ROYAL INSTITUTE
OF TECHNOLOGY

Numerical Linear Algebra Methods for Nonlinear Problems

Bärbel Janssen

Computational Technology Laboratory
KTH Royal Institute of Technology, Stockholm, Sweden

October 16, 2014

Overview

Iterative methods for eigenvalues problems

Methods for Nonlinear Problems

Overview

Iterative methods for eigenvalues problems

Methods for Nonlinear Problems

Remarks to Arnoldi method

Remark

Typically, the Ritz eigenvalues converge to the extreme (maximal) eigenvalues of A. If one is interested in the smallest eigenvalues, i.e. those which are closest to zero, the methos has to be applied to the inverse matrix A^{-1}, similar to the approach used in the Inverse Iteration.

In this case the main work goes into the generation of the Krylov space $K_m = \text{span}\{q, A-1q, \dots, (A^{-1})^{m-1}q\}$ which requires the successive solution of linear systems

$$v^0 := q, \quad Av^1 = v^0, \quad \cdots, \quad Av^m = v^{m-1}.$$

Remarks to Arnoldi method

Remark

Typically, the Ritz eigenvalues converge to the extreme (maximal) eigenvalues of A. If one is interested in the smallest eigenvalues, i.e. those which are closest to zero, the methos has to be applied to the inverse matrix A^{-1}, similar to the approach used in the Inverse Iteration.

In this case the main work goes into the generation of the Krylov space $K_m = \text{span}\{q, A{-}1q, \ldots, (A^{-1})^{m-1}q\}$ which requires the successive solution of linear systems

$$v^0 := q, \quad Av^1 = v^0, \quad \cdots, \quad Av^m = v^{m-1}.$$

Remark

Due to practical storage consideration, common implementation of Arnoldi methods typically restart after some number of iterations. Theoretical results have shown that convergence improves with an increase in the Krylov subspace dimension m.

However, an a priori value of m which would lead to optimal convergence is not known. Recently, a dynamic switching strategy has been proposed which fluctuates the dimension m before each restart and thus leads to acceleration of convergence.

Remark

Due to practical storage consideration, common implementation of Arnoldi methods typically restart after some number of iterations. Theoretical results have shown that convergence improves with an increase in the Krylov subspace dimension m.
However, an a priori value of m which would lead to optimal convergence is not known. Recently, a dynamic switching strategy has been proposed which fluctuates the dimension m before each restart and thus leads to acceleration of convergence.

Remark

The modified Gram-Schmidt algorithm can also be used for the stable orthonormalization of a general basis $\{v^1, \ldots, v^m\} \subset \mathbb{C}^n$:

$$u^1 = \frac{v^1}{\left\|v^1\right\|_2}, \quad t = 2, \ldots, m: \quad j = 1, \ldots, t-1:$$

$$u^{t,1} = v^t$$

$$u^{t,j+1} = u^{t,j} - \operatorname{proj}_{u^j}(u^{t,j}), \quad u^t = \frac{u^{t,t}}{\left\|u^{t,t}\right\|_2}.$$

Both algorithms have the same arithmetical complexity. In each step a vector is determined orthogonal to its preceding one and also orthogonal to any errors introduced in the computation, which enhances stability. This supported by the following stability estimate for the resulting orthonormal matrix $U = (u^1, \ldots, u^m)$

$$\left\|\bar{U}^T U - I\right\|_2 \leq \frac{c_1 \operatorname{cond}_2(A)}{1 - c_2 \operatorname{cond}_2(A)} \epsilon.$$

Remark

The modified Gram-Schmidt algorithm can also be used for the stable orthonormalization of a general basis $\{v^1, \dots, v^m\} \subset \mathbb{C}^n$:

$$u^1 = \frac{v^1}{\left\|v^1\right\|_2}, \quad t = 2, \dots, m : \quad j = 1, \dots, t-1 :$$

$$u^{t,1} = v^t$$

$$u^{t,j+1} = u^{t,j} - \operatorname{proj}_{u^j}(u^{t,j}), \quad u^t = \frac{u^{t,t}}{\left\|u^{t,t}\right\|_2}.$$

