An introduction to optimization on smooth manifolds

Nicolas Boumal

This version compiled March 18, 2022.

A more recent version is available at the following address:
http://www.nicolasboumal.net/book

Contents

Notation

The following lists typical uses of symbols. Local exceptions are documented in place. For example, c typically denotes a curve, but sometimes denotes a real constant. Symbols defined and used locally only are omitted.

$\mathbb{R}, \mathbb{C}$	Real and complex numbers
$\mathbb{R}_+$	Positive reals ($x > 0$)
$\mathbb{R}^{m\times n}$	Real matrices of size $m \times n$
$\mathbb{R}_r^{m\times n}$	Real matrices of size $m \times n$ and rank r
$\mathrm{Sym}(n), \mathrm{Skew}(n)$	Symmetric and skew-symmetric real matrices of size n
$\mathrm{sym}(M), \mathrm{skew}(M)$	Symmetric and skew-symmetric parts of a matrix M
$\mathrm{Sym}(n)^+$	Symmetric positive definite real matrices of size n
$\mathrm{Tr}(M), \det(M)$	Trace, determinant of a square matrix M
$\mathrm{diag}(M)$	Vector of diagonal entries of a matrix M
$\mathrm{diag}(u_1, \ldots, u_n)$	Diagonal matrix of size n with given diagonal entries
$M^\dagger$	Moore–Penrose pseudoinverse of matrix M
I_d	Identity matrix of size d
I	Subset of $\mathbb{R}$ (often open with $0 \in I$) or identity matrix
Id	Identity operator
$\lvert a\rvert$	Modulus of $a \in \mathbb{C}$ (absolute value if $a \in \mathbb{R}$)
$\lvert A\rvert$	Cardinality of a set A
$\mathcal{E}, \mathcal{E}', \mathcal{F}$	Linear spaces, often with a Euclidean structure
$\mathcal{M}, \mathcal{M}', \overline{\mathcal{M}}, \mathcal{N}$	Smooth manifolds, often with a Riemannian structure
S^{d-1}	Unit sphere, in a Euclidean space of dimension d
$\mathrm{OB}(d, n)$	Oblique manifold (product of S^{d-1} copied n times)
$\mathrm{O}(d), \mathrm{SO}(d)$	Orthogonal and special orthogonal groups in $\mathbb{R}^{d\times d}$
$\mathrm{St}(n, p)$	Stiefel manifold embedded in $\mathbb{R}^{n\times p}$
$\mathrm{Gr}(n, p)$	Grassmann manifold as the quotient $\mathrm{St}(n, p)/\mathrm{O}(p)$
$\mathrm{GL}(n)$	General linear group (invertible matrices in $\mathbb{R}^{n\times n}$)
H^n	Hyperbolic space as hyperboloid embedded in $\mathbb{R}^{n+1}$
$\dim \mathcal{M}$	Dimension of $\mathcal{M}$
x, y, z	Points on a manifold
u, v, w, s, ξ, ζ	Tangent vectors
p, q	Integers or polynomials or points on a manifold
$\mathrm{T}\mathcal{M}$	Tangent bundle of $\mathcal{M}$

$\mathrm{T}_x\mathcal{M}$	Tangent space to $\mathcal{M}$ at $x \in \mathcal{M}$
$\mathrm{N}_x\mathcal{M}$	Normal space (orthogonal complement of $\mathrm{T}_x\mathcal{M}$)
$\mathrm{Proj}_x, \mathrm{Proj}_x^\perp$	Orthogonal projector to $\mathrm{T}_x\mathcal{M}$, $\mathrm{N}_x\mathcal{M}$
$\mathrm{H}_x, \mathrm{V}_x$	Horizontal and vertical space at x for a quotient manifold
$\mathrm{Proj}_x^\mathrm{H}, \mathrm{Proj}_x^\mathrm{V}$	Orthogonal projectors to $\mathrm{H}_x, \mathrm{V}_x$
lift_x	Horizontal lift operator for quotient manifolds
$\mathrm{R}_x(v)$	Retraction R evaluated at $(x, v) \in \mathrm{T}\mathcal{M}$
$\mathrm{Exp}_x(v)$	Exponential map Exp evaluated at $(x, v) \in \mathrm{T}\mathcal{M}$
$\mathrm{Log}_x(y)$	Vector v such that $\mathrm{Exp}_x(v) = y$ (see Definition 10.20)
$\exp, \log$	Scalar or matrix exponential and logarithm
$\mathcal{O}, \mathcal{O}_x$	Domain of Exp (subset of $\mathrm{T}\mathcal{M}$), Exp_x (subset of $\mathrm{T}_x\mathcal{M}$); Can also denote these domains for a non-global retraction.
$\mathrm{inj}(\mathcal{M}), \mathrm{inj}(x)$	Injectivity radius of a manifold, at a point
$\langle\cdot,\cdot\rangle_x$	Riemannian inner product on $\mathrm{T}_x\mathcal{M}$
$\langle\cdot,\cdot\rangle$	Euclidean inner product; Sometimes denotes $\langle\cdot,\cdot\rangle_x$ with subscript omitted.
$\\|\cdot\\|, \\|\cdot\\|_x$	Norms associated to $\langle\cdot,\cdot\rangle$ and $\langle\cdot,\cdot\rangle_x$
$\\|\cdot\\|$	Also denotes operator norm for linear maps
$[\cdot,\cdot]$	Lie bracket
∇	Affine connection (often Riemannian) on a manifold
$\frac{\mathrm{d}}{\mathrm{d}t}$	Classical derivative with respect to t
$\frac{\mathrm{D}}{\mathrm{d}t}$	Covariant derivative induced by a connection ∇
$\frac{\partial}{\partial x_i}$	Partial derivative with respect to real variable x_i
x_i	Often the ith coordinate of a vector $x \in \mathbb{R}^n$
x_k	Often the kth element of a sequence $x_0, x_1, x_2, \ldots \in \mathcal{M}$
f, g	Real-valued functions
f_{low}	A real number such that $f(x) \geq f_{\mathrm{low}}$ for all x
h	Often a local defining function with values in $\mathbb{R}^k$
$\mathrm{grad} f, \mathrm{Hess} f$	Riemannian gradient and Hessian of f; Euclidean gradient, Hessian if domain of f is Euclidean.
$\nabla f, \nabla^2 f$	First and second covariant derivatives of f as tensor fields
c, γ	Curves
$c', \gamma', \dot{c}, \dot{\gamma}$	Velocity vector fields of curves c, γ
c'', γ''	Intrinsic acceleration vector fields of c, γ
$\ddot{c}, \ddot{\gamma}$	Extrinsic acceleration vector fields of c, γ
L, L_g, L_H	Lipschitz constants (nonnegative reals)
$L(c)$	Length of a curve c
$\mathrm{dist}(x, y)$	Distance (often Riemannian) between two points x, y
$B(x, r)$	Open ball $\{v \in \mathrm{T}_x\mathcal{M} : \\|v\\|_x < r\}$ or $\{y \in \mathcal{E} : \\|y - x\\| < r\}$
$\bar{B}(x, r)$	Closed ball as above
$\mathcal{A}, \mathcal{L}$	Linear maps
$\mathcal{A}, \mathcal{A}^+$	Atlas, maximal atlas
$\mathrm{im}\,\mathcal{L}, \ker\mathcal{L}$	Range space (image) and null space (kernel) of $\mathcal{L}$

