Set 10: Nonlinear Equations Part 2

Kyle A. Gallivan

Department of Mathematics

Florida State University

Foundations of Computational Math 1

Fall 2012

Systems of Nonlinear Equations

Let $F(x) : \mathbb{R}^n \to \mathbb{R}^n$ be a vector-valued function of n variables.

The problem is to find $x^* \in \mathbb{R}^n$ such that

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

i.e., solve n equations in n unknowns.

Many of the algorithms we have discussed for linear systems and nonlinear scalar equations have generalizations to solve nonlinear systems.

Some Types of Methods

- Generalized Linear Methods
- Newton-like Methods
- Quasi-Newton Methods (Secant Methods, Modification Methods)
- Minimization Methods
- Continuation Methods (Homotopy Methods)
- Nonlinear Conjugate Gradient Methods

Fixed Point Methods

To find a root, $x^* \in \mathbb{R}^n$, of $F(x) : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ we can employ generalizations of the fixed point methods discussed for scalar nonlinear equations and linear sytems of equations.

We define the methods by defining

$$G(x) : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$x^{(k+1)} = G(x^{(k)})$$

$$x^* = G(x^*)$$

must have $\quad x^{(k)} \rightarrow x^*$

Fixed Point/Relaxation Methods

Definition 10.1. Let $G(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. Denote the components of $G(x)$ and x as

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, \quad \text{and} \quad G(x) = \begin{pmatrix} \gamma_1(x) \\ \gamma_2(x) \\ \vdots \\ \gamma_n(x) \end{pmatrix} = \begin{pmatrix} \gamma_1(\xi_1, \xi_2, \ldots, \xi_n) \\ \gamma_2(\xi_1, \xi_2, \ldots, \xi_n) \\ \vdots \\ \gamma_n(\xi_1, \xi_2, \ldots, \xi_n) \end{pmatrix}$$

The Jacobian of G, denoted $J_G(x)$, is a matrix whose elements are functions of x defined by

$$\gamma_{ij} = e_i^T J_G(x) e_j = \frac{\partial \gamma_i}{\partial \xi_j}(x)$$

(Note the use of double subscripts to denote partial derivatives of the corresponding function with a single subscript.)

Convergence

Definition 10.2. $G(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a contraction mapping on D if $\exists 0 < L < 1$ such that $\forall x,\ y \in D$

$$\|G(x) - G(y)\| \leq L\|x - y\|$$

Theorem 10.1. *If the continuous function $G(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a contraction mapping on a closed set $D_0 \subset D$ for which $x \in D_0 \to G(x) \in D_0$ then*

- *$\exists p \in D_0$ such that $p = G(p)$ and p is unique.*
- $\forall x^{(0)} \in D_0,\ \ p = \lim_{k\to\infty} x^{(k)} = \lim_{k\to\infty} G(x^{(k-1)})$

Convergence

The scalar mean value theorem we used earlier does not generalize simply so we generalize the earlier local result around a root:

Theorem 10.2. *Suppose the continuously differentiable function* $G(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *has a fixed point* x^* *on the interior of* D*. If* $\rho(J_G(x^*)) < 1$ *then there exists a neighborhood of* x^**,* $S \subset D$*, such that the sequence* $x^{(k+1)} = G(x^{(k)})$ *with* $x^{(0)} \in S$*, lies in* D *and converges to* x^*.

Corollary 10.3. *If* $\|J_G(x^*)\| < 1$ *then there exists a neighborhood of* x^**,* $S \subset D$*, such that the sequence* $x^{(k+1)} = G(x^{(k)})$ *with* $x^{(0)} \in S$*, lies in* D *and converges to* x^*.

