# Optimization Techniques on Riemannian Manifolds

Steven T. Smith

Harvard University
Division of Applied Sciences
Cambridge, Massachusetts 02138

Abstract. The techniques and analysis presented in this paper provide new methods to solve optimization problems posed on Riemannian manifolds. A new point of view is offered for the solution of constrained optimization problems. Some classical optimization techniques on Euclidean space are generalized to Riemannian manifolds. Several algorithms are presented and their convergence properties are analyzed employing the Riemannian structure of the manifold. Specifically, two apparently new algorithms, which can be thought of as Newton's method and the conjugate gradient method on Riemannian manifolds, are presented and shown to possess, respectively, quadratic and superlinear convergence. Examples of each method on certain Riemannian manifolds are given with the results of numerical experiments. Rayleigh's quotient defined on the sphere is one example. It is shown that Newton's method applied to this function converges cubically, and that the Rayleigh quotient iteration is an efficient approximation of Newton's method. The Riemannian version of the conjugate gradient method applied to this function gives a new algorithm for finding the eigenvectors corresponding to the extreme eigenvalues of a symmetric matrix. Another example arises from extremizing the function  $\operatorname{tr} \Theta^{\mathrm{T}} Q \Theta N$  on the special orthogonal group. In a similar example, it is shown that Newton's method applied to the sum of the squares of the off-diagonal entries of a symmetric matrix converges cubically.

**Keywords.** Optimization, constrained optimization, Riemannian manifolds, Lie groups, homogeneous spaces, steepest descent, Newton's method, conjugate gradient method, eigenvalue problem, Rayleigh's quotient, Rayleigh quotient iteration, Jacobi methods, numerical methods.

#### 1 Introduction

The preponderance of optimization techniques address problems posed on Euclidean spaces. Indeed, several fundamental algorithms have arisen from the desire to compute the minimum of quadratic forms on Euclidean space. However, many optimization problems are posed on non-Euclidean spaces. For example, finding the largest eigenvalue of a symmetric matrix may be posed as the maximization of Rayleigh's quotient defined on the sphere. Optimization problems subject to nonlinear differentiable equality constraints on Euclidean space also

lie within this category. Many optimization problems share with these examples the structure of a differentiable manifold endowed with a Riemannian metric. This is the subject of this paper: the extremization of functions defined on Riemannian manifolds.

The minimization of functions on a Riemannian manifold is, at least locally, equivalent to the smoothly constrained optimization problem on a Euclidean space, because every  $C^{\infty}$  Riemannian manifold can be isometrically imbedded in some Euclidean space [46, Vol. V]. However, the dimension of the Euclidean space may be larger than the dimension of the manifold; practical and aesthetic considerations suggest that one try to exploit the intrinsic structure of the manifold. Elements of this spirit may be found throughout the field of numerical methods, such as the emphasis on unitary (norm preserving) transformations in numerical linear algebra [22], or the use of feasible direction methods [18,21,38].

An intrinsic approach leads one from the extrinsic idea of vector addition to the exponential map and parallel translation, from minimization along lines to minimization along geodesics, and from partial differentiation to covariant differentiation. The computation of geodesics, parallel translation, and covariant derivatives can be quite expensive. For an n-dimensional manifold, the computation of geodesics and parallel translation requires the solution of a system of 2n nonlinear and n linear ordinary differential equations. Nevertheless, many optimization problems are posed on manifolds that have an underlying algebraic structure that may be exploited to greatly reduce the complexity of these computations. For example, on a real compact semisimple Lie group endowed with its natural Riemannian metric, geodesics and parallel translation may be computed via matrix exponentiation [24]. Several algorithms are available to perform this computation [22, 32]. This algebraic structure may be found in the problems posed by Brockett [8, 9, 10], Bloch et al. [3, 4], Smith [45], Faybusovich [17], Lagarias [30], Chu et al. [13,14], Perkins et al. [35], and Helmke [25]. This approach is also applicable if the manifold can be identified with a symmetric space or, excepting parallel translation, a reductive homogeneous space [29, 33]. Perhaps the simplest nontrivial example is the sphere, where geodesics and parallel translation can be computed at low cost with trigonometric functions and vector addition. Furthermore, Brown and Bartholomew-Biggs [11] show that in some cases function minimization by following the solution of a system of ordinary differential equations can be implemented such that it is competitive with conventional techniques.

The outline of the paper is as follows. In Section 2, the optimization problem is posed and conventions to be held throughout the paper are established. The method of steepest descent on a Riemannian manifold is described in Section 3. To fix ideas, a proof of linear convergence is given. The examples of Rayleigh's quotient on the sphere and the function  $\operatorname{tr} \Theta^{\scriptscriptstyle{\text{T}}} Q\Theta N$  on the special orthogonal group are presented. In Section 4, Newton's method on a Riemannian manifold is derived. As in Euclidean space, this algorithm may be used to compute the extrema of differentiable functions. It is proved that this method converges quadratically. The example of Rayleigh's quotient is continued, and it

is shown that Newton's method applied to this function converges cubically, and is approximated by the Rayleigh quotient iteration. The example considering  $\operatorname{tr} \Theta^{\mathsf{T}} Q \Theta N$  is continued. In a related example, it is shown that Newton's method applied to the sum of the squares of the off-diagonal elements of a symmetric matrix converges cubically. This provides an example of a cubically convergent Jacobi-like method. The conjugate gradient method is presented in Section 5 with a proof of superlinear convergence. This technique is shown to provide an effective algorithm for computing the extreme eigenvalues of a symmetric matrix. The conjugate gradient method is applied to the function  $\operatorname{tr} \Theta^{\mathsf{T}} Q \Theta N$ .

## 2 Preliminaries

This paper is concerned with the following problem.

**Problem 2.1.** Let M be a complete Riemannian manifold, and f a  $C^{\infty}$  function on M. Compute

$$\min_{p \in M} f(p)$$
.

There are many well-known algorithms for solving this problem in the case where M is a Euclidean space. This paper generalizes several of these algorithms to the case of complete Riemannian manifolds by replacing the Euclidean notions of straight lines and ordinary differentiation with geodesics and covariant differentiation. These concepts are reviewed in the following paragraphs. We follow Helgason's [24] and Spivak's [46] treatments of covariant differentiation, the exponential map, and parallel translation. Details may be found in these references.

Let M be a complete n-dimensional Riemannian manifold with Riemannian structure g and corresponding Levi-Civita connection  $\nabla$ . Denote the tangent plane at p in M by  $T_p$  or  $T_pM$ . For every p in M, the Riemannian structure g provides an inner product on  $T_p$  given by the nondegenerate symmetric bilinear form  $g_p \colon T_p \times T_p \to \mathbf{R}$ . The notation  $\langle X, Y \rangle = g_p(X, Y)$  and  $\|X\| = g_p(X, X)^{1/2}$ , where  $X, Y \in T_p$ , is often used. The distance between two points p and q in M is denoted by d(p,q). The gradient of a real-valued  $C^{\infty}$  function f on M at p, denoted by  $(\operatorname{grad} f)_p$ , is the unique vector in  $T_p$  such that  $df_p(X) = \langle (\operatorname{grad} f)_p, X \rangle$  for all X in  $T_p$ .

Denote the set of  $C^{\infty}$  functions on M by  $C^{\infty}(M)$  and the set of  $C^{\infty}$  vector fields on M by  $\mathbf{X}(M)$ . An affine connection on M is a function  $\nabla$  which assigns to each vector field  $X \in \mathbf{X}(M)$  an  $\mathbf{R}$ -linear map  $\nabla_X : \mathbf{X}(M) \to \mathbf{X}(M)$  which satisfies

(i) 
$$\nabla_{fX+gY} = f\nabla_X + g\nabla_Y$$
, (ii)  $\nabla_X(fY) = f\nabla_X Y + (Xf)Y$ ,

for all  $f, g \in C^{\infty}(M)$ ,  $X, Y \in \mathbf{X}(M)$ . The map  $\nabla_X$  may be applied to tensors of arbitrary type. Let  $\nabla$  be an affine connection on M and  $X \in \mathbf{X}(M)$ . Then

there exists a unique R-linear map  $A\mapsto \nabla_X A$  of  $C^\infty$  tensor fields into  $C^\infty$  tensor fields which satisfies

- (i)  $\nabla_X f = Xf$ , (iv)  $\nabla_X$  preserves the type of tensors,
- (ii)  $\nabla_X Y$  is given by  $\nabla$ , (v)  $\nabla_X$  commutes with contractions,
- (iii)  $\nabla_X$  is a derivation:  $\nabla_X(A \otimes B) = \nabla_X A \otimes B + A \otimes \nabla_X B$ ,

where  $f \in C^{\infty}(M)$ ,  $Y \in \mathbf{X}(M)$ , and A, B are  $C^{\infty}$  tensor fields. If A is of type (k,l), then  $\nabla_X A$ , called the covariant derivative of A along X, is of type (k,l), and  $\nabla A: X \mapsto \nabla_X A$ , called the covariant differential of A, is of type (k,l+1).

Let M be a differentiable manifold with affine connection  $\nabla$ . Let  $\gamma\colon I\to M$  be a smooth curve with tangent vectors  $X(t)=\dot{\gamma}(t)$ , where  $I\subset\mathbf{R}$  is an open interval. The curve  $\gamma$  is called a geodesic if  $\nabla_{\!X}X=0$  for all  $t\in I$ . Let  $Y(t)\in T_{\gamma(t)}$   $(t\in I)$  be a smooth family of tangent vectors defined along  $\gamma$ . The family Y(t) is said to be parallel along  $\gamma$  if  $\nabla_{\!X}Y=0$  for all  $t\in I$ .

For every p in M and  $X \neq 0$  in  $T_p$ , there exists a unique geodesic  $t \mapsto \gamma_X(t)$  such that  $\gamma_X(0) = p$  and  $\dot{\gamma}_X(0) = X$ . We define the exponential map  $\exp_p: T_p \to M$  by  $\exp_p(X) = \gamma_X(1)$  for all  $X \in T_p$  such that 1 is in the domain of  $\gamma_X$ . Oftentimes the map  $\exp_p$  will be denoted by "exp" when the choice of tangent plane is clear, and  $\gamma_X(t)$  will be denoted by  $\exp tX$ . A neighborhood  $N_p$  of p in M is a normal neighborhood if  $N_p = \exp N_0$ , where  $N_0$  is a star-shaped neighborhood of the origin in  $T_p$  and  $\exp$  maps  $N_0$  diffeomorphically onto  $N_p$ . Normal neighborhoods always exist.

Given a curve  $\gamma: I \to M$  such that  $\gamma(0) = p$ , for each  $Y \in T_p$  there exists a unique family  $Y(t) \in T_{\gamma(t)}$   $(t \in I)$  of tangent vectors parallel along  $\gamma$  such that Y(0) = Y. If  $\gamma$  joins the points p and  $\gamma(\alpha) = q$ , the parallelism along  $\gamma$  induces an isomorphism  $\tau_{pq}: T_p \to T_q$  defined by  $\tau_{pq}Y = Y(\alpha)$ .

