Newton's iteration for quadratic equations

David L. Duncan

Contents

1 Overview

In these notes I review Newton's iteration associated to an equation $\mathcal{F}(x) = 0$. I will address convergence of the iteration, and the induced map that is defined when the initial condition is allowed to vary. I am primarily interested in how the various constants depend on the first and second derivatives of $\mathcal{F}$. The primary application is to the case where $\mathcal{F}$ is a linear expression in the curvature of a connection, so I largely focus on the case where the defining equation $\mathcal{F}$ is quadratic.

In many ways my argument below mimics that of [1, Section 5]. Indeed, my original aim was to extend their result (which applies to mapping tori) to the quilted case described in [2] (which applies, more generally, to manifolds with positive first Betti number). Unfortunately, the scaling does not behave favorably in the more general quilted case. I hope to eventually include a more detailed description of this failure in these notes.

2 Banach space set-up

Suppose $\mathcal{X}, \mathcal{Y}$ are Banach spaces and $\mathcal{F} : \mathcal{X} \to \mathcal{Y}$ is a smooth function. Using Taylor's theorem, we can write

$$\mathcal{F}(x + v) = \mathcal{F}(x) + D_x\mathcal{F}(v) + R_x(v)$$

for some quadratically bounded $R_x(v)$. For our applications we will be interested in the case where $\mathcal{F}$ is quadratic, in which case $R_x(v) = D_x^2\mathcal{F}(v, v)$ is a quadratic form (the Hessian).

Lemma 2.1. *For each $x \in \mathcal{X}$ there are constants $d_2 > 0$ and $\epsilon_0 > 0$, possibly depending on x, such that*

$$\|R_x(v) - R_x(w)\|_{\mathcal{Y}} \leq d_2 \left(\|v\|_{\mathcal{X}} + \|w\|_{\mathcal{X}} \right) \|v - w\|_{\mathcal{X}}$$

for all $v, w \in T_x\mathcal{X} = \mathcal{X}$ with $\|v\|_{\mathcal{X}}, \|w\|_{\mathcal{X}} < \epsilon_0$. If $\mathcal{F}$ is quadratic, then the result holds with $\epsilon_0 = \infty$.

Proof. We prove this in the case where $\mathcal{F}$ is quadratic; for the more general case, see any book on the implicit function theorem. In the quadratic case, we have

$$R_x(v) - R_x(w) = D_x^2\mathcal{F}(v, v) - D_x^2\mathcal{F}(w, w) = D_x^2\mathcal{F}(v, v - w) + D_x^2\mathcal{F}(w, v - w).$$

Then the result follows by taking d_2 to be the operator norm of $D_x^2\mathcal{F}$. □

In particular, we have the following bound:

$$\|R_x\,(v)\,\|_{\mathcal{Y}} \leq d_2\|v\|_{\mathcal{X}}^2, \tag{1}$$

for all v satisfying the conditions of the Lemma.

We also need an estimate on the first derivative $D_x\mathcal{F}$. We begin by assuming that $D_x\mathcal{F}$ is surjective, and let M_x be a right inverse:

$$(D_x\mathcal{F})M_x = \mathrm{Id}_{\mathcal{Y}}.$$

We assume that M_x satisfies

$$\|M_x w\|_{\mathcal{X}} \leq d_1\|w\|_{\mathcal{Y}}. \tag{2}$$

In general $d_1 = d_1(x)$ is independent of $w \in T_{\mathcal{F}(x)}\mathcal{Y} = \mathcal{Y}$, but may depend on $x \in \mathcal{X}$.

Theorem 2.2. *(Newton's iteration) Suppose $D_x\mathcal{F}$ is surjective for all $x \in X$, and let d_1, d_2 and ϵ_0 be as above. Assume these constants can be chosen to be independent of $x \in \mathcal{X}$. Let $\delta > 0$ be any constant with $\delta d_1 \leq \epsilon_0$ and $\delta^2 d_1 < 1$. Suppose there is some $x_0 \in \mathcal{X}$ with $\|\mathcal{F}(x_0)\|_{\mathcal{Y}} \leq \delta$. Then there is some $x_\infty \in \mathcal{X}$ with $\mathcal{F}(x_\infty) = 0$, and*

$$\|x_\infty - x_0\|_{\mathcal{X}} \leq \frac{d_1 d_2}{1 - \delta^2 d_1}.$$

Moreover, there is at most one zero x_∞ of $\mathcal{F}$ satisfying

$$\|x_\infty - x_0\| < \min\left\{ \frac{1}{2d_1 d_2}, \epsilon_0 \right\}.$$

This theorem implies the existence of a map

$$\mathcal{T} : \mathcal{F}^{-1}(B_\delta(0)) \longrightarrow \mathcal{X}$$

with the property that $\mathcal{F}(\mathcal{T}(x)) = 0$ for all $x \in \mathcal{F}^{-1}(B_\delta(0))$, and $\mathcal{T}$ is locally injective.