Both algorithms have the same arithmetical complexity. In each step a vector is determined orthogonal to its preceding one and also orthogonal to any errors introduced in the computation, which enhances stability. This supported by the following stability estimate for the resulting orthonormal matrix $U = (u^1, \dots, u^m)$

$$\left\|\bar{U}^T U - I\right\|_2 \leq \frac{c_1 \operatorname{cond}_2(A)}{1 - c_2 \operatorname{cond}_2(A)} \epsilon.$$

Remark

1. Other orthogonalization algorithms use Householder transformations or Givens rotations.
2. The algoritm using Householder transformations are more stable than the stabilized Gram-Schmidt process.
3. On the other hand, the Gram-Schmidt process produces the t^{th} orthogonalized vector after the t^{th} iteration, while orthogonalization using Householder refelctions produces all the vectors only at the end.

This makes only the Gram-Schmidt process applicable for iterative methods like the Arnoldi iteration.

However, in quantum mechanics there are several orthogonalization schemes with characteristics even better suited for applications than the Gram-Schmidt algorithm.

Remark

1. Other orthogonalization algorithms use Householder transformations or Givens rotations.
2. The algoritm using Householder transformations are more stable than the stabilized Gram-Schmidt process.
3. On the other hand, the Gram-Schmidt process produces the t^{th} orthogonalized vector after the t^{th} iteration, while orthogonalization using Householder refelctions produces all the vectors only at the end.

This makes only the Gram-Schmidt process applicable for iterative methods like the Arnoldi iteration.
However, in quantum mechanics there are several orthogonalization schemes with characteristics even better suited for applications than the Gram-Schmidt algorithm.

Hermitian case

Suppose that the matrix A is hermitian. Then, the recurrence formula of the Arnoldi method

$$\tilde{q}^t = Aq^{t-1} - \sum_{j=1}^{t-1} (Aq^{t-1}, q^j)_2 q^j, \quad t = 2, \ldots, m+1,$$

because of $(Aq^{t-1}, q^j)_2 = (q^{t-1}, Aq^j)_2 = 0$, $j = 1, \ldots, t-3$, simplifies to

$$\begin{aligned} \tilde{q}^t &= Aq^{t-1} - (Aq^{t-1}, q^{t-1})_2 q^{t-1} - (Aq^{t-1}, q^{t-2})_2 q^{t-2} \\ &= Aq^{t-1} - \alpha_{t-1} q^{t-1} - \beta_{t-2} q^{t-2}. \end{aligned}$$

Clearly, $\alpha_{t-1} \in \mathbb{R}$ since A is hermitian. Further, multiplying this identity by q^t yields

$$\begin{aligned} \|\tilde{q}^t\| &= (q^t, \tilde{q}^t)_2 \\ &= (q^t, Aq^{t-1} - \alpha_{t-1}q^{t-1} - \beta_{t-2}q^{t-2})_2 \\ &= (q^t, Aq^{t-1})_2 = (Aq^t, q^{t-1})_2 = \beta_{t-1}. \end{aligned}$$

This implies that also $\beta_{t-1} \in \mathbb{R}$ and $\beta_{t-1}q^t = \tilde{q}^t$. Collecting the foregoing relations, we obtain

$$Aq^{t-1} = \beta_{t-1}q^t + \alpha_{t-1}q^{t-1} + \beta_{t-2}q^{t-2}, \quad t = 2, \ldots, m+1.$$

These equations can be written in matrix form as follows

$$AQ^{(m)} = Q^{(m)} \begin{pmatrix} \alpha_1 & \beta_2 & 0 & \cdots & \cdots & 0 \\ \beta_2 & \alpha_2 & \beta_3 & 0 & \cdots & 0 \\ 0 & \beta_3 & \alpha_3 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \beta_{m-1} & 0 \\ \vdots & \ddots & 0 & \beta_{m-1} & \alpha_{m-1} & \beta_m \\ 0 & \cdots & \cdots & 0 & \beta_m & \alpha_m \end{pmatrix} + \beta_m \begin{pmatrix} 0 \\ \vdots \\ 0 \\ q^{m+1} \end{pmatrix}$$