$\mathrm{rank}(M), \mathrm{rank}(\mathcal{L})$	Rank of a matrix or linear map
$M^\top, M^*$	Transpose or Hermitian conjugate-transpose of matrix M
$\mathcal{L}^*$	Adjoint of a linear map $\mathcal{L}$ between Euclidean spaces
$\mathcal{A} \succeq 0, \mathcal{A} \succ 0$	States $\mathcal{A} = \mathcal{A}^*$ is positive semidefinite or positive definite
$\mathrm{span}(u_1, \ldots, u_m)$	Linear subspace spanned by vectors $u_1, \ldots, u_m$
F, G, H	Maps, usually to and from linear spaces or manifolds
$F \colon A \to B$	A map defined on the whole domain A
$F\vert_U$	Restriction of the map F to the domain U
$F(\cdot, y)$	For a map $(x, y) \mapsto F(x, y)$, this is the map $x \mapsto F(x, y)$
$F \circ G$	Composition of maps: $(F \circ G)(x) = F(G(x))$
$\mathrm{D}F(x)[v]$	Differential of F at x along $[v]$
U, V, W, X, Y, Z	Vector fields on a manifold, or
	Matrices which could be tangent vectors or points on $\mathcal{M}$;
Y, Z	Can also be vector fields along a curve.
U, V, W, O	Can also be open sets, usually in a linear space.
$\bar{f}, \bar{F}, \bar{V}, \ldots$	Smooth extensions or lifts of $f, F, V, \ldots$
$\bar{u}, \bar{U}$	Can also denote complex conjugation of u, U
T	Tensor field
T_s	Differential of retraction $\mathrm{DR}_x(s)$
$\mathcal{U}, \mathcal{V}$	Open sets in a manifold
$\mathcal{W}$	Weingarten map
II	Second fundamental form
fV	Vector field $x \mapsto f(x)V(x)$ (with f real valued)
Vf	Real function $x \mapsto \mathrm{D}f(x)[V(x)]$
$\mathfrak{F}(\mathcal{M})$	Set of smooth real-valued functions on $\mathcal{M}$
$\mathfrak{X}(\mathcal{M})$	Set of smooth vector fields on $\mathcal{M}$
$\mathfrak{X}(c)$	Set of smooth vector fields along a curve c
$\mathrm{T}_{y \leftarrow x}$	Vector transport from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_y\mathcal{M}$
$\mathrm{PT}^c_{t_1 \leftarrow t_0}$	Parallel transport along c from $c(t_0)$ to $c(t_1)$
P_s	Parallel transport along $\gamma(t) = \mathrm{Exp}_x(ts)$ from 0 to 1
$\sim$	Equivalence relation
$A/\sim$	Quotient set of A by the relation $\sim$
$[x]$	Equivalence class of x for some equivalence relation
π	Canonical projection $\pi \colon \mathrm{T}\mathcal{M} \to \mathcal{M}$ or $\pi \colon \overline{\mathcal{M}} \to \overline{\mathcal{M}}/\sim$;
	Occasionally denotes the mathematical constant.
$A \subset B$	A is a proper subset of B (the sets are not equal)
$A \subseteq B$	A is a subset of B (the sets may be equal)
$A \cap B$	Intersection of sets A, B
$A \cup B$	Union of sets A, B
$\emptyset$	Empty set.

