Numerical Optimization
Lecture Notes #26 — Nonlinear Equations: Practical Methods

Peter Blomgren,
⟨blomgren.peter@gmail.com⟩

Department of Mathematics and Statistics
Dynamical Systems Group
Computational Sciences Research Center
San Diego State University
San Diego, CA 92182-7720

http://terminus.sdsu.edu/

Fall 2016
Outline

1. Nonlinear Equations...
   - Breakdown of Global Convergence: Blowup / Cycling
   - Toward Increased Robustness

2. Practical Line Search Methods
   - Convergence
   - Algorithm
   - Convergence Rate

3. Practical Trust-Region Methods
   - Fundamentals
   - Algorithm
   - Convergence
Breakdown of Global Convergence

As we have seen, both Newton’s and Broyden’s method with unit step $\alpha_k \equiv 1$, must be started “close enough” to the solution $\bar{r}(\bar{x}^*) = 0$ in order to converge.

Broyden’s method also requires the more restrictive

$$\|B_0 - J(\bar{x}^*)\| \leq \epsilon.$$

When started too far away from the solution, components of the unknowns $\bar{x}_k$, or function vector $\bar{r}(\bar{x}_k)$, or the Jacobian $J(\bar{x}_k)$ may blow up; — this sort of breakdown is easy to identify. (But not necessarily easy to fix...)

A type of breakdown that is not as easy to detect is cycling, where the sequence of iterates $\{\bar{x}_k\}$ repeat, i.e. $\bar{x}_{k+m} = \bar{x}_k$, for some $m \geq 1$. Clearly, the larger $m$ is, the harder it is to detect cycling (especially in finite precision, where we have $\bar{x}_{k+m} \approx \bar{x}_k$).
The function \( r(x) = -x^5 + x^3 + 4x \) has three non-degenerate real roots. Since the roots are non-degenerate, we expect the fixed point iteration defined by the Newton iteration

\[
x \leftarrow x - \frac{f(x)}{f'(x)} = x - \frac{-x^5 + x^3 + 4x}{-5x^4 + 3x^2 + 4}
\]

to converge quadratically.
Figure: If we start the iteration in $x_0 = 1$, then the Newton iteration cycles $\{1, -1, 1, -1, \ldots\}$ (left figure). On the right we see the rapid convergence to the root $x^* = 0$, for the iteration started at $x_0 = 0.999$. 
If we start the iteration in $x_0 = 1.001$, then the Newton iteration escapes out to the root $1.6004$.

Cycling is quite an exotic occurrence.
Sidenote: Convergence of Newton’s Method can be Surprisingly Complex

Figure: A Newton 4th power fractal. Credit: fractalfoundation.org, Annamarie M.

Figure: The Julia set (in white) for Newton’s method applied to $f(z) = z^3 - 2z + 2$. Start values in the cyan, pink, yellow shaded regions converge to one of the three zeros of $f(z)$. Values from the red/black regions do not converge, they are attracted by a cycle of period 2. Credit: Wikimedia commons.
Increased Robustness

We can make both Newton’s and Broyden’s method more robust by using them in a line-search or trust-region framework. However, in order to use these frameworks, we must define a scalar-valued **merit function** with which we measure progress toward the solution.

The most widely used merit function is the sum-of-squares,

\[
f(\bar{x}) = \frac{1}{2} \|\bar{r}(\bar{x})\|^2 = \frac{1}{2} \sum_{i=1}^{n} r_i^2(\bar{x}).\]

Root of \(\bar{r}(\bar{x}) = 0 \Rightarrow \text{Local minimizer of } f(\bar{x}).\)

Local minimizer of \(f(\bar{x}) \not\Rightarrow \text{Root of } \bar{r}(\bar{x}) = 0.\)
Consider the non-linear function \( r(x) = \sin(5x) - x \) (pictured to the left) and the associated sum-of-squares objective \( f(x) = \frac{1}{2}(\sin(5x) - x)^2 \) (pictured to the right). In this range we have gone from three roots, to seven local minima.
Other Merit Functions

*Figure:* The $l_1$-norm (sum-of-absolute-values) gives us an alternative objective, with its own problems — the derivative does not exist in the optimal points...

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩

Nonlinear Equations: Practical Methods — (10/28)
Practical Line Search Methods

We can build algorithms with global convergence properties by applying the line search approach to the sum-of-squares merit function $f(\bar{x}) = \frac{1}{2} ||\bar{r}(\bar{x})||^2$.

**Note:** Convergence is global in the sense that we guarantee convergence to a stationary point for $f(\bar{x})$, i.e. a point $\bar{x}^*$ such that $\nabla f(\bar{x}^*) = 0$.