$$= Q^{(m)} T^{(m)} + \beta_m (0, \ldots, 0, q^{m+1}),$$

where the matrix $T^{(m)} \in \mathbb{R}^{m \times m}$ is real symmetric.
From this Lanczos relation we finally obtain

$$Q^{(m)T} A Q^{(m)} = T^{(m)}.$$

These equations can be written in matrix form as follows

$$AQ^{(m)} = Q^{(m)} \begin{pmatrix} \alpha_1 & \beta_2 & 0 & \cdots & \cdots & 0 \\ \beta_2 & \alpha_2 & \beta_3 & 0 & \cdots & 0 \\ 0 & \beta_3 & \alpha_3 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \beta_{m-1} & 0 \\ \vdots & \ddots & 0 & \beta_{m-1} & \alpha_{m-1} & \beta_m \\ 0 & \cdots & \cdots & 0 & \beta_m & \alpha_m \end{pmatrix} + \beta_m \begin{pmatrix} 0 \\ \vdots \\ 0 \\ q^{m+1} \end{pmatrix}$$

$$= Q^{(m)} T^{(m)} + \beta_m(0, \ldots, 0, q^{m+1}),$$

where the matrix $T^{(m)} \in \mathbb{R}^{m \times m}$ is real symmetric.
From this Lanczos relation we finally obtain

$$Q^{(m)T} A Q^{(m)} = T^{(m)}.$$

Lanczos Algorithm

For a hermitian matrix $A \in \mathbb{C}^{n \times n}$ the Lanczos method determines a set of orthonormal vectors $\{q^1, \dots, q^m\}$, $m \ll n$ by applying the modified Gram-Schmidt method to the basis $\{q, Aq, \dots, A^{m-1}q\}$ of the Krylov space K_m.

$$\begin{array}{ll}
\text{Starting vector:} & q^1 = \frac{q}{\|q\|_2},\ q^0 = 1, \beta_1 = 0. \\
\text{Iterate for } 1 \le t \le m-1: & r^t = Aq^t, \alpha_t = (r^t, q^t)_2, \\
 & s^t = r^t - \alpha_t q^t - \beta_t q^{t-1}, \\
 & \beta^{t+1} = \left\|s^t\right\|_2, \quad q^{t+1} = \frac{s^t}{\beta_{t+1}}, \\
 & r^m = Aq^m, \quad \alpha_m = (r^m, q^m)_2.
\end{array}$$

Lanczos Algorithm

For a hermitian matrix $A \in \mathbb{C}^{n\times n}$ the Lanczos method determines a set of orthonormal vectors $\{q^1, \dots, q^m\}$, $m \ll n$ by applying the modified Gram-Schmidt method to the basis $\{q, Aq, \dots, A^{m-1}q\}$ of the Krylov space K_m.

Starting vector: $q^1 = \frac{q}{\|q\|_2},\ q^0 = 1, \beta_1 = 0.$

Iterate for $1 \le t \le m-1$:

$$\begin{aligned}
&r^t = Aq^t, \alpha_t = (r^t, q^t)_2,\\
&s^t = r^t - \alpha_t q^t - \beta_t q^{t-1},\\
&\beta^{t+1} = \left\|s^t\right\|_2, \quad q^{t+1} = \frac{s^t}{\beta_{t+1}},\\
&r^m = Aq^m, \quad \alpha_m = (r^m, q^m)_2.
\end{aligned}$$

1. After the matrix $T^{(m)}$ is calculated, one can compute its eigenvalues λ_i and their corresponding eigenvectors w^i, e.g. by the QR alpgorithm.
2. The eigenvalue and eigenvectors of $T^{(m)}$ can be obtained in as little as $\mathcal{O}(m^2)$ work.
3. It can be proved that the eigenvalues are approximate eigenvalues of the original matrix A.

1. After the matrix $T^{(m)}$ is calculated, one can compute its eigenvalues λ_i and their corresponding eigenvectors w^i, e.g. by the QR alpgorithm.
2. The eigenvalue and eigenvectors of $T^{(m)}$ can be obtained in as little as $\mathcal{O}(m^2)$ work.
3. It can be proved that the eigenvalues are approximate eigenvalues of the original matrix A.

