MATH: Operations Research **2014-15 First Term**

Handout 12: Newton's method and variants

Instructor: Junfeng Yang *December 8, 2014*

Contents

12.1 Newton's method for nonlinear equations

Newton's method Newton's method is an iterative method for solving systems of nonlinear equations, i.e.,

$$\text{find } x^* \text{ such that } \quad F(x^*) = 0,$$

where $F : R^n \to R^n$ is a nonlinear mapping. Suppose the current point is x_k, Newton's method determines the next point x_{k+1} via solving the linear approximation of F at x_k, i.e.,

$$F(x) \approx F(x_k) + F'(x_k)(x - x_k) = 0,$$

which yields

$$x_{k+1} = x_k - (F'(x_k))^{-1} F(x_k).$$

Here $F'(x)$ denotes the Jacobian matrix of F at x.

12.1.1 Univariate case

We begin with the most simple case, i.e., solving a nonlinear equation in one unknown.

Example 12.1.1 *Suppose we want to find the square root of* 3. *This problem can be viewed as finding a root of* $f(x) = x^2 - 3$, $x \in R$. *We start at* $x_0 \in R$. *Then Newton's method generates a sequence of points via*

$$x_{k+1} = x_k - \frac{x_k^2 - 3}{2x_k}, \quad k = 0, 1, 2, \dots$$

If $x_0 = 2$, *then* $x_1 = 1.75$, $x_2 = 1.7321428$, $x_3 = 1.7320508$, ... *converges very fast to* $\sqrt{3}$*!* $\lim_{k\to\infty} x_k = \sqrt{3}$ *if* $x_0 > 0$ *and* $\lim_{k\to\infty} x_k = -\sqrt{3}$ *if* $x_0 < 0$. *What if* $x_0 = 0$*? Newton's method breaks down!*

Example 12.1.2 *Suppose we use Newton's method to find a root of*

$$f(x) = x/\sqrt{1 + x^2}, \quad x \in R.$$

Clearly $x^* = 0$. *Let* $x_0 \in R$ *be the initial point. The Newton's method iterates as follows*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} = -x_k^3, \quad k = 0, 1, 2, \dots$$

For any $k \geq 1$, *it hold that* $x_k = (-1)^k x_0^{3k}$*; If* $|x_0| < 1$, *then* $\lim_{k\to\infty} x_k = 0$*; Oscillate if* $|x_0| = 1$*; Diverge if* $|x_0| > 1$.

Theorem 12.1 (Local convergence) *Let* $f : R \to R$ *be a nonlinear mapping satisfying the following assumptions:*

1. f *is continuously differentiable, and* f' *is Lipschitz continuous on* R, *i.e., for some* $L > 0$,

$$|f'(x) - f'(y)| \leq L|x - y|, \quad \forall x, y \in R.$$

2. *There exists* $\rho > 0$ *such that* $|f'(x)| \geq \rho$ *for all* $x \in R$.
3. *There exists* $x^* \in R$ *such that* $f(x^*) = 0$.

If the initial point x_0 is sufficiently close to x^*, then the sequence $\{x_k\}$ generated by

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \ldots$$

is well defined and converges to x^*. Moreover, for $k = 0, 1, 2, \ldots$

$$|x_{k+1} - x^*| \leq \frac{L}{2\rho}|x_k - x^*|^2.$$

Proof: Since $|f'(x)| \geq \rho > 0$ for all $x \in R$, the sequence is well defined. For $k = 0, 1, 2, \ldots$, it holds that

$$\begin{aligned} |x_{k+1} - x^*| &= \left| x_k - \frac{f(x_k)}{f'(x_k)} - x^* \right| \\ &= \left| \frac{1}{f'(x_k)} \left(f(x^*) - f(x_k) - f'(x_k)(x^* - x_k) \right) \right| \\ &\leq \frac{L|x_k - x^*|^2}{2\rho}. \end{aligned}$$

