Expansion and Estimation of the Range of Nonlinear Functions *

S. M. Rump, Hamburg

Abstract

Many verification algorithms use an expansion $f(x) \in f(\tilde{x}) + S \cdot (x - \tilde{x})$, $f : \mathrm{IR}^n \to \mathrm{IR}$ for $x \in X$, where the set of matrices S is usually computed as a gradient or by means of slopes. In the following, an expansion scheme is described frequently yielding sharper inclusions for S. This allows also to compute sharper inclusions for the range of f over a domain. Roughly speaking, f has to be given by means of a computer program. The process of expanding f can then be fully automized.

The function f may be non-differentiable. For locally convex functions special improvements are described. Moreover, in contrast to other methods $\tilde{x} \cap X$ may be empty without implying large overestimations for S. This may be advantageous in practical applications.

AMS Subject Classification: 65G10

0 Notation

IIR denotes the set of real intervals

$$X \in \mathrm{IIR} \;\Rightarrow\; X = [\inf(X), \sup(X)] = \{\, x \in \mathrm{IR} \mid \inf(X) \le x \le \sup(X)\,\}.$$

IPT denotes the power set over a given set T, and $\mathrm{IIR} \subseteq \mathrm{IPIR}$. IIR^n denotes the set of n-dimensional interval vectors, i.e.

$$X \in \mathrm{IIR}^n \quad\Rightarrow\quad X = \{\,(x_i) \mid x_i \in X_i\,\} \quad \text{with } X_i \in \mathrm{IIR},\ 1 \le i \le n.$$

*published in Mathematics of Computation, 65(216):1503–1512, 1996

Interval vectors are closed and bounded. Interval operations and power set operations are defined in the usual way. Details can be found in standard books on interval analysis, among others [9], [2], [10]. If not explicitly noted otherwise, all operations are power set operations.

1 Expansion of nonlinear functions

A differentiable function $f : D \subseteq \mathrm{IR}^n \to \mathrm{IR}$ can be locally expanded by its gradient. For $\widetilde{x} \in D$, $X \subseteq D$, and $[g] \in \mathrm{IIR}^n$ with $\nabla f(\widetilde{x} \underline{\cup} X) \subseteq [g]$ holds

$$\forall\, x \in X\ \exists\, g \in [g] : \quad f(x) - f(\widetilde{x}) = g^T \cdot (x - \widetilde{x}). \tag{1.0}$$

The gradient, also for interval argument, can be computed using automatic differentiation [4], [11]. This process is fully automized. This approach bears three disadvantages:

1) f needs to be differentiable,

2) $[g]$ expands f w.r.t. *every* $\widetilde{x} \in \widetilde{x} \underline{\cup} X$ rather than w.r.t. some *specific* $\widetilde{x} \in X$,

3) $\widetilde{x} \underline{\cup} X$ has to be used enlarging $[g]$.

Number 2) means that (1.0) still holds if $\widetilde{x}$ in (1.0) is replaced by any $y \in \widetilde{x} \underline{\cup} X$. Number 3) expresses that according to the n-dimensional Mean-Value-Theorem for all $x \in X$ some $\zeta^{(i)} \in \widetilde{x} \underline{\cup} x$ exists with $(f(x) - f(\widetilde{x}))_i = \zeta^{(i)T} \cdot (x - \widetilde{x})$. Using $[g] \supseteq \Delta f(\widetilde{x} \underline{\cup} X)$ assures (1.0).

The three problems can be solved by means of so-called slopes. They have been introduced and described in [12], [6], [8], [10]. In the following, we give some generalization and improvement for slopes.

We start with a 1-dimensional function and will see that the approach easily extends to the n-dimensional case. The first steps very much follow the treatise in [10].

