Newton Method on Riemannian Manifolds: Covariant Alpha-Theory.

Jean-Pierre Dedieu[a], Pierre Priouret[a], and Gregorio Malajovich[c]

[a]MIP. Département de Mathématique, Université Paul Sabatier, 31062 Toulouse cedex 04, France (dedieu@mip.ups-tlse.fr).

[b]MIP. Département de Mathématique, Université Paul Sabatier, 31062 Toulouse cedex 04, France (priouret@mip.ups-tlse.fr).

[c]Departamento de Matemática Aplicada, Universidade Federal de Rio de Janeiro, Caixa Postal 68530, CEP 21945-970, Rio de Janeiro, RJ, Brazil (gregorio@labma.ufrj.br)

In this talk we consider quantitative aspects of Newton method for finding zeros of analytic mappings $f : M_n \to R^n$ and analytic vector fields $X : M_n \to TM_n$ defined on a real complete analytic Riemannian manifold M_n; n is its dimension and TM_n denotes its tangent bundle. We extend to this case Smale's alpha-theory introduced in [15], [16] and [17].

The Newton operator associated with f is defined by

$$N_f(z) = \exp_z(-Df(z)^{-1}f(z))$$

where $\exp_z : T_zM_n \to M_n$ denotes the exponential map. When, instead of a mapping we consider a vector field X, in order to define Newton method, we resort to an object studied in differential geometry; namely, the covariant derivative of vector fields denoted here by DX. We let

$$N_X(z) = \exp_z(-DX(z)^{-1}X(z)).$$

These definitions coincide with the usual one when $M_n = R^n$ because $\exp_z(u) = z + u$ and also because, in this context, the covariant derivative is just the usual derivative.

The first to consider Newton method on a manifold is Rayleigh 1899 [7] who defined what we call today "Rayleigh Quotient Iteration" which is in fact a Newton iteration for a vector field on the sphere. Then Shub 1986 [8] defined Newton's method for the problem of finding the zeros of a vector field on a manifold and used retractions to send a neighborhood of the origin in the tangent space onto the manifold itself. In our paper we do not use general retractions but exponential maps. Udriste 1994 [19] studied Newton's method to find the zeros of a gradient vector field defined on a Riemannian manifold; Owren and Welfert 1996 [6] defined Newton iteration for solving the equation $F(x) = 0$ where F is a map from a Lie group to its corresponding Lie algebra; Smith 1994 [18] and Edelman-Arias-Smith 1998 [3] developed Newton and conjugate gradient algorithms on the Grassmann and Stiefel manifolds. These authors define Newton's method via the exponential map like we do here. Shub 1993 [9], Shub and Smale 1993-1996 [10], [11], [12], [13], [14], Malajovich 1994 [5], Dedieu and Shub 2000 [2] introduce and study Newton's method on projective spaces and their products. Another paper about this subject is Adler-Dedieu-Margulies-Martens-Shub 2001 [1] where qualitative aspects of Newton method on Riemannian manifolds are investigated for both mappings and vector fields with an application to a geometric model for the human spine represented as a $18-$tuple of 3×3 orthogonal matrices. Recently Ferreira-Svaiter [4] gave a Kantorovich like theorem for Newton method for vector fields defined on Riemannian manifolds.

Definition 1. *Let given a map $f : M_n \to R^n$ and a vector field $X : M_n \to TM_n$. For any point $z \in M_n$ we let*

$$\gamma(f,z) = \sup_{k\geq 2} \left\| Df(z)^{-1} \frac{D^k f(z)}{k!} \right\|_z^{1/k-1} \quad \text{and} \quad \gamma(X,z) = \sup_{k\geq 2} \left\| DX(z)^{-1} \frac{D^k X(z)}{k!} \right\|_z^{1/k-1}.$$

We also let $\gamma(f,z) = \infty$ when $Df(z)$ is not invertible, idem for $\gamma(X,z)$.

Here $D^k f(z)$ (resp. $D^k X(z)$) denotes the $k-$th covariant derivative of f (resp. X) at z, $\|u\|_z$ is the norm of $u \in T_z M_n$ with respect to the Riemannian structure and $\|Df(z)^{-1} D^k f(z)\|_z$ (resp. $\|DX(z)^{-1} D^k X(z)\|_z$) is the norm of this $k-$th multilinear operator $(T_z M_n)^k \to T_z M_n$ and d the Riemannian distance on M_n.