Solving Nonlinear Systems

Let $F(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. Denote the components of $F(x)$ and x as

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, \quad \text{and} \quad F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix} = \begin{pmatrix} f_1(\xi_1, \xi_2, \ldots, \xi_n) \\ f_2(\xi_1, \xi_2, \ldots, \xi_n) \\ \vdots \\ f_n(\xi_1, \xi_2, \ldots, \xi_n) \end{pmatrix}$$

Nonlinear Jacobi

To advance $x^{(k)} \to x^{(k+1)}$:

for $i = 1, \ldots, n$
solve the nonlinear scalar equation for τ:

$$f_i(\xi_1^{(k)}, \xi_2^{(k)}, \ldots, \xi_{i-1}^{(k)}, \tau, \xi_{i+1}^{(k)}, \ldots, \xi_n^{(k)})$$

$$\xi_i^{(k+1)} \leftarrow \tau$$

- all components of x other than ξ_i are kept at their values from $x^{(k)}$
- We can use any nonlinear scalar equation solver to find $\xi_i^{(k+1)}$
- equations can be ordered in any manner before applying method

Nonlinear Gauss-Seidel

To advance $x^{(k)} \to x^{(k+1)}$:

for $i = 1, \ldots, n$
solve the nonlinear scalar equation for τ:

$$f_i(\xi_1^{(k+1)}, \xi_2^{(k+1)}, \ldots, \xi_{i-1}^{(k+1)}, \tau, \xi_{i+1}^{(k)}, \ldots, \xi_n^{(k)})$$
$$\xi_i^{(k+1)} \leftarrow \tau$$

- All ξ_j with $j < i$ use iteration $k+1$ values.
- We can use any nonlinear scalar equation solver to find $\xi_i^{(k+1)}$
- equations can be ordered in any manner before applying method

Inner Nonlinear Iteration

- Given the choice for the inner or secondary iteration, the method is called:
  - nonlinear Jacobi-Newton
  - nonlinear Jacobi-secant
  - nonlinear Jacobi-Regula-Falsi
- These can be iterated to convergence but more typically one step (or a small number of steps) are done.
- Any customized scalar fixed point method can be used also.
- This can be generalized to a block Outer/Inner iteration.

Nonlinear one-step Jacobi-Newton

$$f_i(\xi_1^{(k)}, \xi_2^{(k)}, \ldots, \xi_{i-1}^{(k)}, \tau, \xi_{i+1}^{(k)}, \ldots, \xi_n^{(k)}) = 0$$

$\tau^{(0)} = \xi_i^{(k)}$ take one step of scalar Newton's method

$$\xi_i^{(k+1)} = \xi_i^{(k)} - \frac{f_i(\xi_1^{(k)}, \xi_2^{(k)}, \ldots, \xi_{i-1}^{(k)}, \xi_i^{(k)}, \xi_{i+1}^{(k)}, \ldots, \xi_n^{(k)})}{f_{ii}(\xi_1^{(k)}, \xi_2^{(k)}, \ldots, \xi_{i-1}^{(k)}, \xi_i^{(k)}, \xi_{i+1}^{(k)}, \ldots, \xi_n^{(k)})}$$

$$1 \le i \le n, \quad k = 0, 1, \ldots$$

Nonlinear one-step Gauss-Seidel-Newton

$$f_i(\xi_1^{(k+1)}, \xi_2^{(k+1)}, \ldots, \xi_{i-1}^{(k+1)}, \tau, \xi_{i+1}^{(k)}, \ldots, \xi_n^{(k)}) = 0$$

$\tau^{(0)} = \xi_i^{(k)}$ take one step of scalar Newton's method

$$\xi_i^{(k+1)} = \xi_i^{(k)} - \frac{f_i(\xi_1^{(k+1)}, \xi_2^{(k+1)}, \ldots, \xi_{i-1}^{(k+1)}, \xi_i^{(k)}, \xi_{i+1}^{(k)}, \ldots, \xi_n^{(k)})}{f_{ii}(\xi_1^{(k+1)}, \xi_2^{(k+1)}, \ldots, \xi_{i-1}^{(k+1)}, \xi_i^{(k)}, \xi_{i+1}^{(k)}, \ldots, \xi_n^{(k)})}$$

$$1 \le i \le n, \quad k = 0, 1, \ldots$$

Example

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sin(\xi_1\xi_2) - \frac{\xi_2}{4\pi} - \frac{\xi_1}{2} \\ (1-\frac{1}{4\pi})(e^{2\xi_1} - e) + \frac{e\xi_2}{\pi} - 2e\xi_1 \end{pmatrix}$$

$$\begin{pmatrix} \xi_1^* \\ \xi_2^* \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \pi \end{pmatrix}$$

There are other roots.