Definition 1. *Let $f : D \subseteq \mathrm{IR} \to \mathrm{IR}$ and $\widetilde{X}, X \subseteq D$ be given. The tripel $(f_c, f_r, f_s) \subseteq \mathrm{IR} \times \mathrm{IR} \times \mathrm{IR}$ expands f in X w.r.t. $\widetilde{X}$ if*

$$\begin{array}{lcl} \forall\, \widetilde{x} \in \widetilde{X} & : & f(\widetilde{x}) \in f_c \\ \forall\, x \in X & : & f(x) \in f_r \\ \forall\, \widetilde{x} \in X \quad \forall\, x \in X \quad \exists\, \widetilde{f}_s \in f_s & : & f(x) - f(\widetilde{x}) = \widetilde{f}_s \cdot (x - \widetilde{x}). \end{array}$$

Furthermore, the slope of f in X w.r.t. $\widetilde{X}$ is defined by

$$\mathrm{slope}\,(f) = \mathrm{slope}(f, \widetilde{X}, X) := \left\{ \frac{f(x) - f(\widetilde{x})}{x - \widetilde{x}} \mid \widetilde{x} \in \widetilde{X}, x \in X, x \neq \widetilde{x} \right\}.$$

For Theorem 3 we also need the following definition. If $\widetilde{X}, X \subseteq D$ either consist of a single point $\widetilde{x}, x \in D$, resp., then slope $(f) = (f(x) - f(\widetilde{x}))/(x - \widetilde{x})$ provided $\widetilde{x} \neq x$. We define

$$\underline{\text{slope}}\,(f, \widetilde{x}, x) := \begin{cases} \text{slope}\,(f, \widetilde{x}, x) & \text{if } \widetilde{x} \neq x \\ \overline{\lim_{\varepsilon \to 0^+}} \dfrac{f(x+\varepsilon) - f(x)}{\varepsilon} & \text{if } \widetilde{x} = x \end{cases}$$

and

$$\overline{\text{slope}}\,(f, \widetilde{x}, x) := \begin{cases} \text{slope}\,(f, \widetilde{x}, x) & \text{if } \widetilde{x} \neq x \\ \underline{\lim}_{\varepsilon \to 0^-} \dfrac{f(x+\varepsilon) - f(x)}{\varepsilon} & \text{if } \widetilde{x} = x \end{cases}$$

This definition is only needed for convex or concave f. The values for $\underline{\text{slope}}$, $\overline{\text{slope}}$ are allowed in $[-\infty, +\infty]$.

Instead of (f_c, f_r, f_s) we sometimes write $(f_c(\widetilde{X}, X), f_r(\widetilde{X}, X), f_s(\widetilde{X}, X))$ in order to emphasize the dependency on $\widetilde{X}$ and X. Clearly, if (f_c, f_r, f_s) expands f in X w.r.t. $\widetilde{X}$ then

$$f(\widetilde{X}) \subseteq f_c, \quad f(X) \subseteq f_r, \quad \text{slope}\ (f, \widetilde{X}, X) \subseteq f_s,$$

and $(f(\widetilde{X}), f(X), \text{slope}\ f(\widetilde{x}, X))$ expands f in X w.r.t. $\widetilde{X}$.

The following theorem can basically be found in [10].

Theorem 2. A constant $c \in \mathbb{R}$ and $f(x) \equiv x$ is expanded in X w.r.t. $\widetilde{X}$ by $(c, c, 0)$ and $(\widetilde{X}, X, 1)$, respectively. Let $f, g : D \subseteq \mathbb{R} \to \mathbb{R}$ and $\widetilde{X}, X \subseteq D$ be given. If (f_c, f_r, f_s), (g_c, g_r, g_s) expand f, g in X w.r.t. $\widetilde{X}$, resp., then (h_c, h_r, h_s) expands $f \circ g$ for $\circ \in \{+, -, \cdot, /\}$ in X w.r.t. $\widetilde{X}$ where

$$\begin{array}{ll} h_c := f_c \circ f_c, \quad h_r := f_r \circ g_r & \text{for } \circ \in \{+, -, \cdot, /\}, \\ h_s := f_s \circ g_s & \text{for } \circ \in \{+, -\}, \\ h_s := (f_s \cdot g_r + f_c \cdot g_s) \cap (f_r \cdot g_s + f_s \cdot g_c) & \text{for } \circ = \cdot, \\ h_s := (f_s - h_c \cdot g_s)/g_r \cap (f_s - h_r g_s)/g_c & \text{for } \circ = /, \end{array}$$

provided no division by zero occurs. The same holds for $h = g(f)$ with

$$\begin{array}{ll} h_c := & g(f_c), \quad h_r := g(f_r) \quad \text{and} \\ h_s := & \text{slope}(g, f_c, f_r) \cdot f_s. \end{array}$$