Definition 2. *We denote by $\mathbf{r}_z > 0$ the radius of injectivity of the exponential map at z. Thus $\exp_z : B_{T_z}(0, \mathbf{r}_z) \to B_{M_n}(z, \mathbf{r}_z)$ is one to one ($B(u,r)$ is the open ball about u with radius r).*

Definition 3. *For any $\zeta \in M_n$*

$$K_\zeta = \sup \frac{d(\exp_z(u), \exp_z(v))}{\|u - v\|_z}$$

where the supremum is taken for all $z \in B_{M_n}(\zeta, \mathbf{r}_\zeta)$, and u, $v \in T_z M_n$ with $\|u\|_z$ and $\|v\|_z \leq \mathbf{r}_\zeta$), with $\mathbf{r}_\zeta$ the radius of injectivity at ζ.

Our first main theorem relates the size of the quadratic attraction basin of a zero ζ of f to the invariants $\mathbf{r}_z > 0$, $\gamma(f,\zeta)$ and K_ζ.

Theorem 1. *Let $f : M_n \to R^n$ be analytic . Suppose that $f(\zeta) = 0$ and $Df(\zeta)$ is an isomorphism. Let*

$$R(f,\zeta) = \min \left(\mathbf{r}_\zeta, \frac{K_\zeta + 2 - \sqrt{K_\zeta^2 + 4K_\zeta + 2}}{2\gamma(f,\zeta)} \right).$$

If $d(z,\zeta) \leq R(f,\zeta)$ then Newton sequence $z_k = N_f^{(k)}(z)$ is defined for all $k \geq 0$, and

$$d(z_k, \zeta) \leq \left(\frac{1}{2}\right)^{2^k - 1} d(z,\zeta).$$

When $M_n = R^n$ equipped with the usual metric structure, the radius of injectivity is $\mathbf{r}_\zeta = \infty$ and $K_\zeta = 1$. Thus $R(f,\zeta) = (3 - \sqrt{7})/2\gamma(f,\zeta)$ like in Smale's point estimates. The second theorem presented here is the following

Definition 4. *We let $\beta(f,z) = \|Df(z)^{-1} f(z)\|_z$ and $\alpha(f,z) = \beta(f,z)\gamma(f,z)$. We give to $\beta(f,z)$ and $\alpha(f,z)$ the value ∞ when $Df(z)$ is singular.*

Theorem 2. *Let $f : M_n \to R^n$ be analytic. Let us denote $\sigma = \sum_{k=0}^{\infty} 2^{1-2^k} = 1.632\ldots$ Let $z \in M_n$ be such that*

$$\beta(f,z) \leq \mathbf{r}_y \quad \text{for all} \quad y \in B_{M_n}(z, \sigma\beta(f,z)).$$

There is a universal constant $\alpha_0 > 0$ with the following property: if $\alpha(f,z) < \alpha_0$ then Newton sequence $z_0 = z$, $z_{k+1} = N_f(z_k)$ is defined for all integers $k \geq 0$ and converges to a zero ζ of f. Moreover

$$d(z_{k+1}, z_k) \leq \left(\frac{1}{2}\right)^{2^k - 1} \beta(f,z)$$

and

$$d(\zeta, z) \leq \sigma\beta(f,z).$$

The case of vector fields is treated similarly, we have:

Theorem 3. *Let $X : M_n \to TM_n$ be an analytic vector field. Suppose that $X(\zeta) = 0$ and $DX(\zeta)$ is an isomorphism. Let*

$$R(X,\zeta) = \min\left(\mathbf{r}_\zeta, \frac{K_\zeta + 2 - \sqrt{K_\zeta^2 + 4K_\zeta + 2}}{2\gamma(X,\zeta)}\right).$$

If $d(z,\zeta) \leq R(X,\zeta)$ then Newton sequence $z_k = N_X^{(k)}(z)$ is defined for all $k \geq 0$, and

$$d(z_k, \zeta) \leq \left(\frac{1}{2}\right)^{2^k - 1} d(z,\zeta).$$

The invariants β and α for vector fields are defined similarly:

Definition 5. *We let*

$$\beta(X,z) = \|DX(z)^{-1}X(z)\|_z$$

and

$$\alpha(X,z) = \beta(X,z)\gamma(X,z).$$

We give to $\beta(X,z)$ and $\alpha(X,z)$ the value ∞ when $DX(z)$ is singular.