Example

$$F(x) = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sin(\xi_1\xi_2) - \frac{\xi_2}{4\pi} - \frac{\xi_1}{2} \\ (1 - \frac{1}{4\pi})(e^{2\xi_1} - e) + \frac{e\xi_2}{\pi} - 2e\xi_1 \end{pmatrix}$$

Partial derivatives needed for this ordering, i.e., f_1 used to update ξ_1 and f_2 used to update ξ_2, are

$$\frac{\partial f_1}{\partial \xi_1} = f_{11} = 0.5(\xi_2 \cos(\xi_1\xi_2) - 1)$$

$$\frac{\partial f_2}{\partial \xi_2} = f_{22} = e/\pi$$

Example Using Nonlinear Jacobi-Newton (one-step)

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	0.4000000000000	3.0000000000000	0.0423500623420
1	-0.2267625048348	3.0374309455933	1.9943566443881
2	0.4366354612399	0.7909068337050	1.9904165772672
3	-0.4390818245352	3.0876410928077	2.9839220525587
4	-3.4570867970408	-0.3092792384391	16.1751680301500
5	0.3586161952341	-18.8309812670440	18.8911411461812
6	0.4814738130303	2.9654659715017	0.1438362083198
7	0.5444314561956	3.1303689472401	0.0318727064360
8	0.5080484185447	3.1520456534478	0.0077834526673
9	0.4987636192426	3.1452402270620	0.0037130353460
10	0.4994216121485	3.1409656301001	0.0004457394573

Example Using Nonlinear Jacobi-Newton (one-step)

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
8	0.5080484185447	3.1520456534478	0.0077834526673
9	0.4987636192426	3.1452402270620	0.0037130353460
10	0.4994216121485	3.1409656301001	0.0004457394573
11	0.5001020779275	3.1413015257463	0.0002973197317
12	0.5000463497582	3.1416436322892	0.0000363488313
13	0.4999919011184	3.1416158160445	0.0000236487155
14	0.4999963136769	3.1415886037697	0.0000028868968
15	0.5000006446415	3.1415908103497	0.0000018819764

Example Using Nonlinear Gauss-Seidel-Newton (one-step)

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	0.4000000000000	3.0000000000000	0.0423500623420
1	-0.2267625048348	0.7909068337050	0.0387511555355
2	-0.5762196655678	-1.0649063919241	0.6607678058899
3	0.1302442337308	2.3296481526139	0.1011152197474
4	0.2955694419489	2.8275172703977	0.0018990298269
5	0.2998072093702	2.8377916110249	0.0001730192982
6	0.2994100158056	2.8368345201176	0.0000186907406
7	0.2994528063993	2.8369376885012	0.0000019877794

Notice different root found for x_0 and the behavior of $\|f(x^{(k)})\|_2$ for both iterations.

Example Using Nonlinear Gauss-Seidel-Newton (one-step)

$(1/2, \pi)$ found for $x_0 = (0.7, 4.0)$

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	0.7000000000000	4.0000000000000	1.0177129773898
1	0.4899654309342	3.1359969213410	0.0051690636118
2	0.5015479596432	3.1423527615522	0.0008413429180
3	0.4998925537886	3.1415388637238	0.0000579703628
4	0.5000086274078	3.1415969668632	0.0000046571573
5	0.4999993139493	3.1415923105617	0.0000003703213

