition and scalar multiplication are defined pointwise for each $x \in R^n$. A function is Lipschitz continuous, denoted $f \in Lip(R^n; R^m)$, if there exists a real number $L_f > 0$, such that

$$\|f(x) - f(y)\| \leq L_f \|x - y\|, \quad \forall x, y \in R^n.$$

Differentiation in R^n

We define the partial derivative of a function $f : R^n \to R^m$ as

$$\frac{\partial f_i}{\partial x_j}(x) = \lim_{h \to 0} \frac{f_i(x_1, ..., x_j + h, ..., x_n) - f_i(x_1, ..., x_j, ..., x_n)}{h},$$

for i the index of the component of the function $f = (f_1, ..., f_m)$, and j the index of the coordinate $x = (x_1, ..., x_n)$. The *Jacobian matrix* $f' \in R^{m \times n}$ at $x \in R^n$, is defined as

$$f'(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix} = \begin{bmatrix} (\nabla f_1)^T(x) \\ \vdots \\ (\nabla f_m)^T(x) \end{bmatrix},$$

with the *gradient*

$$\nabla f_i(x) = \left(\frac{\partial f_i}{\partial x_1}(x), ..., \frac{\partial f_i}{\partial x_n}(x) \right)^T.$$

Vector spaces

The vector space of continuous functions $f : R^n \to R^m$ with also continuous partial derivatives up to the order k is denoted by $C^k(R^n; R^m)$, with $C^0(R^n; R^m) = C(R^n; R^m)$ and $C^\infty(R^n; R^m)$ the vector space of continuous functions with continuous derivatives of arbitrary order. If $m = n$, we write simply $C^k(R^n)$, and similar for the other vector spaces.

Taylor's theorem extends to functions $f \in C^2(R^n; R^m)$, for which an affine (or linear) approximation near $\bar{x} \in R^n$ takes the form

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}),$$

with an approximation error of the order of $\|x - \bar{x}\|^2$.

Linear approximation

Example 8.11. The affine function

$$f(x) = (x_1 - 2x_2 - x_3, 3x_1 - x_2 - x_3, 2x_1 + x_2 + x_3 + 1)^T$$

can be expressed as $f(x) = Ax + b$, with

$$A = \begin{bmatrix} 1 & -2 & -1 \\ 3 & -1 & -1 \\ 2 & 1 & 1 \end{bmatrix},$$

and $b = (0, 0, 1)^T$. The function is an element of the vector space $C^\infty(R^3)$, with the Jacobian matrix $f'(x) = A$.

Linear approximation

Example 8.12. The nonlinear function

$$f(x) = (x_1 - 2x_2 - x_3, x_1^2 - x_2^2 - x_3, -x_1^2 - 2x_2^2 - x_3 + 2)^T$$

is an element of the vector space $C^\infty(R^3)$, with the Jacobian matrix

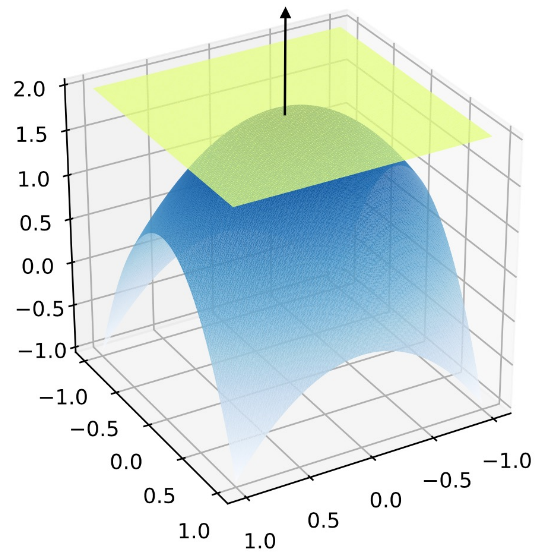
$$f'(x) = \begin{bmatrix} 1 & -2 & -1 \\ 2x_1 & -2x_2 & -1 \\ -2x_1 & -4x_2 & -1 \end{bmatrix}.$$

Hence, by Taylor's theorem, near the point $\bar{x} \in R^3$, we can use the approximation

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}).$$

Linear approximation – tangent hyperplane

If $m = 1$, then $f : R^n \to R$ is a scalar function, and the Jacobian matrix is reduced to the gradient of the function $f'(x) = \nabla f(x)$. Further, near a point $\bar{x} \in R^3$ the hypersurface defined by the equation $f(x) = 0$ is approximated by the *tangent hyperplane*, defined by the normal vectors $\alpha \nabla f(\bar{x})$, for $\alpha \neq 0$ a real number.