Let M be a manifold with an affine connection  $\nabla$ , and  $N_p$  a normal neighborhood of  $p \in M$ . Define the vector field  $\tilde{X}$  on  $N_p$  adapted to the tangent vector X in  $T_p$  by putting  $\tilde{X}_q = \tau_{pq} X$ , the parallel translation of X along the unique geodesic segment joining p and q.

Given a Riemannian structure g on M, there exists a unique affine connection  $\nabla$  on M, called the Levi-Civita connection, which for all  $X, Y \in \mathbf{X}(M)$  satisfies

- (i)  $\nabla_X Y \nabla_Y X = [X, Y]$  ( $\nabla$  is symmetric or torsion-free),
- (ii)  $\nabla g = 0$  (parallel translation is an isometry).

Length minimizing curves on M are geodesics of the Levi-Civita connection. We shall use this connection throughout the paper.

Unless otherwise specified, all manifolds, vector fields, and functions are assumed to be smooth. When considering a function f to be minimized, the assumption that f is differentiable of class  $C^{\infty}$  can be relaxed throughout the paper, but f must be continuously differentiable at least beyond the derivatives that appear. As the results of this paper are local ones, the assumption that M be complete may also be relaxed in certain instances.

We will use the the following definitions to compare the convergence rates of various algorithms.

**Definition 2.2.** Let  $\{p_i\}$  be a Cauchy sequence in M that converges to  $\hat{p}$ . (i) The sequence  $\{p_i\}$  is said to converge (at least) linearly if there exists an integer N and a constant  $\theta \in [0,1)$  such that  $d(p_{i+1},\hat{p}) \leq \theta d(p_i,\hat{p})$  for all  $i \geq N$ . (ii) The sequence  $\{p_i\}$  is said to converge (at least) quadratically if there exists an integer N and a constant  $\theta \geq 0$  such that  $d(p_{i+1},\hat{p}) \leq \theta d^2(p_i,\hat{p})$  for all  $i \geq N$ . (iii) The sequence  $\{p_i\}$  is said to converge (at least) cubically if there exists an integer N and a constant  $\theta \geq 0$  such that  $d(p_{i+1},\hat{p}) \leq \theta d^3(p_i,\hat{p})$  for all  $i \geq N$ . (iv) The sequence  $\{p_i\}$  is said to converge superlinearly if it converges faster than any sequence that converges linearly.

## 3 Steepest descent on Riemannian manifolds

The method of steepest descent on a Riemannian manifold is conceptually identical to the method of steepest descent on Euclidean space. Each iteration involves a gradient computation and minimization along the geodesic determined by the gradient. Fletcher [18], Botsaris [5, 6, 7], and Luenberger [31] describe this algorithm in Euclidean space. Gill and Murray [21] and Sargent [38] apply this technique in the presence of constraints. In this section we restate the method of steepest descent described in the literature and provide an alternative formalism that will be useful in the development of Newton's method and the conjugate gradient method on Riemannian manifolds.

**Algorithm 3.1** (The method of steepest descent). Let M be a complete Riemannian manifold with Riemannian structure g and Levi-Civita connection  $\nabla$ , and let  $f \in C^{\infty}(M)$ .

Step 0. Select  $p_0 \in M$ , compute  $G_0 = -(\operatorname{grad} f)_{p_0}$ , and set i = 0.

Step 1. Compute  $\lambda_i$  such that

$$f(\exp_{p_i} \lambda_i G_i) \le f(\exp_{p_i} \lambda G_i)$$

for all  $\lambda > 0$ .

Step 2. Set

$$p_{i+1} = \exp_{p_i} \lambda_i G_i,$$
  
$$G_{i+1} = -(\operatorname{grad} f)_{p_{i+1}},$$

increment i, and go to Step 1.

It is easy to verify that  $\langle G_{i+1}, \tau G_i \rangle = 0$ , for  $i \geq 0$ , where  $\tau$  is the parallelism with respect to the geodesic from  $p_i$  to  $p_{i+1}$ . By assumption, the function  $\lambda \mapsto f(\exp \lambda G_i)$  is minimized at  $\lambda_i$ . Therefore, we have  $0 = (d/dt)|_{t=0}$   $f(\exp(\lambda_i + t)G_i) = df_{p_{i+1}}(\tau G_i) = \langle (\operatorname{grad} f)_{p_{i+1}}, \tau G_i \rangle$ . Thus the method of steepest descent on a Riemannian manifold has the same deficiency as its counterpart on a Euclidean space, i.e., it makes a ninety degree turn at every step.

The convergence of Algorithm 3.1 is linear. To prove this fact, we will make use of a standard theorem of the calculus, expressed in differential geometric

language. The covariant derivative  $\nabla_X f$  of f along X is defined to be Xf. For  $k = 1, 2, \ldots$ , define  $\nabla_X^k f = \nabla_X \circ \cdots \circ \nabla_X f$  (k times), and let  $\nabla_X^0 f = f$ .

**Remark 3.2** (Taylor's formula). Let M be a manifold with an affine connection  $\nabla$ ,  $N_p$  a normal neighborhood of  $p \in M$ , the vector field  $\tilde{X}$  on  $N_p$  adapted to X in  $T_p$ , and f a  $C^{\infty}$  function on M. Then there exists an  $\epsilon > 0$  such that for every  $\lambda \in [0, \epsilon)$ 

$$f(\exp_{p} \lambda X) = f(p) + \lambda(\nabla_{\tilde{X}} f)(p) + \dots + \frac{\lambda^{n-1}}{(n-1)!} (\nabla_{\tilde{X}}^{n-1} f)(p) + \frac{\lambda^{n}}{(n-1)!} \int_{0}^{1} (1-t)^{n-1} (\nabla_{\tilde{X}}^{n} f)(\exp_{p} t \lambda X) dt.$$
(1)

**Proof.** Let  $N_0$  be a star-shaped neighborhood of  $0 \in T_p$  such that  $N_p = \exp N_0$ . There exists  $\epsilon > 0$  such that  $\lambda X \in N_0$  for all  $\lambda \in [0, \epsilon)$ . The map  $\lambda \mapsto f(\exp \lambda X)$  is a real  $C^{\infty}$  function on  $[0, \epsilon)$  with derivative  $(\nabla_{\bar{X}} f)(\exp \lambda X)$ . The statement follows by repeated integration by parts.

The following special cases of Remark 3.2 will be particularly useful. When n=2, Eq. (1) yields

$$f(\exp_p \lambda X) = f(p) + \lambda(\nabla_{\tilde{X}} f)(p) + \lambda^2 \int_0^1 (1 - t)(\nabla_{\tilde{X}}^2 f)(\exp_p t \lambda X) dt.$$
 (2)

Furthermore, when n=1, Eq. (1) applied to the function  $\tilde{X}f=\nabla_{\tilde{X}}f$  yields

$$(\tilde{X}f)(\exp_p \lambda X) = (\tilde{X}f)(p) + \lambda \int_0^1 (\nabla_{\tilde{X}}^2 f)(\exp_p t\lambda X) dt.$$
 (3)

The convergence proofs require a characterization of the second order terms of f near a critical point. Consider the second covariant differential  $\nabla \nabla f = \nabla^2 f$  of a smooth function  $f: M \to \mathbf{R}$ . If  $(U, x^1, \dots, x^n)$  is a coordinate chart on M, then at  $p \in U$  this (0,2) tensor takes the form

$$(\nabla^2 f)_p = \sum_{i,j} \left( \left( \frac{\partial^2 f}{\partial x^i \partial x^j} \right)_p - \sum_k \Gamma_{ji}^k \left( \frac{\partial f}{\partial x^k} \right)_p \right) dx^i \otimes dx^j \tag{4}$$

where  $\Gamma_{ij}^k$  are the Christoffel symbols at p. If  $\hat{p}$  in U is a critical point of f, then  $(\partial f/\partial x^k)_{\hat{p}}=0,\ k=1,\ldots,n$ . Therefore  $(\nabla^2 f)_{\hat{p}}=(d^2 f)_{\hat{p}}$ , where  $(d^2 f)_{\hat{p}}$  is the Hessian of f at the critical point  $\hat{p}$ . Furthermore, for  $p\in M, X,Y\in T_p$ , and  $\tilde{X}$  and  $\tilde{Y}$  vector fields adapted to X and Y, respectively, on a normal neighborhood  $N_p$  of p, we have  $(\nabla^2 f)(\tilde{X},\tilde{Y})=\nabla_{\tilde{Y}}\nabla_{\tilde{X}}f$  on  $N_p$ . Therefore the coefficient of the second term of the Taylor expansion of  $f(\exp tX)$  is  $(\nabla^2_{\tilde{X}}f)_p=(\nabla^2 f)_p(X,X)$ . Note that the bilinear form  $(\nabla^2 f)_p$  on  $T_p\times T_p$  is symmetric if and only if  $\nabla$  is symmetric, which true of the Levi-Civita connection by definition.

**Theorem 3.3.** Let M be a complete Riemannian manifold with Riemannian structure g and Levi-Civita connection  $\nabla$ . Let  $f \in C^{\infty}(M)$  have a nondegenerate critical point at  $\hat{p}$  such that the Hessian  $(d^2f)_{\hat{p}}$  is positive definite. Let  $p_i$  be a sequence of points in M converging to  $\hat{p}$  and  $H_i \in T_{p_i}$  a sequence of tangent vectors such that

(i) 
$$p_{i+1} = \exp_{p_i} \lambda_i H_i$$
 for  $i = 0, 1, ...,$ 

(ii) 
$$\langle -(\operatorname{grad} f)_{p_i}, H_i \rangle \ge c \|(\operatorname{grad} f)_{p_i}\| \|H_i\| \text{ for } c \in (0, 1],$$

where  $\lambda_i$  is chosen such that  $f(\exp \lambda_i H_i) \leq f(\exp \lambda H_i)$  for all  $\lambda \geq 0$ . Then there exists a constant E and a  $\theta \in [0,1)$  such that for all  $i=0,1,\ldots$ ,

$$d(p_i, \hat{p}) \leq E\theta^i$$
.

**Proof.** The proof is a generalization of the one given in Polak [36, p. 242ff] for the method of steepest descent on Euclidean space.