Suppose $r_0 := |x_0 - x^*| < 2\rho/L$ and let $\theta := \frac{Lr_0}{2\rho} < 1$, then

$$|x_{k+1} - x^*| \leq \frac{L|x_k - x^*|^2}{2\rho} \leq \theta|x_k - x^*| < 2\rho/L, \quad k = 0, 1, 2, \ldots$$

The convergence of $\{x_k\}$ to x^* follows from $|x_k - x^*| \leq \theta^k |x_0 - x^*|$. ■

12.1.2 Multivariate case

Newton's method for systems of nonlinear equations Let $F : R^n \to R^n$ be a continuously differentiable nonlinear mapping. The task is to find a root of F, i.e.,

$$\text{find } x^* \in R^n \text{ such that } F(x^*) = 0.$$

Algorithm 1 (Newton's method) *Initialization: choose $x_0 \in R^n$ and set $k = 0$.*

1. *Compute $F'(x_k)$ (an n-by-n matrix);*
2. *Solve $F'(x_k)d = -F(x_k)$ for d_k;*
3. *Update by $x_{k+1} = x_k + d_k$;*
4. *Set $k = k + 1$ and repeat if necessary.*

Example 12.1.3 *Let $n = 2$. F and F' are given by*

$$F(x) = \begin{bmatrix} x_1 + x_2 - 3 \\ x_1^2 + x_2^2 - 9 \end{bmatrix} \quad \text{and} \quad F'(x) = \begin{bmatrix} 1 & 1 \\ 2x_1 & 2x_2 \end{bmatrix}.$$

Clearly F has two roots $(3, 0)$ and $(0, 3)$. We start with $x_0 = (1, 5)$. The first few iterations of Newton's method are

$$\begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix} d_0 = -\begin{bmatrix} 3 \\ 17 \end{bmatrix}, \; d_0 = \begin{bmatrix} -\frac{13}{8} \\ -\frac{11}{8} \end{bmatrix}, \; x_1 = x_0 + d_0 = \begin{bmatrix} -0.625 \\ 3.625 \end{bmatrix};$$

$$\begin{bmatrix} 1 & 1 \\ -\frac{5}{4} & \frac{29}{4} \end{bmatrix} d_1 = -\begin{bmatrix} 0 \\ \frac{145}{32} \end{bmatrix}, \; d_1 = \begin{bmatrix} \frac{145}{272} \\ -\frac{145}{272} \end{bmatrix}, \; x_2 = x_1 + d_1 = \begin{bmatrix} -0.092 \\ 3.092 \end{bmatrix},$$

$$x_3 = \begin{bmatrix} -0.00265 \\ 3.00265 \end{bmatrix}, \; x_4 = \begin{bmatrix} -0.0000023426 \\ 3.0000023426 \end{bmatrix}, \; x_5 = \begin{bmatrix} -1.8 \times 10^{-12} \\ 3 + 1.8 \times 10^{-12} \end{bmatrix}.$$

Observations:

- converges very fast from good starting points;
- exact at each iteration for any affine component functions of F;
- requires $F'(x_k)$ at each iteration;
- need to solve a linear system at each iteration; can be difficult if $F'(x_k)$ is singular or nearly singular (ill-conditioned);
- not globally convergent in general (e.g., $x_0 = (1, 1)$ for this example).

Lemma 12.2 *Let $\|\cdot\|$ be any matrix norm on $R^{n\times n}$ that obeys $\|I\| = 1$ and*

$$\|AB\| \le \|A\|\|B\|, \quad \forall A, B \in R^{n\times n}.$$

Let $E \in R^{n\times n}$. If $\|E\| < 1$, then $(I - E)^{-1}$ exists and

$$\|(I - E)^{-1}\| \le \frac{1}{1 - \|E\|}.$$

If A is nonsingular and $\|A^{-1}\Delta A\| < 1$, then $A + \Delta A$ is nonsingular and

