Chapter 4
The Newton Bracketing Method for Convex Minimization: Convergence Analysis

Adi Ben-Israel and Yuri Levin

Summary: Let f be a convex function bounded below with infimum $f_{\min}$ attained. A bracket is an interval $[L,U]$ containing $f_{\min}$. The Newton Bracketing (NB) method for minimizing f, introduced in [8], is an iterative method that at each iteration transforms a bracket $[L,U]$ into a strictly smaller bracket $[L_+,U_+]$ with $L \le L_+ < U_+ \le U$. We show, under certain conditions on f, that an upper bound on the bracket ratio $(U_+ - L_+)/(U - L)$ can be guaranteed by the selection of the method parameters.

Key words: Newton Bracketing method, directional Newton method, convex functions, unconstrained minimization, Fermat–Weber location problem.
AMS 2010 Subject Classification: 52A41, 90C25, 49M15, 90B85

Introduction

The **Newton Bracketing** (or **NB**) method, introduced in [8], is an iterative method for convex minimization. It was applied to location problems [9], semi-definite programming [4], and linearly-constrained convex programs, [1].

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function of n variables with an attained infimum $f_{\min}$. An optimal solution $\mathbf{x}_{\min}$ (unique if f is strictly convex) can be approximated iteratively by a gradient method

$$\mathbf{x}_+ := \mathbf{x} - c\,\nabla f(\mathbf{x}),\ c > 0.$$

Adi Ben-Israel
RUTCOR–Rutgers Center for Operations Research, Rutgers University, 640 Bartholomew Rd., Piscataway, NJ 08854-8003, USA, e-mail: adi.benisrael@gmail.com

Yuri Levin
School of Business, Queen's University, 143 Union Str., Kingston, ON, K7L 3N6, CANADA e-mail: ylevin@business.queensu.ca

Gradient methods often suffer from slow convergence near $\mathbf{x}_{\min}$. They also lack a natural stopping rule.

The problem can also be solved by approximating the optimal value $f_{\min}$. A **bracket** is a closed interval $[L,U]$ with

$$L \leq f_{\min} \leq U. \tag{4.1}$$

The length of the bracket $[L,U]$ is denoted $\Delta := U - L$. A **bracketing method** generates a sequence of nested brackets, shrinking to a point. The brackets are defined iteratively by

$$[L_+, U_+] := \Psi([L,U])$$

where Ψ maps intervals to intervals,

$$L \leq L_+ \leq f_{\min} \leq U_+ \leq U, \text{ and } \Delta_+ := U_+ - L_+ < \Delta.$$

Using fixed point terminology, the optimal $\mathbf{x}_{\min}$ is a fixed point of the mapping

$$\Phi(\mathbf{x}) := \mathbf{x} - c\nabla f(\mathbf{x}),\, c > 0\,,$$

while the optimal value $f_{\min}$ (viewed as a degenerate interval) is a fixed point of Ψ.

The bracket size is a natural stopping criterion, stopping the iterations when

$$U - L < \varepsilon \tag{4.2}$$

for a given tolerance $\varepsilon > 0$. For fast convergence it is desirable to have large reductions of successive brackets, i.e., small values of the **bracket ratios**

$$\frac{\Delta_+}{\Delta} = \frac{U_+ - L_+}{U - L}\,, \tag{4.3}$$

and an upper bound on (4.3) translates to a guaranteed reduction. We study conditions that imply such upper bounds for the NB method.

First, a description of the method for $n = 1$. An iteration begins with a current solution x where $f'(x) \neq 0$, and a bracket $[L, U := f(x)]$ containing $f_{\min}$ (an initial lower bound L on $f_{\min}$ is assumed given.) An intermediate value

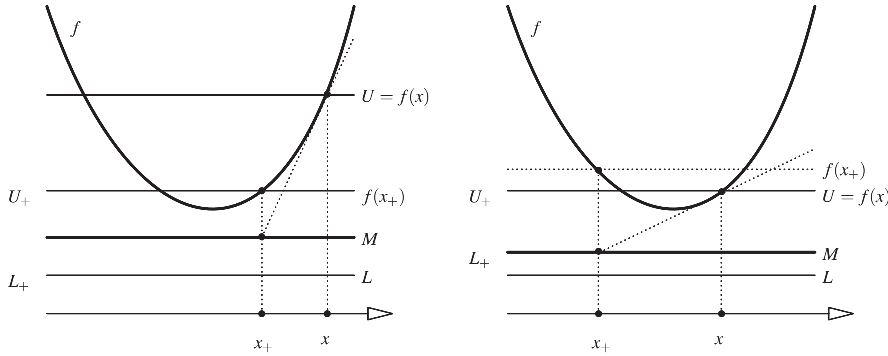
$$M := \alpha U + (1-\alpha)L \tag{4.4}$$

is selected for some $0 < \alpha < 1$, and one Newton iteration for solving

$$f(x) = M \tag{4.5}$$

is carried out, giving

$$x_+ := x - \frac{f(x) - M}{f'(x)}, \tag{4.6}$$



(a) Case 1: $f(x_+) < f(x)$ (b) Case 2: $f(x_+) \geq f(x)$

Fig. 4.1 Illustration of the 2 cases of the NB method

and two cases, illustrated in Fig. 4.1.