Fixed point iteration

Now consider a system of nonlinear equations: find $x \in R^n$ such that

$$f(x) = 0,$$

where $f \in Lip(R^n)$, for which we form the fixed point iteration

$$x^{(k+1)} = g(x^{(k)}) = x^{(k)} + Af(x^{(k)}),$$

with $g \in Lip(R^n)$ and $A \in R^{n \times n}$ a matrix which is chosen as part of the method. We refer to the function $f(x^{(k)})$ as the residual for the approximation $x^{(k)}$, and we use the norm of the residual as the stopping criterion for the fixed point iteration.

Fixed point iteration

ALGORITHM 8.3. x = fixed_point_iteration(f, x0, A).
Input: a function **f**, an initial guess **x0**, an nxn matrix **A**.
Output: an approximate root **x**.

```
x[:] = x0[:]
while norm(f(x)) > TOL do
    x[:] = x[:] + matrix_vector_product(A, f(x))
end while
return x
```

Banach fixed point theorem

Theorem 8.13 (Banach fixed point theorem). *If* $g \in Lip(R^n)$ *is Lipschitz continuous with Lipschitz constant* $L_g < 1$, *then the fixed point iteration* $x^{(k+1)} = g(x^{(k)})$ *converges to a unique solution to the equation* $x = g(x)$.

Banach fixed point theorem

Proof. For $k > 1$ we have that

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

and for $m > n$,

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

Because $L_g < 1$, $\{x^{(n)}\}_{n=1}^{\infty}$ is a Cauchy sequence, and since R^n is a Banach space there exists an $x \in R^n$ such that

$$\lim_{n\to\infty} \|x - x^{(n)}\| = 0.$$

Uniqueness follows from assuming that there exists another solution $y \in R^n$ such that $y = g(y)$, which then implies that

$$\|x - y\| = \|g(x) - g(y)\| \leq L_g \|x - y\| \Leftrightarrow (1 - L_g)\|x - y\| \leq 0,$$

and hence $x = y$ is the unique solution to the equation $x = g(x)$.

Newton's method

Newton's method to solve the equation $f(x) = 0$ in R^n is analogous to the case of the scalar equation, but where the inverse of the derivative is replaced by the inverse of the Jacobian matrix $(f'(x^{(k)}))^{-1}$. If the Jacobian is not available in analytical form we compute an approximation, for example, by a finite difference approximation based on the function $f(x)$.

For a large system the Jacobian may be too expensive to compute or to hold in memory, in which case we use a less expensive approximation. Methods based on approximations of the Jacobian are referred to as *quasi-Newton methods*. The inverse Jacobian matrix is typically not constructed explicitly, instead a system of linear equations is solved for the increment

$$\Delta x^{(k+1)} = x^{(k+1)} - x^{(k)},$$

with the system matrix $f'(x^{(k)})$ and vector $-f(x^{(k)})$.

Newton's method

ALGORITHM 8.4. x = newtons_method_system(f, x0).
Input: a function **f**, and an initial guess **x0**.
Output: an approximate root **x**.

```
x[:] = x0[:]
while norm(f(x)) > TOL do
    Df = jacobian(f, x)
    dx = solve_linear_system(Df, -f(x))
    x[:] = x[:] + dx[:]
end while
return x
```