$$\|(A + \Delta A)^{-1}\| \le \frac{\|A^{-1}\|}{1 - \|A^{-1}\Delta A\|}.$$

Proof: Define $\{S_k = \sum_{j=0}^{k} E^j : k = 0, 1, 2, \ldots\}$. First show that $\{S_k\}$ is a Cauchy sequence in $(R^{n\times n}, \|\cdot\|)$ and thus converges. Clearly $(I - E)S_k = I - E^{k+1}$. Take limit on both sides, we see that $I - E$ is nonsingular and

$$(I - E)^{-1} = \lim_{k\to\infty} S_k = \sum_{j=0}^{\infty} E^j.$$

Furthermore, $\|(I - E)^{-1}\| \le \sum_{j=0}^{\infty} \|E\|^j = \frac{1}{1-\|E\|}$.

$\|A^{-1}\Delta A\| < 1$ implies that $(I + A^{-1}\Delta A)^{-1}$ exists and

$$\|(A + \Delta A)^{-1}\| = \|(I + A^{-1}\Delta A)^{-1}A^{-1}\| \le \|(I + A^{-1}\Delta A)^{-1}\|\|A^{-1}\| \le \frac{\|A^{-1}\|}{1 - \|A^{-1}\Delta A\|}.$$

Theorem 12.3 (Local convergence) *Let $F : R^n \to R^n$ be a nonlinear mapping satisfying the following assumptions:*

1. *There exists $x^* \in R^n$ such that $F(x^*) = 0$.*
2. *F is continuously differentiable, and there exists $r > 0$ such that $J = F'$ is Lipschitz continuous on $N(x^*, r)$ (open ball centered at x^* with radius r), i.e., for some $L > 0$,*

$$\|J(x) - J(y)\| \le L||x - y||, \quad \forall x, y \in N(x^*, r).$$

3. *There exists $\beta > 0$ such that $\|J(x^*)^{-1}\| \le \beta$.*

If the initial point x_0 is sufficiently close to x^*, then the sequence $\{x_k\}$ generated by

$$x_{k+1} = x_k - J(x_k)^{-1}F(x_k), \quad k = 0, 1, 2, \ldots$$

is well defined and converges to x^*. Moreover, for $k = 0, 1, 2, \ldots$

$$\|x_{k+1} - x^*\| \leq \beta L \|x_k - x^*\|^2.$$

Proof: Let $\delta := \min\{r, \frac{1}{2\beta L}\} > 0$. Suppose $x_0 \in N(x^*, \delta)$. $J(x_0)$ is nonsingular because

$$\|J(x^*)^{-1}(J(x_0) - J(x^*))\| \leq \|J(x^*)^{-1}\| \|J(x_0) - J(x^*)\| \leq \beta L \|x_0 - x^*\| \leq \beta L \delta \leq 1/2.$$

Furthermore,

$$\|J(x_0)^{-1}\| \leq \frac{\|J(x^*)^{-1}\|}{1 - \|J(x^*)^{-1}(J(x_0) - J(x^*))\|} \leq 2\|J(x^*)^{-1}\| \leq 2\beta.$$

$$\begin{aligned}
\|x_1 - x^*\| &= \|x_0 - x^* - J(x_0)^{-1}F(x_0)\| \\
&= \|J(x_0)^{-1}[F(x^*) - F(x_0) - J(x_0)(x^* - x_0)]\| \\
&\leq 2\beta \times \frac{L}{2}\|x_0 - x^*\|^2 = \beta L \|x_0 - x^*\|^2.
\end{aligned}$$

It follows that $\|x_1 - x^*\| \leq \frac{1}{2}\|x_0 - x^*\|$ and $x_1 \in N(x^*, \delta)$. The proof is completed by induction. ■

12.2 Newton's method for unconstrained optimization

12.2.1 Framework of the method

Consider $\min_{x \in R^n} f(x)$, where $f : R^n \to R$ is twice continuously differentiable.