The existence of a convergent sequence is guaranteed by the smoothness of f. If  $p_j = \hat{p}$  for some integer j, the assertion becomes trivial; assume otherwise. By the smoothness of f, there exists an open neighborhood U of  $\hat{p}$  such that  $(\nabla^2 f)_p$  is positive definite for all  $p \in U$ . Therefore, there exist constants k > 0 and  $K \ge k > 0$  such that for all  $X \in T_p$  and all  $p \in U$ ,

$$k||X||^2 \le (\nabla^2 f)_p(X, X) \le K||X||^2.$$
(5)

Define  $X_i \in T_{\hat{p}}$  by the relations  $\exp X_i = p_i$ ,  $i = 0, 1, \ldots$  By assumption,  $df_{\hat{p}} = 0$  and from Eq. (2), we have

$$f(p_i) - f(\hat{p}) = \int_0^1 (1 - t)(\nabla_{\tilde{X}_i}^2 f)(\exp_{\hat{p}} tX_i) dt.$$
 (6)

Combining this equality with the inequalities of (5) yields

$$\frac{1}{2}kd^2(p_i,\hat{p}) \le f(p_i) - f(\hat{p}) \le \frac{1}{2}Kd^2(p_i,\hat{p}). \tag{7}$$

Similarly, we have by Eq. (3)

$$(\tilde{X}_i f)(p_i) = \int_0^1 (\nabla_{\tilde{X}_i}^2 f)(\exp_{\hat{p}} tX_i) dt.$$

Next, use (6) with Schwarz's inequality and the first inequality of (7) to obtain

$$kd^{2}(p_{i},\hat{p}) = k||X_{i}||^{2} \leq \int_{0}^{1} (\nabla_{\tilde{X}_{i}}^{2} f)(\exp_{\hat{p}} tX_{i}) dt = (\tilde{X}_{i}f)(p_{i})$$

$$= df_{p_{i}}((\tilde{X}_{i})_{p_{i}}) = df_{p_{i}}(\tau X_{i}) = \langle (\operatorname{grad} f)_{p_{i}}, \tau X_{i} \rangle$$

$$\leq ||(\operatorname{grad} f)_{p_{i}}|| \, ||\tau X_{i}|| = ||(\operatorname{grad} f)_{p_{i}}|| \, d(p_{i},\hat{p}).$$

Therefore,

$$\|(\operatorname{grad} f)_{p_i}\| \ge kd(p_i, \hat{p}). \tag{8}$$

Define the function  $\Delta: T_p \times \mathbf{R} \to \mathbf{R}$  by the equation  $\Delta(X, \lambda) = f(\exp_p \lambda X) - f(p)$ . By Eq. (2), the second order Taylor formula, we have

$$\Delta(H_i, \lambda) = \lambda(\tilde{H}_i f)(p_i) + \frac{1}{2}\lambda^2 \int_0^1 (1 - t)(\nabla_{\tilde{H}_i}^2 f)(\exp_{p_i} \lambda H_i) dt.$$

Using assumption (ii) of the theorem along with (5) we establish for  $\lambda \geq 0$ 

$$\Delta(H_i, \lambda) \le -\lambda c \|(\operatorname{grad} f)_{p_i}\| \|H_i\| + \frac{1}{2}\lambda^2 K \|H_i\|^2.$$
 (9)

We may now compute an upper bound for the rate of linear convergence  $\theta$ . By assumption (i) of the theorem,  $\lambda$  must be chosen to minimize the right hand side of (9). This corresponds to choosing  $\lambda = c \|(\operatorname{grad} f)_{p_i}\|/K\|H_i\|$ . A computation reveals that

$$\Delta(H_i, \lambda_i) \le -\frac{c^2}{2K} \|(\operatorname{grad} f)_{p_i}\|^2.$$

Applying (7) and (8) to this inequality and rearranging terms yields

$$f(p_{i+1}) - f(\hat{p}) \le \theta \left( f(p_i) - f(\hat{p}) \right), \tag{10}$$

where  $\theta = (1 - (ck/K)^2)$ . By assumption,  $c \in (0,1]$  and  $0 < k \le K$ , therefore  $\theta \in [0,1)$ . (Note that Schwarz's inequality bounds c below unity.) From (10) it is seen that  $(f(p_i) - f(\hat{p})) \le E\theta^i$  where  $E = (f(p_0) - f(\hat{p}))$ . From (7) we conclude that for  $i = 0, 1, \ldots$ ,

$$d(p_i, \hat{p}) \le \sqrt{\frac{2E}{k}} \left(\sqrt{\theta}\right)^i. \quad \blacksquare$$
 (11)

**Corollary 3.4.** If Algorithm 3.1 converges to a local minimum, it converges linearly.

The choice  $H_i = -(\text{grad}f)_{p_i}$  yields c = 1 in the second assumption the Theorem 3.3, which establishes the corollary.

**Example 3.5** (Rayleigh's quotient on the sphere). Let  $S^{n-1}$  be the imbedded sphere in  $\mathbf{R}^n$ , i.e.,  $S^{n-1} = \{x \in \mathbf{R}^n : x^{\mathrm{T}}x = 1\}$ , where  $x^{\mathrm{T}}y$  denotes the standard inner product on  $\mathbf{R}^n$ , which induces a metric on  $S^{n-1}$ . Geodesics on the sphere are great circles and parallel translation along geodesics is equivalent to rotating the tangent plane along the great circle. Let  $x \in S^{n-1}$  and  $h \in T_x$  have unit length, and  $v \in T_x$  be any tangent vector. Then

$$\begin{split} \exp_x th &= x \cos t + h \sin t, \\ \tau h &= h \cos t - x \sin t, \\ \tau v &= v - (h^{ \mathrm{\scriptscriptstyle T} } v) \big( x \sin t + h (1 - \cos t) \big), \end{split}$$

where  $\tau$  is the parallelism along the geodesic  $t \mapsto \exp th$ . Let Q be an n-by-n positive definite symmetric matrix with distinct eigenvalues and define  $\rho: S^{n-1} \to \mathbf{R}$  by  $\rho(x) = x^{\mathrm{T}}Qx$ . A computation shows that

$$\frac{1}{2}(\operatorname{grad}\rho)_x = Qx - \rho(x)x. \tag{12}$$

The function  $\rho$  has a unique minimum and maximum point at the eigenvectors corresponding to the smallest and largest eigenvalues of Q, respectively. Because  $S^{n-1}$  is geodesically complete, the method of steepest descent in the opposite direction of the gradient converges to the eigenvector corresponding to the smallest eigenvalue of Q; likewise for the eigenvector corresponding to the largest eigenvalue. Chu [13] considers the continuous limit of this problem. A computation shows that  $\rho(x)$  is maximized along the geodesic  $\exp_x th$  (||h|| = 1) when  $a \cos 2t - b \sin 2t = 0$ , where  $a = 2x^TQh$  and  $b = \rho(x) - \rho(h)$ . Thus  $\cos t$  and  $\sin t$  may be computed with simple algebraic functions of a and b (which appear below in Algorithm 5.5). The results of a numerical experiment demonstrating the convergence of the method of steepest descent applied to maximizing Rayleigh's quotient on  $S^{20}$  are shown in Figure 1 on page 133.

Example 3.6 (Brockett [9,10]). Consider the function  $f(\Theta) = \operatorname{tr} \Theta^{\mathsf{T}} Q \Theta N$  on the special orthogonal group SO(n), where Q is a real symmetric matrix with distinct eigenvalues and N is a real diagonal matrix with distinct diagonal elements. It will be convenient to identify tangent vectors in  $T_{\Theta}$  with tangent vectors in  $T_{I} \cong \mathbf{so}(n)$ , the tangent plane at the identity, via left translation. The gradient of f (with respect to the negative Killing form of  $\mathbf{so}(n)$ , scaled by 1/(n-2)) at  $\Theta \in SO(n)$  is  $\Theta[H,N]$ , where  $H = \operatorname{Ad}_{\Theta^{\mathsf{T}}}(Q) = \Theta^{\mathsf{T}} Q \Theta$ . The group SO(n) acts on the set of symmetric matrices by conjugation; the orbit of Q under the action of SO(n) is an isospectral submanifold of the symmetric matrices. We seek a  $\hat{\Theta}$  such that  $f(\hat{\Theta})$  is maximized. This point corresponds to a diagonal matrix whose diagonal entries are ordered similarly to those of N. A related example is found in Smith [45], who considers the homogeneous space of matrices with fixed singular values, and in Chu [14].

The Levi-Civita connection on SO(n) is bi-invariant and invariant with respect to inversion; therefore, geodesics and parallel translation may be computed via matrix exponentiation of elements in  $\mathbf{so}(n)$  and left (or right) translation [24, Ch. II, Ex. 6]. The geodesic emanating from the identity in SO(n) in direction  $X \in \mathbf{so}(n)$  is given by the formula  $\exp_I tX = e^{tX}$ , where the right hand side denotes regular matrix exponentiation. The expense of geodesic minimization may be avoided if instead one uses Brockett's estimate [10] for the step size. Given  $\Omega \in \mathbf{so}(n)$ , we wish to find t > 0 such that  $\phi(t) = \operatorname{tr} \operatorname{Ad}_{e^{-t\Omega}}(\operatorname{ad}_{\Omega} H)N$  is minimized. Differentiating  $\phi$  twice shows that  $\phi'(t) = -\operatorname{tr} \operatorname{Ad}_{e^{-t\Omega}}(\operatorname{ad}_{\Omega} H)N$  and  $\phi''(t) = -\operatorname{tr} \operatorname{Ad}_{e^{-t\Omega}}(\operatorname{ad}_{\Omega} H) \operatorname{ad}_{\Omega} N$ , where  $\operatorname{ad}_{\Omega} A = [\Omega, A]$ . Hence,  $\phi'(0) = 2\operatorname{tr} H\Omega N$  and, by Schwarz's inequality and the fact that Ad is an isometry,  $|\phi''(t)| \leq \|\operatorname{ad}_{\Omega} H\| \|\operatorname{ad}_{\Omega} N\|$ . We conclude that if  $\phi'(0) > 0$ , then  $\phi'$  is nonnegative on the interval

$$0 \le t \le \frac{2 \operatorname{tr} H\Omega N}{\|\operatorname{ad}_{\Omega} H\| \|\operatorname{ad}_{\Omega} N\|},\tag{13}$$

which provides an estimate for the step size of Step 1 in Algorithm 3.1. The results of a numerical experiment demonstrating the convergence of the method of steepest descent (ascent) in SO(20) using this estimate are shown in Figure 2.

## 4 Newton's method on Riemannian manifolds

As in the optimization of functions on Euclidean space, quadratic convergence can be obtained if the second order terms of the Taylor expansion are used appropriately. In this section we present Newton's algorithm on Riemannian manifolds, prove that its convergence is quadratic, and provide examples. Whereas the convergence proof for the method of steepest descent relies upon the Taylor expansion of the function f, the convergence proof for Newton's method will rely upon the Taylor expansion of the one-form df. Note that Newton's method has a counterpart in the theory of constrained optimization, as described by, e.g., Fletcher [18], Bertsekas [1, 2], or Dunn [15, 16]. The Newton method presented in this section has only local convergence properties. There is a theory of global Newton methods on Euclidean space and computational complexity; see the work of Hirsch and Smale [27], Smale [43, 44], and Shub and Smale [40, 41].

Let M be an n-dimensional Riemannian manifold with Riemannian structure g and Levi-Civita connection  $\nabla$ , let  $\mu$  be a  $C^{\infty}$  one-form on M, and let p in M be such that the bilinear form  $(\nabla \mu)_p : T_p \times T_p \to \mathbf{R}$  is nondegenerate. Then, by abuse of notation, we have the pair of isomorphisms

$$T_p \xleftarrow{(\nabla \mu)_p} T_p^*$$

with the forward map defined by  $X \mapsto (\nabla_X \mu)_p = (\nabla \mu)_p(\cdot, X)$ , which is nonsingular. The notation  $(\nabla \mu)_p$  will henceforth be used for both the bilinear form defined by the covariant differential of  $\mu$  evaluated at p and the homomorphism from  $T_p$  to  $T_p^*$  induced by this bilinear form. In case of an isomorphism, the inverse can be used to compute a point in M where  $\mu$  vanishes, if such a point exists. The case  $\mu = df$  will be of particular interest, in which case  $\nabla \mu = \nabla^2 f$ . Before expounding on these ideas, we make the following remarks.