Case 1: $$f(x_+) < f(x), \tag{4.7a}$$

and the bracket is updated,

$$U_+ := f(x_+),\ L_+ := L. \tag{4.7b}$$

Case 2: $$f(x_+) \geq f(x), \tag{4.8a}$$

in which case the bracket is updated, keeping x,

$$U_+ := U,\ L_+ := M,\ x_+ := x. \tag{4.8b}$$

The iteration is summarized as follows:

1 Stopping rule. If $U - L < \varepsilon$, **stop** with x as solution.
2 Select a value $M := \alpha U + (1-\alpha)L$, for some $0 < \alpha < 1$.
3 Do one Newton iteration $x_+ := x - \dfrac{f(x) - M}{f'(x)}$.
4 Case 1: If $f(x_+) < f(x)$ then update U: $U_+ := f(x_+)$ and leave $L_+ := L$. Go to **1**.
5 Case 2: If $f(x_+) \geq f(x)$ then update L: $L_+ := M$ and leave $U_+ := U,\ x_+ := x$. Go to **1**. (NB)

For $n > 1$, the only change is in step 3 :

3 Do one directional Newton iteration
$$\mathbf{x}_+ := \mathbf{x} - \frac{f(\mathbf{x}) - M}{\|\nabla f(\mathbf{x})\|^2} \nabla f(\mathbf{x}). \tag{NB.n}$$

using the directional Newton method, [7]. The iterate $\mathbf{x}_+$ satisfies

$$
\begin{aligned}
f(\mathbf{x}_+) &\geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{x}_+ - \mathbf{x}), \text{ since } f \text{ is convex},\\
&= f(\mathbf{x}) + \nabla f(\mathbf{x})^T \left(-\frac{f(\mathbf{x}) - M}{\|\nabla f(\mathbf{x})\|^2} \nabla f(\mathbf{x}) \right) = f(\mathbf{x}) - (f(\mathbf{x}) - M).
\end{aligned}
$$

$$\therefore\ f(\mathbf{x}_+) \geq M, \text{ and} \tag{4.9a}$$

$$f(\mathbf{x}_+) > M, \text{ if } f \text{ is strictly convex and } f(\mathbf{x}) \neq M. \tag{4.9b}$$

The NB method is valid if (4.1) holds throughout the iterations, i.e., if the new interval $[L_+, U_+]$ also contains $f_{\min}$. This is clearly the case for $n = 1$ (the picture is the proof), but not in general for $n > 1$. Sufficient conditions for validity were given in [8], in particular, the method is valid for the quadratic function,

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T Q \mathbf{x} + \mathbf{c}^T \mathbf{x} + \gamma,\ Q \text{ positive definite}, \tag{4.10}$$

if Q is well-conditioned,

$$\frac{\lambda_n}{\lambda_1} \geq 7 - \sqrt{48} \approx 0.071796768, \tag{4.11}$$

where λ_n and λ_1 are respectively the smallest and largest eigenvalues of Q.

In Case 2 the point $\mathbf{x}$ does not change (with the bonus that $\nabla f(\mathbf{x})$ need not be recomputed.) Since the bracket decreases in every iteration of the NB method, a convergence analysis of the method must therefore be based on the brackets $[L, U]$ in $\mathbb{R}$, rather than on the iterates $\{\mathbf{x}\}$ in $\mathbb{R}^n$.

Example 4.1. One way the NB method may fail is illustrated by the function

$$f(x) = \tfrac{1}{2} x^2 + 1 \tag{4.12}$$

with initial $x := 1, U := f(1) = \frac{3}{2}$ and $L := 0$. For $\alpha = \frac{1}{3}$ we get $M = \frac{1}{2}$ by (4.4), and

$$x_+ := x - \frac{f(x) - M}{f'(x)} = 1 - \frac{\frac{3}{2} - \frac{1}{2}}{1} = 0$$

which is the optimal solution. However, if $\varepsilon < 1$ the NB method does not stop since the bracket size is 1, but it also cannot continue since $f'(0) = 0$. This issue can be resolved by adding a derivative based stopping rule, but as a practical matter it can be ignored.