Algorithm 2 (Newton's method) *Initialization: choose $x_0 \in R^n$ and set $k = 0$.*

1. *Compute $\nabla^2 f(x_k)$;*
2. *Solve $\nabla^2 f(x_k) d = -\nabla f(x_k)$ for d_k;*
3. *Update $x_{k+1} = x_k + d_k$;*
4. *Set $k = k + 1$, determine to stop or to repeat.*

12.2.2 Motivation

Motivation 1. Solving $\min_{x \in R^n} f(x)$ reduces to solving $\nabla f(x) = 0$. Apply Newton's method for solving system of nonlinear equations:

$$\nabla f(x) \approx \nabla f(x_k) + \nabla^2 f(x_k)(x - x_k) = 0,$$

resulting iteration formula

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k).$$

Motivation 2. Approximate f by a quadratic function:

$$f(x) \approx f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2}(x - x_k)^T \nabla^2 f(x_k)(x - x_k).$$

By minimizing the right-hand side to generate x_{k+1}, we obtain the same iteration formula.

12.2.3 Convergence

Theorem 12.4 (Local convergence) *Consider* $\min_{x \in R^n} f(x)$*, where* $f : R^n \to R$*. Assume the following conditions are satisfied:*

1. f *has a local minimum* x^**;*
2. $f \in C_M^{2,2}(R^n)$*;* [1]
3. $\nabla^2 f(x^*) \succeq \mu I_n$ *for some* $\mu > 0$*.*

If $r_0 := \|x_0 - x^*\| < \bar{r} := \frac{2\mu}{3M}$*, then the Newton sequence is well defined,* $\|x_k - x^*\| < \bar{r}$ *for all* k*,* x_k *converges to* x^* *and obeys* [2]

$$\|x_{k+1} - x^*\| \leq \frac{3M}{2\mu} \|x_k - x^*\|^2.$$

Proof. It follows from $f \in C_M^{2,2}(R^n)$ that

$$\nabla^2 f(x_0) \succeq \nabla^2 f(x^*) - M r_0 I_n \succeq \frac{1}{3} \mu I_n.$$

Therefore, $[\nabla^2 f(x_0)]^{-1}$ exists, $\|[\nabla^2 f(x_0)]^{-1}\| \leq 3/\mu$, and x_1 is well defined. Furthermore,

$$\begin{aligned}
\|x_1 - x^*\| &= \|x_0 - x^* - [\nabla^2 f(x_0)]^{-1} \nabla f(x_0)\| \\
&= \|[\nabla^2 f(x_0)]^{-1} \left(\nabla f(x^*) - \nabla f(x_0) - \nabla^2 f(x_0)(x^* - x_0)\right)\| \\
&\leq \frac{3M}{2\mu} \|x_0 - x^*\|^2 \leq \frac{3M}{2\mu} r_0 \|x_0 - x^*\| < \|x_0 - x^*\|.
\end{aligned}$$

(Note that $\frac{3M}{2\mu} r_0 < 1$.) By induction, the Newton sequence is well defined, $\|x_k - x^*\| \leq \bar{r}$ for all k, and x_k converges to x^* quadratically, i.e.,

$$\|x_{k+1} - x^*\| \leq \frac{3M}{2\mu} \|x_k - x^*\|^2, \quad k = 0, 1, 2, \ldots$$

12.2.4 Pros and cons

Pros: Quadratic local convergence, extremely fast!

Cons: (1) need to compute $\nabla^2 f(x_k)$ at each iteration, which is very costly for large problems; (2) need to solve a linear system of order n at each iteration; (3) $\nabla^2 f(x_k)$ could be (nearly) singular when x_k is far away from solution; (4) Newton direction $-[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$ is not necessarily a descent direction; (5) global convergence is not guaranteed.