Newton's method

For $x^{(k)}$ close to x the quadratic order of convergence for Newton's method follows from Taylor's formula in R^n, assuming that $f \in C^2(R^n)$,

$$f(x) = f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + E_f(x - x^{(k)}),$$

with $E_f(x - x^{(k)})$ a vector in R^n, for which

$$\|E_f(x - x^{(k)})\| = \mathcal{O}(\|x - x^{(k)}\|^2).$$

Since $f(x) = 0$, and assuming that the Jacobian matrix $f'(x^{(k)})$ is nonsingular,

$$x - (x^{(k)} - f'(x^{(k)})^{-1} f(x^{(k)})) = E_f(x - x^{(k)}),$$

and with $x^{(k+1)} = x^{(k)} - f'(x^{(k)})^{-1} f(x^{(k)})$,

$$\|e^{(k+1)}\| = \mathcal{O}(\|e^{(k)}\|^2),$$

for the error $e^{(k)} = x - x^{(k)}$.

Recurrence relations

A *recurrence relation* of order k defines each element $x_n \in X$ of a sequence through the iteration

$$x_n = f(x_{n-1}, ..., x_{n-k}),$$

for a function $f : X \to X$ and some set X, together with k initial values $x_0, ..., x_{k-1}$.

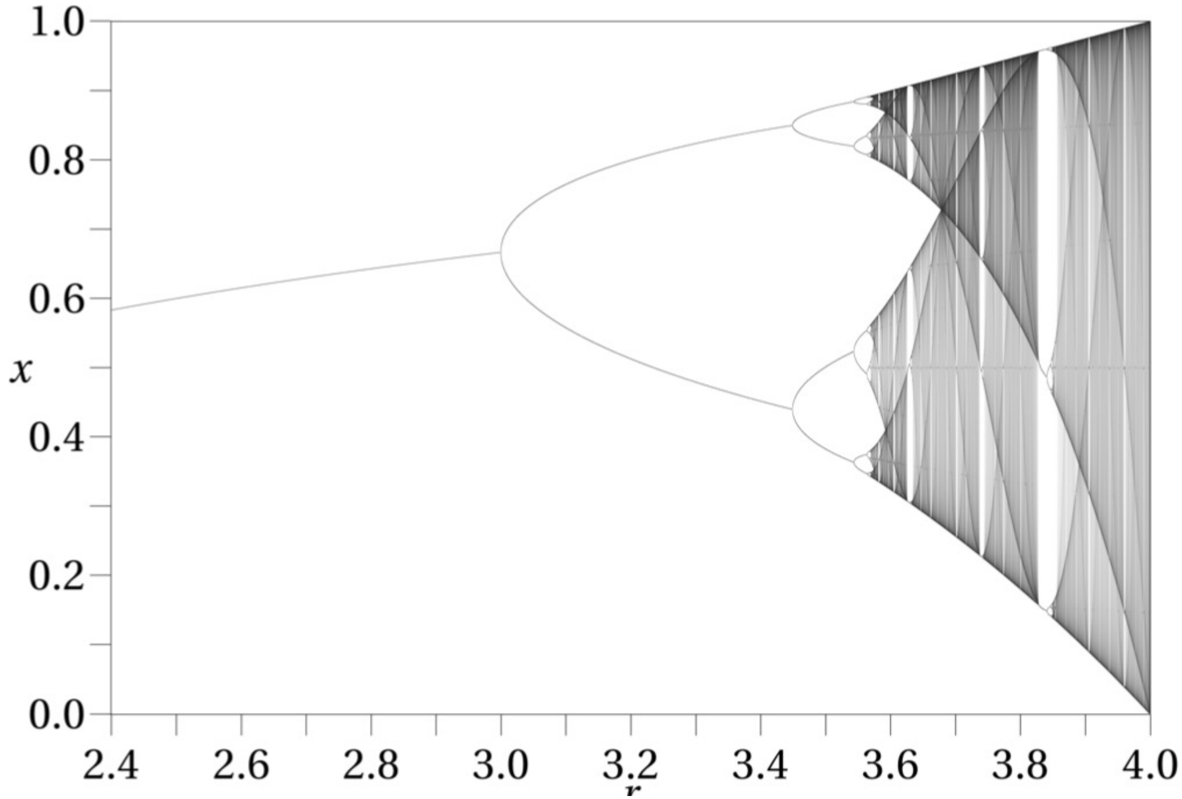
Example 8.14 (Fibonacci numbers). The Fibonacci numbers x_n are natural numbers generated by the second order linear recurrence relation

$$x_n = x_{n-1} + x_{n-2},$$

with initial values $x_0 = 0$ and $x_1 = 1$. Notable is that the famous *golden ratio* φ is obtained as the limit

