1. LOCAL ALGORITHMS FOR SOLVING NONLINEAR EQUATIONS

Here we discuss local methods for nonlinear equations

$$r(x) = 0.$$ 

These methods are Newton, inexact Newton and quasi-Newton. We will show that the Newton method converges $Q$-superlinearly if the Jacobian $J$ is continuous and $J(x^*)$ is nonsingular at the solution point $x^*$. Furthermore, the convergence is $Q$-quadratic if the Jacobian is Lipschitz-continuous. However, the Newton method is not globally convergent.

In inexact Newton methods, the step is made not all the way to the solution of the linear model (recall the 1d case where at each step we jump to the point where the tangent line intersects the $x$-axis), but just some distance along this direction, making sure that the residual is sufficiently reduced. This strategy allows us to avoid solving a linear system of equations that is expensive to do in high dimensions. Contrary to inexact Newton methods, the quasi-Newton methods (e.g., Broyden’s method) do not use the exact Jacobian but an approximation to it. This approximation is built in the process of iterations. This is very handy in the case where the exact Jacobian is hard to calculate.

All these methods exhibit a fast super linear convergence under certain conditions but suffer the same shortcoming: their behavior can be erratic unless the initial approximation is close enough to the solution.

However, one can improve their behavior dramatically by keeping track of some merit function, often chosen to be

$$f(x) = \frac{1}{2} \sum_{j=1}^{n} r_j^2(x),$$

and using the step sizes wisely. Then the problem of finding a zero of $r(x)$ is reduced to finding a minimum of $f(x)$, hence to a minimization problem. We will discuss it in the chapter “Optimization”.

2. NEWTON METHOD FOR SOLVING NONLINEAR EQUATIONS

In the Newton method, as we have done it in 1d, at every iteration, we define linear model

$$M_k := r(x_k) + J(x_k)p.$$ 

Then we make a step $p_k$ right into the zero of the linear model, i.e., the Newton step $p_{k+1}$ is the solution of

$$J(x_k)p_k = -r(x_k).$$

Thus the algorithm is the following.

**Algorithm 1: Newton Method**

Choose $x_0$;

**While** $\|r(x_k)\| > tol$

- Solve $J(x_k)p = -r(x_k)$ for $p_k$;
Theorem 1. Suppose that \( r \) is continuously differentiable in a convex open set \( D \subset \mathbb{R}^n \). Let \( x^* \in D \) be a non degenerate solution of \( r(x) = 0 \), and let \( \{x_k\} \) be the sequence of iterates generated by Algorithm 1. Then when \( x_k \in D \) is sufficiently close to \( x^* \), we have
\[
x_{k+1} - x^* = o(\|x_k - x^*\|),
\]
indicating local \( Q \)-superlinear convergence. When \( r \) is Lipschitz-continuously differentiable near \( x^* \), we have for all \( x_k \) sufficiently close to \( x^* \) that
\[
x_{k+1} - x^* = O(\|x_k - x^*\|^2),
\]
indicating local \( Q \)-quadratic convergence.

Proof. The proof is conducted in three steps. First we show that \( r(x_k) = J(x_k)(x_k - x^*) + o(\|x_k - x^*\|) \). Then we show that \( x_{k+1} - x^* = x_k + p_k - x^* = x_k + J^{-1}(x_k)r(x_k) - x^* = o(\|x_k - x^*\|) \). Finally we will show that the Lipschitz continuity of \( J \) leads to \( Q \)-quadratic convergence.

1. Since \( r(x^*) = 0 \) we have
\[
r(x_k) = r(x_k) - r(x^*)
= J(x_k)(x_k - x^*) + \int_0^1 \left[ J(x_k + t(x^* - x_k)) - J(x_k) \right] (x_k - x^*) dt
= J(x_k)(x_k - x^*) + o(\|x_k - x^*\|).
\]
Let us explain why
\[
\int_0^1 \left[ J(x_k + t(x^* - x_k)) - J(x_k) \right] (x_k - x^*) dt = o(\|x_k - x^*\|).
\]
We can bound this integral from above in absolute value as
\[
\left\| \int_0^1 \left[ J(x_k + t(x^* - x_k)) - J(x_k) \right] (x_k - x^*) dt \right\| 
\leq \int_0^1 \left\| J(x_k + t(x^* - x_k)) - J(x_k) \right\| (x_k - x^*) \| dt.
\]
Set \( p_k := x_k - x^* \). Since \( J(x) \) is continuous we have
\[
\lim_{\|p\| \to 0} \frac{\int_0^1 \| J(x_k + tp_k) - J(x_k) \| \| p_k \| dt}{\| p_k \|} = 0.
\]
This implies Eq. (1).