These statements remain valid if h_r is replaced by

$$h_r := h_r \cap \{h_c + h_s \cdot (X - \widetilde{X})\}.$$

The **proof** is demonstrated for $f \cdot g$, f/g, and $g(f)$. The others follow similarly. Computation of h_c and h_r is obvious. For $h = f \cdot g$

$$\text{slope}(f \cdot g) = \left\{ \frac{f(x) \cdot g(x) - f(\widetilde{x}) \cdot g(\widetilde{x})}{x - \widetilde{x}} \mid \widetilde{x} \in \widetilde{X}, x \in X, x \neq \widetilde{x} \right\}.$$

For fixed but arbitrary $\widetilde{x} \in X$, $x \in X$ there exist $\widetilde{f}_s \in f_s$, $\widetilde{g}_s \in g_s$ with

$$\begin{aligned} f(x) \cdot g(x) - f(\widetilde{x}) \cdot g(\widetilde{x}) &= [f(\widetilde{x}) + \widetilde{f}_s \cdot (x - \widetilde{x})] \cdot [g(\widetilde{x}) + \widetilde{g}_s \cdot (x - \widetilde{x})] - f(\widetilde{x}) \cdot g(\widetilde{x}) = \\ &= [\widetilde{f}_s \cdot g(x) + f(\widetilde{x}) \cdot \widetilde{g}_s] \cdot (x - \widetilde{x}). \end{aligned}$$

Hence slope $(f \cdot g) \subseteq f_s \cdot g_r + f_c \cdot g_s$ and therefore slope $(f \cdot g) =$ slope $(g \cdot f) \subseteq f_r \cdot g_s + f_s \cdot g_c$. For fixed but arbitrary $\widetilde{x} \in \widetilde{X}$, $x \in X$ there exist $\widetilde{f}_s \in f_s$, $\widetilde{g}_s \in g_s$ with

$$\begin{aligned} \frac{f(x)}{g(x)} - \frac{f(\widetilde{x})}{g(\widetilde{x})} &= \frac{[f(\widetilde{x}) + \widetilde{f}_s \cdot (x - \widetilde{x})] \cdot g(\widetilde{x}) - f(\widetilde{x}) \cdot [g(\widetilde{x}) + \widetilde{g}_s \cdot (x - \widetilde{x})]}{g(x) \cdot g(\widetilde{x})} \\ &= \frac{\widetilde{f}_s \cdot (g(x) - \widetilde{g}_s \cdot (x - \widetilde{x})) - (f(x) - \widetilde{f}_s \cdot (x - \widetilde{x})) \cdot \widetilde{g}_s}{g(x) \cdot g(\widetilde{x})} \cdot (x - \widetilde{x}) \\ &= \left\{ \frac{\widetilde{f}_s}{g(\widetilde{x})} - \frac{f(x) \cdot \widetilde{g}_s}{g(x) g(\widetilde{x})} \right\} \cdot (x - \widetilde{x}) \end{aligned}$$

proving the second part of the formula for h_s for division. The first part can be derived similarly, cf. also [10]. For $h = g(f)$ holds

$$\text{slope}\,(h) = \left\{ \frac{g(f(x)) - g(f(\widetilde{x}))}{x - \widetilde{x}} \mid \widetilde{x} \in \widetilde{X},\ x \in X,\ x \neq \widetilde{x} \right\}.$$

For $f(x) \neq f(\widetilde{x})$, holds

$$g(f(x)) - g(f(\widetilde{x})) = \frac{g(f(x)) - g(f(\widetilde{x}))}{f(x) - f(\widetilde{x})} \cdot \frac{f(x) - f(\widetilde{x})}{x - \widetilde{x}} \in \text{slope}\,(g, f_c, f_r) \cdot f_s.$$

If $f(x) = f(\widetilde{x})$ for $x \neq \widetilde{x}$, then $0 \in f_s$, together yielding slope$(g(f)) \subseteq$ slope$(g, f_c, f_r) \cdot f_s$. ■

Note that intersection for multiplication and division is possible because we are in the 1-dimensional case. This is no longer possible in n dimensions because the slope is no longer unique.