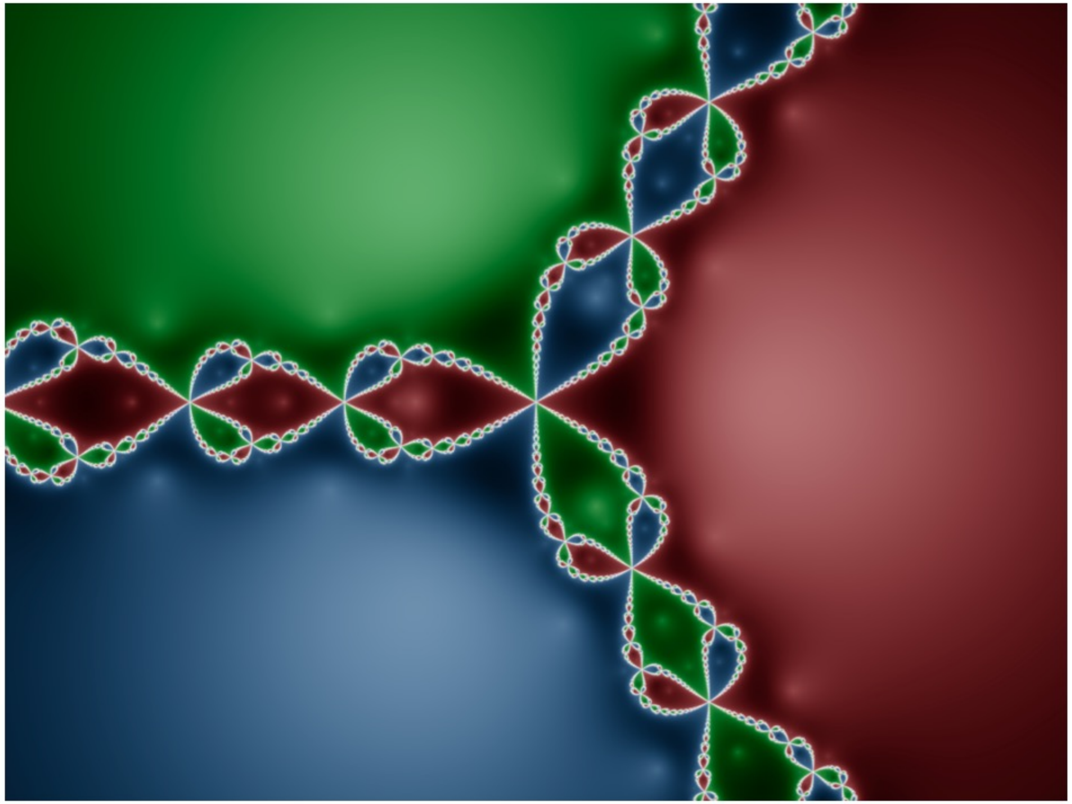
$$\varphi = \lim_{n\to\infty} \frac{x_n}{x_{n-1}}.$$

Logistic map



$$x_n = r x_{n-1}(1 - x_{n-1})$$

Julia set / Newton fractal



Points in the complex plane for which the follwing recurrence relation does not converge:

$$z_n = \frac{1 + 2z_{n-1}^3}{3z_{n-1}^2}$$

Corresponds to Newton's method for the equation:

$$z^3 = 1$$

Nonlinear dynamical system

Now consider a nonlinear dynamical system for the state vector $x \in R^n$ over a time interval $I = [0, T]$, described by the formula

$$x^{(k+1)} = x^{(k)} + \alpha f(x^{(k)}), \tag{8.6}$$

with $f : R^n \to R^n$ a nonlinear function, and where the parameter $\alpha = T/N$ represents the time step length for a partition of the interval I into N subintervals. We note the similarity with the fixed point iteration (8.5) for the matrix $A = \alpha I$, where the solution x^* to the equation

$$f(x^*) = 0$$

represents a steady state, or equilibrium point, of the dynamical system (8.6).