Example 4.2. To illustrate why the NB method requires the attainment of the infimum $f_{\min}$, consider the function

$$f(x) = e^{-x}$$

with initial $x := 0, U := f(0) = 1$, and $L := -1$. The NB method, for any choices of α, has only iterations of Case 1, the iterates $\rightarrow \infty$, and the bracket size remains ≥ 1.

Example 4.2 is special in that one case, Case 1, occurs in all iterations. In normal circumstances, cases 1 and 2 alternate, with large [small] α making Case 1 [Case 2] more likely.

The bracket ratio (4.3) of the NB method is

$$\frac{\Delta_+}{\Delta} = \begin{cases} \dfrac{f(\mathbf{x}_+) - L}{U - L} , & \text{in case 1 ,} \\ \dfrac{U - M}{U - L} = 1 - \alpha , & \text{in case 2 .} \end{cases} \tag{4.13}$$

We study here the convergence of the NB method in terms of the bracket ratio. These results are stated for $n = 1$ in Section 4, for $n > 1$ in Section 4, and are applied in Section 4 to the Fermat–Weber location problem.

Remark 4.1. For similar ideas and more general results see Kim et al [6], and Cegielski [2].

Bracket reduction in the NB method for $n = 1$

The results of this section, for the simple case $n = 1$, pave the way for the more interesting case, $n > 1$, in the next section.

Let $f : \mathbb{R} \to \mathbb{R}$ be convex and differentiable as needed, and consider the Newton method for solving the equation

$$f(x) = M. \tag{4.5}$$

Let x be a point where $f'(x) \neq 0$, and let

$$x_+ := x - \frac{f(x) - M}{f'(x)} \tag{4.14}$$

be the next Newton iterate.

Remark 4.2. In what follows, there appear expressions like $|f(x_+) - M|$ where the absolute value sign is not necessary (because of (4.9)) but is given pro forma.

Let X_0 be the interval with endpoints x, x_+. If $f \in C_N^{1,1}(X_0)$ (i.e $f'(x)$ is Lipschitz on X_0 with Lipschitz constant N) and $|f'(x)|$ is sufficiently large, then the value of $|f(x_+) - M|$ is bounded as follows.

Lemma 4.1. *Using the above notation, let $f \in C_N^{1,1}(X_0)$ and $f(x) > M$. If for some $\beta > 0$,*

$$|f'(x)|^2 \geq \beta N |f(x) - M|, \tag{4.15}$$

then

$$|f(x_+) - M| \le \frac{1}{2\beta}|f(x) - M| \ . \tag{4.16}$$

Proof. Since $f \in C_N^{1,1}(X_0)$,

$$|f(x^+) - f(x) - f'(x)(x^+ - x)| \le \frac{N}{2}(x^+ - x)^2,$$

by the descent lemma, [11, 3.2.12, page 73]. Therefore,

$$\begin{aligned} |f(x^+) - M| &\le \frac{N}{2}\frac{(f(x) - M)^2}{(f'(x))^2}\,, \text{by (4.14)} \\ &\le \frac{1}{2\beta}|f(x) - M|\,, \text{by (4.15).} \end{aligned}$$

■

Remark 4.3. To guarantee decrease in (4.16), it is required that

$$\beta > \tfrac{1}{2} \ . \tag{4.17}$$

Recall that $U := f(x)$. In order to apply Lemma 4.1 to an NB iteration, we note that

$$\begin{aligned} f(x) - M &= U - M \\ &= (1-\alpha)\,(U - L), \quad \text{by (4.4)}, \\ &= (1-\alpha)\,\Delta, \end{aligned} \tag{4.18}$$

and inequality (4.15) can be written as

$$(1-\alpha)\,\beta \le \frac{|f'(x)|^2}{N\Delta}, \tag{4.19}$$

relating α and β for any given x. We observe that β can be chosen arbitrarily, with α then constrained by (4.19).

The next theorem guarantees an upper bound on the bracket ratio (4.13).

Theorem 4.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be convex, x a point where $f'(x) \ne 0$, X_0 the interval with endpoints x, x_+, and assume $f \in C_N^{1,1}(X_0)$. Let $[L,U]$ a bracket for $f_{\min}$ (i.e. $L \le f_{\min} \le U$), and let β satisfy*

$$\beta > \frac{|f'(x)|^2}{N\Delta}. \tag{4.20}$$

Then for

$$\alpha := 1 - \frac{|f'(x)|^2}{\beta N \Delta}, \tag{4.21}$$

the NB iteration results in a bracket ratio

$$\frac{\Delta_+}{\Delta} \leq \max \left\{ \frac{1}{2\beta}(1-\alpha) + \alpha, 1-\alpha \right\}. \tag{4.22}$$