From a point $\bar{x}_k$, the search direction $\bar{p}_k$ must be a descent direction for $f(\bar{x})$, i.e.

$$\cos \theta_k = \frac{-\bar{p}_k^T \nabla f(\bar{x}_k)}{||\bar{p}_k|| \ ||\nabla f(\bar{x}_k)||} > 0.$$  

Then we use a line search procedure to identify a step $\alpha_k$, satisfying e.g. the **Wolfe conditions**.
Theorem

Suppose that $J(\bar{x})$ is Lipschitz continuous in a neighborhood $D$ of the level set $\mathcal{L}(\bar{x}_0) = \{ \bar{x} \in \mathbb{R}^n : f(\bar{x}) \leq f(\bar{x}_0) \}$. Suppose that a line search algorithm is applied and that the search directions $\bar{p}_k$ satisfy $\cos \theta_k > 0$, and the step lengths $\alpha_k$ satisfy the Wolfe conditions. Then the Zoutendijk condition holds, i.e.

$$\sum_{k=0}^{\infty} \cos^2 \theta_k \| J_k^T \bar{r}(\bar{x}_k) \|^2 < \infty$$

As long as we can bound $\cos \theta_k \geq \delta > 0$, this guarantees that $\| J_k^T \bar{r}(\bar{x}_k) \| \to 0$.

Further, if $\| J(\bar{x})^{-1} \|$ is bounded on $D$, then $\bar{r}(\bar{x}_k) \to 0$. 
We take a look at the search directions generated by Newton and inexact Newton line-search methods — is the condition \( \cos \theta_k \geq \delta > 0 \) satisfied???

When the Newton-step is well defined, it is a descent direction for \( f(\cdot) \) whenever \( \bar{r}(\bar{x}_k) \neq 0 \), since

\[
\bar{p}_k^T \nabla f(\bar{x}_k) = -\bar{p}_k^T J(\bar{x}_k)^T \bar{r}(\bar{x}_k) = -\|\bar{r}(\bar{x}_k)\|^2 < 0,
\]

and we have

\[
\cos \theta_k = -\frac{\bar{p}_k^T \nabla f(\bar{x}_k)}{\|\bar{p}_k\| \|\nabla f(\bar{x}_k)\|} = \frac{\|\bar{r}(\bar{x}_k)\|^2}{\|J(\bar{x}_k)^{-1}\bar{r}(\bar{x}_k)\| \|J(\bar{x}_k)^T \bar{r}(\bar{x}_k)\|} \geq \frac{1}{\|J(\bar{x}_k)^{-1}\| \|J(\bar{x}_k)^T\|} = \frac{1}{\kappa(J(\bar{x}_k))} = \frac{\lambda_{\min}}{\lambda_{\max}}.
\]

If the **condition number** \( \kappa(J(\bar{x}_k)) \) is uniformly bounded, we have \( \cos \theta_k \geq \delta > 0 \). When \( \kappa(J(\bar{x}_k)) \) is large, the Newton direction may cause poor performance, since \( \cos \theta_k \rightarrow 0 \).
If $J(\bar{x})$ is **ill-conditioned** (close to singular), then we must modify the Newton step in order to ensure that $\cos \theta_k \geq \delta > 0$ holds. For instance, we can add a $\tau_k I$ to $J(\bar{x}_k)^T J(\bar{x}_k)$, and define the modified Newton step to be

$$\bar{p}_k = - \left[ J(\bar{x}_k)^T J(\bar{x}_k) + \tau_k I \right]^{-1} J(\bar{x}_k)^T \bar{r}(\bar{x}_k)$$

Usually, we do not want to do this explicitly. Instead we use the fact that the Cholesky factor of $J(\bar{x}_k)^T J(\bar{x}_k) + \tau_k I$ is identical to $R^T$, where $R$ is the upper triangular factor of the **QR-factorization** of the matrix

$$\begin{bmatrix} J(\bar{x}_k) \\ \sqrt{\tau} I \end{bmatrix}.$$ 

This factorization can be implemented in such a way that repeating the factorization for an updated value of $\tau$ is cheap.
The inexactness does not compromise the global convergence behavior:  
For an inexact Newton step, $\bar{p}_k$, we have,

$$\|\bar{r}(\bar{x}_k) + J(\bar{x}_k)\bar{p}_k\| \leq \eta_k \|\bar{r}(\bar{x}_k)\|.$$  