Preface

Optimization problems on smooth manifolds arise in science and engineering as a result of natural geometry (e.g., the set of orientations of physical objects in space is a manifold), latent data simplicity (e.g., high-dimensional data points lie close to a low-dimensional linear subspace, leading to low-rank data matrices), symmetry (e.g., observations are invariant under rotation, translation or other group actions, leading to quotients) and positivity (e.g., covariance matrices and diffusion tensors are positive definite). This has led to successful applications notably in machine learning, computer vision, robotics, scientific computing, dynamical systems and signal processing.

Accordingly, optimization on manifolds has garnered increasing interest from researchers and engineers alike. Building on fifty years of research efforts that have recently intensified, it is now recognized as a wide, beautiful and effective generalization of unconstrained optimization on linear spaces.

Yet, engineering programs seldom include training in differential geometry: the field of mathematics concerned with smooth manifolds. Moreover, existing textbooks on this topic usually align with the interests of mathematicians more than with the needs of engineers and applied mathematicians. This creates a significant but avoidable barrier to entry for optimizers.

One of my goals in writing this book is to offer a different, if at times unorthodox, introduction to differential geometry. Definitions and tools are introduced in a need-based order for optimization. We start with a restricted setting—that of embedded submanifolds of linear spaces—which allows us to define all necessary concepts in direct reference to their usual counterparts from linear spaces. This covers a wealth of applications.

In what is perhaps the clearest departure from standard exposition, charts and atlases are not introduced until quite late. The reason for doing so is twofold: pedagogically, charts and atlases are more abstract than what is needed to work on embedded submanifolds; and pragmatically, charts are seldom if ever useful in practice. It would be unfortunate to give them center stage.

Of course, charts and atlases are the right tool to provide a unified treatment of all smooth manifolds in an intrinsic way. They are introduced eventually, at which point it becomes possible to discuss quotient manifolds: a powerful language to understand symmetry in optimization. Perhaps this abstraction is necessary to fully appreciate the depth of optimization on manifolds as more than just a fancy tool for constrained optimization in linear spaces, and truly a mathematically natural setting for *unconstrained* optimization in a wider sense.

Time-tested optimization algorithms are introduced immediately after the early chapters about embedded geometry. Crucially, the design and analysis of these methods remain unchanged whether we are optimizing on a manifold which is embedded in a linear space or not. This makes it possible to get to algorithms early on, without sacrificing generality. It also underlines the conceptual point that the algorithms truly operate on the manifolds intrinsically.

The last two chapters visit more advanced topics that are not typically necessary for simple applications. The first one delves deeper into geometric tools. The second one introduces the basics of *geodesic* convexity: a broad generalization of convexity, which is one of the most fruitful structures in classical optimization.

Intended audience

This book is intended for students and researchers alike. The material has proved popular with applied mathematicians and mathematically inclined engineering and computer science students at the graduate and advanced undergraduate levels.

Readers are assumed to be comfortable with linear algebra and multivariable calculus. Central to the raison d'être of this book, there are no prerequisites in differential geometry or optimization. For computational aspects, it is helpful to have notions of numerical linear algebra, for which I recommend the approachable textbook by Trefethen and Bau [TB97].

Building on these expectations, the aim is to give full proofs and intuition for all concepts that are introduced, at least for submanifolds of linear spaces. The hope is to equip readers to pursue research projects in (or using) optimization on manifolds, involving both mathematical analysis and efficient implementation.

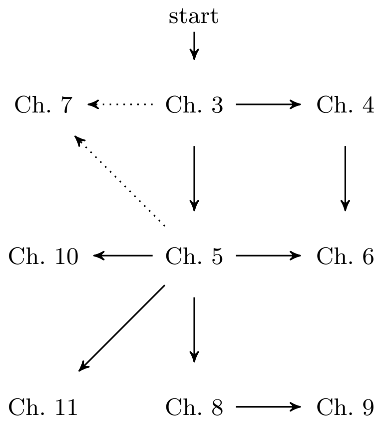
How to use this book

The book is self-contained and should suit both self-learners and instructors.

Chapters 3 and 5 can serve as a standalone introduction to differential and Riemannian geometry. They focus on embedded submanifolds of linear spaces, with proofs. Chapter 7 details examples of manifolds: it is meant for on-and-off reading in parallel with Chapters 3 and 5. These chapters do not involve charts, and they aim to convey the fact that geometric tools are computational tools.

From there, the expected next step is to work through Chapters 4 and 6 about optimization algorithms. Readers may also choose to embark on Chapter 8 to see how embedded manifolds fit into the general theory of smooth manifolds. That is a useful (though not fully necessary) stepping stone toward Chapter 9 about quotient manifolds. Alternatively, they may decide to learn about further geometric tools in Chapter 10 or about a Riemannian notion of convexity in Chapter 11.

These chapter dependencies are summarized in the diagram below, where an arrow from A to B means it is preferable to read A before B.