Proof. In Case 2 we get from (4.13)

$$\frac{\Delta_+}{\Delta} = 1 - \alpha,$$

and in Case 1,

$$\begin{aligned}
|f(x_+) - M| &\leq \frac{1}{2\beta}|U - M|, \text{ by Lemma 4.1, and } U = f(x), \\
&= \frac{1}{2\beta}(1-\alpha)|U - L|, \text{ by (4.4).} \\
M - L &= \alpha(U - L), \text{ by (4.4).} \\
\therefore\ U_+ - L_+ &= (f(x_+) - M) + (M - L), \text{ by (4.7b),} \\
&\leq \left(\frac{1}{2\beta}(1-\alpha) + \alpha \right)(U - L),
\end{aligned}$$

completing the proof. ■

Example 4.3. For

$$f(x) = \tfrac{1}{2}x^2 + 1 \tag{4.12}$$

we have $f'(x) = x$ and $f'' = 1$, giving $N = 1$. Then (4.20) and (4.21) become

$$\beta > \frac{x^2}{\Delta}, \text{ and } \alpha = 1 - \frac{x^2}{\beta \Delta}.$$

For the initial $x := 1, U := f(x) = \frac{3}{2}, L := 0$ and initial bracket size $\Delta = U - L = \frac{3}{2}$,

$$\beta > \tfrac{2}{3}, \text{ and } \alpha = 1 - \frac{2}{3\beta}.$$

Table 4.1 lists some admissible values of β, and the corresponding bracket ratios. The guaranteed ratios in the last column are pessimistic because of the high frequency of iterations of Case 2, with ratios given in the penultimate column.

Theorem 4.1 concerns a single iteration, and its application requires selecting an admissible β and recalculating α in each iteration.

To simplify matters, we fix the parameter β throughout the iterations (it may no longer be admissible in some iterations) and impose the following constraint on α,

β	Case 1 $\frac{1}{2\beta}(1-\alpha)+\alpha$	Case 2 $(1-\alpha)$	Upper bound on bracket ratio, RHS(4.22)
0.8	0.6875	0.833333333	0.833333333
1	0.666666667	0.666666667	0.666666667
1.5	0.703703704	0.444444444	0.703703704
2	0.75	0.333333333	0.75
2.5	0.786666667	0.266666667	0.786666667

Table 4.1 Some admissible β's and the corresponding bracket ratios

$$\alpha_{\min} \leq \alpha \leq \alpha_{\max} \tag{4.23}$$

with given bounds $\{\alpha_{\min}, \alpha_{\max}\}$. The parameter α is computed in each iteration as the point in the interval $[\alpha_{\min}, \alpha_{\max}]$ that is closest to (4.21).

Example 4.4. Consider the quadratic

$$f(x) = \tfrac{1}{2}x^2 + 1 \tag{4.12}$$

with initial $x = 1, U := f(1) = \frac{3}{2}, L := 0$, and initial bracket size $\Delta = \frac{3}{2}$.

(a) Using $\beta = \frac{2}{3},\ \alpha_{\min} = 0.2,$ and $\alpha_{\max} = 0.9$, the NB method (with $\varepsilon = 10^{-6}$) stopped after 16 iterations, an average reduction per iteration

$$\frac{\Delta_+}{\Delta} = 0.41.$$

The reductions in each iterations are shown in Figure 4(a). Case 1 occurred in 11 iterations and Case 2 in 5.

(b) Using $\beta = 0.2$ (an inadmissible value), $\alpha_{\min} = 0.3$ and $\alpha_{\max} = 0.7$, the NB method (with $\varepsilon = 10^{-9}$) required 56 iterations, 15 of case 1, and 41 of case 2. The reductions in each iteration are shown in Fig. 4(b). The average reduction per iteration is

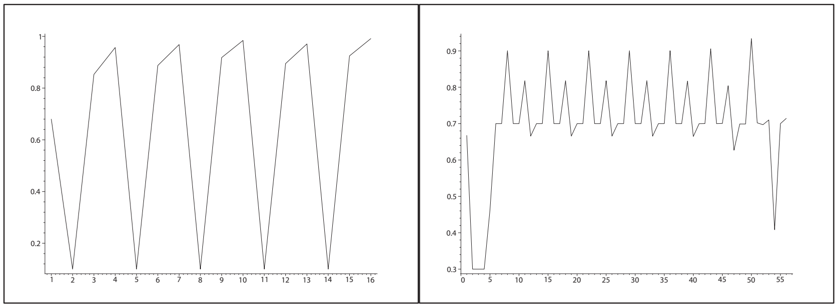
$$\frac{\Delta_+}{\Delta} = .68.$$

Remark 4.4. The periodic reductions in Figs. 4(a)–(b) are explained by self–similarity of the NB method for this particular function and the selected values of β.