Remark 12.2.1

- *Newton's method is locally quadratic convergent, which means that when the current point is good enough it will be improved rapidly.*
- *Unfortunately, it is not unusual to expend significant computational efforts in "getting close enough".*
- *The strategies for getting close constitute the major part of the program and the programming effort, and they can be sensitive to small difference in implementation.*

---

[1] can be relaxed to $f \in C_M^{2,2}(N(x^*, r))$ for some $r > 0$.
[2] this type of convergence is called quadratic convergence.

- **General recommendation:**
  - *use a robust and stable method at the initial stage in order to get close to a solution;*
  - *the final job should be performed by the Newton's method, if ever viable, to generate a high accuracy solution.*

12.3 About different convergence rates

The following three types of convergence rates are most well-known:

1. Sublinear rate. $r_k \leq \frac{c}{k^t}$ for some $c, t > 0$.
   - To ensure $r_k \leq \epsilon$, the upper complexity bound is $\frac{c^{1/t}}{\epsilon^{1/t}}$.
   - Sublinear convergence is rather slow. In terms of complexity, each new right digit takes the amount of computations comparable with the total amount of the previous work:
   $$\frac{c}{(k+k_0)^t} \approx \frac{c}{10k^t} \;\Rightarrow\; k_0 \approx (10^{1/t}-1)k.$$
   - The constant c plays a significant role in the corresponding complexity estimate.
2. Linear rate. $r_k \leq c(1-q)^k$ for some $c > 0$ and $0 < q < 1$.
   - Upper complexity bound to ensure $r_k \leq \epsilon$:
   $$\frac{\ln c + \ln(1/\epsilon)}{\ln(1-q)^{-1}} \approx \frac{\ln c + \ln(1/\epsilon)}{q}.$$
   - Linear convergence is fast. Each new right digit takes a constant amount of computations:
   $$c(1-q)^{k+k_0} \approx \frac{1}{10}c(1-q)^k \;\Rightarrow\; k_0 \approx \frac{\ln 10}{\ln(1-q)^{-1}}.$$
   - The dependence of the complexity estimate in c is much weaker.
3. Quadratic rate. $r_{k+1} \leq c r_k^2$ for some $c > 0$ and $r_0 < 1/c$.
   - By induction, it holds that $r_k \leq c^{2^k-1} r_0^{2^k}$ for all $k \geq 1$.
   - Upper complexity bound to ensure $r_k \leq \epsilon$ is approximately
   $$\log_2 \log_2(1/\epsilon).$$
   - Each iteration doubles the number of right digits in the answer.
   - The constant c is important only for the starting stage of the quadratic convergence.

More about convergence rates can be found in *"Iterative solution of nonlinear equations in several variables"* by Ortega and Rheinboldt (1970).

12.4 Globally convergent modifications of Newton's method

12.4.1 Truncated Newton's method

Truncated Newton's method Suppose $f \in \mathcal{S}_\mu^2(R^n)$, or at least uniformly convex over

$$L(x_0) := \{x \in R^n : f(x) \le f(x_0)\}.$$

The following is the framework of truncated Newton's method.

Algorithm 3 (truncated Newton's method) *Initialization.*

1. *Compute Newton step* $d_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$.
2. *Compute a step length* α_k *via either*

$$f(x_k + \alpha_k d_k) = \min_{\alpha \ge 0} f(x_k + \alpha d_k)$$

   or ($\delta > 0$ *is a constant)*

$$f(x_k) - f(x_k + \alpha_k d_k) \ge \delta \|g_k\|^2 \cos^2 \langle d_k, -g_k \rangle.$$

3. *update* $x_{k+1} = x_k + \alpha_k d_k$*;*
4. *determine to stop or to repeat.*

Since f is strongly convex, the Hessian matrix $\nabla^2 f(x)$ is always positive definite and the Newton direction is a descent direction. The truncated Newton's method has the following convergence.