1. After the matrix $T^{(m)}$ is calculated, one can compute its eigenvalues λ_i and their corresponding eigenvectors w^i, e.g. by the QR alpgorithm.
2. The eigenvalue and eigenvectors of $T^{(m)}$ can be obtained in as little as $\mathcal{O}(m^2)$ work.
3. It can be proved that the eigenvalues are approximate eigenvalues of the original matrix A.

Overview

Iterative methods for eigenvalues problems

Methods for Nonlinear Problems

The Newton method in $\mathbb{R}$

Let f be a function on the interval $[a, b]$ which is continuously differentiable. To approximate a zero of f, we compute the tangent to $f(x)$ to determine its zero. The tangent is given as

$$T(x) = f'(x_t)(x - x_t) + f(x_t).$$

Its zero x_{t+1} is determined through

$$x_{t+1} = x_t - \frac{f(x_t)}{f'(x_t)}.$$

This iteration is only possible if the values of $f'(x_t)$ do not become too small. It allows us to approximate simple zeros of f.

Theorem (Newton method)

The function $f \in C^2[a,b]$ is assumed to have a zero z in the inner of $[a,b]$. Further, let

$$m := \min_{a \le x \le b} \left| f'(x) \right| > 0, \quad M := \max_{a \le x \le b} \left| f''(x) \right| .$$

Let $\rho > 0$ be chosen such that

$$q := \frac{M}{2m} \rho < 1, \quad K_\rho(z) := \{ x \in \mathbb{R} : |x - z| \le \rho \} \subset [a,b].$$

Then, the Newton iterates $x_t \in K_\rho(z)$ are defined and converge towards the zero z for any starting value $x_0 \in K_\rho(z)$. There holds the a priori estimate

$$|x_t - z| \le \frac{2m}{M} q^{(2^t)}, \quad t \in \mathbb{N},$$

and the a posteriori estimate

$$|x_t - z| \le \frac{1}{m} \left| f(x_t) \right| \le \frac{M}{2m} |x_t - x|^2 , \quad t \in \mathbb{N}.$$

Theorem (Newton method)

The function $f \in C^2[a,b]$ is assumed to have a zero z in the inner of $[a,b]$. Further, let

$$m := \min_{a \le x \le b} \left| f'(x) \right| > 0, \quad M := \max_{a \le x \le b} \left| f''(x) \right| .$$

Let $\rho > 0$ be chosen such that

$$q := \frac{M}{2m}\rho < 1, \quad K_\rho(z) := \{x \in \mathbb{R} : |x - z| \le \rho\} \subset [a,b].$$

Then, the Newton iterates $x_t \in K_\rho(z)$ are defined and converge towards the zero z for any starting value $x_0 \in K_\rho(z)$. There holds the a priori estimate

$$|x_t - z| \le \frac{2m}{M} q^{(2^t)}, \quad t \in \mathbb{N},$$

and the a posteriori estimate

$$|x_t - z| \le \frac{1}{m}\left| f(x_t) \right| \le \frac{M}{2m}|x_t - x|^2 , \quad t \in \mathbb{N}.$$

Remark

There exists a $K_\rho(z)$ in case that $f \in C^2[a, b]$ and $f(z) = 0$ but $f'(z) \neq 0$. It might be very small. The biggest problem for the Newton method is to find a good starting value. If there happens to be one, the Newton method converges extremely fast towards the zero of f.

Compute the root

The n^{th} root of a number $a > 0$ is the root of the function $f(z) = x^n - a$. The method for its approximation is given as

$$x_{t+1} = x_t - \frac{x_t^n - a}{n x_t^{n-1}} = \frac{1}{n}\left((n-1)\,x_t + \frac{a}{x_t^{n-1}}\right)$$

In case $n = 2$, we get the following estimates for a starting value

$$\frac{1}{1\sqrt{2}}\left|x_0 - \sqrt{a}\right| < 1 \quad \text{and} \quad \left|x_0 - \sqrt{a}\right| < 2\sqrt{a}.$$