The intersection for h_r in Theorem 2 combines naive interval evaluation with centered forms. For example, $(\widetilde{X}, X, 1)$ expands the function $f(x) \equiv x$ in X w.r.t. $\widetilde{X}$. Defining $g(x) := x - x$ we obtain $g(X) \subseteq X - X = [-\operatorname{diam}(X), +\operatorname{diam}(X)]$ by naive interval calculation. The intersection $g_r := g_r \cap (g_c + g_s(X - \widetilde{X}))$ yields $h_r = g_r \cap (g_c + 0 \cdot (X - \widetilde{X})) = g_c = 0$.

It has already been observed in [10] that slope (g, f_c, f_r) can be replaced by $g'(f_c \underline{\cup} f_r)$ if g is differentiable. The disadvantage is that this set may be big. It covers slope $(g, f_c \underline{\cup} f_r, f_c \underline{\cup} f_r)$ thus expanding g in $f_c \underline{\cup} f_r$ with respect to each $\widetilde{x} \in f_c \underline{\cup} f_r$. In special cases, slope (g, f_c, f_r) can be computed explicitly. For example, let $g(x) = x^2$. Then for every $\widetilde{y} \in f_c$, $y \in f_r$, $y \neq \widetilde{y}$ holds

$$\frac{g(y) - g(\widetilde{y})}{y - \widetilde{y}} = \frac{y^2 - \widetilde{y}^2}{y - \widetilde{y}} = y + \widetilde{y} \quad \Rightarrow \quad \text{slope}\,(g, f_c, f_r) \subseteq f_c + f_r. \tag{1}$$

For $g(x) = \sqrt{x}$ holds

$$\frac{g(y) - g(\widetilde{y})}{y - \widetilde{y}} = \frac{\sqrt{y} - \sqrt{\widetilde{y}}}{y - \widetilde{y}} = (\sqrt{y} + \sqrt{\widetilde{y}})^{-1} \Rightarrow \text{slope}\,(g, f_c, f_r) \subseteq (g_c + g_r)^{-1}. \tag{2}$$

A similar principle can be extended to locally convex or concave functions. This may sharpen the inclusion interval for slopes significantly.

Theorem 3. Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$, $\widetilde{X}, X \subseteq D$ be given and $(f_c, f_r, f_s) \subseteq \mathbb{IR} \times \mathbb{IR} \times \mathbb{IR}$ expanding f in X w.r.t. $\widetilde{X}$. Let $g : D' \subseteq \mathbb{R} \to \mathbb{R}$, $\overline{f_c \underline{\cup} f_r} \subseteq D'$ be given and define $h_c := g(f_c)$, $h_r := g(f_r)$. If g is convex on $\overline{f_c \underline{\cup} f_r}$, then (h_c, h_r, h_s) expands $g(f)$ on X w.r.t. $\widetilde{X}$ when

$$h_s \supseteq \left[\underline{\text{slope}}\,(g, \inf(f_c), \inf(f_r)), \quad \overline{\text{slope}}\,(g, \sup(f_c), \sup(f_r))\right] \cdot f_s. \tag{3}$$

If g is concave on $\overline{f_c \underline{\cup} f_r}$, then the same holds for

$$h_s \supseteq \left[\overline{\text{slope}}\,(g, \sup(f_c), \sup(f_r)), \quad \underline{\text{slope}}\,(g, \inf(f_c), \inf(f_r))\right] \cdot f_s. \tag{4}$$

Proof. Let g be convex on $\overline{f_c \underline{\cup} f_r}$. We prove that slope (g, y_1, y_2) increases when y_1 or y_2 increase. Let $y_1 < y < y_2$ with $y = \alpha y_1 + (1-\alpha) y_2$, $0 < \alpha < 1$, $y_1, y, y_2 \in \overline{f_c \underline{\cup} f_r}$. Then due to convexity $g(\alpha y_1 + (1-\alpha) y_2) \leq \alpha g(y_1) + (1-\alpha) g(y_2)$ and

$$\frac{g(y_2) - g(y_1)}{y_2 - y_1} = \alpha \cdot \frac{g(y_2) - g(y_1)}{y_2 - y} \leq \frac{\alpha g(y_2) - g(y) + (1-\alpha) g(y_2)}{y_2 - y} = \frac{g(y_2) - g(y)}{y_2 - y}.$$