Squaring this inequality gives

$$2\bar{p}_k^T J(\bar{x}_k)^T \bar{r}(\bar{x}_k) + \|\bar{r}(\bar{x}_k)\|^2 + \|J(\bar{x}_k)\bar{p}_k\|^2 \leq \eta^2 \|\bar{r}(\bar{x}_k)\|^2$$

$$\Rightarrow \bar{p}_k^T \nabla \bar{r}(\bar{x}_k) = \bar{p}_k^T J(\bar{x}_k)^T \bar{r}(\bar{x}_k) \leq \left[ \frac{\eta^2 - 1}{2} \right] \|\bar{r}(\bar{x}_k)\|^2.$$  

We also have,

$$\|\bar{p}_k\| \leq \|J(\bar{x}_k)^{-1}\| \left[ \|\bar{r}(\bar{x}_k) + J(\bar{x}_k)\bar{p}_k\| + \|\bar{r}(\bar{x}_k)\| \right] \leq (\eta+1) \|J(\bar{x}_k)^{-1}\| \|\bar{r}(\bar{x}_k)\|,$$

$$\|\nabla \bar{r}(\bar{x}_k)\| = \|J(\bar{x}_k)^T \bar{r}(\bar{x}_k)\| \leq \|J(\bar{x}_k)\| \|\bar{r}(\bar{x}_k)\|.$$  

Putting it all together...
We can now write down an estimate for $\cos \theta_k$ for the inexact Newton directions

$$
\cos \theta_k = -\frac{\bar{p}_k^T \nabla \bar{r}(\bar{x}_k)}{\| \bar{p}_k \| \| \nabla \bar{r}(\bar{x}_k) \|} \geq \frac{1 - \eta^2}{2(1 + \eta) \| J(\bar{x}_k) \| \| J(\bar{x}_k)^{-1} \|} \geq \frac{1 - \eta}{2\kappa(J(\bar{x}_k))}.
$$

This is the same bound (with a different constant) as the bound for Newton’s method.

— Hence, inexact Newton converges when Newton’s method does.
Given $\delta \in (0, 1)$ and $c_1, c_2$ with $0 < c_2 < c_1 < \frac{1}{2}$, and $\bar{x}_0 \in \mathbb{R}^n$:

**Algorithm: Line Search Newton for Nonlinear Equations**

while( $\|\bar{r}(\bar{x}_k)\| > \epsilon$ )
  if $\bar{p} = -J(\bar{x}_k)^{-1}\bar{r}(\bar{x}_k)$ satisfies $\cos \theta_k \geq \delta$
    Accept $\bar{p}_k = \bar{p}$
  else
    Search for $\bar{p}_k(\tau_k)$ satisfying $\cos \theta_k(\tau_k) \geq \delta$
    $\bar{p}_k(\tau_k) = -[J(\bar{x}_k)^TJ(\bar{x}_k) + \tau_k I]^{-1}J(\bar{x}_k)^T\bar{r}(\bar{x}_k)$
  endif
  if $\alpha = 1$ satisfies the Wolfe conditions
    $\alpha_k = 1$
  else
    Perform a line-search to find $\alpha_k > 0$ satisfying the Wolfe conditions.
  endif
  $\bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{p}_k$
endwhile( $k = k + 1$ )
Theorem

Suppose that a line search algorithm that uses Newton search directions yields a sequence \( \{\bar{x}_k\} \) that converges to \( \bar{x}^* \), where \( \bar{r}(\bar{x}^*) = 0 \) and \( J(\bar{x}^*) \) is non-singular. Suppose also that there is an open neighborhood \( D \) of \( \bar{x}^* \) such that the components \( r_i(\bar{x}) \) are twice differentiable, with \( \|\nabla r_i(\bar{x})\| \) bounded for \( \bar{x} \in D \). If the unit step length \( \alpha_k \) is accepted whenever it satisfies the Wolfe conditions, with \( c_1 < \frac{1}{2} \), then the convergence is Q-quadratic; that is \( \|\bar{x}_{k+1} - \bar{x}^*\| = O(\|\bar{x}_k - \bar{x}^*\|^2) \).

Note: This theorem applies to any algorithm which eventually uses the Newton search direction.
The most commonly used trust-region method for nonlinear equations is simply “standard trust-region” applied to the merit function $f(\bar{x}) = \frac{1}{2}||\bar{r}(\bar{x})||^2$, using $B_k = J(\bar{x}_k)^T J(\bar{x}_k)$ as the approximate Hessian in the model function $m_k(\bar{p})$. (Levenberg-Marquardt style...)

Global convergence follows directly from previously proved theorems for convergence of trust-region methods.

Rapid local convergence can be shown under the assumption that the Jacobian $J(\bar{x})$ is Lipschitz continuous.