Further, we have the relationship

$$\frac{a}{x_t} \leq \sqrt{a} \leq x_t,$$

which can be translated into a stopping criterion

$$0 \leq e_t := x_t - \frac{a}{x_t} \leq \varepsilon.$$

Compute the root

The n^{th} root of a number $a > 0$ is the root of the function $f(z) = x^n - a$. The method for its approximation is given as

$$x_{t+1} = x_t - \frac{x_t^n - a}{n x_t^{n-1}} = \frac{1}{n}\left((n-1)\,x_t + \frac{a}{x_t^{n-1}}\right)$$

In case $n = 2$, we get the following estimates for a starting value

$$\frac{1}{1\sqrt{2}}\left|x_0 - \sqrt{a}\right| < 1 \quad \text{and} \quad \left|x_0 - \sqrt{a}\right| < 2\sqrt{a}.$$

Further, we have the relationship

$$\frac{a}{x_t} \leq \sqrt{a} \leq x_t,$$

which can be translated into a stopping criterion

$$0 \leq e_t := x_t - \frac{a}{x_t} \leq \varepsilon.$$

Compute the root

The n^{th} root of a number $a > 0$ is the root of the function $f(z) = x^n - a$. The method for its approximation is given as

$$x_{t+1} = x_t - \frac{x_t^n - a}{n x_t^{n-1}} = \frac{1}{n}\left((n-1)\,x_t + \frac{a}{x_t^{n-1}}\right)$$

In case $n = 2$, we get the following estimates for a starting value

$$\frac{1}{1\sqrt{2}}\left|x_0 - \sqrt{a}\right| < 1 \quad \text{and} \quad \left|x_0 - \sqrt{a}\right| < 2\sqrt{a}.$$

Further, we have the relationship

$$\frac{a}{x_t} \leq \sqrt{a} \leq x_t,$$

which can be translated into a stopping criterion

$$0 \leq e_t := x_t - \frac{a}{x_t} \leq \varepsilon.$$

Damped Newton method

Often, the range of possible starting values is very small. In this case, we use a damped version of the Newton method which is defined by

$$x_{t+1} = x_t - \lambda_t \frac{f(x_t)}{f'(x_t)},$$

with a damping paramter $\lambda_t \in (0, 1]$. The determination of this damping parameter λ_t is a whole new other story.

More than one root

In the critical case that the root of $f(z)$ is also a root of $f'(z) = 0$. Let $f''(z) \neq 0$. Then the Newton method reads

$$x_{t+1} = x_t - \frac{f(x_t) - f(z)}{f'(x_t) - f'(z)} = x_t - \frac{f'(\zeta_t)}{f''(\mu_t)}$$

with points $\zeta_t, \mu_t \in [x_t, z]$.
This idea can be used for higher order roots as well.

More than one root

In the critical case that the root of $f(z)$ is also a root of $f'(z) = 0$. Let $f''(z) \neq 0$. Then the Newton method reads

$$x_{t+1} = x_t - \frac{f(x_t) - f(z)}{f'(x_t) - f'(z)} = x_t - \frac{f'(\zeta_t)}{f''(\mu_t)}$$

with points $\zeta_t, \mu_t \in [x_t, z]$.
This idea can be used for higher order roots as well.

Simplified Newton Method

If the computation of $f'(x_t)$ in each iteration step is costly, a simplified version of the Newton method can be used. It reads

$$x_{t+1} = x_t - \frac{f(x_t)}{f'(c)},$$

where c is suitably chosen.

This iteration is a version of the general fixed point iteration

$$x_{t+1} = x_t + \sigma f(x_t)$$

with σ chosen suitably.

Simplified Newton Method

If the computation of $f'(x_t)$ in each iteration step is costly, a simplified version of the Newton method can be used. It reads

$$x_{t+1} = x_t - \frac{f(x_t)}{f'(c)},$$

where c is suitably chosen.
This iteration is a version of the general fixed point iteration

$$x_{t+1} = x_t + \sigma f(x_t)$$

with σ chosen suitably.