1. Introduction
2. Applications
3. First-order geometry
4. First-order optimization
5. Second-order geometry
6. Second-order optimization
7. Examples of manifolds
8. General manifolds
9. Quotient manifolds
10. Additional geometric tools
11. Geodesic convexity

In a graduate course at Princeton University in 2019 and 2020 (24 lectures of 80 minutes each), I covered much of Chapters 1–6 and select parts of Chapter 7 before the midterm break, then much of Chapters 8–9 and select parts of Chapters 10–11 after the break. At EPFL in 2021, I discussed mostly Chapters 1–8 in 13 lectures of 90 minutes supplemented with exercise sessions.

The numerous exercises in the book have wide ranging difficulty levels. Some are included in part as a way to convey information while skipping technicalities.

Starred sections can be skipped safely for a first encounter with the material. Chapters end with references and notes that many readers may find relevant but ⋆ which would otherwise break the flow. Did the mark in the margin catch your attention? That is its purpose. You may see a couple of those in the book.

What is new, or different, or hard to find elsewhere

The de facto reference for optimization on manifolds is the landmark 2008 book *Optimization algorithms on matrix manifolds* by Pierre-Antoine Absil, Robert Mahony and Rodolphe Sepulchre [AMS08]. It is an important source for the present book as well, with significant overlap of topics. In the years since, the field has evolved, and with it the need for an entry point catering to a broader audience. In an effort to address these needs, I aim to:

1. **Provide a different, self-contained introduction to the core concepts.** This includes a “charts last” take on differential geometry with proofs adapted accordingly; a somewhat unusual yet equivalent definition of connections that (I believe) is more intuitive from an optimizer’s point of view; and an account of optimization on quotient manifolds which benefits from years of hindsight. This introduction is informed by the pains I had entering the field.

2. **Discuss new topics that have grown in importance since 2008.** This includes a replacement of asymptotic convergence results in favor of worst-case, non-asymptotic iteration complexity results; a related take on Lipschitz continuity for Riemannian gradients and Hessians paired with their effect on Taylor expansions on manifolds; an explicit construction of geometric tools necessary for optimization over matrices of fixed rank; a simple study of metric projection retractions; an extrinsic view of the Riemannian Hessian for submanifolds through the Weingarten map and second fundamental form; a discussion of the smooth invertibility of retractions and of the domain of the inverse of the exponential map; transporters as a natural alternative to vector and parallel transports; finite differences of gradients to approximate Hessians; and an introduction to geodesic convexity (not restricted to Hadamard manifolds) with a gradient algorithm for the strongly convex case. Many of these build on research papers referenced in text.
3. **Share tricks of the trade that are seldom if ever spelled out.** This includes several examples of manifolds worked out in full detail; pragmatic instructions for how to derive expressions for gradients and Hessians of matrix functions, and how to check them numerically; explicit formulas for geometric tools on product manifolds (mostly given as exercises); and a number of comments informed by ten years of software development in the field.

The main differential geometry references I used are the fantastic books by Lee [Lee12, Lee18], O'Neill [O'N83], and Brickell and Clark [BC70]. Definitions of geometric concepts in this book, though at times stated differently, are fully compatible with Absil et al.'s book. This is also compatible with Lee's textbooks with one exception: Riemannian *sub*manifolds to us are understood to be embedded submanifolds, whereas Lee also allows them to be merely immersed submanifolds. Moreover, we use the word "manifold" to mean "smooth manifold," that is, C^∞. Most results extend to manifolds and functions of class C^k.

There is much to say about the impact of curvature on optimization. This is an active research topic that has not stabilized yet. Therefore, I chose to omit curvature entirely from this book, save for a few brief comments in the last two chapters. Likewise, optimization on manifolds is proving to be a particularly fertile ground for benign non-convexity and related phenomena. There are only a few hints to that effect throughout the book: the research continues.

Software and online resources

Little to no space is devoted to existing software packages for optimization on manifolds, or to numerical experiments. Yet, such packages significantly speed up research and development in the field. The reader may want to experiment with Manopt (Matlab), PyManopt (Python) or Manopt.jl (Julia), all available from manopt.org.

In particular, the Matlab implementations of most manifolds discussed in this book are listed in Table 7.1 on p156. Gradient descent (Algorithm 4.1) with backtracking line-search (Algorithm 4.2) is available as `steepestdescent`. The trust-region method (Algorithm 6.3) with the truncated conjugate gradient sub-problem solver (Algorithm 6.4) is available as `trustregions`. These implementations include a wealth of tweaks and tricks that are important in practice: many are explained here, some are only documented in the code. The Python and Julia versions offer similar features.

The following webpage collects further resources related to this book:

nicolasboumal.net/book

In particular, errata and potential video lectures will be listed there.

Thanks

A number of people offered decisive comments. I thank Pierre-Antoine Absil (who taught me much of this) and Rodolphe Sepulchre for their input at the early stages of planning for this book, as well as (in no particular order) Eitan Levin, Chris Criscitiello, Quentin Rebjock, Razvan-Octavian Radu, Joe Kileel, Bart Vandereycken, Bamdev Mishra, Suvrit Sra, Stephen McKeown, John M. Lee and Sándor Z. Németh for numerous conversations that led to direct improvements. I am also indebted to the mathematics departments at Princeton University and EPFL for supporting me while I was writing. Finally, I thank Katie Leach at Cambridge University Press for her enthusiasm and candid advice that helped shape this project into its final form.