Linear stability analysis

To investigate the stability of the steady state we add a small perturbation φ to x^* and compare the evolution of $y^{(k)}$ from the perturbed initial state

$$y^{(0)} = x^* + \varphi,$$

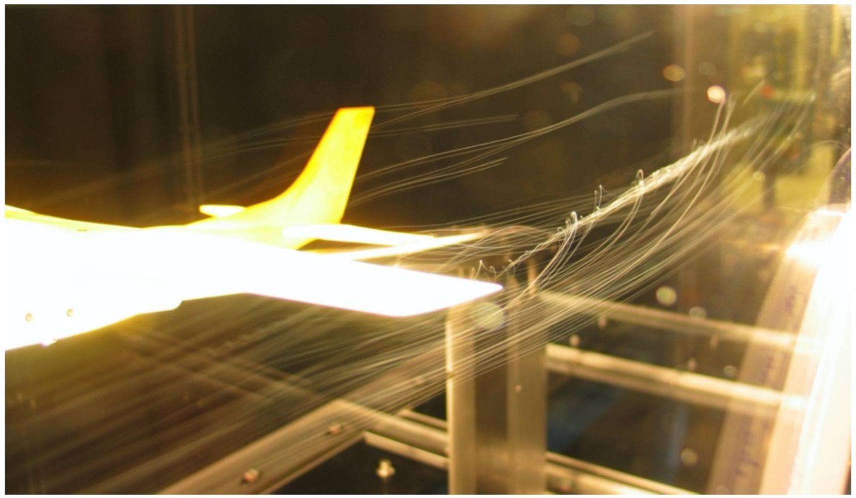
to the evolution of $x^{(k)}$ with initial state $x^{(0)} = x^*$. Then by Taylor's formula the evolution of the perturbation $\varphi^{(k)} = y^{(k)} - x^{(k)}$ can be approximated by the linear dynamical system

$$\varphi^{(k+1)} = \varphi^{(k)} + \alpha(f(y^{(k)}) - f(x^{(k)})) \approx \varphi^{(k)} + \alpha f'(x^*)\varphi^{(k)} = (I + \alpha f'(x^*))\varphi^{(k)}, \quad (8.7)$$

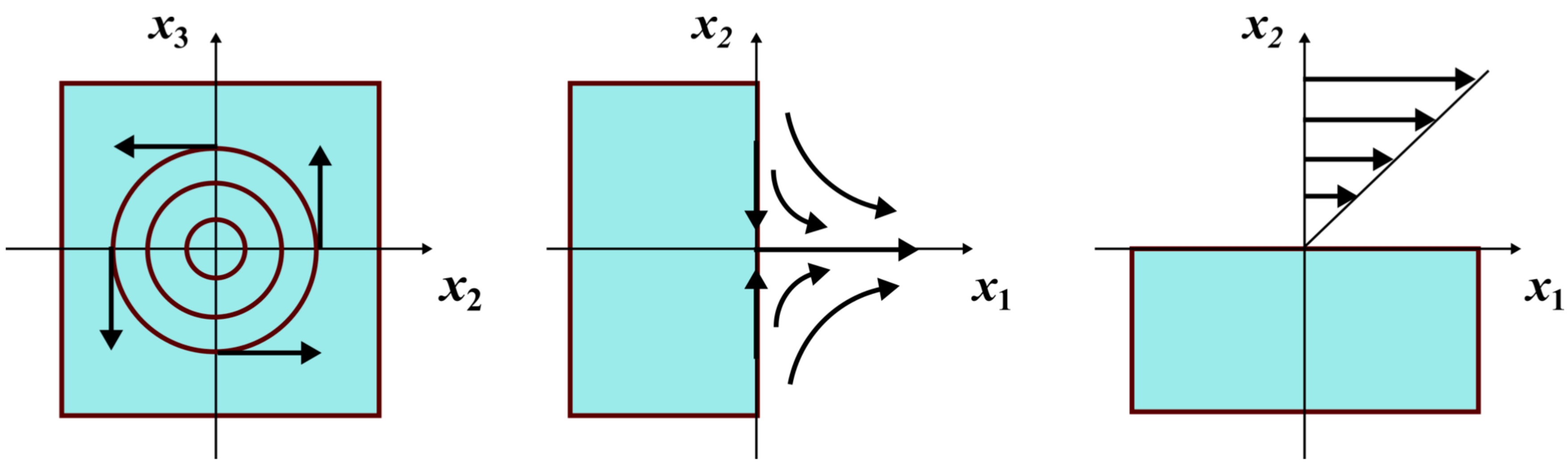
with the Jacobian $f'(x^*)$ linearized at the steady state x^*, and $(I + \alpha f'(x^*))$ acting as the state transition matrix analogous to equation (7.19).

