# Interval Krawczyk and Newton method

February 20, 2007

## 1 Interval Newton method

In the presentation of the method we follow [A].

**Theorem 1** *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a $C^1$ function. Let $X = \Pi_{i=1}^n [a_i, b_i]$, $a_i < b_i$. Assume the interval enclosure of $Df(X)$, denoted here by $[Df(X)]$, is invertible. Let $x_0 \in X$ and we define*

$$N(x_0, X) = -[Df(X)]^{-1} f(x_0) + x_0 \tag{1}$$

*Then*

**0.** *if $x_1, x_2 \in X$ and $f(x_1) = f(x_2)$, then $x_1 = x_2$*

**1.** *if $N(x_0, X) \subset X$, then $\exists! x^* \in X$ such that $f(x^*) = 0$*

**2.** *if $x_1 \in X$ and $f(x_1) = 0$, then $x_1 \in N(x_0, X)$*

**3.** *if $N(x_0, X) \cap X = \emptyset$, then $f(x) \neq 0$ for all $x \in X$*

**Proof** We will show first that invertibility of $[Df(X)]$ implies that $f$ is injective on $X$, which implies assertion **0** and the uniqueness part in assertion **1**. We have for any $x_0, x_1 \in X$

$$f(x_1) - f(x_0) = \int_0^1 \frac{df}{dt} f(x_0 + t(x_1 - x_0)) dt = \int_0^1 \frac{\partial f}{\partial x}(x_0 + t(x_1 - x_0)) dt \cdot (x_1 - x_0) \tag{2}$$

Let us denote by $J(x_1, x_0) = \int_0^1 \frac{\partial f}{\partial x}(x_0 + t(x_1 - x_0)) dt$. Obviously $J(x_1, x_0) \in [Df(X)]$, hence $J(x_1, x_0)$ is invertible for any choice of $x_1, x_0 \in X$.

We can rewrite (2) as follows

$$f(x_1) - f(x_0) = J(x_1, x_0)(x_1 - x_0), \quad \text{for any } x_0, x_1 \in X \tag{3}$$

We are now ready to show that if $f(x_1) = f(x_0)$ and $x_0, x_1 \in X$ then $x_1 = x_0$. From (3) it follows that $J(x_1, x_0)(x_1 - x_0) = 0$. Hence from invertibility of $J(x_1, x_0)$ it follows that $x_1 - x_0 = 0$.

We will prove **1** now. Consider the map $P$

$$X \ni x \longrightarrow P(x) = -J(x, x_0)^{-1} f(x_0) + x_0$$

From the above considerations it follows that $P$ is well defined. Observe that $P(x) \in N(x_0, X) \subset X$, so $P(X) \subset X$. By the Brouwer theorem it follows that $P$ has a fixed point $x^* \in X$.

We show that $f(x^*) = 0$. We have

$$\begin{aligned} x^* &= -J(x^*, x_0)^{-1} f(x_0) + x_0 \\ -f(x_0) = J(x^*, x_0)(x^* - x_0) &= f(x^*) - f(x_0) \\ f(x^*) &= 0 \end{aligned}$$

Redoing the above transformations backwards, shows that each zero of $f$ in $X$ is the fixed point of $P$, hence must be in $N(x_0, X)$. This proves assertions **2** and **3**. ∎

# 2 Krawczyk method

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a $C^1$-function. We would like to solve the equation

$$F(x) = 0 \tag{4}$$

## 2.1 Motivation, heuristic derivation

We begin by explaining the basic idea of the Krawczyk method. The Newton method is given by

$$N(x) = x - dF(x)^{-1} F(x). \tag{5}$$

It is well know that if $F(x^*) = 0$ and $dF(x^*)$ is nonsingular, then $x^*$ is an attracting fixed point for $N(x)$.

It turns out that the same is true if we replace $dF(x)^{-1}$ by a fixed matrix $C$, which is sufficiently close to $dF(x^*)^{-1}$. The modified Newton operator is given by

$$N_m(x) = x - CF(x). \tag{6}$$

Now let us turn the things around and ask how can we use $N_m$ as a way to prove the existence of solution of (4).