Nicolas Boumal Lausanne, Switzerland, March 2022
`nicolas.boumal@epfl.ch`

1 Introduction

Optimization is a staple of mathematical modeling. In this rich framework, we consider a set S called the *search space*—it contains all possible answers to our problem, good and bad—and a *cost function* $f \colon S \to \mathbb{R}$ which associates a cost $f(x)$ to each element x of S. The goal is to find $x \in S$ such that $f(x)$ is as small as possible: a best answer. We write

$$\min_{x \in S} f(x)$$

to represent both the optimization problem and the minimal cost (if it exists). Occasionally, we wish to denote specifically the subset of S for which the minimal cost is attained; the standard notation is:

$$\arg\min_{x \in S} f(x),$$

bearing in mind that this set might be empty. We will discuss a few simple applications which can be modeled in this form.

Rarely, optimization problems admit an analytical solution. Typically, we need numerical algorithms to (try to) solve them. Often, the best algorithms exploit mathematical structure in S and f.

An important special case arises when S is a linear space such as $\mathbb{R}^n$. Minimizing a function f in $\mathbb{R}^n$ is called *unconstrained optimization* because the variable x is free to move around $\mathbb{R}^n$, unrestricted.

If f is sufficiently differentiable and $\mathbb{R}^n$ is endowed with an inner product (that is, if we make it into a Euclidean space), then we have a notion of gradient and perhaps even a notion of Hessian for f. These objects give us a firm understanding of how f behaves locally around any given point. Famous algorithms such as gradient descent and Newton's method exploit these objects to move around $\mathbb{R}^n$ efficiently in search of a solution.

Notice however that the Euclidean structure of $\mathbb{R}^n$ and the smoothness of f are irrelevant to the definition of the optimization problem itself: they are merely structures that we may (and as experience shows: we should) use algorithmically to our advantage.

Subsuming linearity, we focus on *smoothness* as the key structure to exploit: we assume the set S is a *smooth manifold* and the function f is smooth on S. This calls for precise definitions, constructed first in Chapter 3. For a first intuition, one can think of smooth manifolds as surfaces in $\mathbb{R}^n$ that do not have kinks or boundaries, such as a plane, a sphere, a torus, or a hyperboloid for example.

We could think of optimization over such surfaces as *constrained*, in the sense that x is not allowed to move freely in $\mathbb{R}^n$: it is constrained to remain on the surface. Alternatively, and this is the viewpoint favored here, we can think of this as unconstrained optimization, in a world where the smooth surface is the only thing that exists: like an ant walking on a large ball might feel unrestricted in its movements, aware only of the sphere it lives on; or like the two-dimensional inhabitants of Flatland [Abb84] find it hard to imagine that there exists such a thing as a third dimension, feeling thoroughly free in their own subspace.

A natural question then is: can we generalize the Euclidean algorithms from unconstrained optimization to handle the broader class of optimization over smooth manifolds? The answer is essentially yes, going back to the 70s [Lue72, Lic79], the 80s [Gab82] and the 90s [Udr94, Smi94, HM96, Rap97, EAS98] and sparking a significant amount of research in the past two decades.

To generalize algorithms such as gradient descent and Newton's method, we need a proper notion of gradient and Hessian on smooth manifolds. In the linear case, this required the introduction of an inner product: a Euclidean structure. In our more general setting, we leverage the fact that smooth manifolds can be linearized locally around every point. The linearization at x is called the *tangent space* at x. By endowing each tangent space with its own inner product (varying smoothly with x, in a sense to be made precise), we construct what is called a *Riemannian structure* on the manifold: it becomes a *Riemannian manifold*.

A Riemannian structure is sufficient to define gradients and Hessians on the manifold, paving the way for optimization. There exist several Riemannian structures on each manifold: our choice may impact algorithmic performance. In that sense, identifying a useful structure is part of the algorithm design—as opposed to being part of the problem formulation, which ended with the definition of the search space (as a crude set) and the cost function.

Chapter 2 covers a few simple applications, mostly to give a sense of how manifolds come up. We then go on to define smooth manifolds in a restricted[1] setting in Chapter 3, where manifolds are *embedded* in a linear space, much like the unit sphere in three-dimensional space. In this context, we define notions of smooth functions, smooth vector fields, gradients and *retractions* (a means to move around on a manifold). These tools are sufficient to design and analyze a first optimization algorithm in Chapter 4: Riemannian gradient descent. As readers progress through these chapters, it is the intention that they also read bits of Chapter 7 from time to time: useful embedded manifolds are studied there in detail. Chapter 5 provides more advanced geometric tools for embedded manifolds, including the notions of Riemannian *connections* and Hessians. These

[1] Some readers may know Whitney's celebrated embedding theorems, which state that any smooth manifold can be embedded in a linear space [BC70, p82]. The mere existence of an embedding, however, is of little use for computation.

are put to good use in Chapter 6 to design and analyze Riemannian versions of Newton's method and the trust-region method.

The linear *embedding space* is useful for intuition, to simplify definitions, and to design tools. Notwithstanding, all the tools and concepts we define in the restricted setting are *intrinsic*, in the sense that they are well defined regardless of the embedding space. We make this precise much later, in Chapter 8, where all the tools from Chapters 3 and 5 are redefined in the full generality of standard treatments of differential geometry. This is also the time to discuss topological issues to some extent. Generality notably makes it possible to discuss a more abstract class of manifolds called *quotient manifolds* in Chapter 9. They offer a beautiful way to harness symmetry, so common in applications.