**Remark 4.1** (The mean value theorem). Let M be a manifold with affine connection  $\nabla$ ,  $N_p$  a normal neighborhood of  $p \in M$ , the vector field  $\tilde{X}$  on  $N_p$  adapted to  $X \in T_p$ ,  $\mu$  a one-form on  $N_p$ , and  $\tau_{\lambda}$  the parallelism with respect to  $\exp tX$  for  $t \in [0, \lambda]$ . Denote the point  $\exp \lambda X$  by  $p_{\lambda}$ . Then there exists an  $\epsilon > 0$  such that for every  $\lambda \in [0, \epsilon)$ , there is an  $\alpha \in [0, \lambda]$  such that

$$\tau_{\lambda}^{-1}\mu_{p_{\lambda}} - \mu_{p} = \lambda(\nabla_{\tilde{X}}\mu)_{p_{\alpha}} \circ \tau_{\alpha}.$$

**Proof.** As in the proof of Remark 3.2, there exists an  $\epsilon > 0$  such that  $\lambda X \in N_0$  for all  $\lambda \in [0, \epsilon)$ . The map  $\lambda \mapsto (\tau_{\lambda}^{-1}\mu_{p_{\lambda}})(A)$ , for any A in  $T_p$ , is a  $C^{\infty}$  function on  $[0, \epsilon)$  with derivative  $(d/dt)(\tau_t^{-1}\mu_{p_t})(A) = (d/dt)\mu_{p_t}(\tau_t A) = \nabla_{\tilde{X}}(\mu_{p_t}(\tau_t A)) = (\nabla_{\tilde{X}}\mu)_{p_t}(\tau_t A) + \mu_{p_t}(\nabla_{\tilde{X}}(\tau_t A)) = (\nabla_{\tilde{X}}\mu)_{p_t}(\tau_t A)$ . The lemma follows from the mean value theorem of real analysis.

This remark can be generalized in the following way.

**Remark 4.2** (Taylor's theorem). Let M be a manifold with affine connection  $\nabla$ ,  $N_p$  a normal neighborhood of  $p \in M$ , the vector field  $\tilde{X}$  on  $N_p$  adapted

to  $X \in T_p$ ,  $\mu$  a one-form on  $N_p$ , and  $\tau_{\lambda}$  the parallelism with respect to  $\exp tX$  for  $t \in [0, \lambda]$ . Denote the point  $\exp \lambda X$  by  $p_{\lambda}$ . Then there exists an  $\epsilon > 0$  such that for every  $\lambda \in [0, \epsilon)$ , there is an  $\alpha \in [0, \lambda]$  such that

$$\tau_{\lambda}^{-1}\mu_{p_{\lambda}} = \mu_{p} + \lambda(\nabla_{\tilde{X}}\mu)_{p} + \dots + \frac{\lambda^{n-1}}{(n-1)!}(\nabla_{\tilde{X}}^{n-1}\mu)_{p} + \frac{\lambda^{n}}{n!}(\nabla_{\tilde{X}}^{n}\mu)_{p_{\alpha}} \circ \tau_{\alpha}. \quad (14)$$

The remark follows by applying Remark 4.1 and the Taylor's theorem of real analysis to the function  $\lambda \mapsto (\tau_{\lambda}^{-1} \mu_{p_{\lambda}})(A)$  for any A in  $T_p$ .

Remarks 4.1 and 4.2 can be generalized to  $C^{\infty}$  tensor fields, but we will only require Remark 4.2 for case n=2 to make the following observation.

Let  $\mu$  be a one-form on M such that for some  $\hat{p}$  in M,  $\mu_{\hat{p}} = 0$ . Given any p in a normal neighborhood of  $\hat{p}$ , we wish to find X in  $T_p$  such that  $\exp_p X = \hat{p}$ . Consider the Taylor expansion of  $\mu$  about p, and let  $\tau$  be the parallel translation along the unique geodesic joining p to  $\hat{p}$ . We have by our assumption that  $\mu$  vanishes at  $\hat{p}$ , and from Eq. (14) for n = 2,

$$0 = \tau^{-1} \mu_{\hat{p}} = \tau^{-1} \mu_{\exp_n X} = \mu_p + (\nabla \mu)_p(\cdot, X) + \text{h.o.t.}$$

If the bilinear form  $(\nabla \mu)_p$  is nondegenerate, the tangent vector X may be approximated by discarding the higher order terms and solving the resulting linear equation

$$\mu_p + (\nabla \mu)_p(\cdot, X) = 0$$

for X, which yields

$$X = -(\nabla \mu)_p^{-1} \mu_p.$$

This approximation is the basis of the following algorithm.

**Algorithm 4.3** (Newton's method). Let M be a complete Riemannian manifold with Riemannian structure g and Levi-Civita connection  $\nabla$ , and let  $\mu$  be a  $C^{\infty}$  one-form on M.

Step 0. Select  $p_0 \in M$  such that  $(\nabla \mu)_{p_0}$  is nondegenerate, and set i = 0.

Step 1. Compute

$$H_i = -(\nabla \mu)_{p_i}^{-1} \mu_{p_i}$$
$$p_{i+1} = \exp_{p_i} H_i,$$

(assume that  $(\nabla \mu)_{p_i}$  is nondegenerate), increment i, and repeat.

It can be shown that if  $p_0$  is chosen suitably close (within the so-called domain of attraction) to a point  $\hat{p}$  in M such that  $\mu_{\hat{p}} = 0$  and  $(\nabla \mu)_{\hat{p}}$  is non-degenerate, then Algorithm 4.3 converges quadratically to  $\hat{p}$ . The following theorem holds for general one-forms; we will consider the case where  $\mu$  is exact.

**Theorem 4.4.** Let  $f \in C^{\infty}(M)$  have a nondegenerate critical point at  $\hat{p}$ . Then there exists a neighborhood U of  $\hat{p}$  such that for any  $p_0 \in U$ , the iterates of Algorithm 4.3 for  $\mu = df$  are well defined and converge quadratically to  $\hat{p}$ .

The proof of this theorem is a generalization of the corresponding proof for Euclidean spaces, with an extra term containing the Riemannian curvature tensor (which of course vanishes in the latter case).

**Proof.** If  $p_j = \hat{p}$  for some integer j, the assertion becomes trivial; assume otherwise. Define  $X_i \in T_{p_i}$  by the relations  $\hat{p} = \exp X_i$ ,  $i = 0, 1, \ldots$ , so that  $d(p_i, \hat{p}) = ||X_i||$  (n.b. this convention is opposite that used in the proof of Theorem 3.3). Consider the geodesic triangle with vertices  $p_i$ ,  $p_{i+1}$ , and  $\hat{p}$ , and sides  $\exp tX_i$  from  $p_i$  to  $\hat{p}$ ,  $\exp tH_i$  from  $p_i$  to  $p_{i+1}$ , and  $\exp tX_{i+1}$  from  $p_{i+1}$  to  $\hat{p}$ , for  $t \in [0,1]$ . Let  $\tau$  be the parallelism with respect to the side  $\exp tH_i$  between  $p_i$  and  $p_{i+1}$ . There exists a unique tangent vector  $\Xi_i$  in  $T_{p_i}$  defined by the equation

$$X_i = H_i + \tau^{-1} X_{i+1} + \Xi_i \tag{15}$$

 $(\Xi_i)$  may be interpreted as the amount by which vector addition fails). If we use the definition  $H_i = -(\nabla^2 f)_{p_i}^{-1} df_{p_i}$  of Algorithm 4.3, apply the isomorphism  $(\nabla^2 f)_{p_i}: T_{p_i} \to T_{p_i}^*$  to both sides of Eq. (15), we obtain the equation

$$(\nabla^2 f)_{p_i}(\tau^{-1} X_{i+1}) = df_{p_i} + (\nabla^2 f)_{p_i} X_i - (\nabla^2 f)_{p_i} \Xi_i.$$
(16)

By Taylor's theorem, there exists an  $\alpha \in [0,1]$  such that

$$\tau_1^{-1} df_{\hat{p}} = df_{p_i} + (\nabla_{\tilde{X}_i} df)_{p_i} + \frac{1}{2} (\nabla_{\tilde{X}_i}^2 df)_{p_\alpha} \circ \tau_\alpha$$
 (17)

where  $\tau_t$  is the parallel translation from  $p_i$  to  $p_t = \exp tX_i$ . The trivial identities  $(\nabla_{\tilde{X}_i} df)_{p_i} = (\nabla^2 f)_{p_i} X_i$  and  $(\nabla_{\tilde{X}_i}^2 df)_{p_\alpha} = (\nabla^3 f)_{p_\alpha} (\tau_\alpha, \tau_\alpha X_i, \tau_\alpha X_i)$  will be used to replace the last two terms on the right hand side of Eq. (17). Combining the assumption that  $df_{\hat{p}} = 0$  with Eqs. (16) and (17), we obtain

$$(\nabla^{2} f)_{p_{i}}(\tau^{-1} X_{i+1}) = -\frac{1}{2} (\nabla^{2}_{\tilde{X}_{i}} df)_{p_{\alpha}} \circ \tau_{\alpha} - (\nabla^{2} f)_{p_{i}} \Xi_{i}.$$
(18)

By the smoothness of f and g, there exists an  $\epsilon > 0$  and constants  $\delta'$ ,  $\delta''$ ,  $\delta'''$ , all greater than zero, such that whenever p is in the convex normal ball  $B_{\epsilon}(\hat{p})$ ,

(i) 
$$\|(\nabla^2 f)_p(\cdot, X)\| \ge \delta' \|X\|$$
 for all  $X \in T_p$ ,

(ii) 
$$\|(\nabla^2 f)_p(\cdot, X)\| \le \delta'' \|X\|$$
 for all  $X \in T_p$ ,

(iii) 
$$\|(\nabla^3 f)_p(\cdot, X, X)\| \le \delta''' \|X\|^2$$
 for all  $X \in T_p$ ,

where the induced norm on  $T_p^*$  is used in all three cases. Taking the norm of both sides of Eq. (18), applying the triangle inequality to the right hand side, and using the fact that parallel translation is an isometry, we obtain the inequality

$$\delta' d(p_{i+1}, \hat{p}) \le \delta''' d^2(p_i, \hat{p}) + \delta'' \|\Xi_i\|.$$
 (19)

The length of  $\Xi_i$  can be bounded by a cubic expression in  $d(p_i, \hat{p})$  by considering the distance between the points  $\exp(H_i + \tau^{-1}X_{i+1})$  and  $\exp X_{i+1} = \hat{p}$ . Given  $p \in M$ ,  $\epsilon > 0$  small enough, let  $a, v \in T_p$  be such that  $||a|| + ||v|| \le \epsilon$ , and

let  $\tau$  be the parallel translation with respect to the geodesic from p to  $q=\exp_p a$ . Karcher [28, App. C2.2] shows that

$$d(\exp_n(a+v), \exp_a(\tau v)) \le ||a|| \cdot \text{const.} (\max |K|) \cdot \epsilon^2, \tag{20}$$

where K is the sectional curvature of M along any section in the tangent plane at any point near p.