This is quite obvious. Namely, if $U$ is homeomorphic to a closed finite-dimensional ball and if

$$N_m(U) \subset U, \tag{7}$$

then from the Brouwer Theorem it follows, that there exists $x_0 \in U$ such that $N_m(x_0) = x_0$. Since $C$ is invertible we obtain that $F(x_0) = 0$. To obtain the uniqueness it is enough show that $N_m$ is a contraction on $U$.

Observe that it is impossible to verify in a single interval evaluation of the formula (6), that for some interval set $[x]$ holds $N_m([x]) \subset [x]$, because the computed diameter of $[x] - CF[x]$ is greater than or equal to $\operatorname{diam}([x]) + \operatorname{diam}(CF([x]))$.

It turns out the middle value form of $N_m$ can cure this deficiency. If $x_0 \in [x]$, then

$$\begin{aligned} N_m([x]) \subset N_m(x_0) + [dN_m([x])]_I \cdot ([x] - x_0) = \\ x_0 - CF(x_0) + (Id - C[df([X])]_I)([x] - x_0) = K(x_0, [x], F). \end{aligned}$$

This explains why the requirement $K(x_0, [x], F) \subset [x]$ has something to do with zeros of $F(x)$.

## 2.2 The Krawczyk method

A method proposed by Krawczyk for finding zero's of $F$:

- $[x] \subset \mathbb{R}^n$ be an interval set (i.e. product of intervals),
- $x_0 \in [x]$
- $C \in \mathbb{R}^{n \times n}$ be a linear isomorphism

The Krawczyk operator is given by

$$K(x_0, [x], F) := x_0 - CF(x_0) + (Id - C\,[dF([x])]_I)([x] - x_0). \tag{8}$$

**Theorem 2** **1.** *If $x^* \in [x]$ and $F(x^*) = 0$, then $x^* \in K(x_0, [x], F)$.*

**2.** *If $K(x_0, [x], F) \subset \operatorname{int}[x]$, then there exists in $[x]$ exactly one solution of equation $F(x) = 0$.*

**Proof of 1.**

$$N_m(x^*) = x^* \in K(x_0, [x]).$$

∎

Before we prove second assertion we will need several lemmas.

**Lemma 3** *Assume $f_0 \in \mathbb{R}^n$, $X, Y \subset \mathbb{R}^n$, $X = -X$, $Y = -Y$ and $X, Y$ are convex. If*

$$f_0 + AX \subset Y, \tag{9}$$

*then $AX \subset Y$.*

**Proof:** Let $x_1 \in X$. We have

$$\begin{aligned} & & f_0 + Ax_1 \in Y, \\ f_0 + A(-x_1) \in Y, & \quad \Rightarrow \quad & Ax_1 - f_0 \in Y. \end{aligned}$$

Hence $Ax_1 = \frac{1}{2}(Ax_1 + f_0) + \frac{1}{2}(Ax_1 - f_0) \in Y$. ∎

**Lemma 4** *Assume $f_0, x_s, y_s \in \mathbb{R}^n$, $X, Y \subset \mathbb{R}^n$, $X = -X$, $Y = -Y$ and $X, Y$ are convex. If*

$$f_0 + A(x_s + X) \subset y_s + Y, \tag{10}$$

*then $AX \subset Y$.*

**Proof:** Since

$$(f_0 + Ax_s - y_s) + AX \subset Y, \tag{11}$$

and the assertion follows directly from Lemma 3. ∎

**Proof of Theorem 2:** Since $N_m([x]) \subset K(x_0, [x]) \subset \text{int}\,[x]$, hence the existence of $x^*$ such that $F(x^*) = 0$ follows from (6) and the Brouwer Theorem.

We show now that $N_m$ is a contraction on $[x]$ in a suitable norm.