Remark 2.3. *(a) More generally, suppose the image* $\operatorname{im} D_x\mathcal{F} \subset \mathcal{Y}$ *is closed as a Banach subspace that is independent of* x*. Then by replacing* $\mathcal{Y}$ *with this image we obtain a situation where the linearization is surjective. However, if this image depends on* x *then Newton's iteration breaks down, in general.*

(b) This theorem has various straight-forward extensions to the case where any of d_1, d_2 *or* ϵ_0 *depends on* $x \in \mathcal{X}$*. For example, in our proof we use* δ *to measure the smallness of* $\|\mathcal{F}(x_0)\|_{\mathcal{Y}}$ *and this translates to a measure of the closeness of* x_0 *and* x_∞*. A more sophisticated treatment would measure these using separate parameters. The precise way in which one would do this depends on the specific application (e.g., how good of an initial guess* x_0 *is relative to the value of the constants* d_1, d_2*), so here we only address the simplest case where the constants do not depend on* x*.*

Proof of Theorem 2.2. Given x_0, define a sequence $\{x_n\}$ recursively by

$$x_{n+1} := x_n - M_{x_n}\mathcal{F}(x_n).$$

We will prove the following by induction on n:

(A) $\|x_{n+1} - x_n\|_{\mathcal{X}} \leq d_1\|\mathcal{F}(x_n)\|_{\mathcal{Y}}$

(B) $\|\mathcal{F}(x_{n+1})\|_{\mathcal{Y}} \leq d_1 d_2\|\mathcal{F}(x_n)\|_{\mathcal{Y}}^2$

(C) $\|x_{n+1} - x_n\|_{\mathcal{X}} \leq d_1 d_2(\delta^2 d_1)^n$

(D) $\|\mathcal{F}(x_n)\|_{\mathcal{Y}} \leq d_2(\delta^2 d_1)^n$.

The proof of (A) is immediate from the definition of the x_n. For (B), we use the Taylor expansion of $\mathcal{F}$:

$$\begin{aligned}
\|\mathcal{F}(x_{n+1})\|_{\mathcal{Y}} &= \|\mathcal{F}(x_n - M_{x_n}\mathcal{F}(x_n))\|_{\mathcal{Y}} \\
&= \|\mathcal{F}(x_n) - D_{x_n}\mathcal{F}\left(M_{x_n}\mathcal{F}(x_n)\right) + R_{x_n}\left(-M_{x_n}\mathcal{F}(x_n)\right)\|_{\mathcal{Y}} \\
&= \|R_{x_n}\left(-M_{x_n}\mathcal{F}(x_n)\right)\|_{\mathcal{Y}} \\
&\leq d_1 d_2\|\mathcal{F}(x_n)\|_{\mathcal{Y}}^2.
\end{aligned}$$

In the inequality we used (1), and the assumption $\delta_0 d_1 \leq \epsilon_0$. Now we use (A) and (B) to prove (C) and (D):

$$\begin{aligned}
\|x_{n+1} - x_n\|_{\mathcal{X}} &\leq d_1\|\mathcal{F}(x_n)\|_{\mathcal{Y}} \\
&\leq d_1^2 d_2\|\mathcal{F}(x_{n-1})\|_{\mathcal{Y}}^2 \\
&\vdots \\
&\leq d_1^{n+1} d_2\|\mathcal{F}(x_0)\|_{\mathcal{Y}}^{2n} \\
&\leq d_1 d_2(\delta^2 d_1)^n.
\end{aligned}$$

This finishes the induction.

It follows from (C) that $\{x_n\}$ is Cauchy:

$$\begin{aligned}\|x_{n+k} - x_n\|_{\mathcal{X}} &\leq \sum_{j=0}^{k-1} \|x_{n+j+1} - x_{n+j}\|_{\mathcal{X}} \\ &\leq d_1 d_2 (\delta^2 d_1)^n \sum_{j=0}^{k-1} (\delta^2 d_1)^j \\ &\leq d_1 d_2 (\delta^2 d_1)^n \frac{1}{1-\delta^2 d_1},\end{aligned}$$

which goes to zero in n, since $\delta_0^2 d_1 < 1$. We set $x_\infty = \lim_{n\to\infty} x_n$. Taking $n = 0$, and taking k to ∞, the above computation also shows

$$\|x_\infty - x_0\|_{\mathcal{X}} \leq \frac{d_1 d_2}{1-\delta^2 d_1}.$$

Similarly, it follows from (D) that $\mathcal{F}(x_\infty) = 0$.