There exists a constant c > 0 such that  $\|\Xi_i\| \le c d(\hat{p}, \exp(H_i + \tau^{-1}X_{i+1}))$ . By (20), we have  $\|\Xi_i\| \le \text{const.} \|H_i\|\epsilon^2$ . Taking the norm of both sides of the Taylor formula  $df_{p_i} = -\int_0^1 (\nabla_{\tilde{X}_i} df)(\exp tX_i) dt$  and applying a standard integral inequality and inequality (ii) from above yields  $\|df_{p_i}\| \le \delta'' \|X_i\|$  so that  $\|H_i\| \le \text{const.} \|X_i\|$ . Furthermore, we have the triangle inequality  $\|X_{i+1}\| \le \|X_i\| + \|H_i\|$ , therefore  $\epsilon$  may be chosen such that  $\|H_i\| + \|X_{i+1}\| \le \epsilon \le \text{const.} \|X_i\|$ . By (20) there exists  $\delta^{\text{iv}} > 0$  such that  $\|\Xi_i\| \le \delta^{\text{iv}} d^3(p_i, \hat{p})$ .

**Corollary 4.5.** If  $(\nabla^2 f)_{\hat{p}}$  is positive (negative) definite and Algorithm 4.3 converges to  $\hat{p}$ , then Algorithm 4.3 converges quadratically to a local minimum (maximum) of f.

Example 4.6 (Rayleigh's quotient on the sphere). Let  $S^{n-1}$  and  $\rho(x) = x^{\mathrm{T}}Qx$  be as in Example 3.5. It will be convenient to work with the coordinates  $x^1, \ldots, x^n$  of the ambient space  $\mathbf{R}^n$ , treat the tangent plane  $T_x S^{n-1}$  as a vector subspace of  $\mathbf{R}^n$ , and make the identification  $T_x S^{n-1} \cong T_x^* S^{n-1}$  via the metric. In this coordinate system, geodesics on the sphere obey the second order differential equation  $\ddot{x}^k + x^k = 0, k = 1, \ldots, n$ . Thus the Christoffel symbols are given by  $\Gamma^k_{ij} = \delta_{ij} x^k$ , where  $\delta_{ij}$  is the Kronecker delta. The ijth component of the second covariant differential of  $\rho$  at x in  $S^{n-1}$  is given by (cf. Eq. (4))

$$\left( (\nabla^2 \rho)_x \right)_{ij} = 2Q_{ij} - \sum_{k,l} \delta_{ij} x^k \cdot 2Q_{kl} x^l = 2\left( Q_{ij} - \rho(x) \delta_{ij} \right),$$

or, written as matrices,

$$\frac{1}{2}(\nabla^2 \rho)_x = Q - \rho(x)I. \tag{21}$$

Let u be a tangent vector in  $T_xS^{n-1}$ . A linear operator  $A: \mathbf{R}^n \to \mathbf{R}^n$  defines a linear operator on the tangent plane  $T_xS^{n-1}$  for each x in  $S^{n-1}$  such that

$$A \cdot u = Au - (x^{\mathrm{T}}Au)x = (I - xx^{\mathrm{T}})Au$$

If A is invertible as an endomorphism of the ambient space  $\mathbb{R}^n$ , the solution to the linear equation  $A \cdot u = v$  for u, v in  $T_x S^{n-1}$  is

$$u = A^{-1} \left( v - \frac{(x^{\mathrm{T}}A^{-1}v)}{(x^{\mathrm{T}}A^{-1}x)} x \right). \tag{22}$$

For Newton's method, the direction  $H_i$  in  $T_x S^{n-1}$  is the solution of the equation

$$(\nabla^2 \rho)_{x_i} \cdot H_i = -(\operatorname{grad} \rho)_{x_i}.$$

Combining Eqs. (12), (21), and (22), we obtain

$$H_i = -x_i + \alpha_i (Q - \rho(x_i)I)^{-1} x_i$$

where  $\alpha_i = 1/x_i^{\mathrm{T}}(Q - \rho(x_i)I)^{-1}x_i$ . This gives rise to the following algorithm for computing eigenvectors of the symmetric matrix Q.

**Algorithm 4.7** (Newton-Rayleigh quotient method). Let Q be a real symmetric n-by-n matrix.

Step 0. Select  $x_0$  in  $\mathbf{R}^n$  such that  $x_0^T x_0 = 1$ , and set i = 0.

Step 1. Compute

$$y_i = (Q - \rho(x_i)I)^{-1}x_i$$

and set  $\alpha_i = 1/x_i^{\mathrm{T}} y_i$ .

Step 2. Compute

$$H_i = -x_i + \alpha_i y_i, \quad \theta_i = ||H_i||,$$
  
$$x_{i+1} = x_i \cos \theta_i + H_i \sin \theta_i / \theta_i,$$

increment i, and go to Step 1.

The quadratic convergence guaranteed by Theorem 4.4 is in fact too conservative for Algorithm 4.7. As evidenced by Figure 1, Algorithm 4.7 converges cubically.

**Proposition 4.8.** If  $\lambda$  is a distinct eigenvalue of the symmetric matrix Q, and Algorithm 4.7 converges to the corresponding eigenvector  $\hat{x}$ , then it converges cubically.

**Proof 1.** In the coordinates  $x^1, \ldots, x^n$  of the ambient space  $\mathbf{R}^n$ , the ijkth component of the third covariant differential of  $\rho$  at  $\hat{x}$  is  $-2\lambda\hat{x}^k\delta_{ij}$ . Let  $X\in T_{\hat{x}}S^{n-1}$ . Then  $(\nabla^3\rho)_{\hat{x}}(\cdot,X,X)=0$  and the second order terms on the right hand side of Eq. (18) vanish at the critical point. The proposition follows from the smoothness of  $\rho$ .

**Proof 2.** The proof follows Parlett's [34, p. 72ff] proof of cubic convergence for the Rayleigh quotient iteration. Assume that for all i,  $x_i \neq \hat{x}$ , and denote  $\rho(x_i)$  by  $\rho_i$ . For all i, there is an angle  $\psi_i$  and a unit length vector  $u_i$  defined by the equation  $x_i = \hat{x} \cos \psi_i + u_i \sin \psi_i$ , such that  $\hat{x}^{\text{T}} u_i = 0$ . By Algorithm 4.7

$$\begin{split} x_{i+1} &= \hat{x}\cos\psi_{i+1} + u_{i+1}\sin\psi_{i+1} = x_i\cos\theta_i + H_i\sin\theta_i/\theta_i \\ &= \hat{x}\bigg(\frac{\alpha_i\sin\theta_i}{(\lambda - \rho_i)\theta_i} + \beta_i\bigg)\cos\psi_i + \bigg(\frac{\alpha_i\sin\theta_i}{\theta_i}(Q - \rho_iI)^{-1}u_i + \beta_iu_i\bigg)\sin\psi_i, \end{split}$$

where  $\beta_i = \cos \theta_i - \sin \theta_i / \theta_i$ . Therefore,

$$|\tan \psi_{i+1}| = \frac{\left\| \frac{\alpha_i \sin \theta_i}{\theta_i} (Q - \rho_i I)^{-1} u_i + \beta_i u_i \right\|}{\left| \frac{\alpha_i \sin \theta_i}{(\lambda - \rho_i)\theta_i} + \beta_i \right|} \cdot |\tan \psi_i|. \tag{23}$$

The following equalities and low order approximations in terms of the small quantities  $\lambda - \rho_i$ ,  $\theta_i$ , and  $\psi_i$  are straightforward to establish:  $\lambda - \rho_i = (\lambda - \rho(u_i)) \times \sin^2 \psi_i$ ,  $\theta_i^2 = \cos^2 \psi_i \sin^2 \psi_i + \text{h.o.t.}$ ,  $\alpha_i = (\lambda - \rho_i) + \text{h.o.t.}$ , and  $\beta_i = -\theta_i^2/3 + \text{h.o.t.}$  Thus, the denominator of the large fraction in Eq. (23) is of order unity and the numerator is of order  $\sin^2 \psi_i$ . Therefore, we have

$$|\psi_{i+1}| = \text{const.} |\psi_i|^3 + \text{h.o.t.}$$

Remark 4.9. If Algorithm 4.7 is simplified by replacing Step 2 with

Step 2! Compute

$$x_{i+1} = y_i / ||y_i||,$$

increment i, and go to Step 1.

then we obtain the Rayleigh quotient iteration. These two algorithms differ by the method in which they use the vector  $y_i = (Q - \rho(x_i)I)^{-1}x_i$  to compute the next iterate on the sphere. Algorithm 4.7 computes the point  $H_i$  in  $T_{x_i}S^{n-1}$  where  $y_i$  intersects this tangent plane, then computes  $x_{i+1}$  via the exponential map of this vector (which "rolls" the tangent vector  $H_i$  onto the sphere). The Rayleigh quotient iteration computes the intersection of  $y_i$  with the sphere itself and takes this intersection to be  $x_{i+1}$ . The latter approach approximates Algorithm 4.7 up to quadratic terms when  $x_i$  is close to an eigenvector. Algorithm 4.7 is more expensive to compute than—though of the same order as—the Rayleigh quotient iteration; thus, the RQI is seen to be an efficient approximation of Newton's method.

If the exponential map is replaced by the chart  $v \in T_x \mapsto (x+v)/\|x+v\| \in S^{n-1}$ , Shub [39] shows that a corresponding version of Newton's method is equivalent to the RQI.

**Example 4.10** (The function  $\operatorname{tr} \Theta^{\mathrm{T}} Q \Theta N$ ). Let  $\Theta$ , Q,  $H = \operatorname{Ad}_{\Theta^{\mathrm{T}}}(Q)$ , and  $\Omega$  be as in Example 3.6. The second covariant differential of  $f(\Theta) = \operatorname{tr} \Theta^{\mathrm{T}} Q \Theta N$  may be computed either by polarization of the second order term of  $\operatorname{tr} \operatorname{Ad}_{e^{-t\Omega}}(H)N$ , or by covariant differentiation of the differential  $df_{\Theta} = -\operatorname{tr}[H, N]\Theta^{\mathrm{T}}(\cdot)$ :

$$(\nabla^2 f)_{\Theta}(\Theta X, \Theta Y) = -\frac{1}{2}\operatorname{tr}([H, \operatorname{ad}_X N] - [\operatorname{ad}_X H, N])Y,$$

where  $X, Y \in \mathbf{so}(n)$ . To compute the direction  $\Theta X \in T_{\Theta}, X \in \mathbf{so}(n)$ , for Newton's method, we must solve the equation  $(\nabla^2 f)_{\Theta}(\Theta, \Theta X) = df_{\Theta}$ , which yields the linear equation

$$L_{\Theta}(X) \stackrel{\text{def}}{=} [H, \operatorname{ad}_X N] - [\operatorname{ad}_X H, N] = 2[H, N].$$

The linear operator  $L_{\Theta} : \mathbf{so}(n) \to \mathbf{so}(n)$  is self-adjoint for all  $\Theta$  and, in a neighborhood of the maximum, negative definite. Therefore, standard iterative techniques in the vector space  $\mathbf{so}(n)$ , such as the classical conjugate gradient method, may be used to solve this equation near the maximum. The results of a numerical experiment demonstrating the convergence of Newton's method in SO(20) are shown in Figure 2. As can be seen, Newton's method converged within round-off error in two iterations.