Theorem 12.5 (Global convergence of truncated Newton's method for strongly convex function) *Suppose f is twice continuously differentiable and, for any x_0, f is strongly convex over $L(x_0)$ with constant $\mu > 0$. Let $\{x_k\}$ be the sequence generated by the truncated Newton's method presented in the last slide. Then, x_k converges to the unique global minimizer of f.*

Proof: Check Theorems 3.2.3 and 3.2.4 in Yuan-Sun's book. ■

12.4.2 Other modifications

If $\nabla^2 f(x_k)$ is not positive definite, then the quadratic model does not necessarily have minimizer (unbounded below if Hessian has negative eigenvalue). Let $d_k^N = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$.

Modification 1: Take negative gradient as search direction if the Newton direction is unsatisfactory:

$$d_k = \begin{cases} d_k^N, & \text{if } \langle d_k^N, -\nabla f(x_k) \rangle \le \frac{\pi}{2} - \mu \text{ for some } \mu > 0; \\ -\nabla f(x_k), & \text{otherwise.} \end{cases}$$

(of course, incorporate line search at each iteration.)

Modification 2: Modify $\nabla^2 f(x_k)$ when it is not positive definite. Also, modify it as slight as possible.

Algorithm 4 (Goldfeld's modification) *Let $G_k := \nabla^2 f(x_k)$ in the following. The kth iteration is as follows:*

1. *Compute $\bar{G}_k = G_k + \nu_k I$, where $\nu_k = 0$ is G_k is positive definite and $\nu_k > 0$ (computed according to certain rule) if otherwise.*

2. *Compute the Cholesky factorization of $\bar{G}_k$: $\bar{G}_k = L_k D_k L_k^T$, where L_k is lower triangular.*

3. *Solve $\bar{G}_k d = -g_k$ for d_k;*

4. *Update x_k via $x_{k+1} = x_k + d_k$ or $x_{k+1} = x_k + \alpha_k d_k$, where $\alpha_k > 0$ is obtained via line search.*

Note that the quadratic convergence is maintained only when $\alpha_k \equiv 1$, i.e., the full Newton step is accepted, in the final stage.

Many other variants: simplified Newton's method, Newton-like methods, inexact Newton's method, etc.

12.4.3 Trust region method

Trust region model problem at the kth iteration:

$$\begin{array}{ll} \min_{s \in R^n} & q^{(k)}(s) := f(x_k) + g_k^T s + \frac{1}{2} s^T \nabla^2 f(x_k) s \\ s.t. & \qquad\qquad \|s\| \le h_k, \end{array}$$

where $s = x - x_k$, and $h_k > 0$ is called the trust region radius. Let s_k be the solution to the model trust region problem. Note that

$$f(x_k + s_k) \le f(x_k)$$

always holds. The choice of h_k depends on

$$r_k = \frac{\text{True decrease}}{\text{Model decrease}} = \frac{f(x_k) - f(x_k + s_k)}{f(x_k) - q^{(k)}(s_k)}.$$

Algorithm 5 (Trust region algorithmic framework) *Initialization: choose $x_0 \in R^n$ and set $h_0 = \|g_0\|$. The kth iteration has the following form:*

1. *compute g_k and $\nabla^2 f(x_k)$;*

2. *solve trust region problem for s_k;*

3. *compute $f(x_k + s_k)$ and r_k;*

4. *compute h_{k+1} via*

$$h_{k+1} = \begin{cases} \|s_k\|/4, & \text{if } r_k < 0.25; \\ 2h_k, & \text{if } r_k > 0.75 \text{ and } \|s_k\| = h_k; \\ h_k, & \text{otherwise.} \end{cases}$$

5. *If $r_k \le 0$, set $x_{k+1} = x_k$; otherwise set $x_{k+1} = x_k + s_k$;*

Convergence: Under mild assumptions, trust region method is globally convergent to a point x^* satisfying both the first and the second order necessary conditions. Moreover, if $\nabla^2 f(x^*)$ is positive definite, then the convergence rate is quadratic.