In closing, Chapter 10 offers a limited treatment of more advanced geometric tools such as the Riemannian distance, geodesics, the exponential map and its inverse, parallel transports and transporters, notions of Lipschitz continuity, finite differences, and covariant differentiation of tensor fields. Then, Chapter 11 covers elementary notions of convexity on Riemannian manifolds with simple implications for optimization. This topic has been around since the 90s, and has been gaining traction in research lately.

More than 150 years ago, Riemann invented a new kind of geometry for the abstract purpose of understanding curvature in high-dimensional spaces. Today, this geometry plays a central role in the development of efficient algorithms to tackle technological applications Riemann himself—arguably—could have never envisioned. Through this book, I would like to invite you to enjoy this singularly satisfying success of mathematics, with an eye to turn geometry into algorithms.

2 Simple examples

Before formally defining what manifolds are, and before introducing any particular algorithms, this chapter surveys simple problems that are naturally modeled as optimization on manifolds. These problems are motivated by applications in various scientific and technological domains. We introduce them chiefly to illustrate how manifolds arise and to motivate the mathematical abstractions in subsequent chapters.

The first two examples lead to optimization on linear spaces: they fall within the scope of optimization on manifolds, but one can also handle them with more traditional tools. Subsequently, we encounter optimization on spheres, products of spheres, orthonormal matrices, the set of all linear subspaces, rotation matrices, fixed-rank matrices, positive definite matrices and certain quadratic surfaces. Through those, we get a glimpse of the wide reach of optimization on manifolds.

Below, we use a few standard concepts from linear algebra and calculus that are revisited in Section 3.1.

2.1 Logistic regression: optimization on a linear space

Given a large number of images, determine automatically which ones contain the digit 7, and which do not. Given numerical data about many patients, determine which are at risk of a certain health hazard, and which are not. Given word counts and other textual statistics from a large number of e-mails, identify which ones are spam. In all cases, examples are elements of a linear space $\mathcal{E}$ (for grayscale images of $n \times n$ pixels, $\mathcal{E} = \mathbb{R}^{n\times n}$; for other scenarios, $\mathcal{E} = \mathbb{R}^n$) and each example takes on one of two possible labels: 0 or 1, corresponding to "seven or not," "at risk or safe," "spam or legitimate." These tasks are called *binary classification*.

One popular technique to address binary classification is *logistic regression*. Equip the linear space $\mathcal{E}$ with an inner product $\langle \cdot, \cdot \rangle$. The model (the assumption) is that there exists an element $\theta \in \mathcal{E}$ such that, given an example $x \in \mathcal{E}$, the probability that its label y is either 0 or 1 is given as follows:

$$\begin{aligned} &\mathrm{Prob}[y = 1|x, \theta] = \sigma(\langle \theta, x \rangle), \text{ and} \\ &\mathrm{Prob}[y = 0|x, \theta] = 1 - \sigma(\langle \theta, x \rangle) = \sigma(- \langle \theta, x \rangle), \end{aligned}$$

where $\sigma\colon \mathbb{R} \to (0,1)$ is the *logistic function* (a sigmoid function):

$$\sigma(z) = \frac{1}{1+e^{-z}}.$$

We can rewrite both identities in one, as a function of $y \in \{0,1\}$:

$$\mathrm{Prob}[y|x,\theta] = \sigma(\langle \theta, x\rangle)^y \sigma(-\langle \theta, x\rangle)^{1-y}.$$

In other words, if we have θ (and if the logistic model is accurate...), then we can easily compute the probability that a new example x belongs to either class. The task is to learn θ from labeled examples.

Let $x_1, \ldots, x_m \in \mathcal{E}$ be given examples, and let $y_1, \ldots, y_m \in \{0,1\}$ be their labels. For any candidate θ, we can evaluate how compatible it is with the data under our model by evaluating the corresponding *likelihood function*. Assuming the observed examples are independent, the likelihood of θ is:

$$L(\theta) = \prod_{i=1}^m \mathrm{Prob}[y_i|x_i,\theta] = \prod_{i=1}^m \sigma(\langle \theta, x_i\rangle)^{y_i} \sigma(-\langle \theta, x_i\rangle)^{1-y_i}.$$

Intuitively, if $L(\theta)$ is large, then θ is doing a good job at classifying the examples. The maximizer of the function over $\theta \in \mathcal{E}$ is the *maximum likelihood estimator* for θ. Equivalently, we may maximize the logarithm of L, known as the *log-likelihood*. Still equivalently, it is traditional to minimize the negative of the log-likelihood:

$$\ell(\theta) = -\log(L(\theta)) = -\sum_{i=1}^m y_i \log(\sigma(\langle \theta, x_i\rangle)) + (1-y_i)\log(\sigma(-\langle \theta, x_i\rangle)).$$

For reasons that we do not get into, it is important to penalize θs that are too large, for example in the sense of the norm induced by the inner product: $\|\theta\| = \sqrt{\langle \theta, \theta\rangle}$. The competing desires between attaining a small negative log-likelihood and a small norm for θ are balanced with a regularization weight $\lambda > 0$ to be chosen by the user. Overall, logistic regression comes down to solving the following optimization problem:

$$\min_{\theta \in \mathcal{E}} \ell(\theta) + \lambda\|\theta\|^2.$$

This problem falls within our framework as the search space $\mathcal{E}$ is a *linear manifold*: admittedly the simplest example of a manifold. The cost function $f(\theta) = \ell(\theta) + \lambda\|\theta\|^2$ is smooth. Furthermore, it is *convex*: a highly desirable property, but a rare occurrence in our framework.