Example

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sin(\xi_1\xi_2) - \frac{\xi_2}{4\pi} - \frac{\xi_1}{2} \\ (1-\frac{1}{4\pi})(e^{2\xi_1}-e) + \frac{e\xi_2}{\pi} - 2e\xi_1 \end{pmatrix}$$

$$\xi_1 = \gamma_1(\xi_1, \xi_2) = \sin(\xi_1\xi_2) - \frac{\xi_2}{2\pi}$$

$$\xi_2 = \gamma_2(\xi_1, \xi_2) = 2\pi\xi_1 - (\pi - \frac{1}{4})(e^{2\xi_1 - 1} - 1)$$

$$\xi_1^{(k+1)} = \gamma_1(\xi_1^{(k)}, \xi_2^{(k)}) = \sin(\xi_1^{(k)}\xi_2^{(k)}) - \frac{\xi_2^{(k)}}{2\pi}$$

$$\xi_2^{(k+1)} = \gamma_2(\xi_1^{(k)}, \xi_2^{(k)}) = 2\pi\xi_1^{(k)} - (\pi - \frac{1}{4})(e^{2\xi_1^{(k)} - 1} - 1)$$

Jacobi Schedule on a splitting that has ξ_i on both sides.

Example using Nonlinear Splitting with Jacobi Schedule

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	0.4000000000000	3.0000000000000	0.0423500623420
1	0.4545742566915	3.0374309455933	0.0643322652770
2	0.4985710524398	3.1072995231908	0.0292338530055
3	0.5052249367379	3.1408663824161	0.0037989127996
4	0.4999868281924	3.1440466898242	0.0021374458581
5	0.4996087251728	3.1415860666826	0.0002556151918
6	0.5000002887899	3.1413961310246	0.0001708710026
7	0.5000312727988	3.1415927979842	0.0000206029822
8	0.4999999721856	3.1416082843332	0.0000135923909
9	0.4999975122601	3.1415926396826	0.0000016378601
10	0.5000000021828	3.1415914096840	0.0000010816803

Example

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sin(\xi_1\xi_2) - \frac{\xi_2}{4\pi} - \frac{\xi_1}{2} \\ (1-\frac{1}{4\pi})(e^{2\xi_1} - e) + \frac{e\xi_2}{\pi} - 2e\xi_1 \end{pmatrix}$$

$$\xi_1 = \gamma_1(\xi_1, \xi_2) = \sin(\xi_1\xi_2) - \frac{\xi_2}{2\pi}$$

$$\xi_2 = \gamma_2(\xi_1, \xi_2) = 2\pi\xi_1 - (\pi - \frac{1}{4})(e^{2\xi_1 - 1} - 1)$$

$$\xi_1^{(k+1)} = \gamma_1(\xi_1^{(k)}, \xi_2^{(k)}) = \sin(\xi_1^{(k)}\xi_2^{(k)}) - \frac{\xi_2^{(k)}}{2\pi}$$

$$\xi_2^{(k+1)} = \gamma_2(\xi_1^{(k+1)}, \xi_2^{(k)}) = 2\pi\xi_1^{(k+1)} - (\pi - \frac{1}{4})(e^{2\xi_1^{(k+1)} - 1} - 1)$$

Gauss-Seidel Schedule on a splitting that has ξ_i on both sides.

Example Using Nonlinear Gauss-Seidel Schedule

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	0.4000000000000	3.0000000000000	0.0423500623420
1	0.4545742566915	3.1072995231908	0.0191903422181
2	0.4929549411276	3.1377844315468	0.0036814651710
3	0.5003178714697	3.1417510048546	0.0001718273675
4	0.4999742167347	3.1415797581127	0.0000139159103
5	0.5000020485554	3.1415936778432	0.0000011057972

Newton-like Methods

Recall, we used the idea of replacing a scalar root finding problem with a series of linear root finding problems that were trivially solved in order to converge to a scalar root x^*.

This idea can be used to approximate locally $F(x)$ and generate several methods.

An integral generalization of the scalar mean value theorem provides an error expression that is useful conceptually and theoretically.