**Remark 4.11.** If Newton's method applied to the function  $f(\Theta) = \operatorname{tr} \Theta^{\mathrm{T}} Q \Theta N$  converges to the point  $\hat{\Theta}$  such that  $\operatorname{Ad}_{\hat{\Theta}^{\mathrm{T}}}(Q) = H_{\infty} = \alpha N, \ \alpha \in \mathbf{R}$ , then it converges cubically.

**Proof.** By covariant differentiation of  $\nabla^2 f$ , the third covariant differential of f at  $\Theta$  evaluated at the tangent vectors  $\Theta X$ ,  $\Theta Y$ ,  $\Theta Z \in T_{\Theta}$ , X, Y,  $Z \in \mathbf{so}(n)$ , is

$$(\nabla^{3} f)_{\Theta}(\Theta X, \Theta Y, \Theta Z) = -\frac{1}{4} \operatorname{tr} ([\operatorname{ad}_{Y} \operatorname{ad}_{Z} H, N] - [\operatorname{ad}_{Z} \operatorname{ad}_{Y} N, H] + [H, \operatorname{ad}_{\operatorname{ad}_{Y} Z} N] - [\operatorname{ad}_{Y} H, \operatorname{ad}_{Z} N] + [\operatorname{ad}_{Y} N, \operatorname{ad}_{Z} H]) X.$$

If  $H = \alpha N$ ,  $\alpha \in \mathbf{R}$ , then  $(\nabla^3 f)_{\Theta}(\cdot, \Theta X, \Theta X) = 0$ . Therefore, the second order terms on the right hand side of Eq. (18) vanish at the critical point. The remark follows from the smoothness of f.

This remark illuminates how rapid convergence of Newton's method applied to the function f can be achieved in some instances. If  $E_{ij} \in \mathbf{so}(n)$  is a matrix with entry +1 at element (i,j), -1 at element (j,i), and zero elsewhere,  $X = \sum_{i < j} x^{ij} E_{ij}$ ,  $H = \operatorname{diag}(h_1, \ldots, h_n)$ , and  $N = \operatorname{diag}(\nu_1, \ldots, \nu_n)$ , then

$$(\nabla^3 f)_{\Theta}(\Theta E_{ij}, \Theta X, \Theta X) =$$

$$-2 \sum_{k \neq i,j} x^{ik} x^{jk} \left( (h_i \nu_j - h_j \nu_i) + (h_j \nu_k - h_k \nu_j) + (h_k \nu_i - h_i \nu_k) \right).$$

If the  $h_i$  are close to  $\alpha \nu_i$ ,  $\alpha \in \mathbf{R}$ , for all i, then  $(\nabla^3 f)_{\Theta}(\cdot, \Theta X, \Theta X)$  may be small, yielding a fast rate of quadratic convergence.

**Example 4.12** (Jacobi's method). Let  $\pi$  be the projection of a square matrix onto its diagonal, and let Q be as above. Consider the maximization of the function  $f(\Theta) = \operatorname{tr} H\pi(H)$ ,  $H = \operatorname{Ad}_{\Theta^T}(Q)$ , on the special orthogonal group. This is equivalent to minimizing the sum of the squares of the off-diagonal elements of H (Golub and Van Loan [22] derive the classical Jacobi method). The gradient of this function at  $\Theta$  is  $2\Theta[H, \pi(H)]$  [14]. By repeated covariant differentiation of f, we find

$$\begin{split} (\nabla f)_I(X) &= -2\operatorname{tr}[H, \pi(H)]X \\ (\nabla^2 f)_I(X,Y) &= -\operatorname{tr}\big([H,\operatorname{ad}_X \pi(H)] - [\operatorname{ad}_X H, \pi(H)] - 2[H, \pi(\operatorname{ad}_X H)]\big)Y \\ (\nabla^3 f)_I(X,Y,Z) &= -\frac{1}{2}\operatorname{tr}\big([\operatorname{ad}_Y \operatorname{ad}_Z H, \pi(H)] - [\operatorname{ad}_Z \operatorname{ad}_Y \pi(H), H] \\ &+ [H,\operatorname{ad}_{\operatorname{ad}_Y Z} \pi(H)] - [\operatorname{ad}_Y H, \operatorname{ad}_Z \pi(H)] + [\operatorname{ad}_Y \pi(H), \operatorname{ad}_Z H] \\ &+ 2[H, \pi(\operatorname{ad}_Y \operatorname{ad}_Z H)] + 2[H, \pi(\operatorname{ad}_Z \operatorname{ad}_Y H)] \\ &+ 2[\operatorname{ad}_Y H, \pi(\operatorname{ad}_Z H)] - 2[H, \operatorname{ad}_Y \pi(\operatorname{ad}_Z H)] \\ &+ 2[\operatorname{ad}_Z H, \pi(\operatorname{ad}_Y H)] - 2[H, \operatorname{ad}_Z \pi(\operatorname{ad}_Y H)])X \end{split}$$

where I is the identity matrix and  $X, Y, Z \in \mathbf{so}(n)$ . It is easily shown that if  $[H, \pi(H)] = 0$ , i.e., if H is diagonal, then  $(\nabla^3 f)_{\Theta}(\cdot, \Theta X, \Theta X) = 0$  (n.b.  $\pi(\mathrm{ad}_X H) = 0$ ). Therefore, by the same argument as the proof of Remark 4.11, Newton's method applied to the function  $\mathrm{tr} H \pi(H)$  converges cubically.

## 5 Conjugate gradient on Riemannian manifolds

The method of steepest descent provides an optimization technique which is relatively inexpensive per iteration, but converges relatively slowly. Each step requires the computation of a geodesic and a gradient direction. Newton's method provides a technique which is more costly both in terms of computational complexity and memory requirements, but converges relatively rapidly. Each step requires the computation of a geodesic, a gradient, a second covariant differential, and its inverse. In this section we describe the conjugate gradient method, which has the dual advantages of algorithmic simplicity and superlinear convergence.

Hestenes and Stiefel [26] first used conjugate gradient methods to compute the solutions of linear equations, or, equivalently, to compute the minimum of a quadratic form on  $\mathbb{R}^n$ . This approach can be modified to yield effective algorithms to compute the minima of nonquadratic functions on  $\mathbb{R}^n$ . In particular, Fletcher and Reeves [19] and Polak and Ribière [36] provide algorithms based upon the assumption that the second order Taylor expansion of the function to be minimized sufficiently approximates this function near the minimum. In addition, Davidon, Fletcher, and Reeves developed the variable metric methods [18, 36], but these will not be discussed here. One noteworthy feature of conjugate gradient algorithms on  $\mathbb{R}^n$  is that when the function in question is quadratic, they compute its minimum in no more than n steps.

The conjugate gradient method on Euclidean space is uncomplicated. Given a function  $f: \mathbf{R}^n \to \mathbf{R}$  with continuous second derivatives and a local minimum at  $\hat{x}$ , and an initial point  $x_0 \in \mathbf{R}^n$ , the algorithm is initialized by computing the (negative) gradient direction  $G_0 = H_0 = -(\operatorname{grad} f)_{x_0}$ . The recursive part of the algorithm involves (i) a line minimization of f along the affine space  $x_i + tH_i$ ,  $t \in \mathbf{R}$ , where the minimum occurs at, say,  $t = \lambda_i$ , (ii) computation of the step  $x_{i+1} = x_i + \lambda_i H_i$ , (iii) computation of the (negative) gradient  $G_{i+1} = -(\operatorname{grad} f)_{x_{i+1}}$ , and (iv) computation of the next direction for line minimization,

$$H_{i+1} = G_{i+1} + \gamma_i H_i, (24)$$

where  $\gamma_i$  is chosen such that  $H_i$  and  $H_{i+1}$  conjugate with respect to the Hessian matrix of f at  $\hat{x}$ . When f is a quadratic form represented by the symmetric positive definite matrix Q, the conjugacy condition becomes  $H_i^T Q H_{i+1} = 0$ ; therefore,  $\gamma_i = -H_i^T Q G_{i+1}/H_i^T Q H_i$ . It can be shown in this case that the sequence of vectors  $G_i$  are all mutually orthogonal and the sequence of vectors  $H_i$  are all mutually conjugate with respect to Q. Using these facts, the computation of  $\gamma_i$  may be simplified with the observation that  $\gamma_i = \|G_{i+1}\|^2/\|G_i\|^2$  (Fletcher-Reeves) or  $\gamma_i = (G_{i+1} - G_i)^T G_{i+1}/\|G_i\|^2$  (Polak-Ribière). When f is not quadratic, it is assumed that its second order Taylor expansion sufficiently approximates f in a neighborhood of the minimum, and the  $\gamma_i$  are chosen so that  $H_i$  and  $H_{i+1}$  are conjugate with respect to the matrix  $(\partial^2 f/\partial x^i \partial x^j)(x_{i+1})$  of second partial derivatives of f at  $x_{i+1}$ . It may be desirable to "reset" the algorithm by setting  $H_{i+1} = G_{i+1}$  every rth step (frequently, r = n) because the conjugate gradient method does not, in general, converge in n steps if the

function f is nonquadratic. However, if f is closely approximated by a quadratic function, the reset strategy may be expected to converge rapidly, whereas the unmodified algorithm may not be.

Many of these ideas have straightforward generalizations in the geometry of Riemannian manifolds; several of them have already appeared. We need only make the following definition.

**Definition 5.1.** Given a tensor field  $\omega$  of type (0,2) on M such that for p in M,  $\omega_p: T_p \times T_p \to \mathbf{R}$  is a symmetric bilinear form, the tangent vectors X and Y in  $T_p$  are said to be  $\omega_p$ -conjugate or conjugate with respect to  $\omega_p$  if  $\omega_p(X,Y) = 0$ .