In closing, we note that, to address possible centering issues in the data, it is standard to augment the logistic model slightly, as:

$$\mathrm{Prob}[y = 1|x,\theta] = \sigma(\theta_0 + \langle \theta, x\rangle),$$

with $\theta_0 \in \mathbb{R}$ and $\theta \in \mathcal{E}$. This can be accommodated seamlessly in the above notation by replacing the linear space with $\mathbb{R} \times \mathcal{E}$, examples with $(1, x_i)$, the parameter with (θ_0, θ), and adapting the metric accordingly.

2.2 Sensor network localization from directions: a linear subspace

Consider n sensors located at unknown positions $t_1, \ldots, t_n$ in $\mathbb{R}^d$. We aim to locate the sensors, that is, estimate the positions t_i, based on some directional measurements. Specifically, for some pairs of sensors $(i, j) \in G$, we receive a noisy direction measurement from t_i to t_j:

$$v_{ij} \approx \frac{t_i - t_j}{\|t_i - t_j\|},$$

where $\|x\| = \sqrt{x_1^2 + \cdots + x_d^2}$ is the Euclidean norm on $\mathbb{R}^d$ induced by the inner product $\langle u, v\rangle = u^\top v = u_1 v_1 + \cdots + u_d v_d$.

There are two fundamental ambiguities in this task. First, directional measurements reveal nothing about the global location of the sensors: translating the sensors as a whole does not affect pairwise directions. Thus, we may assume without loss of generality that the sensors are centered:

$$t_1 + \cdots + t_n = 0.$$

Second, the measurements reveal nothing about the global scale of the sensor arrangement. Specifically, scaling all positions t_i by a scalar $\alpha > 0$ as αt_i has no effect on the directions separating the sensors, so that the true scale cannot be recovered from the measurements. It is thus legitimate to fix the scale in some arbitrary way. One fruitful way is to assume the following [HLV18]:

$$\sum_{(i,j)\in G} \langle t_i - t_j, v_{ij}\rangle = 1.$$

Given a tentative estimator $\hat{t}_1, \ldots, \hat{t}_n \in \mathbb{R}^d$ for the locations, we may assess its compatibility with the measurement v_{ij} by computing

$$\|(\hat{t}_i - \hat{t}_j) - \left\langle \hat{t}_i - \hat{t}_j, v_{ij}\right\rangle v_{ij}\|.$$

Indeed, if $\hat{t}_i - \hat{t}_j$ and v_{ij} are aligned in the same direction, this evaluates to zero. Otherwise, it evaluates to a positive number, growing as alignment degrades. Combined with the symmetry-breaking conditions, this suggests the following formulation for sensor network localization from pairwise direction measurements:

$$\min_{\hat{t}_1,\ldots,\hat{t}_n\in\mathbb{R}^d} \sum_{(i,j)\in G} \|(\hat{t}_i - \hat{t}_j) - \left\langle \hat{t}_i - \hat{t}_j, v_{ij}\right\rangle v_{ij}\|^2$$
$$\text{subject to } \ \hat{t}_1 + \cdots + \hat{t}_n = 0 \ \text{ and } \sum_{(i,j)\in G} \left\langle \hat{t}_i - \hat{t}_j, v_{ij}\right\rangle = 1.$$

The role of the affine constraint is clear: it excludes $\hat{t}_1 = \cdots = \hat{t}_n = 0$, which would otherwise be optimal.

Grouping the variables as the columns of a matrix, we find that the search space for this problem is an affine subspace of $\mathbb{R}^{d\times n}$: this too is a linear manifold. It is also an *embedded submanifold* of $\mathbb{R}^{d\times n}$. Hence, it falls within our framework.

With the simple cost function as above, this problem is in fact a convex quadratic minimization problem on an affine subspace. As such, it admits an explicit solution which merely requires solving a linear system. Optimization algorithms can be used to solve this system implicitly. More importantly, the power of optimization algorithms lies in the flexibility that they offer: alternative cost functions may be used to improve robustness against specific noise models for example [HLV18].

2.3 Single extreme eigenvalue or singular value: spheres

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix: $A = A^\top$. By the spectral theorem, A admits n real eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$ and corresponding real, orthonormal eigenvectors $v_1, \ldots, v_n \in \mathbb{R}^n$, where orthonormality is assessed with respect to the standard inner product over $\mathbb{R}^n$: $\langle u, v \rangle = u_1 v_1 + \cdots + u_n v_n$.

For now, we focus on computing one extreme eigenpair of A: (λ_1, v_1) or (λ_n, v_n) will do. Let $\mathbb{R}^n_*$ denote the set of nonzero vectors in $\mathbb{R}^n$. It is well known that the *Rayleigh quotient*,

$$r\colon \mathbb{R}^n_* \to \mathbb{R}\colon x \mapsto r(x) = \frac{\langle x, Ax \rangle}{\langle x, x \rangle},$$

attains its extreme values when x is aligned with $\pm v_1$ or $\pm v_n$, and that the corresponding value of the quotient is λ_1 or λ_n. We will rediscover such properties through the prism of optimization on manifolds as a running example in this book.