Jacobian Matrix

Definition 10.3. Let $F(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. Denote the components of $F(x)$ and x as

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, \quad \text{and} \quad F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix} = \begin{pmatrix} f_1(\xi_1, \xi_2, \ldots, \xi_n) \\ f_2(\xi_1, \xi_2, \ldots, \xi_n) \\ \vdots \\ f_n(\xi_1, \xi_2, \ldots, \xi_n) \end{pmatrix}$$

The Jacobian of F, denoted $J_F(x)$, is a matrix whose elements are functions of x defined by

$$f_{ij} = e_i^T J_F(x) e_j = \frac{\partial f_i}{\partial \xi_j}(x)$$

Jacobian Matrix

Definition 10.4. Let $F(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. F is Lipschitz continuously differentiable if $\forall x, \ \ y \in D$ there exists a constant L such that

$$\|J_F(x) - J_F(y)\| \leq L\|x - y\|$$

i.e., the Jacobian of F is Lipschitz continuous on D.

Local Linear Approximations

Theorem 10.4 (Nocedal and Wright,1999)**.** *Suppose the function $F(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable on a convex open set D, $x \in D$ and $x + p \in D$. Define a local model*

$$M_k(p) = F(x^{(k)}) + J_F(x^{(k)})p.$$

We then have the following expression for the difference

$$F(x^{(k)} + p) - M_k(p) = \int_0^1 \left(J_F(x^{(k)} + \tau p) - J_F(x^{(k)}) \right) p d\tau$$

$$\forall 0 \leq \tau \leq 1 \quad \lim_{p \to 0} \| \int_0^1 \left(J_F(x^{(k)} + \tau p) - J_F(x^{(k)}) \right) p d\tau \| = 0$$

If in addition, J_F is Lipschitz continuous then

$$\| \int_0^1 \left(J_F(x^{(k)} + \tau p) - J_F(x^{(k)}) \right) p d\tau \| = O(\|p\|^2)$$

Newton's Method

Newton's Method:

Choose $x^{(0)}$

loop over k until convergence

Solve $J_F(x^{(k)})p_k = -F(x^{(k)})$

$x^{(k+1)} = x^{(k)} + p_k$

end

Newton's Method Example

$$F(x) = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sin(\xi_1\xi_2) - \frac{\xi_2}{4\pi} - \frac{\xi_1}{2} \\ (1 - \frac{1}{4\pi})(e^{2\xi_1} - e) + \frac{e\xi_2}{\pi} - 2e\xi_1 \end{pmatrix} \rightarrow x^* = \begin{pmatrix} \frac{1}{2} \\ \pi \end{pmatrix}$$

$$J_F = \begin{pmatrix} 0.5(\xi_2\cos(\xi_1\xi_2) - 1) & 0.5\xi_1\cos(\xi_1\xi_2) - 1/(4\pi) \\ (2 - (1/(2\pi)))e^{2\xi_1} - 2e & e/\pi \end{pmatrix}$$

Newton's Method Example

Different root found for $x_0 = (0.4, 3.0)$

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	0.4000000000000	3.0000000000000	0.0423500623420
1	-0.4305398475234	1.7514947665888	1.7634895849962
2	-0.2454702651118	0.7331660836104	0.0393269819353
3	-0.2613873006594	0.6189008465340	0.0009518879232
4	-0.2606005650094	0.6225252774941	0.0000015686338
5	-0.2605992900257	0.6225308965998	0.0000000000040

Newton's Method Example

$(1/2, \pi)$ root found for $x_0 = (0.7, 4.0)$

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	0.7000000000000	4.0000000000000	1.0177129773898
1	0.6426820605004	3.1104442441014	0.1164311300807
2	0.5147124140743	3.2635739576641	0.1022549934586
3	0.5040997590098	3.1442111438842	0.0023795097510
4	0.5000756408858	3.1417246093516	0.0000947423645
5	0.5000000377836	3.1415927055406	0.0000000367207