Bracket reduction in the NB method for $n > 1$

Theorem 4.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and differentiable, let $\mathbf{x}$ be a point where $\nabla f(\mathbf{x}) \neq 0$, and let*

$$\mathbf{x}_+ := \mathbf{x} - \frac{f(\mathbf{x}) - M}{\|\nabla f(\mathbf{x})\|^2}\nabla f(\mathbf{x}), \tag{4.24}$$



(a) $\beta = \frac{2}{3}$, $\alpha_{min} = 0.2$, $\alpha_{\max} = 0.9$ (b) $\beta = 0.2$, $\alpha_{min} = 0.3$, $\alpha_{\max} = 0.7$

Fig. 4.2 Illustration of Example 4.4: Reduction per iteration for $f(x) = \frac{1}{2}x^2 + 1$, $L = 0$ and $x_0 = 1$

be the next Newton iterate with step

$$\mathbf{h} := -\frac{f(\mathbf{x}) - M}{\|\nabla f(\mathbf{x})\|^2} \nabla f(\mathbf{x}) . \tag{4.25}$$

Consider the ball

$$X_0 := \{\xi : \|\xi - \mathbf{x}_+\| \ \leq \ \|\mathbf{h}\|\} ,$$

and assume that $f \in C_N^{1,1}(X_0)$*, i.e* ∇f *is Lipschitz in* X_0 *with Lipschitz constant* N*. Finally, assume* $\mathbf{x}$ *satisfies,*

$$\|\nabla f(\mathbf{x})\|^2 \geq \beta N |f(\mathbf{x}) - M| , \tag{4.26}$$

for some $\beta > 0$*. Then:*

(a) $\|\nabla f(\mathbf{x}_+)\| \geq \dfrac{\beta - 1}{\beta} \|\nabla f(\mathbf{x})\|$.

(b) $|f(\mathbf{x}_+) - M| \leq \dfrac{1}{2\beta} |f(\mathbf{x}) - M|$.

Proof. See Appendix A. ■

Remark 4.5. For part (a) to be useful, it is required that $\beta > 1$.

As in (4.19), the inequality (4.26) can be written as,

$$(1 - \alpha)\beta \leq \frac{\|\nabla f(\mathbf{x})\|^2}{N\Delta}. \tag{4.27}$$

An admissible β can thus be selected at will, and α then determined by

$$\alpha = 1 - \frac{\|\nabla f(\mathbf{x})\|^2}{\beta N \Delta} \tag{4.28}$$

the analog of (4.21).

Example 4.5. Consider the quadratic function $f : \mathbb{R}^n \to \mathbb{R}$,

$$f(\mathbf{x}) = \tfrac{1}{2}\,\mathbf{x}^T Q\,\mathbf{x} + 1,\ Q \text{ positive definite}, \tag{4.29}$$

with $f_{\min} = 1$. Let Q have the eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0, \quad \text{and corresponding eigenvectors} \quad \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\,.$$

Then $\nabla f(\mathbf{x}) = Q\,\mathbf{x}$, and $f''(\mathbf{x}) = Q$. The norm of Q corresponding to the Euclidean norm is $\|Q\| = \lambda_1$, which is taken as the Lipschitz constant N in Theorem 4.2. The inequality (4.27) becomes,

$$(1-\alpha)\,\beta \leq \frac{\|Q\mathbf{x}\|^2}{\lambda_1\,\Delta}, \tag{4.30}$$

giving for $\mathbf{x} = \mathbf{v}_n$,

$$(1-\alpha)\,\beta \leq \frac{\lambda_n^2}{\lambda_1\,\Delta}.$$

The following theorem is the analog of Theorem 4.1, giving an upper bound on the bracket ratio in a single iteration.

Theorem 4.3. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and differentiable, let $\mathbf{x}$ be a point where $\nabla f(\mathbf{x}) \neq \mathbf{0}$, and let $[L, U]$ be a bracket for $f_{\min}$. Let β satisfy,*