Linear stability analysis

Example 8.17 (Navier-Stokes equations). Fluid dynamics is governed by the Navier-Stokes equations, a nonlinear dynamical system that describes the evolution of the scalar pressure $p(x) \in R$ and the velocity vector $u(x,t) \in R^3$ for each spatial point $x \in R^3$ and time $t > 0$.



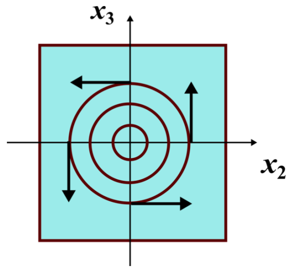
Linear stability analysis



Linear stability analysis

With x_1 the main flow direction, a vortex normal to the flow can be described in the x_2x_3 plane by the velocity vector

$$u_{vortex}(x) = (-x_3, x_2)^T,$$



which leads to the Jacobian

$$u'_{vortex}(x) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

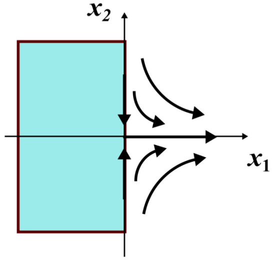
which has purely imaginary eigenvalues, hence, a stable structure with no perturbation growth.

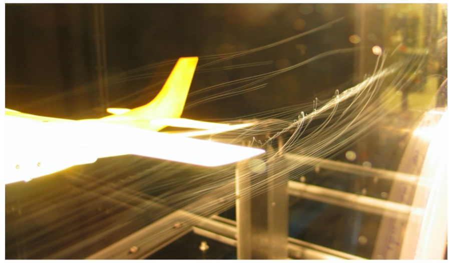
Linear stability analysis

Due to the incompressibility of the flow, corresponding to a zero trace Jacobian, an immersed object causes local acceleration and retardation, specifically at attachment and separation of the flow. Separation of the flow from an object like the wing of an airplane could be assumed to locally be an ideal two dimensional flow of the form

$$u_{separation}(x) = (x_1, -x_2)^T,$$

$$u'_{separation}(x) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$



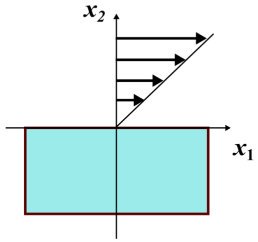


But since one of the eigenvalues of the Jacobian is positive, this two dimensional flow is unstable and will never manifest itself. Instead a pattern of stable vortices establish at separation, see the trailing edge of the wing in Figure 8.7.

Linear stability analysis

Shear flow in the x_1x_2 plane takes the form

$$u_{shear}(x) = (x_2, 0)^T,$$
$$u'_{shear}(x) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$



a defective matrix with transient growth of perturbations. Therefore, in shear flow we can experience the phenomenon of transition to turbulence, where perturbations slowly grow in a laminar shear flow until the accumulative effect is that the flow transitions into a chaotic turbulent flow, illustrated in Figure 8.7 by the rising smoke from a candle.