An outline of the conjugate gradient method on Riemannian manifolds may now be given. Let M be an n-dimensional Riemannian manifold with Riemannian structure g and Levi-Civita connection  $\nabla$ , and let  $f \in C^{\infty}(M)$  have a local minimum at  $\hat{p}$ . As in the conjugate gradient method on Euclidean space, choose an initial point  $p_0$  in M and compute the (negative) gradient directions  $G_0 = H_0 = -(\operatorname{grad} f)_{p_0}$  in  $T_{p_0}$ . The recursive part of the algorithm involves minimizing f along the geodesic  $t \mapsto \exp_{p_i} tH_i$ ,  $t \in \mathbf{R}$ , making a step along the geodesic to the minimum point  $p_{i+1} = \exp \lambda_i H_i$ , computing  $G_{i+1} = -(\operatorname{grad} f)_{p_{i+1}}$ , and computing the next direction in  $T_{p_{i+1}}$  for geodesic minimization. This direction is given by the formula

$$H_{i+1} = G_{i+1} + \gamma_i \tau H_i, \tag{25}$$

where  $\tau$  is the parallel translation with respect to the geodesic step from  $p_i$  to  $p_{i+1}$ , and  $\gamma_i$  is chosen such that  $\tau H_i$  and  $H_{i+1}$  are  $(\nabla^2 f)_{p_{i+1}}$ -conjugate, i.e.,

$$\gamma_i = -\frac{(\nabla^2 f)_{p_{i+1}}(\tau H_i, G_{i+1})}{(\nabla^2 f)_{p_{i+1}}(\tau H_i, \tau H_i)}.$$
(26)

Eq. (26) is, in general, expensive to use because the second covariant differential of f appears. However, we can use the Taylor expansion of df about  $p_{i+1}$  to compute an efficient approximation of  $\gamma_i$ . By the fact that  $p_i = \exp_{p_{i+1}}(-\lambda_i \tau H_i)$  and by Eq. (14), we have

$$\tau df_{p_i} = \tau df_{\exp_{p_{i+1}}(-\lambda_i\tau H_i)} = df_{p_{i+1}} - \lambda_i(\nabla^2 f)_{p_{i+1}}(\cdot,\tau H_i) + \text{h.o.t.}$$

Therefore, the numerator of the right hand side of Eq. (26) multiplied by the step size  $\lambda_i$  can be approximated by the equation

$$\lambda_i(\nabla^2 f)_{p_{i+1}}(\tau H_i, G_{i+1}) = df_{p_{i+1}}(G_{i+1}) - (\tau df_{p_i})(G_{i+1})$$
$$= -\langle G_{i+1} - \tau G_i, G_{i+1} \rangle$$

because, by definition,  $G_i = -(\operatorname{grad} f)_{p_i}$ ,  $i = 0, 1, \ldots$ , and for any X in  $T_{p_{i+1}}$ ,  $(\tau df_{p_i})(X) = df_{p_i}(\tau^{-1}X) = \langle (\operatorname{grad} f)_{p_i}, \tau^{-1}X \rangle = \langle \tau(\operatorname{grad} f)_{p_i}, X \rangle$ . Similarly, the denominator of the right hand side of Eq. (26) multiplied by  $\lambda_i$  can be approximated by the equation

$$\lambda_i(\nabla^2 f)_{p_{i+1}}(\tau H_i, \tau H_i) = df_{p_{i+1}}(\tau H_i) - (\tau df_{p_i})(\tau H_i)$$
$$= \langle G_i, H_i \rangle$$

because  $\langle G_{i+1}, \tau H_i \rangle = 0$  by the assumption that f is minimized along the geodesic  $t \mapsto \exp tH_i$  at  $t = \lambda_i$ . Combining these two approximations with Eq. (26), we obtain a formula for  $\gamma_i$  that is relatively inexpensive to compute:

$$\gamma_i = \frac{\langle G_{i+1} - \tau G_i, G_{i+1} \rangle}{\langle G_i, H_i \rangle}.$$
 (27)

Of course, as the connection  $\nabla$  is compatible with the metric g, the denominator of Eq. (27) may be replaced, if desired, by  $\langle \tau G_i, \tau H_i \rangle$ .

The conjugate gradient method may now be presented in full.

**Algorithm 5.2** (Conjugate gradient method). Let M be a complete Riemannian manifold with Riemannian structure g and Levi-Civita connection  $\nabla$ , and let f be a  $C^{\infty}$  function on M.

Step 0. Select  $p_0 \in M$ , compute  $G_0 = H_0 = -(\operatorname{grad} f)_{p_0}$ , and set i = 0.

Step 1. Compute  $\lambda_i$  such that

$$f(\exp_{p_i} \lambda_i H_i) \le f(\exp_{p_i} \lambda H_i)$$

for all  $\lambda \geq 0$ .

Step 2. Set  $p_{i+1} = \exp_{p_i} \lambda_i H_i$ .

Step 3. Set

$$G_{i+1} = -(\operatorname{grad} f)_{p_{i+1}},$$
  

$$H_{i+1} = G_{i+1} + \gamma_i \tau H_i, \qquad \gamma_i = \frac{\langle G_{i+1} - \tau G_i, G_{i+1} \rangle}{\langle G_i, H_i \rangle},$$

where  $\tau$  is the parallel translation with respect to the geodesic from  $p_i$  to  $p_{i+1}$ . If  $i \equiv n-1 \pmod{n}$ , set  $H_{i+1} = G_{i+1}$ . Increment i, and go to Step 1.

**Theorem 5.3.** Let  $f \in C^{\infty}(M)$  have a nondegenerate critical point at  $\hat{p}$  such that the Hessian  $(d^2f)_{\hat{p}}$  is positive definite. Let  $p_i$  be a sequence of points in M generated by Algorithm 5.2 converging to  $\hat{p}$ . Then there exists a constant  $\theta > 0$  and an integer N such that for all  $i \geq N$ ,

$$d(p_{i+n}, \hat{p}) \leq \theta d^2(p_i, \hat{p}).$$

Note that linear convergence is already guaranteed by Theorem 3.3.

**Proof.** If  $p_j = \hat{p}$  for some integer j, the assertion becomes trivial; assume otherwise. Recall that if  $X_1, \ldots, X_n$  is some basis for  $T_{\hat{p}}$ , then the map  $\exp_{\hat{p}}(a^1X_1 + \cdots + a^nX_n) \stackrel{\nu}{\to} (a^1, \ldots, a^n)$  defines a set of normal coordinates at  $\hat{p}$ . Let  $N_{\hat{p}}$  be a normal neighborhood of  $\hat{p}$  on which the normal coordinates  $\nu = (x^1, \ldots, x^n)$  are defined. Consider the map  $\nu_* f \stackrel{\text{def}}{=} f \circ \nu^{-1} \colon \mathbf{R}^n \to \mathbf{R}$ . By the

smoothness of f and exp,  $\nu_* f$  has a critical point at  $0 \in \mathbf{R}^n$  such that the Hessian matrix of  $\nu_* f$  at 0 is positive definite. Indeed, by the fact that  $(d \exp)_0 = \mathrm{id}$ , the ijth component of the Hessian matrix of  $\nu_* f$  at 0 is given by  $(d^2 f)_{\bar{p}}(X_i, X_j)$ .

Therefore, there exists a neighborhood U of  $0 \in \mathbf{R}^n$ , a constant  $\theta' > 0$ , and an integer N, such that for any initial point  $x_0 \in U$ , the conjugate gradient method on Euclidean space (with resets) applied to the function  $\nu_* f$  yields a sequence of points  $x_i$  converging to 0 such that for all  $i \geq N$ ,

$$||x_{i+n}|| \le \theta' ||x_i||^2$$
.

See Polak [36, p. 260ff] for a proof of this fact. Let  $x_0 = \nu(p_0)$  in U be an initial point. Because exp is not an isometry, Algorithm 5.2 yields a different sequence of points in  $\mathbf{R}^n$  than the classical conjugate gradient method on  $\mathbf{R}^n$  (upon equating points in a neighborhood of  $\hat{p} \in M$  with points in a neighborhood of  $0 \in \mathbf{R}^n$  via the normal coordinates).

Nevertheless, the amount by which exp fails to preserve inner products can be quantified via the Gauss Lemma and Jacobi's equation; see, e.g., Cheeger and Ebin [12], or the appendices of Karcher [28]. Let t be small, and let  $X \in T_{\hat{p}}$  and  $Y \in T_{tX}(T_{\hat{p}}) \cong T_{\hat{p}}$  be orthonormal tangent vectors. The amount by which the exponential map changes the length of tangent vectors is approximated by the Taylor expansion

$$||d\exp(tY)||^2 = t^2 - \frac{1}{3}Kt^4 + \text{h.o.t.}$$

where K is the sectional curvature of M along the section in  $T_{\hat{p}}$  spanned by X and Y. Therefore, near  $\hat{p}$  Algorithm 5.2 differs from the conjugate gradient method on  $\mathbf{R}^n$  applied to the function  $\nu_* f$  only by third order and higher terms. Thus both algorithms have the same rate of convergence. The theorem follows.

**Example 5.4** (Rayleigh's quotient on the sphere). Applied to Rayleigh's quotient on the sphere, the conjugate gradient method provides an efficient technique to compute the eigenvectors corresponding to the largest or smallest eigenvalue of a real symmetric matrix. Let  $S^{n-1}$  and  $\rho(x) = x^{\mathrm{T}}Qx$  be as in Examples 3.5 and 4.6. From Algorithm 5.2, we have the following algorithm.

**Algorithm 5.5** (CG method for the extreme eigenvalue/eigenvector). Let Q be a real symmetric n-by-n matrix.

Step 0. Select  $x_0$  in  $\mathbf{R}^n$  such that  $x_0^T x_0 = 1$ , compute  $G_0 = H_0 = (Q - \rho(x_0)I)x_0$ , and set i = 0.

Step 1. Compute c, s, and  $v = 1 - c = s^2/(1+c)$ , such that  $\rho(x_ic + h_is)$  is maximized, where  $c^2 + s^2 = 1$  and  $h_i = H_i/\|H_i\|$ . This can be accomplished by geodesic minimization, or by the formulae

$$c = (\frac{1}{2}(1+b/r))^{\frac{1}{2}}$$
 if  $b \ge 0$ , or  $s = (\frac{1}{2}(1-b/r))^{\frac{1}{2}}$  if  $b \le 0$ ,  $c = a/(2rs)$ 

where  $a = 2x_i^{\mathrm{T}}Qh_i$ ,  $b = x_i^{\mathrm{T}}Qx_i - h_i^{\mathrm{T}}Qh_i$ , and  $r = \sqrt{(a^2 + b^2)}$ .

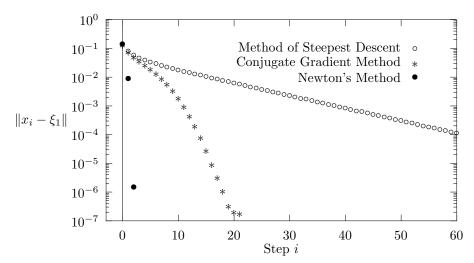


Figure 1: Maximization of Rayleigh's quotient  $x^TQx$  on  $S^{20} \subset \mathbf{R}^{21}$ , where  $Q = \operatorname{diag}(21,\ldots,1)$ . The *i*th iterate is  $x_i$ , and  $\xi_1$  is the eigenvector corresponding to the largest eigenvalue of Q. Algorithm 4.7 was used for Newton's method and Algorithm 5.5 was used for the conjugate gradient method.

Step 2. Set

$$x_{i+1} = x_i c + h_i s$$
,  $\tau H_i = H_i c - x_i || H_i || s$ ,  $\tau G_i = G_i - (h_i^T G_i)(x_i s + h_i v)$ .