Interpolation methods

The aim of interpolation methods is to avoid evaluation of the derivative of f but be more efficient than interval bisection methods. Instead of using the tangent to the function f as in the Newton method, we use a secant through the points $(x_{t-1}, f(x_{t-1}))$ and $(x_t, f(x_t))$. The secant is given by

$$s(x) = f(x_t) + (x - x_t)\frac{f(x_t) - f(x_{t-1})}{x_t - x_{t-1}},$$

and the secant iteration reads

$$x_{t+1} = x_t - f(x_t)\frac{x_t - x_{t-1}}{f(x_t) - f(x_{t-1})}.$$

Fibonacci numbers

The Fibonacci numbers defines as

$$\gamma_0 = \gamma_1 = 1, \quad \gamma_{t+1} = \gamma_t + \gamma_{t-1}, \quad t \in \mathbb{N},$$

play an important role in the convergence of the secant method.

Theorem (Secant method)

Let f be in $C^2[a,b]$ and assume that f has a zero in the inner of $[a,b]$. Moreover, let

$$m := \min_{a \leq x \leq b} \left| f'(x) \right| > 0, \quad M := \max_{a \leq x \leq b} \left| f''(x) \right| < \infty,$$

and let $\rho > 0$ be chosen such that

$$q = \frac{M}{2m} \rho < 1, \quad K_\rho(z) = \{ x \in \mathbb{R} : |x - z| \leq \rho \} \subset [a,b].$$

Then the secant method is well defined for any two starting values $x_0, x_1 \in K_\rho(z, x_0 \neq x_1)$ and the method converges towards the zero of f. We have the following a priori estimate

$$|x_t - z| \leq \frac{2m}{M} q^{\gamma_t}, \quad t \in \mathbb{N},$$

where $\gamma_t \sim 0.723 \cdot (1.618)^t$, and the a posteriori estimate

$$|x_t - z| \leq \frac{1}{m} \left| f(x_t) \right| \leq \frac{M}{2m} |x_t - x_{t-1}| |x_t - x_{t-2}|, \quad t \in \mathbb{N}.$$

Theorem (Secant method)

Let f be in $C^2[a,b]$ and assume that f has a zero in the inner of $[a,b]$. Moreover, let

$$m := \min_{a \le x \le b} \left| f'(x) \right| > 0, \quad M := \max_{a \le x \le b} \left| f''(x) \right| < \infty,$$

and let $\rho > 0$ be chosen such that

$$q = \frac{M}{2m}\rho < 1, \quad K_\rho(z) = \{x \in \mathbb{R} : |x - z| \le \rho\} \subset [a,b].$$

Then the secant method is well defined for any two starting values $x_0, x_1 \in K_\rho(z, x_0 \ne x_1)$ and the method converges towards the zero of f. We have the following a priori estimate

$$|x_t - z| \le \frac{2m}{M} q^{\gamma_t}, \quad t \in \mathbb{N},$$

where $\gamma_t \sim 0.723 \cdot (1.618)^t$, and the a posteriori estimate

$$|x_t - z| \le \frac{1}{m}\left| f(x_t) \right| \le \frac{M}{2m} |x_t - x_{t-1}| |x_t - x_{t-2}| \,, \quad t \in \mathbb{N}.$$

Successive approximation in $\mathbb{R}^n$

We would like to investigate methods for nonlinear problems in $\mathbb{R}^n$

$$f_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, n,$$

or $f(x) = 0$ with $f = (f1, \dots, f_n)^T$ and $x = (x_1, \dots, x_n)^T$.

Successive approximation method

$$x^{t+1} = x^t + C^{-1} f(x^t), \quad t = 0, 1, 2, \dots$$

with a regular matrix $C \in \mathbb{R}^{n \times n}$.

Successive approximation in $\mathbb{R}^n$

We would like to investigate methods for nonlinear problems in $\mathbb{R}^n$

$$f_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, n,$$

or $f(x) = 0$ with $f = (f1, \dots, f_n)^T$ and $x = (x_1, \dots, x_n)^T$.

Successive approximation method

$$x^{t+1} = x^t + C^{-1} f(x^t), \quad t = 0, 1, 2, \dots$$

with a regular matrix $C \in \mathbb{R}^{n \times n}$.