We proceed similarly for $y_1 \leq y_2 < y$. Thus slope (g, f_c, f_r) achieve its extreme values at the extremes of f_c and f_r and this proves (1.3). Concavity is treated similarly. ■

Note that in the calculation of (1.3) and (1.4) only slopes of g for points, not for intervals are necessary. Theorem 3 yields sharper slopes for many functions. Moreover, it extends the principle to non-differentiable functions such as absolute value.

As an example, consider e^{x^2} for $\widetilde{X} = \{1\}$ and $X = [0.5, 1.5]$. In the following table we compare a standard gradient evaluation $\nabla f(\tilde{x} \underline{\cup} X)$, the slope using slope $(g, f_c, f_r) \subseteq g'(f_c \underline{\cup} f_r)$ and the slope computed by Theorem 3. Results are rounded to 4 figures.

	$\nabla f(\widetilde{X} \underline{\cup} X)$	standard slope	new slope (Theorem 3)
x^2	$[1, 3]$	$[1.5, 2.5]$	$[1.5, 2.5]$
e^{x^2}	$[1.284, 28.46]$	$[1.926, 23.72]$	$[2.869, 13.54]$

Table 1. Expansions for $\widetilde{X} = \{1\}$, $X = [0.5, 1.5]$

Here we set $\widetilde{X} = \{\text{mid}(X)\}$. In practical applications, sometimes one cannot assure $\widetilde{X} \subseteq X$ unless extra function evaluations are spent. If we take the same $X = [0.5, 1.5]$ but $\widetilde{X} = \{2\}$ then Table 1 looks as follows.

	$\nabla f(\widetilde{X} \underline{\cup} X)$	standard slope	new slope (Theorem 3)
x^2	$[1, 4]$	$[2.5, 3.5]$	$[2.5, 3.5]$
e^{x^2}	$[1.284, 218.4]$	$[3.21, 191.1]$	$[35.54, 90.22]$

Table 2. Expansions for $\widetilde{X} = \{2\}$, $X = [0.5, 1.5]$

In order to apply Theorems 2 and 3 to the n-dimensional case, first we generalize Definition 1 in the following way.

Definition 4. Let $F : D \subseteq \mathbb{R} \to \mathbb{IPR}$ and $\widetilde{X}, X \subseteq D$ be given. The tripel $(F_c, F_r, F_s) \subseteq \mathbb{R}^3$ *expands* F in X w.r.t. $\widetilde{X}$ if

$$\begin{array}{l} \forall\, \tilde{x} \in \widetilde{X} : \quad F(\tilde{x}) \subseteq F_c \\ \forall\, x \in X : \quad F(x) \subseteq F_r \\ \forall\, \tilde{x} \in \widetilde{X} \quad \forall\, x \in X \quad \forall\, \tilde{y} \in F(\tilde{x}) \quad \forall\, y \in F(x) \quad \exists\, \tilde{F}_s \in F_s : \quad y - \tilde{y} = \tilde{F}_s \cdot (x - \tilde{x}) \end{array}$$

Furthermore, the slope of F w.r.t. X and $\widetilde{X}$ is defined by

$$\text{slope}\,(F) = \text{slope}\,(F, \widetilde{X}, X) := \left\{ \frac{y - \tilde{y}}{x - \tilde{x}} \mid y \in F(X), \tilde{y} \in F(\widetilde{X}), x \in X, \tilde{x} \in \widetilde{X}, x \neq \tilde{x} \right\}$$

As before we have $F(\widetilde{X}) \subseteq F_c$, $F(X) \subseteq F_r$ and $\text{slope}\,(F, \widetilde{X}, X) \subseteq F_s$ and $(F(\widetilde{X}), F(X),$ $\text{slope}\,(F, \widetilde{X}, X))$ expands F in X w.r.t. $\widetilde{X}$.

Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ and $\widetilde{X}, X \in \mathbb{IIR}^n$ with $\widetilde{X}, X \subseteq D$ be given. We need Definition 4 to apply it to the following "component functions" $F^k : \mathbb{R} \to \mathbb{PIR}$ defined by

$$F^k(y) := f(X_1, \ldots, X_{k-1}, y, \widetilde{X}_{k+1}, \ldots, \widetilde{X}_n) \quad \text{for} \quad 1 \leq k \leq n.$$

Then $(F^k(\widetilde{X}_k), F^k(X_k),\ \text{slope}\,(F^k, \widetilde{X}_k, X_k))$ expands F^k in X_k w.r.t. $\widetilde{X}_k$ and by induction follows for $x \in X$, $\tilde{x} \in \widetilde{X}$ and $0 \leq k \leq n$

$$f(x_1, \ldots, x_k, \tilde{x}_{k+1}, \ldots, \tilde{x}_n) \subseteq f(\tilde{x}) + \sum_{\nu=1}^{k} \text{slope}\,(F^\nu, \widetilde{X}_\nu, X_\nu) \cdot (X_\nu - \widetilde{X}_\nu).$$

Of course, Definition 4 could substitute Definition 1 at the beginning; we split it for didactical purposes. Theorem 2 and 3 can be adapted to Definition 4 and applied to every component function F^k. In the practical application, the vector $V = (V_0, \ldots, V_n) \in \mathbb{IIR}^{nn}$ with

$$f(X_1, \ldots, X_{k-1}, \widetilde{X}_k, \ldots, \widetilde{X}_n) \subseteq V_k \quad \text{for} \quad 0 \leq k \leq n \tag{5}$$

is stored and

$$F^k(\widetilde{X}_k) \subseteq V_{k-1} \quad \text{and} \quad F^k(X_k) \subseteq V_k$$

is used. The difference to Neumaier's approach [10] is that he stores $(f_c, f_r, f_s) \in \mathbb{IIR} \times \mathbb{IIR} \times \mathbb{IIR}^n$ with $f(\widetilde{X}) \subseteq f_c$, $f(X) \subseteq f_r$ and corresponding slope. In the approach described above more information is stored. It is very much in the spirit of Hansen [5], where the concept of componentwise application of the n-dimensional Mean-Value-Theorem is used to improve gradients, see also [1].

2 Implementation and examples

In the following we give some implementation hints and some computational results. We use a Pascal-like notation together with an operator concept. In fact, it is the notation of TPX (Turbo Pascal eXtended, [7]), a precompiler for Turbo Pascal offering these and other features. We use the data structure

```
slope = record
            r : array[0..n] of interval;
```

```
            s : array[1..n] of interval;
        end;
```

$\widetilde{X}, X \in \mathbb{IR}$ are fixed and globally avaible constants. $\widetilde{X}$ is denoted by Xs. Then $f.r$ represents the range vector V as in (1.4) and for all $\widetilde{x} \in \widetilde{X}$, $x \in X$ we have

$$f(x_1, \ldots, x_{k-1}, \widetilde{x}_k, \ldots, \widetilde{x}_n) \in f.r[i] \quad \text{for} \quad 0 \leq k \leq n$$
$$f(x) \in f(\widetilde{x}) + \sum_{\nu=1}^{n} f.s[i] \cdot (x_i - \widetilde{x}_i).$$

As an example we display the algorithm for the multiplication operator.

```
function mul(f, g : expansion): expansion    implements *;
var    i : integer;    R : interval;
begin
    R := f.r[0] * g.r[0];    mul.r[0] := R;
    for i := 1 to n do    begin
        mul.s[i] := intersection(f.r[i] * g.s[i] + g.r[i − 1] * f.s[i],
                                 g.r[i] * f.s[i] + f.r[i − 1] * g.s[i]);
        R := R + mul.s[i] * (X[i] − Xs[i]);
        mul.r[i] := intersection(f.r[i] * g.r[i], R);
    end;
end {mul};
```

