GRADIENT METHODS ON FINSLER MANIFOLDS

G. Bercu

Abstract

The present paper refers to the gradient method on Finsler manifold, showing how to use the direction y for obtaining a suitable descent algorithm.

1991 Mathematics Subject Classification: 49M07, 53C60
Key words: Finsler manifold, geodesics, gradient.

1 Preliminaries

Let M be an n-dimensional, complete, connected C^∞ manifold and TM its tangent bundle. Denote by (x, y) an arbitrary point in TM and by x the corresponding in M.

Definition. The pair (M, F) is called Finsler manifold if the function $F : TM \to \mathbf{R}$ satisfies the axioms:

1) $F(x, y) > 0,\ \forall x \in M, \forall y \neq 0$;

2) $F(x, \lambda y) = |\lambda| \cdot F(x, y), \forall \lambda \in \mathbf{R},\ \forall (x, y) \in TM$;

3) the fundamental tensor $g_{ij}(x, y) = \dfrac{1}{2} \dfrac{\partial^2 F}{\partial y^i \partial y^j}$ is positive definite;

4) F is of C^∞-class at every point $(x, y) \in TM$, where $y \neq 0$, and it is continuous at every point $(x, 0) \in TM$.

Suppose we have a C^2 real function $f : M \to \mathbf{R}$ and we want to find one of its minima.

We consider the 1-form $df(x)$ which has the components $f_i(x) = \dfrac{\partial f}{\partial x^i}(x)$ and the vector field $grad\, f(x, y)$, which has the components $f^i(x, y) = g^{ij}(x, y) f_j(x)$, called the gradient of the function f. We remark that $-grad\, f$ is a vector field orthogonal to the hypersurfaces of constant level of f, which shows at every point $x \in M$ the

Editor Gr.Tsagas *Proceedings of the Workshop on Global Analysis, Differential Geometry and Lie Algebras, 1997*, 230-233

direction and the sense of steepest descent. This suggests that the most suitable *solution of the inequality* $df(x_i)(X_{x_i}) < 0$ is

$$X_{x_i} = -grad\, f(x_i, y), \quad y \in T_{x_i}M.$$

We obtain an iterative process, called *the gradient method.* This method is described by the following algorithm:

1) One considers an initial point x_1 and computes $grad\, f(x_1, y)$, $y \in T_{x_1}M$. If $grad\, f(x_1, y) = 0$ for every $y \in T_{x_1}M$, then stop !

2) If $grad\, f(x_1, y) \neq 0$, then from the point x_1 we pass to another point $x_2 = \exp_{x_1}(-t \cdot grad\, f(x_1, y))$, $t \in [0, \infty)$.
We choose y such that $\max\limits_{\|y\|=1} f(x_2(x_1, y)) < f(x_1)$.
If $gradf(x_2, y) = 0$ for every $y \in T_{x_2}M$, then stop.

3) If $grad\, f(x_2, y) \neq 0$, then from the point x_2 we pass to another point $x_3 = \exp_{x_2}(-t \cdot grad\, f(x_2, y))$.
We choose y such that $\max\limits_{\|y\|=1} f(x_3(x_2(x_1, y))) < f(x_2)$.
If $gradf(x_3, y) = 0$ for every $y \in T_{x_3}M$, then stop.
At "stop" one verifies if the detected point is a minimum point.

4) Generally, if we have $grad\, f(x_i, y) \neq 0$, then we set

$$x_{i+1} = \exp_{x_i}(-tgrad\, f(x_i, y)).$$

We choose y such that $\max\limits_{\|y\|=1} f(x_{i+1}(x_i(\ldots(x_2(x_1, y))\ldots))) < f(x_i)$.

Remark. The real number $t > 0$ is arbitrarily chosen and the same for all iterations, such as the next inequality is verified

$$(*) \qquad f(x) - f(x_i) \leq t \cdot \varepsilon \cdot df(X_i).$$

Here

$$X_i = -grad\, f(x_i, y), \; x = \exp_{x_i}(-t \cdot gradf(x_i, y))$$

and $\varepsilon \in (0; 1)$ is an arbitrarily fixed constant, independent of i.

If the inequality $(*)$ is not satisfied, then we replace t by λt, $\lambda \in (0; 1)$, with λ fixed such as $(*)$ to be satisfied.

In the book [1] the notion of forward (resp. backward) Cauchy sequence is introduced.

Definition 2. A sequence $\{x_i\}$ in M is called *a forward* (resp. *backward*) Cauchy sequence if, for all $\varepsilon > 0$, there exists a positiv integer $N(\varepsilon)$ such that $N \leq i < j \Rightarrow d(x_i, x_j) < \varepsilon$ (resp. $d(x_j, x_i) < \varepsilon$).

2 Main Results

Theorem 1. *Let $\{x_i\}$ be a sequence in a Finsler manifold (M, F). Then the following three statements are equivalent:*

(a) $\{x_i\}$ converges to x in the manifold topology of M.

(b) $d(x, x_i) \to 0$
(c) $d(x_i, x) \to 0$.

Let us refer to the sequence generated by the previous algorithm. If

$$\{x_1, x_2(x_1, y), x_3(x_1, y), \ldots\}$$

is a sequence which uniformly converges forward to x_*, then the sequence of values $f(x_1) > f(x_2) > \ldots f(x_j) > \ldots$ converges to the minimum $f(x_*)$. Also, we can prove the convergence of $\{grad\, f(x_i, y)\}$ to zero.