2. Since \( J(x^*) \) is nonsingular, there exists a ball \( B_\delta(x^*) \) of radius \( \delta \) around \( x^* \) and a positive constant \( \beta \) such that
\[
\| J^{-1}(x) \| \leq \beta, \quad x \in B_\delta(x^*) \cap D.
\]
If \( x_k \in B_\delta(x^*) \) we have
\[
p_k = -J^{-1}(x_k)r(x_k) = -J^{-1}(x_k)(J(x_k)(x_k - x^*) + o(\|x_k - x^*\|)) = -x_k + x^* + o(\|x_k - x^*\|).
\]
Hence
\[
x_{k+1} - x^* = x_k + p_k - x^* = o(\|x_k - x^*\|).
\]
(3) If \( J(x) \) is Lipschitz-continuous we have
\[
\left\| \int_0^1 [J(x_k + t(x^* - x_k)) - J(x_k)](x_k - x^*) dt \right\|
\leq \int_0^1 \|J(x_k + t(x^* - x_k)) - J(x_k)\| \|x_k - x^*\| dt
\leq \int_0^1 Mt\|x_k - x^*\|^2 dt
= O(\|x_k - x^*\|^2).
\]
Therefore
\[
r(x^k) = J(x_k)(x_k - x^*) + O(\|x_k - x^*\|^2).
\]
Hence
\[
p_k = -J^{-1}(x_k)r(x_k) = -J^{-1}(x_k)(J(x_k)(x_k - x^*) + O(\|x_k - x^*\|^2))
= -x_k + x^* + O(\|x_k - x^*\|^2).
\]
Then
\[
x_{k+1} - x^* = x_k + p_k - x^* = O(\|x_k - x^*\|^2),
\]
i.e., the Newton method converges \( Q \)-quadratically.

\(\square\)

2.1. Inexact Newton methods. Solving equation \( J(x_k)p_k = -r(x_k) \) for \( p_k \) exactly can be expensive in high dimensions. Inexact Newton methods use the search directions \( p_k \) that satisfy the condition
\[
\|r_k + J_k p_k\| \leq \eta_k \|r_k\|, \quad \text{for some} \quad \eta_k \in [0, \eta],
\]
where \( \eta \) is a constant. The most important methods in this class make use of iterative techniques for solving linear systems of the form \( Jp = -r \). The convergence theory for these methods depends only on the condition (2) and not on particular technique used to calculate \( p_k \). Its results are summarized in the following theorem.

Theorem 2. Suppose that \( r \) is continuously differentiable in a convex open set \( D \subset \mathbb{R}^n \). Let \( x^* \in D \) be a non-degenerate solution of \( r(x) = 0 \), and let \( x_k \) be a sequence of iterates generated by an inexact Newton method. Then when \( x_k \in D \) is sufficiently close to \( x^* \), the following are true:

- If \( \eta \) in Eq. (2) is sufficiently small, then the convergence of \( \{x_k\} \) to \( x^* \) is \( Q \)-linear.
- If \( \eta_k \to 0 \), the convergence is \( Q \)-superlinear.
If, in addition, \( J \) is Lipschitz-continuous on \( D \) and \( \eta_k = O(\|r_k\|) \), then the convergence is \( Q \)-quadratic.

**Exercise** Prove it.

2.2. **Broyden’s method.** Secant methods, also known as quasi-Newton methods, do not require calculation of the Jacobian \( J(x) \). Instead, they construct their own approximation to \( J(x) \), updating it at each iteration so that it mimics the behavior of the true Jacobian over the step just taken.