$$\beta > \frac{\|\nabla f(\mathbf{x})\|^2}{N\,\Delta}. \tag{4.31}$$

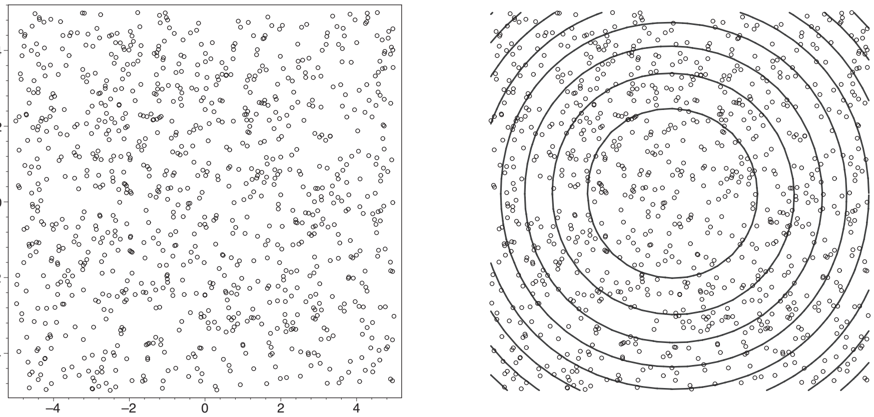
Then for

$$\alpha := 1 - \frac{\|\nabla f(\mathbf{x})\|^2}{\beta\,N\,\Delta}, \tag{4.28}$$

the NB iteration results in a reduction

$$\frac{\Delta_+}{\Delta} \leq \max\left\{\frac{1}{2\beta}\,(1-\alpha) + \alpha, 1-\alpha\right\}. \tag{4.32}$$

Proof. The proof of Theorem 4.1 applies verbatim. ■



(a) Points (b) Level sets of the function (4.33)

Fig. 4.3 Illustration of Example 4.6: 1,000 random points in $[-5,5]^2$

Application to the Fermat–Weber location problem

The **Fermat–Weber location problem** is to find a point $\mathbf{x}$ minimizing the sum of Euclidean distances

$$f(\mathbf{x}) = \sum_{i=1}^{m} \|\mathbf{a}_i - \mathbf{x}\| \tag{4.33}$$

from m given points $\{\mathbf{a}_i : i = 1, \dots, m\}$, see [3], [10], [17] and their references.

For large m, the contours of the function (4.33) are close to circular, see, e.g., Fig. 4.3(b), and the problem is well–conditioned, so the NB method is valid by (4.11).

To apply Theorem 4.3 we need an estimate of $N = \sup f''$, which is problematic for the function (4.33). However, $f(\mathbf{x})$ can be approximated, near the optimal solution, by a quadratic

$$f(\mathbf{x}) \approx \tfrac{1}{2}\,\mathbf{x}^T Q\,\mathbf{x}$$

giving $f'' \approx Q$, $N \approx \lambda_1$ (the largest eigenvalue of Q), and therefore $N = O(m)$, say $N = \frac{m}{2}$.

As in the case $n = 1$, we fix the parameter β throughout the iterations (it may no longer be admissible in some iterations) and impose the constraint,

$$\alpha_{\min} \le \alpha \le \alpha_{\max} \tag{4.23}$$

where the bounds $\{\alpha_{\min}, \alpha_{\max}\}$ are given. The parameter α is computed in each iteration as the point in the interval $[\alpha_{\min}, \alpha_{\max}]$ that is closest to (4.28).

The value of the parameter α depends on the bounds $\{\alpha_{\min}, \alpha_{\max}\}$. The following considerations apply to choosing these bounds (and indirectly α.)

(a) For large values of α, say $\alpha \geq 0.8$, the target M is close to U by (4.4), making Case 1 more likely, and the bracket ratio (see proof of Theorem 4.1),

$$\frac{\Delta_+}{\Delta} \leq \frac{1-\alpha}{2\beta} + \alpha \tag{4.34}$$

is large. However, when Case 2 occurs, the ratio

$$\frac{\Delta_+}{\Delta} = 1 - \alpha \tag{4.35}$$

is small since α is large. We thus alternate between small and large reductions, see, e.g., Fig. 4.4.

(b) Small values of α (say $\alpha \leq 0.2$) make Case 2 more likely, with a large ratio (4.35), i.e., a small reduction. However, in Case 2 the derivative need not be computed, so the overall time may be smaller.

Our numerical experience suggests that convergence is faster for higher values of α, as illustrated in Examples 4.6–4.7 below.

Remark 4.6. For large m, the function (4.33) is very flat near the optimal solution. Using a small ε as a stopping criterion, and stopping the computations with a final bracket $\Delta < \varepsilon$, does not guarantee that the final iterate $\mathbf{x}$ is close to the optimal solution, only that its value $f(\mathbf{x})$ is within ε of the optimal value.

Example 4.6. Consider a problem with $m = 1{,}000$ points, randomly generated in $[-5,5]^2$. The points are shown in Figure 4.3(a), and the level sets of the sum of distances (4.33) in Figure 4.3(b).

The initial lower bound was taken as $L = 0$ (a better lower bound would be the distance between any two points, but the method converges so fast that we do not save much by improving L.) For the needed Lipschitz constant we substitute the value $N = \frac{m}{2} = 500$, for no good reason other than it works.