In the next few slides we take a closer look at the trust-region method for nonlinear equations.
Our model function is given by

\[
m_k(\bar{p}) = \frac{1}{2} \| \bar{r}(\bar{x}_k) + J(\bar{x}_k)\bar{p} \|_2^2
\]

\[
= f(\bar{x}_k) + \bar{p}^T J(\bar{x}_k)^T \bar{r}(\bar{x}_k) + \frac{1}{2} \bar{p}^T J(\bar{x}_k)^T J(\bar{x}_k)\bar{p}.
\]

As usual we generate the step \( \bar{p}_k \) by solving the sub-problem

\[
\bar{p}_k = \arg \min_{\bar{p} \in \mathbb{R}^n} m_k(\bar{p}), \quad \text{subject to } \| \bar{p} \| \leq \Delta_k.
\]

We can express \( \rho_k \), the ratio of actual to predicted reduction as

\[
\rho_k = \frac{\| \bar{r}(\bar{x}_k) \|^2 - \| \bar{r}(\bar{x}_k + \bar{p}_k) \|^2}{\| \bar{r}(\bar{x}_k) \|^2 - \| \bar{r}(\bar{x}_k) + J(\bar{x}_k)\bar{p}_k \|^2}.
\]
Given $\Delta > 0$, $\Delta_0 \in (0, \Delta)$ and $\eta \in [0, \frac{1}{4})$

**Algorithm: Trust Region for Nonlinear Equations**

\[
\text{while} (\|\bar{r}(\bar{x}_k)\| > \epsilon) \\
\bar{p}_k = \arg\min_{\bar{p} \in \mathbb{R}^n} m_k(\bar{p}), \quad \text{subject to } \|\bar{p}\| \leq \Delta_k \\
\rho_k = \frac{\|\bar{r}(\bar{x}_k)\|^2 - \|\bar{r}(\bar{x}_k + \bar{p}_k)\|^2}{\|\bar{r}(\bar{x}_k)\|^2 - \|\bar{r}(\bar{x}_k) + J(\bar{x}_k)\bar{p}_k\|^2} \\
\text{if} (\rho_k < \frac{1}{4}) \\
\Delta_{k+1} = \frac{1}{4} \|\bar{p}_k\| \\
\text{else} \\
\text{if} (\rho_k > \frac{3}{4} \text{ and } \|\bar{p}_k\| = \Delta_k) \\
\Delta_{k+1} = \min(2\Delta_k, \Delta) \\
\text{else} \\
\Delta_{k+1} = \Delta_k \\
\text{endif} \\
\text{endif} \\
\text{if} (\rho_k > \eta) \{ \bar{x}_{k+1} = \bar{x}_k + \bar{p}_k \} \text{ else } \{ \bar{x}_{k+1} = \bar{x}_k \} \text{ endif} \\
\text{endwhile}(k = k + 1)
Trust Region for Nonlinear Equations

We take a closer look at the solution of the subproblem [TR] using the dogleg method.

The **Cauchy point** is given by

\[
\bar{p}_k^C = -\tau_k \left( \frac{\Delta_k}{\|J(\bar{x}_k)^T \bar{r}(\bar{x}_k)\|} \right) J(\bar{x}_k)^T \bar{r}(\bar{x}_k),
\]

where

\[
\tau_k = \min \left\{ 1, \frac{\|J(\bar{x}_k)^T \bar{r}(\bar{x}_k)\|^3}{\Delta_k \bar{r}(\bar{x}_k)^T J(\bar{x}_k)(J(\bar{x}_k)^T J(\bar{x}_k)) J(\bar{x}_k)^T \bar{r}(\bar{x}_k)} \right\}.
\]

For the **full step**, we use the fact that the model Hessian \(B_k = J(\bar{x}_k)^T J(\bar{x}_k)\) is symmetric semi-definite; when \(J(\bar{x}_k)\) has full rank we get

\[
\bar{p}_k^J = - \left[ J(\bar{x}_k)^T J(\bar{x}_k) \right]^{-1} \left[ J(\bar{x}_k)^T \bar{r}(\bar{x}_k) \right] = -J(\bar{x}_k)^{-1} \bar{r}(\bar{x}_k).
\]
The dogleg selection of $\bar{p}_k$ is given by:

Algorithm: Dogleg Selection

Calculate $\bar{p}_k^c$

if( $\|\bar{p}_k^c\| = \Delta_k$ )

   $\bar{p}_k = \bar{p}_k^c$

else

   Calculate $\bar{p}_k^J$

   if( $\|\bar{p}_k^J\| < \Delta_k$ )