Newton's Method Example

Yet another root found for $x_0 = (1, 4)$

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	1.0000000000000	4.0000000000000	2.6136151459905
3	2.2005371775434	-9.3201168573739	52.5271716381808
5	1.6892006744608	-15.2297863459313	2.1290671649331
6	1.6546583732832	-15.7503640566293	0.0679620939460
7	1.6544853014803	-15.8141281148623	0.0021702215055
8	1.6545817935158	-15.8191396416826	0.0000205991963
9	1.6545827186773	-15.8191882276008	0.0000000019375

Another Newton's Method Example

$$F(x) = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} (\xi_1 + 3)(\xi_2^3 - 7) + 18 \\ \sin(\xi_2 e^{\xi_1} - 1) \end{pmatrix} \rightarrow x^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$J_F = \begin{pmatrix} (\xi_2^3 - 7) & 3\xi_2^2(\xi_1 + 3) \\ \xi_2 e^{\xi_1} \cos(\xi_2 e^{\xi_1} - 1) & e^{\xi_1} \cos(\xi_2 e^{\xi_1} - 1) \end{pmatrix}$$

Another Newton's Method Example

k	$\xi_1^{(k)}$	$\xi_2^{(k)}$	$\|f(x^{(k)})\|_2$
0	-0.5000000000000	1.4000000000000	7.3615341974672
1	-0.0553151357177	1.0280665838357	0.5874890107585
2	-0.0001403508964	1.0001574043270	0.0022589653109
3	-0.0000000177908	1.0000000055514	0.0000001571844

Newton's Method

- If $x^{(0)}$ is far from the root the behavior of Newton can be erratic. (Damped Newton, i.e., add scale α_k)
- Requires knowledge of Jacobian.
- Requires evaluation of Jacobian.
- Requires solution of equation with Jacobian.
- Root x^* may be degenerate, i.e., $J_F(x^*)$ may be singular so convergence slows and ill-conditioning may be a local problem increasing the complexity of the linear system solution process.

Newton's Method

Theorem 10.5. *Suppose the function $F(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable on a convex open set D. Let $x^* \in D$ be a nondegenerate solution of $F(x) = 0$. If $x^{(0)} \in D$ is sufficiently close to x^* the Newton's method converges superlinearly. If in addition $F(x)$ is Lipschitz continuously differentiable then for all $x^{(k)}$ sufficiently close to x^* satisfy*

$$\|x^{(k+1)} - x^*\| = O(\|x^{(k)} - x^*\|^2)$$

indicating quadratic convergence.

Inexact Newton Methods

Instead of solving

$$J_F(x^{(k)})p_k = -F(x^{(k)})$$

exactly, inexact Newton methods require the residuals to satisfy

$$\|r(x^{(k)})\| = \|J_F(x^{(k)})p_k + F(x^{(k)})\| \leq \eta_k \|F(x^{(k)})\|, \quad 0 \leq \eta_k \leq \eta$$

- η_k is called the forcing parameter.
- $0 \leq \eta < 1$ is constant
- inexact Jacobians or factorizations, e.g., from previous steps
- iterative methods

Newton's Method

Theorem 10.6. *Suppose the function $F(x) : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable on a convex open set D. Let $x^* \in D$ be a nondegenerate solution of $F(x) = 0$. If $x^{(0)} \in D$ is sufficiently close to x^* then for the inexact Newton's method:*

- *If η is sufficiently small (depends on $\|J_F(x^*)\|$) then convergence to x^* is linear.*
- *If $\eta_k \to 0$ then convergence to x^* is superlinear.*
- *If, in addition, $J_F(x)$ is Lipschitz continuous on D and $\eta_k = O(\|F(x^{(k)})\|)$ then convergence is quadratic.*

Some Types of Methods

An efficient (and popular, especially for optimization) alternative to Newton and Inexact Newton are methods that use an idea of a secant in $\mathbb{R}^n$. They are called, essentially equivalently,

- Secant Methods
- Modification Methods
- Quasi-Newton Methods

See Ortega and Rheinboldt for a detailed discussion of the convergence theory of these methods.