The problem was solved with an initial point chosen randomly in $[-5,5]^2$. Table 4.2 shows the results for the initial $\mathbf{x} = (-.640353501, .937409957)$, with $U := f(\mathbf{x}) = 3809.901722$, and initial $\Delta = 3809.901722$. The value of the parameter β was fixed at $\frac{2}{3}$ and the bounds for α were $\alpha_{\min} = 0.2,\ \alpha_{\max} = 0.95$.

Using a tolerance $\varepsilon = 10^{-3}$, the method converged in 12 iterations. Case 1 occurred in 7 iterations, and Case 2 in 5 iterations (shown in bold numbers in Table 4.2). The lower bound $\alpha_{\min}$ was too low to be activated, but the upper bound $\alpha_{\max}$ applied in all but 3 iterations.

The last row of the table lists the bracket ratios, that are plotted in Fig. 4.4. The average reduction per iteration is

$$\left(\frac{0.0008}{3809.9}\right)^{1/12} = 0.277$$

Iteration	0	1	**2**	3	4	**5**	**6**	7	8	**9**	10	11	**12**
α	0.95	0.95	**0.95**	0.88	0.95	**0.95**	**0.95**	0.83	0.95	**0.95**	0.89	0.95	**0.95**
Case	1	1	**2**	1	1	**2**	**2**	1	1	**2**	1	1	**2**
Δ	3809.9	3771.1	**188.5**	172.3	167.3	**8.37**	**0.418**	0.367	0.355	**0.018**	0.0163	0.016	**0.0008**
Reduction		0.989	**0.05**	0.914	0.971	**0.05**	**0.05**	0.881	0.945	**0.05**	0.905	0.981	**0.05**

Table 4.2 Results for Example 4.6 with $\beta = \frac{2}{3}$, $\alpha_{\min} = 0.2,\ \alpha_{\max} = 0.95$

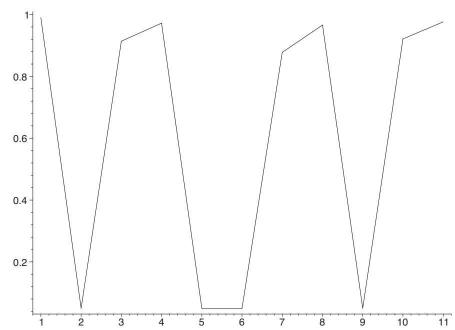


Fig. 4.4 Illustration of Example 4.6: Reduction per iteration for $\beta = \frac{2}{3}$, $\alpha_{min} = 0.2$, $\alpha_{\max} = 0.95$

Example 4.7. The results of Example 4.6 are typical: we solved 20 problems, each with 1000 random points in $[-5,5]^2$, and a random initial solution, using the same parameters as above,

$$\varepsilon = 10^{-3}, L = 0, \beta = \tfrac{2}{3}, \alpha_{\min} = 0.2, \alpha_{\max} = 0.9.$$

Table 4.3 shows the total number of iterations (until a bracket with length $\leq 10^{-3}$ is reached), the number of iterations of Case 2 (5 in all but one problem), and the average reduction per iteration, that is around 30%.

Problem	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Iterations	14	12	16	15	12	13	12	11	13	15	13	14	11	14	13	14	13	13	13	11
Case 2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	6	5
Reduction	.33	.28	.37	.35	.27	.30	.27	.24	.30	.35	.30	.32	.25	.33	.30	.32	.30	.30	.24	.25

Table 4.3 Results for Example 4.7, with 20 random problems , $\beta = \frac{2}{3}$, $\alpha_{\min} = 0.2,\ \alpha_{\max} = 0.95$

Remark 4.7. The NB method for convex minimization is based on the fact that the graph of a convex function f is supported by its tangents, but differentiability of f, i.e. uniqueness of tangents, is not required. The NB method can therefore be translated for the minimization of non-differentiable convex functions, using the subgradient methods of Shor [14], [15]; see also [16], [13], [5], to mention but a few.

Acknowledgements We thank the referees, and Professor A. Cegielski, for their constructive suggestions.