      $\bar{p}_k = \bar{p}_k^J$

   else

      $\bar{p}_k = \bar{p}_k^c + \tau(\bar{p}_k^J - \bar{p}_k^c)$, where $\tau \in [0, 1] : \|\bar{p}_k\| = \Delta_k$

endif

endif
From previous results we know that the exact solution of the subproblem [TR] has the form

$$\bar{p}_k = - \left[ J(\bar{x}_k)^T J(\bar{x}_k) + \lambda_k I \right]^{-1} \left[ J(\bar{x}_k)^T \bar{r}(\bar{x}_k) \right]$$

for some $\lambda_k \geq 0$, and that $\lambda_k = 0$ if $\|\bar{p}_k^J\| \leq \Delta_k$.

Note that this is the same linear system that gives the Levenberg-Marquardt step $\bar{p}_k^{LM}$ in the discussion on nonlinear least squares.

In a sense the LM-approach for non-linear equations is a special case of the LM-approach for nonlinear least squares problems.

We can identify an approximation of $\lambda_k$ using the Cholesky factorization, e.g. `modellhess` in the project code; alternatively we can base the search on the QR-factorization.
Trust Region for Nonlinear Equations

The dogleg method has the **advantage** over methods trying to attain the exact solution to the subproblem in that **only one linear system needs to be solved per iteration**.

Global convergence for the trust-region algorithm is described in the two following theorems (which should look somewhat familiar...): –
Theorem

Let $\eta = 0$ in the trust-region algorithm. Suppose that $J(\bar{x})$ is continuous in a neighborhood $\mathcal{D}$ of the level set $\mathcal{L}(\bar{x}_0) = \{\bar{x} \in \mathbb{R}^n : f(\bar{x}) \leq f(\bar{x}_0)\}$ and that $\|J(\bar{x})\|$ is bounded above on $\mathcal{L}(\bar{x}_0)$. Suppose in addition that all approximate solutions of the trust-region subproblem satisfy $(c_1 > 0, \gamma \geq 1)$

$$m_k(0) - m_k(\bar{p}_k) \geq c_1 \|J(\bar{x}_k)^T \bar{r}(\bar{x}_k)\| \min \left\{ \Delta_k, \frac{J(\bar{x}_k)^T \bar{r}(\bar{x}_k)}{J(\bar{x}_k)^T J(\bar{x}_k)} \right\},$$

$$\|\bar{p}_k\| \leq \gamma \Delta_k.$$

We then have that

$$\lim \inf_{k \to \infty} \|J(\bar{x}_k)^T \bar{r}(\bar{x}_k)\| = 0.$$
Theorem

Let \( \eta \in (0, \frac{1}{4}) \) in the trust-region algorithm. Suppose that \( J(\bar{x}) \) is Lipschitz continuous in a neighborhood \( D \) of the level set \( \mathcal{L}(\bar{x}_0) = \{ \bar{x} \in \mathbb{R}^n : f(\bar{x}) \leq f(\bar{x}_0) \} \) and that \( \| J(\bar{x}) \| \) is bounded above on \( \mathcal{L}(\bar{x}_0) \). Suppose in addition that all approximate solutions of the trust-region subproblem satisfy \( (c_1 > 0, \gamma \geq 1) \)

\[
 m_k(0) - m_k(\bar{p}_k) \geq c_1 \| J(\bar{x}_k)^T \bar{r}(\bar{x}_k) \| \min \left\{ \Delta_k, \frac{J(\bar{x}_k)^T \bar{r}(\bar{x}_k)}{J(\bar{x}_k)^T J(\bar{x}_k)} \right\},
\]

\[ \| \bar{p}_k \| \leq \gamma \Delta_k. \]

We then have that

\[
 \lim_{k \to \infty} \| J(\bar{x}_k)^T \bar{r}(\bar{x}_k) \| = 0
\]
Finally, we state a result regarding the convergence rate. Note that the result requires exact solution of the subproblem.

Theorem

Suppose that the sequence \( \{\overline{x}_k\} \) generated by the trust-region algorithm converges to a non-degenerate solution \( \overline{x}^* \) of the problem \( \overline{r}(\overline{x}) = 0 \). Suppose also that \( J(\overline{x}) \) is Lipschitz continuous in an open neighborhood \( D \) of \( \overline{x}^* \) and that the trust-region subproblem is solved exactly for all sufficiently large \( k \). Then the sequence \( \{\overline{x}_k\} \) converges quadratically to \( \overline{x}^* \).

Thus we can design a globally convergent method which converges quadratically! — Robustness and Speed in the same algorithm!