Theorem 2. *Let M be n-dimensional, complete connected C^∞ Finsler manifold, and $f : M \to \mathbb{R}$ a real lower bounded C^2 function. We denote by X_x and $X_{\bar{x}}$ the tangent vectors at x and $\bar{x}$ respectively to the geodesic which joins the point x and $\bar{x}$. If for any $x, \bar{x} \in M$, the Lipschitz condition*

$$|df(X_{\bar{x}}) - df(X_x)| \leq r \cdot d^2(x, \bar{x}),\ r > 0$$

is satisfied, and if the choice of number $t > 0$ is made as described above, then in the iterative process $x_{i+1} = \exp_{x_i}(-t \cdot grad(x_i, y))$, $i = 1, 2, \ldots$, we have $\lim_{i \to \infty} grad\, f(x_i) = 0$, for any given initial point x_1.

Proof. Let $\gamma_{x\bar{x}} : [0, 1] \to M$ be a geodesic which joins the point $x = \gamma_{x\bar{x}}(0)$ and $\bar{x} = \gamma_{x\bar{x}}(1)$. Since the equation of $\gamma_{x\bar{x}}$ does not depend on y, we infer that the proof is similar as in Riemann case [3]. Thus, since f is of class C^2, we find

$$\begin{aligned} f(\bar{x}) - f(x) &= f(\gamma_{x\bar{x}}(1)) - f(\gamma_{x\bar{x}}(0)) = \int_0^1 \frac{d}{du} f(\gamma_{x\bar{x}}(u))du = \\ &= \int_0^1 df(\dot{\gamma}_{x\bar{x}}(u))du = df(\dot{\gamma}_{x\bar{x}}(u_0)), \end{aligned}$$

and $u_0 \in [0; 1]$. Denoting $z = \gamma_{x\bar{x}}(u_0)$, we can write

$$\begin{aligned} df(\dot{\gamma}_{x\bar{x}}(u_0)) &= df(X_z) = df(X_x) + (df(X_z) - df(X_x)) \leq \\ &\leq df(X_x) + rd^2(x, z). \end{aligned}$$

Since $\gamma_{x\bar{x}}$ is a geodesic, we have $||\dot{\gamma}_{x\bar{x}}(u)|| = ||\dot{\gamma}_{x\bar{x}}(0)|| = \text{const}$. Thus

$$d^2(x, z) \leq \left(\int_0^{u_0} ||\dot{\gamma}_{x\bar{x}}(u)|| du \right)^2 = ||\dot{\gamma}_{x\bar{x}}(0)||^2 \cdot u_0^2 \leq ||\dot{\gamma}_{x\bar{x}}(0)||^2.$$

Putting $X_x = \dot{\gamma}_{x\bar{x}}(0) := -t \cdot grad\, f(x, y),\ t > 0$, it follows

$$\begin{aligned} & f(\bar{x}) - f(x) = df(\dot{\gamma}_{x\bar{x}}(u_0)) \leq -t \cdot ||grad\, f(x, y)||^2 + \\ & +rt^2 \cdot ||grad\, f(x, y)||^2 = t \cdot ||grad\, f(x, y))^2 \cdot (-1 + tr) \end{aligned}$$

It follows the inequality $f(\bar{x}) - f(x) \leq t \cdot b^2 \cdot (-1 + tr)$.

This estimation shows there exist some numbers $t > 0$ such that the inequality $f(x) - f(x_i) \leq \varepsilon \cdot t \cdot df(X_i)$ is satisfied, (where $X_i = -grad\, f(x_i, y)$), namely those for which $b^2(-1+tr) < -\varepsilon \Rightarrow t < \frac{b^2 - \varepsilon}{b^2 r}$.

So, we find that $f(x_{i+1}) - f(x_i) \leq -\varepsilon \cdot t \cdot ||grad\, f(x_i, y)||^2$. If $||grad\, f(x_i, y)||^2 > 0$, then for any $i \in \mathbf{N}^*$ we have $f(x_{i+1}) - f(x_i) < 0$; i.e. the sequence $\{f(x_i)\}$ is decreasing. Moreover, f is lower bounded, hence $\lim_{i\to\infty}(f(x_{i+1}) - f(x_i)) = 0$.

But, from the last inequality we infer that

$$||grad\, f(x_i, y)||^2 \leq \frac{f(x_i) - f(x_{i+1})}{\varepsilon t},$$

where $0 < t < \frac{b^2 - \varepsilon}{b^2 r}$. Hence, we obtain that $\lim_{i\to\infty} ||grad\, f(x_i, y)|| = 0$.

References

[1] D.Bao, S.S. Chern, Z.Shen, *Introduction to the Riemann-Finsler manifolds theory*, Springer Verlag, 1999.

[2] S.T.Smith, *Optimization techniques on Riemannian manifolds*, Field Institute Communications, 3 (1994), 113-136.

[3] C.Udrişte, *Convex functions and optimization methods on Riemannian manifolds*, Kluwer Academic Publishers, 1994.

[4] C.Udrişte, *Optimization methods on Riemannian manifolds*, IRN International Workshop, Monteroduni, Italy, August 8-12, 1995.

[5] C.Udrişte, *Sufficient Decrease Principle on Riemannian Manifolds*, BJGA, 1 2 (1996), 111-123.

Author's address:

G. Bercu
University "Dunărea de Jos"
Department of Mathematics
Galaţi, Romania