References

1. Ben-Israel, A., Levin, Y.: The Newton bracketing method for the minimization of convex functions subject to affine constraints. Discrete Appl. Math. **156**, 1977–1987 (2008)
2. Cegielski, A.: A method of projection onto an acute cone with level control in convex minimization. Math. Prog. **85**, 469–490 (1999)
3. Drezner, Z., Klamroth, K., Schöbel, A., Wesolowsky, G.: The Weber problem. In: Z. Dezner, H.W. Hamacher (eds.) Facility Location: Applications and Theory, Springer (2002)
4. Fortin, C., Wolkowicz, H.: The trust region subproblem and semidefinite programming. Optimization Methods and Software **19**, 41–67 (2004)
5. Held, M., Wolfe, P., Crowder, H.: Validation of subgradient optimization. Math. Prog. **6**, 62–88 (1974)
6. Kim, S., Ahn, H., Cho, S.-C.: Variable target value subgradient method. Math. Prog. **49**, 359–369 (1991)
7. Levin, Y., Ben-Israel, A.: Directional Newton methods in n variables. Math. of Comput. **71**, 251–262 (2001)
8. Levin, Y., Ben-Israel, A.: The Newton bracketing method for convex minimization. Comput. Optimiz. & Appl. **21**, 213–229 (2002)
9. Levin, Y., Ben-Israel, A.: A heuristic method for large–scale multifacility location problems. Computers and Operations Research **31**, 257–272 (2004)
10. Love, R.F., Morris, J.G., Wesolowsky, G.O.: Facilities Location: Models and Methods. North-Holland (1988)
11. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, London (1970)
12. Ostrwoski, A.M.: Solutions of Equations in Euclidean and Banach Spaces, 3rd edn. Academic Press (1973)
13. Poljak, B.T.: Minimization of unsmooth functionals. Z. Vycis. Mat. i Mat. Fiz. **9**, 509–521 (1969)
14. Shor, N.Z.: Application of the Gradient Method for the solution of Network Transportation Problems. Notes, Scientific Seminar on Theory and Application of Cybernetics and Operations Research, Academy of Sciences, Kiev (1962)
15. Shor, N.Z.: Generalized gradient methods for non-smooth functions and their applications to mathematical programming problems. Ekonomika i Matematicheskie Metody **2**, 337–356 (1976).
16. Shor, N.Z., Kiwiel, K., Ruszcayǹski, A.: Minimization methods for non-differentiable functions. Springer (1985)
17. Wesolowsky, G.O.: The Weber problem: its history and perspectives. Location Science **1**, 5–23 (1993)

Appendix A: Proof of Theorem 4.2

Part (a). Proof that

$$\|\nabla f(\mathbf{x}_+)\| \geq \frac{\beta - 1}{\beta} \|\nabla f(\mathbf{x})\| . \tag{A1}$$

Since $f \in C_N^{1,1}(X_0)$,

$$\|\nabla f(\xi) - \nabla f(\mathbf{x})\| \ \leq \ N\|\xi - \mathbf{x}\| \tag{A2}$$

for all $\xi \in X_0$. In particular, for $\mathbf{x}_+$,

$$\begin{aligned} \|\nabla f(\mathbf{x}_+) - \nabla f(\mathbf{x})\| &\leq N\|\mathbf{x}_+ - \mathbf{x}\| = N\,\frac{|f(\mathbf{x}) - M|}{\|\nabla f(\mathbf{x})\|} \\ &\leq \frac{1}{\beta}\,\|\nabla f(\mathbf{x})\|, \text{ by (4.26)}\,. \qquad \text{(A3)} \\ \therefore\ \|\nabla f(\mathbf{x}_+)\| &\geq \|\nabla f(\mathbf{x})\| - \|\nabla f(\mathbf{x}_+) - \nabla f(\mathbf{x})\|\,, \\ &\geq \|\nabla f(\mathbf{x})\| - \frac{1}{\beta}\|\nabla f(\mathbf{x})\|\,, \text{ by (A3)}\,, \\ &= \frac{\beta - 1}{\beta}\|\nabla f(\mathbf{x})\|\,, \text{ proving (A1)}\,. \end{aligned}$$

Part (b). The proof that

$$|f(\mathbf{x}_+) - M| \leq \frac{1}{2\beta}|f(\mathbf{x}) - M| \tag{A4}$$

is analogous to that of Lemma 4.1.

Since $f \in C_N^{1,1}(X_0)$,

$$f(\mathbf{x}^+) - f(\mathbf{x}) \leq \nabla f(\mathbf{x})^T(\mathbf{x}^+ - \mathbf{x}) + \frac{N}{2}||\mathbf{x}^+ - \mathbf{x}||^2,$$

by the descent lemma, [11, 3.2.12, page 73]. Therefore,

$$\begin{aligned} |f(\mathbf{x}^+) - M| &\leq \frac{N}{2}\frac{(f(\mathbf{x}) - M)^2}{||\nabla f(\mathbf{x})||^2}\,, \text{by (4.24)} \\ &\leq \frac{1}{2\beta}|f(\mathbf{x}) - M|\,, \text{by (4.26), proving (A4).} \end{aligned}$$