Algorithm 2.1 Multiplication for strong expansions

The procedure gives enough detail for an implementation of basic operators for the computation of slope expansions. Note that all operations are interval operations. The main point is that replacing the data type *double* by *slope*, automatically creates the slope expansion. We close the implementation hints by giving a procedure for the absolute value, a convex but not everywhere differentiable function. It implements Theorem 3. We assume the function *abs* to be given for interval arguments, that is $abs(X) := \{ |x| \mid x \in X \}$.

```
function abs (f: expansion) : expansion;
var    i : integer;    R : interval;    Sl, Su : double;
begin
    R := abs(f.r[0]); abs.r[0] := R;
    for i := 1 to n do    begin
        if f.r[i].inf = f.r[i − 1].inf then
            if f.r[i].inf >= 0.0 then    Sl := 1.0 else    Sl := −1.0
        else
            Sl := (abs(f.r[i].inf) − abs(f.r[i − 1].inf)/(f.r[i].inf − f.r[i − 1].inf);
```

```
        if  f.r[i].sup = f.r[i − 1].sup  then
            if  f.r[i].sup <= 0.0  then    Su := −1.0    else  Su := 1.0
        else
            Su := (abs(f.r[i].sup) − abs(f.r[i − 1].sup)/(f.r[i].sup − f.r[i − 1].sup);
        abs.s[i] := hull(Sl, Su) ∗ f.s[i];
        R := R + abs.s[i] ∗ (X[i] − Xs[i]);
        abs.r[i] := intersection (abs(f.r[i]), R);
    end;
end  {abs};
```

Algorithm 2.2 Slope expansion of absolute value

For the implementation of not globally convex functions, case distinctions for local convexity have to be used. Next we compare the following three methods for expanding a function:

Method 1: Gradients ∇f
Method 2: Slopes according to [10]
Method 3: Slopes as described above

As an example for comparing these methods we use $f : \mathbb{R}^2 \to \mathbb{R}$ with

$$f(x, y) := e^{xy} - x \quad \text{for } X = [-1, 1],\ Y = [0, 2],\ \widetilde{X} = 0,\ \widetilde{Y} = 1. \tag{6}$$

Method 1 with automatic differentiation yields

	$z_i(X, Y)$	$\nabla z_i(X, Y)$	
$z_1 = X$	$[-1, 1]$	1	0
$z_2 = Y$	$[0, 2]$	0	1
$z_3 = z_1 \cdot z_2$	$[-2, 2]$	$[0, 2]$	$[-1, 1]$
$z_4 = e^{z_3}$	$[0.135,\ 7.390]$	$[0,\ 14.779]$	$[-7.390,\ 7.390]$
$z_5 = z_4 - z_1$	$[-0.865,\ 8.390]$	$[-1,\ 13.779]$	$[-7.390,\ 7.390]$

Table 2.1. Method 1 for (2.1)

The final value z_5 equals the function value at X. Slopes according to [10] compute as follows.

	$(z_i)_c$	$(z_i)_r$	$(z_i)_s$	
$z_1 = X$	0	$[-1, 1]$	1	0
$z_2 = Y$	1	$[0, 2]$	0	1
$z_3 = z_1 \cdot z_2$	0	$[-2, 2]$	1	$[-1, 1]$
$z_4 = e^{z_3}$	1	[0.135, 7.390]	[0.135, 7.390]	[−7.390, 7.390]
$z_5 = z_4 - z_1$	1	[−0.865, 8.390]	[−0.865, 6.390]	[−7.390, 7.390]

Table 2.2. Method 2 (slopes) for (2.1)

The results of Table 2.2 are exactly the same for the componentwise definition of gradients according to Hansen [5]. The estimation for the range $f(X, Y)$ is the same as for method 1. It cannot be improved by using

$$f(X, Y) \subseteq \{f_c + f_s \cdot (X - \widetilde{X}, Y - \widetilde{Y})^T\} \ \cap \ f_r.$$

Finally we give the results for the new method 3 using Theorems 2 and 3.