Theorem 4. *Let $X : M_n \to TM_n$ be an analytic vector field. Let $z \in M_n$ be such that*

$$\beta(X,z) \leq \mathbf{r}_y \quad \text{for all} \quad y \in B_{M_n}(z, \sigma\beta(X,z)).$$

There is a universal constant $\alpha_0 > 0$ such that: if $\alpha(X,z) < \alpha_0$ then Newton sequence $z_0 = z$, $z_{k+1} = N_X(z_k)$ is defined for all integers $k \geq 0$ and converges to a zero ζ of X. Moreover

$$d(z_{k+1}, z_k) \leq \left(\frac{1}{2}\right)^{2^k - 1} \beta(X,z)$$

and

$$d(\zeta, z) \leq \sigma\beta(X,z).$$

Acknowledgments. Part of this work was carried out when J.-P. Dedieu and G. Malajovich were visiting City University of Hong-Kong.

References

[1] Adler R., J.-P. Dedieu, J. Margulies, M. Martens and M. Shub, *Newton Method on Riemannian Manifolds and a Geometric Model for the Human Spine*. To appear in IMA Journal on Numerical Analysis.

[2] Dedieu, J.-P. and M. Shub, *Multihomogeneous Newton's Method*. Mathematics of Computation, 69 (2000) 1071-1098.

[3] Edelman, A., T. Arias and S. Smith, *The Geometry of Algorithms with Orthogonality Constraints*, SIAM J. Matrix Anal. Appl. 20 (1998) 303-353.

[4] Ferreira O., B. Svaiter, *Kantorovich's Theorem on Newton's Method in Riemannian Manifolds*, to appear in: Journal of Complexity.

[5] Malajovich, G., *On Generalized Newton Algorithms*, Theoretical Computer Science, (1994), vol. 133, pp.65-84.

[6] Owren, B. and B. Welfert, *The Newton Iteration on Lie Groups*, Preprint, 1996.

[7] Rayleigh, J. W. Strutt *On the Calculation of the Frequency of Vibration of a System in its Gravest Mode, with Examples from Hydrodynamics.* The Philosophical Magazine 47 (1899) 556-572.

[8] Shub, M., *Some Remarks on Dynamical Systems and Numerical Analysis*, in: Dynamical Systems and Partial Differential Equations, Proceedings of VII ELAM (L. Lara-Carrero and J. Lewowicz eds.), Equinoccio, Universidad Simon Bolivar, Caracas, 1986, 69-92.

[9] Shub, M., *Some Remarks on Bezout's Theorem and Complexity*, in *From Topology to Computation: Proceedings of the Smalefest*, Marsden, J.E., M. W. Hirsch and M. Shub eds. Springer, 1993, pp. 443-455.

[10] Shub, M., S. Smale, *Complexity of Bézout's Theorem I: Geometric Aspects*, J. Am. Math. Soc. (1993) 6 pp. 459-501.

[11] Shub, M., S. Smale, *Complexity of Bézout's Theorem II: Volumes and Probabilities* in: *Computational Algebraic Geometry,* F. Eyssette and A. Galligo eds., em Progress in Mathematics, vol. 109, Birkhäuser, 1993, 267-285.

[12] Shub, M., S. Smale, *Complexity of Bézout's Theorem III: Condition Number and Packing,* J. of Complexity (1993) vol. 9, pp 4-14.

[13] Shub, M., S. Smale, *Complexity of Bezout's Theorem IV: Probability of Success, Extensions*, SIAM J. Numer. Anal. (1996) vol. 33, pp. 128-148.

[14] Shub, M., S. Smale, Complexity of Bézout's Theorem V: Polynomial Time, Theoretical Computer Science, (1994) vol. 133, pp.141-164.

[15] Smale, S., *On the Efficiency of Algorithms of Analysis*, Bull. A.M.S. (1985) vol. 13 pp. 87-121

[16] Smale, S., *Algorithms for Solving Equation*, in: Proceedings of the International Congress of Mathematicians, A.M.S. pp. 172–195 1986.

[17] Smale, S., *Newton's Method Estimates from Data at One Point* in: The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics (R. Ewing, K. Gross, and C. Martin eds.), Springer 1986.

[18] Smith, S., *Optimization Techniques on Riemannian Manifolds*, in: Fields Institute Communications, vol. 3, AMS, 113-146, 1994.

[19] Udriste, C., *Convex Functions and Optimization Methods on Riemannian Manifolds*, Kluwer, 1994.