Finally, suppose $x_\infty, x'_\infty \in \mathcal{X}$ are two distinct zeros of $\mathcal{F}$ satisfying $\|x_\infty - x_0\| < 1/2d_1d_2$ and $\|x'_\infty - x_0\| < 1/2d_1d_2$. Then we have

$$D_{x_0}\mathcal{F}(x'_\infty - x_\infty) = R_{x_0}(x'_\infty - x_0) - R_{x_0}(x_\infty - x_0),$$

so by Lemma 2.1 we get

$$\begin{aligned}\|x'_\infty - x_\infty\|_{\mathcal{X}} &\leq d_1 \|R_{x_0}(x'_\infty - x_0) - R_{x_0}(x_\infty - x_0)\|_{\mathcal{Y}} \\ &\leq d_1 d_2 \left(\|x'_\infty - x_0\|_{\mathcal{X}} + \|x_\infty - x_0\|_{\mathcal{X}}\right) \|x'_\infty - x_\infty\|_{\mathcal{X}} \\ &< \|x'_\infty - x_\infty\|_{\mathcal{X}}\end{aligned}$$

which is impossible. So we must have $x'_\infty = x_\infty$. □

3 Applications to the curvature

Let (Z, g) be a smooth, compact, oriented Riemannian 4-manifold. Fix a principal G-bundle $P \to Z$ and a smooth reference connection on P. Let $\mathcal{X}$ be the $W^{1,2}$-completion of the compactly supported elements of $\Omega^1 := \Omega^1(Z, P(\mathfrak{g}))$, and let $\mathcal{Y}$ be the L^2-completion of the compactly supported elements of $\Omega^+ := \Omega^+(Z, P(\mathfrak{g}))$, where the superscript $+$ denotes self-dual 2-forms. Define $\mathcal{F} : \mathcal{X} \to \mathcal{Y}$ by

$$\mathcal{F}(x) := F^+_{A+x} = F^+_A + d^+_A x + \frac{1}{2}\left[x \wedge x\right]^+.$$

Since $\mathcal{F}$ is quadratic we can take $\epsilon_0 = \infty$ in the statement of Lemma 2.1. Moreover, the quadratic part is $\frac{1}{2}\left[x \wedge x\right]^+$, and so is bounded in norm by $d_2\|x\|^2_{W^{1,2}}$, where d_2 is the Sobolev multiplication constant $W^{1,2} \otimes W^{1,2} \to L^2$. In particular, d_2 is independent of A and x.

To address d_1, we need to assume A has been chosen so that d_A^+ maps surjectively onto $\Omega^+(Z, P(\mathfrak{g}))$ (this can be achieved if A is near the space of instantons, and the function $\mathcal{F}$ is suitably perturbed, but we will ignore these issues). It follows that d_{A+x}^+ is also surjective, provided x is sufficiently small. Moreover, general elliptic theory implies that the formal (L^2) adjoint $d_{A+x}^* : \Omega^+ \to \Omega^1$ is injective, and that the operator $d_{A+x}^+ d_{A+x}^* : \Omega^+ \to \Omega^+$ extends to a Banach space isomorphism from the $W^{2,2}$-completion of the domain to the L^2-completion of the codomain; let $(d_{A+x}^+ d_{A+x}^*)^{-1}$ denote the inverse of this operator. Define

$$M_x := d_{A+x}^* (d_{A+x}^+ d_{A+x}^*)^{-1} : \Omega^+ \longrightarrow \Omega^1,$$

which extends to a bounded map $\mathcal{Y} \longrightarrow \mathcal{X}$. This is a right inverse of $D_x \mathcal{F}$ and, since M_x is bounded, there is constant d_1 with

$$\|M_x w\|_{\mathcal{X}} \leq d_1 \|w\|_{\mathcal{Y}}$$

for all $w \in \mathcal{Y}$. This depends on x, but not on w.

We want to import the theory from the previous section. The only issue is that d_1 does depend on the basepoint x, and so does not quite fit into the formalism of Theorem 2.2; however, see Remark 2.3 (b).

The long-term goal of this section is to understand the dependence of the constants d_1, d_2 on the metric g, and then use the appropriate extension of Theorem 2.2 to translate this into a condition on the smallness of $\|\mathcal{F}(V)\|_{L^2}$ (i.e., via the value of δ). The particular situation we have in mind is when $g = g_\epsilon$, where g_ϵ is a metric as in [2]. Here $\epsilon > 0$ is a parameter that measures the smallness in certain directions, and the goal of [2] is to investigate the limit as ϵ approaches zero. Here we would be interested in seeing how δ depends on ϵ...

References

[1] S. Dostoglou, D. Salamon. Self-dual instantons and holomorphic curves. *Ann. of Math.* (2) 139 (1994), no. 3, 581-640.

[2] D. Duncan. Higher rank instanton Floer theory and the quilted Atiyah-Floer conjecture. arXiv:1311.5609.