	$(R_{z_i})_0$	$(R_{z_i})_1$	$(R_{z_i})_2$	$(S_{z_i})_1$	$(S_{z_i})_2$
$z_1 = X$	0	$[-1, 1]$	$[-1, 1]$	1	0
$z_2 = Y$	1	1	$[0, 2]$	0	1
$z_3 = z_1 \cdot z_2$	0	$[-1, 1]$	$[-2, 2]$	1	$[-1, 1]$
$z_4 = e^{z_3}$	1	[0.367, 2.719]	[0.135, 7.390]	[0.633, 1.719]	[−4.671, 4.671]
$z_5 = z_4 - z_1$	1	[0.281, 1.719]	[−0.865, 6.390]	[−0.367, 0.719]	[−4.671, 4.671]

Table 2.3. Method 4 for (2.1)

Method 3 stores more information and computes better inclusions. The last line of Table 2.3 shows a sharper inclusion [−0.865, 6.390] for the range $f(X, Y)$ and sharper slopes. The true range is [0, 6.390], thus the upper bound is already sharp.

Slope expansions for non-continuous functions like signum (x) or $\lfloor x \rfloor := \max\{\, k \in \mathbb{Z} \mid k \leq x \,\}$ according to Theorem 2 and 3 can easily be implemented along the lines of Algorithms 2.1 and 2.2.

The following example is taken from Broyden's function [3]

$$f(x_1, x_2) := (1 - 1/(4\pi)) \cdot (e^{2x_1} - e) + x_2 \cdot e/\pi - 2ex_1.$$

Setting $\widetilde{X} = (0.5, \pi)$ and $X := \widetilde{X} \cdot [1 - 0.3, 1 + 0.3]$ we obtain the following ranges computed by methods 1), 2), and 3) without and with using Theorem 3.

method	range $f(X)$
method 1	[-3.019, + 3.947]
method 2	[-3.019, + 3.947]
method 3 without Theorem 3	[-1.364, + 1.364]
method 3 with Theorem 3	[-0.761, + 0.762]

We summarize the main properties of our expansion scheme.

- The method is applicable to rather general, non-differentiable and even non-continuous functions and can be used in an automized way similar to automatic differentiation.
- The quality of the inclusions is improved through various intersections and special treatment of locally convex or concave functions.
- In practical applications, expansions may be necessary w.r.t. some $\tilde{x} \in \mathbb{IR}^n$ not being exactly representable on the computer; therefore $\widetilde{X} \in \mathbb{IIR}^n$ is used instead of $\tilde{x}$.
- We do not require $\widetilde{X} \subseteq X$ nor use $\widetilde{X} \,\underline{\cup}\, X$.
- The computational effort and storage as compared to standard slopes increase by about a factor of 2.

References

[1] G. Alefeld. Intervallanalytische Methoden bei nichtlinearen Gleichungen. In S.D. Chatterji et al., editor, *Jahrbuch Überblicke Mathematik 1979*, pages 63–78. Bibliographisches Institut, Mannheim, 1979.

[2] G. Alefeld and J. Herzberger. *Introduction to Interval Computations.* Academic Press, New York, 1983.

[3] C.G. Broyden. A new method of solving nonlinear simultaneous equations. *Comput. J.*, 12:94–99, 1969.

[4] A. Griewank. *On Automatic Differentiation*, volume 88 of *Mathematical Programming.* Kluwer Academic Publishers, Boston, 1989.

[5] E.R. Hansen. On Solving Systems of Equations Using Interval Arithmetic. *Math. Comput. 22*, pages 374–384, 1968.

[6] E.R. Hansen. A generalized interval arithmetic. In K. Nickel, editor, *Interval Mathematics*, volume 29, pages 7–18. Springer, 1975.

[7] D. Husung. Precompiler for Scientific Computation (TPX). Technical Report 91.1, Inst. f. Informatik III, TU Hamburg-Harburg, 1989.

[8] R. Krawczyk and A. Neumaier. Interval Slopes for Rational Functions and Associated Centered Forms. *SIAM J. Numer. Anal.*, 22(3):604–616, 1985.

[9] R.E. Moore. *Interval Analysis.* Prentice-Hall, Englewood Cliffs, N.J., 1966.

[10] A. Neumaier. *Interval Methods for Systems of Equations.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1990.

[11] L.B. Rall. Automatic Differentiation: Techniques and Applications. In *Lecture Notes in Computer Science 120.* Springer Verlag, Berlin-Heidelberg-New York, 1981.

[12] J.W. Schmidt. Die Regula-Falsi für Operatoren in Banachräumen. *Angew. Math. Mech.*, 41:61–63, 1961.