Step 3. Set

$$G_{i+1} = (Q - \rho(x_{i+1})I)x_{i+1},$$
  

$$H_{i+1} = G_{i+1} + \gamma_i \tau H_i, \qquad \gamma_i = \frac{(G_{i+1} - \tau G_i)^{\mathrm{T}} G_{i+1}}{G_i^{\mathrm{T}} H_i}.$$

If  $i \equiv n-1 \pmod{n}$ , set  $H_{i+1} = G_{i+1}$ . Increment i, and go to Step 1.

The convergence rate of this algorithm to the eigenvector corresponding to the largest eigenvalue of Q is given by Theorem 5.3. This algorithm costs one matrix-vector multiplication (relatively inexpensive when Q is sparse), one geodesic minimization or computation of  $\rho(h_i)$ , and 10n flops per iteration. The results of a numerical experiment demonstrating the convergence of Algorithm 5.5 on  $S^{20}$  are shown in Figure 1.

Fuhrmann and Liu [20] provide a conjugate gradient algorithm for Rayleigh's quotient on the sphere that uses an azimuthal projection onto tangent planes.

**Example 5.6** (The function  $\operatorname{tr} \Theta^{\mathrm{T}} Q \Theta N$ ). Let  $\Theta$ , Q, and H be as in Examples 3.6 and 4.10. As before, the natural Riemannian structure of SO(n) is used. Let  $X, Y \in \operatorname{\mathbf{so}}(n)$ . The parallel translation of Y along the geodesic  $e^{tX}$  is given by the formula  $\tau Y = L_{e^{tX}} e^{-(t/2)X} Y e^{(t/2)X}$ , where  $L_g$  denotes left translation

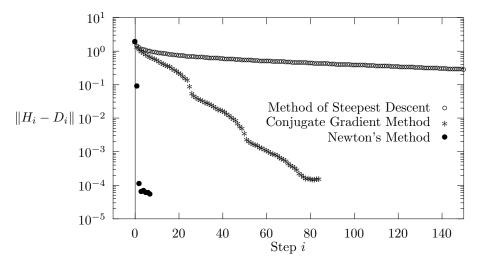


Figure 2: Maximization of  $\operatorname{tr} \Theta^{\mathrm{T}} Q \Theta N$  on SO(20) (dimension SO(20) = 190), where  $N = \operatorname{diag}(20, \ldots, 1)$ . The *i*th iterate is  $H_i = \Theta_i^{\mathrm{T}} Q \Theta_i$ ,  $D_i$  is the diagonal matrix of eigenvalues of  $H_i$ ,  $H_0$  is near N, and  $\|\cdot\|$  is the norm induced by the standard inner product on  $\mathfrak{gl}(n)$ . Geodesics and parallel translation were computed using the algorithm of Ward and Gray [47, 48]; the step sizes for the method of steepest descent and the conjugate gradient method were computed using Brockett's estimate [10].

by g. Brockett's estimate (n.b. Eq. (13)) for the step size may be used in Step 1 of Algorithm 5.2. The results of a numerical experiment demonstrating the convergence of the conjugate gradient method in SO(20) are shown in Figure 2.

Acknowledgments. The author enthusiastically thanks Tony Bloch and the Fields Institute for the invitation to speak at the Fields Institute and for their generous support during his visit. The author also thanks Roger Brockett for his suggestion to investigate conjugate gradient methods on manifolds and for his criticism of this work, and the referee for his helpful suggestions. This work was supported in part by the National Science Foundation under the Engineering Research Center Program NSF D CRD-8803012, the Army Research Office under Grant DAA103-92-G-0164 supporting the Brown, Harvard, and MIT Center for Intelligent Control, and by DARPA under Air Force contract F49620-92-J-0466.

## References

- Bertsekas, D. P. Projected Newton methods for optimization problems with simple constraints, SIAM J. Cont. Opt. 20:221-246, 1982.
- [2] ——. Constrained Optimization and Lagrange Multiplier Methods. New York: Academic Press, 1982.
- [3] BLOCH, A. M., BROCKETT, R. W., and RATIU, T. S. A new formulation of the generalized Toda lattice equations and their fixed point analysis via the momentum map, *Bull. Amer. Math. Soc.* 23 (2): 477–485, 1990.

- [4] ——. Completely integrable gradient flows, Commun. Math. Phys. 147:57-74, 1992.
- [5] BOTSARIS, C. A. Differential gradient methods, J. Math. Anal. Appl. 63:177-198, 1978.
- [6] ——. A class of differential descent methods for constrained optimization, J. Math. Anal. Appl. 79:96–112, 1981.
- [7] ——. Constrained optimization along geodesics, J. Math. Anal. Appl. 79:295–306, 1981.
- [8] BROCKETT, R. W. Least squares matching problems, Lin. Alg. Appl. 122/123/124: 761-777, 1989.
- [9] ——. Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems, Lin. Alq. Appl. 146:79–91, 1991.
- [10] ——. Differential geometry and the design of gradient algorithms, Proc. Symp. Pure Math. R. Green and S. T. Yau, eds. Providence, RI: Amer. Math. Soc., to appear.
- [11] Brown, A. A. and Bartholomew-Biggs, M. C. Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations, J. Optim. Theory Appl. 62 (2): 211–224, 1989.
- [12] CHEEGER, J. and EBIN, D. G. Comparison Theorems in Riemannian Geometry. Amsterdam: North-Holland Publishing Company, 1975.
- [13] CHU, M. T. Curves on S<sup>n-1</sup> that lead to eigenvalues or their means of a matrix, SIAM J. Alq. Disc. Meth. 7 (3): 425-432, 1986.
- [14] CHU, M. T. and DRIESSEL, K. The projected gradient method for least squares matrix approximations with spectral constraints, SIAM J. Numer. Anal. 27 (4):1050–1060, 1990.
- [15] Dunn, J. C. Newton's method and the Goldstein step length rule for constrained minimization problems, SIAM J. Cont. Opt. 18:659-674, 1980.
- [16] ——. Global and asymptotic convergence rate estimates for a class of projected gradient processes, SIAM J. Cont. Opt. 19:368–400, 1981.
- [17] FAYBUSOVICH, L. Hamiltonian structure of dynamical systems which solve linear programming problems, Phys. D 53:217-232, 1991.
- [18] FLETCHER, R. Practical Methods of Optimization, 2d ed. New York: Wiley & Sons, 1987.
- [19] FLETCHER, R. and REEVES, C. M. Function minimization by conjugate gradients, Comput. J. 7 (2): 149–154, 1964.
- [20] FUHRMANN, D. R. and LIU, B. An iterative algorithm for locating the minimal eigenvector of a symmetric matrix, Proc. IEEE ICASSP 84 pp. 45.8.1–4, 1984.
- [21] GILL, P. E. and MURRAY, W. Newton-type methods for linearly constrained optimization, in *Numerical Methods for Constrained Optimization*. P. E. Gill and W. Murray, eds. London: Academic Press, Inc., 1974.
- [22] GOLUB, G. H. and VAN LOAN, C. Matrix Computations. Baltimore, MD: Johns Hopkins University Press, 1983.
- [23] GOLUBITSKY, M. and GUILLEMIN, V. Stable Mappings and Their Singularities. New York: Springer-Verlag, 1973.
- [24] HELGASON, S. Differential Geometry, Lie Groups, and Symmetric Spaces. New York: Academic Press, 1978.
- [25] HELMKE, U. Isospectral flows on symmetric matrices and the Riccati equation, Systems & Control Lett. 16:159–165, 1991.
- [26] HESTENES, M. R. and STIEFEL, E. Methods of conjugate gradients for solving linear systems, J. Res. Nat. Bur. Stand. 49:409-436, 1952.
- [27] HIRSCH, M. W. and SMALE, S. On algorithms for solving f(x) = 0, Comm. Pure Appl. Math. 32:281-312,1979.
- [28] KARCHER, H. Riemannian center of mass and mollifier smoothing, Comm. Pure Appl. Math. 30: 509–541, 1977.
- [29] KOBAYASHI, S. and NOMIZU, K. Foundations of Differential Geometry, Vol. 2. New York: Wiley Interscience Publishers, 1969.
- [30] LAGARIAS, J. C. Monotonicity properties of the Toda flow, the QR-flow, and subspace iteration, SIAM J. Numer. Anal. Appl. 12 (3):449–462, 1991.
- [31] LUENBERGER, D. G. Introduction to Linear and Nonlinear Programming. Reading, MA: Addison-Wesley, 1973.

- [32] MOLER, C. and VAN LOAN, C. Nineteen dubious ways to compute the exponential of a matrix, SIAM Rev. 20 (4): 801–836, 1978.
- [33] Nomizu, K. Invariant affine connections on homogeneous spaces. Amer. J. Math. 76: 33–65, 1954.
- [34] PARLETT, B. The Symmetric Eigenvalue Problem. Englewood Cliffs, NJ: Prentice-Hall, 1980
- [35] PERKINS, J. E., HELMKE, U., and MOORE, J. B. Balanced realizations via gradient flow techniques, Systems & Control Lett. 14:369–380, 1990.
- [36] POLAK, E. Computational Methods in Optimization. New York: Academic Press, 1971.
- [37] Rudin, W. Principles of Mathematical Analysis, 3d ed. New York: McGraw-Hill, 1976.
- [38] SARGENT, R. W. H. Reduced gradient and projection methods for nonlinear programming, in *Numerical Methods for Constrained Optimization*. P. E. Gill and W. Murray, eds. London: Academic Press, Inc., 1974.
- [39] Shub, M. Some remarks on dynamical systems and numerical analysis, in *Dynamical Systems and Partial Differential Equations: Proc. VII ELAM.* L. Lara-Carrero and J. Lewowicz, eds. Caracas: Equinoccio, U. Simón Bolívar, pp. 69–92, 1986.
- [40] SHUB, M. and SMALE, S. Computational complexity: On the geometry of polynomials and a theory of cost, Part I, Ann. scient. Éc. Norm. Sup. 4 (18): 107–142, 1985.
- [41] —. Computational complexity: On the geometry of polynomials and a theory of cost, Part II, SIAM J. Comput. 15 (1):145–161, 1986.
- [42] On the existence of generally convergent algorithms, J. Complex. 2:2–11, 1986.
- [43] SMALE, S. The fundamental theorem of algebra and computational complexity, Bull. Amer. Math. Soc. 4 (1): 1–36, 1981.
- [44] ——. On the efficiency of algorithms in analysis, Bull. Amer. Math. Soc. 13 (2):87–121, 1985.
- [45] SMITH, S. T. Dynamical systems that perform the singular value decomposition, Systems & Control Lett. 16:319–327, 1991.
- [46] SPIVAK, M. A Comprehensive Introduction to Differential Geometry, 2d ed. Vols. 1, 2, 5, Houston, TX: Publish or Perish, Inc., 1979.
- [47] WARD, R. C. and GRAY, L. J. Eigensystem computation for skew-symmetric matrices and a class of symmetric matrices, ACM Trans. Math. Softw. 4(3):278–285, 1978.
- [48] WARD, R. C. and GRAY, L. J. Algorithm 530: An algorithm for computing the eigensystem of skew-symmetric matrices and a class of symmetric matrices, ACM Trans. Math. Softw. 4 (3): 286–289, 1978. See also Collected Algorithms from ACM, Vol. 3. New York: Assoc. Comput. Mach., 1978.