**Example** Secant method in 1D is defined by the recurrence relation

\[
x_{n+1} = x_n - f(x_n) \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}.
\]

**Exercise** Prove that it converges with the order

\[
\alpha = \frac{1 + \sqrt{5}}{2} \approx 1.618.
\]

Let \( B_k \) be the Jacobian approximation at iteration \( k \). Assuming that \( B \) is nonsingular we define the step \( p_k \) to the next iterate as the solution of

\[
B_k p_k + r(x_k) = 0, \quad x_{k+1} = x_k + p_k.
\]

The standard notations are

\[
s_k = x_{k+1} - x_k, \quad y_k = r(x_{k+1}) - r(x_k).
\]

At this point, \( s_k = p_k \), but in the practical methods where a merit function is being monitored, usually a step is done in the direction of \( p_k \) but of different length. The length is chosen to satisfy Wolfe’s conditions that we will discuss later. We have that \( s_k \) and \( y_k \) are related via

\[
y_k = \int_0^1 J(x_k + ts_k)s_k dt \approx J(x_{k+1})s_k + o(\|s_k\|).
\]

We require the updated Jacobian to satisfy the following equation known as the secant equation

\[
y_k = B_{k+1} s_k,
\]

which insures that \( B_{k+1} \) and \( J(x_{k+1}) \) have similar behavior along the direction \( s_k \). Eq. (5) does not determine \( B_{k+1} \) uniquely for \( n > 1 \): It is a system of \( n \) equations with \( n^2 \) unknowns.

Various quasi-Newton methods differ in the way how they determine \( B_{k+1} \) satisfying Eq. (5). The most successful practical algorithm is Broyden’s method for which the update formula is

\[
B_{k+1} = B_k + \frac{(y - B_k s_k) s_k^T}{s_k^T s_k}.
\]

The Broyden update makes the smallest possible change to the Jacobian measured by the Euclidean norm \( \|B_k - B_{k+1}\| \)
Theorem 3. Among all maurices $B$ satisfying $Bs_k = y_k$, the matrix $B_{k+1}$ defined Eq. (6) minimizes the difference $\|B - B_k\|$.

Proof.

$$\|B_{k+1} - B_k\| = \left| \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k} \right| = \left| \frac{(B - B_k) s_k^T s_k^T}{s_k^T s_k} \right| \leq \|B - B_k\| \left| \frac{s_k^T s_k^T}{s_k^T s_k} \right| = \|B - B_k\|.$$ 

Here we have used the property of the norm that $\|AB\| \leq \|A\|\|B\|$ and the fact that $\left| \frac{s_k^T s_k^T}{s_k^T s_k} \right| = 1$.

This is easy to check. Since $s_k^T s_k$ is symmetric, we have

$$\|s_k^T s_k^T\| = \max_{\|x\|=1} x^T s_k^T s_k^T x = \|s_k\|^2 \cdot s_k^T s_k = 1.$$ 

\[\square\]

Broyden’s algorithm work as follows.

Algorithm: Broyden’s method
Choose $x_0$ and $B_0$;
for $k = 0, 1, 2, \ldots$
    Solve $B_k p_k = -r(x_k)$ for $p_k$;
    Choose $\alpha_k$ by performing a line search along direction $p_k$;
    $x_{k+1} = x_k + \alpha_k p_k$;
    $s_k = x_{k+1} - x_k$;
    $y_k = r(x_{k+1}) - r(x_k)$;
    $B_{k+1} = B_k + \frac{(y - B_k s_k) s_k^T}{s_k^T s_k}$;
end

Under certain assumptions, Broyden’s method converges super linearly, that is,

$$\|x_{k+1} - x^*\| = o(\|x_k - x^*\|).$$

Theorem 4. Suppose that $r$ is continuously differentiable in a convex open set $D \subset \mathbb{R}^n$. Let $x^* \in D$ be a non degenerate solution of $r(x) = 0$. Then there are positive constants $\epsilon$ and $\delta$ such that if the starting point $x_0$ and the initial approximate Jacobian $B_0$ satisfy

$$\|x_0 - x^*\| \leq \delta, \quad \|B_0 - J(x^*)\| \leq \epsilon,$$

the sequence $\{x_k\}$ generated by

$$x_{k+1} = x_k - B_k^{-1} r(x_k), \quad s_k = x_{k+1} - x_k, \quad B_{k+1} = B_k + \frac{(y - B_k s_k) s_k^T}{s_k^T s_k}.$$
is well-defined and converges $Q$-superlinearly to $x^*$.

References