Theorem (Successive approximation)

Let $G \subset \mathbb{R}^n$ be a non empty, closed set and let $g : G \to G$ defined by $g(x) \coloneqq x + C^{-1} f(x)$ be a contraction. Then there exists a uniquely defined fixed point $z \in G$ and the Successive approximation converges for any starting point $x^0 \in G$.

There hold the a priori and a posteriori estimates

$$\left\| x^t - z \right\| \leq \frac{q}{1-q} \left\| x^t - x^{t-1} \right\| \leq \frac{q^t}{1-q} \left\| x^1 - x^0 \right\|,$$

where q is the Lipschitz constant in

$$\left\| g(x) - g(y) \right\| \leq q \left\| x - y \right\|, \quad x, y \in G, \quad q < 1.$$

Theorem (Successive approximation)

Let $G \subset \mathbb{R}^n$ be a non empty, closed set and let $g : G \to G$ defined by $g(x) := x + C^{-1} f(x)$ be a contraction. Then there exists a uniquely defined fixed point $z \in G$ and the Successive approximation converges for any starting point $x^0 \in G$.
There hold the a priori and a posteriori estimates

$$\left\|x^t - z\right\| \leq \frac{q}{1-q}\left\|x^t - x^{t-1}\right\| \leq \frac{q^t}{1-q}\left\|x^1 - x^0\right\|,$$

where q is the Lipschitz constant in

$$\left\|g(x) - g(y)\right\| \leq q\|x - y\|, \quad x, y \in G, \quad q < 1.$$

Newton method in $\mathbb{R}^n$

The Newton method for the solution of nonlinear systems with a differentiable function $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ reads

$$x^{t+1} = x^t - f'(x^t)^{-1} f(x^t), \quad t = 0, 1, 2, \dots,$$

with the Jacobi matrix $f'(\cdot)$ of f.
In each iteration step a system of the form

$$f'(x^t) x^{t+1} = f'(x^t) x^t - f(x^t), \quad t = 0, 1, 2, \dots,$$

has to be solved.

Newton method in $\mathbb{R}^n$

The Newton method for the solution of nonlinear systems with a differentiable function $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ reads

$$x^{t+1} = x^t - f'(x^t)^{-1} f(x^t), \quad t = 0, 1, 2, \ldots,$$

with the Jacobi matrix $f'(\cdot)$ of f.
In each iteration step a system of the form

$$f'(x^t) x^{t+1} = f'(x^t) x^t - f(x^t), \quad t = 0, 1, 2, \ldots,$$

has to be solved.

Newton as a defect correction method

The Newton method can equivalently be written as

$$f'(x^t)\delta x^t = -f(x^t), \quad x^{t+1} = x^t + \delta x^t, \quad t = 0, 1, 2, \dots,$$

Some examples

1. Compute the root of $f(x) = x^n - a$, $a > 0$.
2. Compute the zero of $f(x) = \sin(x)$ with an error smaller than 10^{-6}. Compare
    - 2.1 the fixed point iteration $x_t = x_{t-1} + f(x_{t-1})$ with starting value $x_0 = 4$.
    - 2.2 the Newton iteration $x_t = x_{t-1} - f'(x_{t-1})^{-1} f(x_{t-1})$ with starting value $x_0 = 4$.
3. Solve the equation $x + \ln(x) = 0$ with
    - 3.1 $x_t = \exp(-x_{t-1})$,
    - 3.2 $x_t = \frac{1}{2}(x_{t-1} + \exp(-x_{t-1}))$
    - 3.3 Can you find an even better iteration?

More examples

1. Compute the root $A^{\frac{1}{2}} \in \mathbb{R}^{n \times n}$ of a positive definite matrix $A \in \mathbb{R}^{n \times n}$ via the funtion

$$g(X) = \frac{1}{2}(X^2 + B), \quad B = I - A.$$

2. The eigenvalue problem $Ax = \lambda x$ for a matrix $A \in \mathbb{R}^{n \times n}$ is equivalent to solving the nonlinear problem

$$\begin{aligned} Ax - \lambda x &= 0, \\ \|x\|_2^2 - 1 &= 0, \end{aligned}$$

of $n + 1$ equations and $n + 1$ unknowns $x_1, \ldots, x_n, \lambda$.