Let $A \in Id - C[dF([x])]_I$ and $f_0 = x_0 - CF(x_0)$. We have

$$f_0 + A([x] - x_0) \subset \text{int}\,[x] \tag{12}$$

$$(f_0 - x_0) + A([x] - x_0) \subset \text{int}\ ([x] - x_0) \tag{13}$$

Since $[x] - x_0$ is a product of intervals, hence there exists $x_s \in \mathbb{R}^n$ and $X_s = \Pi_{i=1}^n [-z_i, z_i]$, such that

$$[x] - x_0 = x_s + X_s.$$

Since $X_s = -X_s$ and $X_s$ is convex we can apply Lemma 4 to equation (13) with $Y = x_s + \text{int}\,X_s$ to obtain

$$AX_s \subset \text{int}\,X_s. \tag{14}$$

Since the set $Id - C[dF([x])]_I$ is compact, then there exists $\alpha < 1$ such that

$$(Id - C[dF([x])]_I)X_s \subset \alpha X_s. \tag{15}$$

This show that in the norm in which $X_s$ is a ball we have

$$|Id - C[dF([x])]_I| \leq \alpha < 1. \tag{16}$$

Hence $N_m$ is a contraction on $[x]$ and has at most one fixed point. ∎

If we consider a fixed point problem $x = P(x)$, then the Krawczyk operator is given by

$$K(x_0, [x], Id-P) = x_0 - C(x_0 - P(x_0)) + (Id - C(Id - [dP([x])]))([x] - x_0) \tag{17}$$

and it makes sense to chose $C \approx (Id - dP(x_0))^{-1}$.

# 3 What is better $C^0$-tools or $C^1$ for fixed points

$f : \mathbb{R}^n \to \mathbb{R}^n$, $C^1$. Assume that we apparently have $x_0$, $f(x_0) = x_0$ and $\det(df(x_0)) \neq 0$. $x_0$ is an isolated fixed point.

Generally $C^1$ tools are easier to apply and in fact they are faster.

**Advantages and disadvantages of $C^0$-methods**(Brouwer Thm, Miranda Thm, covering relations)

+ require $C^0$ computation
- conditions to check differ depend on the dynamical type of $x_0$ , this may result in the need of multiple computations
- the set on which the conditions are checked has to be carefully chosen (based on the diagonalization of $df(x_0)$)

**Advantages and disadvantages of $C^1$-methods**(interval Newton method, Krawczyk method)

- require $C^1$ computation, but for ODEs this almost as fast as $C^0$-computations (use $C^1$-Lohner algorithm)
+ conditions to check differ **do not depend** on the dynamical type of $x_0$
+ as the test sets we can always chose a small box around numerical approximation of $x_0$

## 4 Continuous families of solutions

Kapela, Simo

Periodic orbits for ODEs in the presence of first integrals

**Theorem 5** *Let* $X \subset \mathbb{R}^{k+m}$ *and* $(z_0, c_0) \in \mathbb{R}^k \times \mathbb{R}^m$*. Let* $G : X \to \mathbb{R}^{m+k}$ *and* $J : X \to \mathbb{R}^m$ *be* $C^1$ *functions, such that*

$$\pi_z G(z_0, c_0) = z_0, \quad J(z_0, c_0) = J(G(z_0, c_0)).$$

*Let* $Z \subset \mathbb{R}^k$ *and* $C \subset \mathbb{R}^m$ *be interval sets such that*

$$z_0 \in Z, \quad c_0 \in C, \quad [\pi_c(G(Z, c_0))] \subset C.$$

*Then if the interval matrix* $\left[\frac{\partial J}{\partial c}(Z, C)\right]$ *is invertible, then* $G(z_0, c_0) = (z_0, c_0)$

## References


[A] G. Alefeld *Inclusion methods for systems of nonlinear equations - the interval Newton method and modifications.* in *Topics in Validated Computations* J. Herzberger (Editor), 1994 Elsevier Science B.V.

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

[N] A. Neumeier, *Interval methods for systems of equations.* Cambridge University Press, 1990