Say we are interested in the smallest eigenvalue, λ_1. Then, we must solve the following optimization problem:

$$\min_{x \in \mathbb{R}^n_*} \frac{\langle x, Ax \rangle}{\langle x, x \rangle}.$$

The set $\mathbb{R}^n_*$ is open in $\mathbb{R}^n$: it is an *open submanifold* of $\mathbb{R}^n$. Optimization over an open set has its challenges (more on this later). Fortunately, we can easily circumvent these issues in this instance.

Since the Rayleigh quotient is invariant under scaling, that is, since $r(\alpha x) = r(x)$ for all nonzero real α, we may fix the scale arbitrarily. Given the denominator of r, one particularly convenient way is to restrict our attention to unit-norm vectors: $\|x\|^2 = \langle x, x \rangle = 1$. The set of such vectors is the *unit sphere* in $\mathbb{R}^n$:

$$\mathrm{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}.$$

This is an *embedded submanifold* of $\mathbb{R}^n$. Our problem becomes:

$$\min_{x \in \mathrm{S}^{n-1}} \langle x, Ax \rangle. \tag{2.1}$$

This is perhaps the simplest non-trivial instance of an optimization problem on a manifold: we use it recurringly to illustrate concepts as they occur.

Similarly to the above, we may compute the largest singular value of a matrix $M \in \mathbb{R}^{m\times n}$ together with associated left- and right-singular vectors by solving

$$\max_{x\in \mathrm{S}^{m-1}, y\in \mathrm{S}^{n-1}} \langle x, My\rangle . \tag{2.2}$$

This is the basis of *principal component analysis*: see also below. The search space is a Cartesian product of two spheres: this too is a manifold; specifically, an embedded submanifold of $\mathbb{R}^m \times \mathbb{R}^n$. In general:

Products of manifolds are manifolds.

This is an immensely useful property.

2.4 Dictionary learning: products of spheres

JPEG and its more recent version JPEG 2000 are some of the most commonly used compression standards for photographs. At their core, these algorithms rely on basis expansions: discrete cosine transforms for JPEG, and wavelet transforms for JPEG 2000. That is, an image (or rather, each patch of the image) is written as a linear combination of a fixed collection of basis images. To fix notation, say an image is represented as a vector $y \in \mathbb{R}^d$ (its pixels rearranged into a single column vector) and the basis images are $b_1, \ldots, b_d \in \mathbb{R}^d$ (each of unit norm). There exists a unique set of coordinates $c \in \mathbb{R}^d$ such that:

$$y = c_1 b_1 + \cdots + c_d b_d.$$

Since the basis images are fixed (and known to anyone creating or reading image files), it is equivalent to store y or c.

The basis is designed carefully with two goals in mind. First, the transform between y and c should be fast to compute (one good starting point to that effect is orthogonality). Second, images encountered in practice should lead to many of the coefficients c_i being zero, or close to zero. Indeed, to recover y, it is only necessary to record the nonzero coefficients. To compress further, we may also decide not to store the small coefficients: if so, y can still be reconstructed approximately. Beyond compression, another benefit of sparse expansions is that they can reveal structural information about the contents of the image, which in turn may be beneficial for tasks such as classification.

In *dictionary learning*, we focus on the second goal. As a key departure from the above, the idea here is not to design a basis by hand, but rather to learn a good basis from data automatically. This way, we may exploit structural properties of images that come up in a particular application. For example, it may be the case that photographs of faces can be expressed more sparsely in a dedicated basis, compared to a standard wavelet basis. Pushing this idea further, we relax the requirement of identifying a basis, instead allowing ourselves to pick more than d images for our expansions. The collection of images $b_1, \ldots, b_n \in \mathbb{R}^d$ forms a *dictionary*. Its elements are called *atoms*, and they normally span $\mathbb{R}^d$ in an overcomplete way, meaning any image y can be expanded into a linear combination of atoms in more than one way. The aim is that at least one of these expansions should be sparse, or have many small coefficients. For the magnitudes of coefficients to be meaningful, we further require all atoms to have the same norm: $\|b_i\| = 1$ for all i.

Thus, given a collection of k images $y_1, \ldots, y_k \in \mathbb{R}^d$, the task in dictionary learning is to find a dictionary $b_1, \ldots, b_n \in \mathbb{R}^d$ such that (as much as possible) each image y_i is a sparse linear combination of atoms. Collect the input images as the columns of a data matrix $Y \in \mathbb{R}^{d \times k}$, and the atoms into a matrix $D \in \mathbb{R}^{d \times n}$ (to be determined). Expansion coefficients for the images in this dictionary form the columns of a matrix $C \in \mathbb{R}^{n \times k}$ so that

$$Y = DC.$$

Typically, many choices of C are possible. We aim to pick D such that there exists a valid (or approximately valid) choice of C with numerous zeros. Let $\|C\|_0$ denote the number of entries of C different from zero. Then, one possible formulation of dictionary learning balances both aims with a parameter $\lambda > 0$ as (with $b_1, \ldots, b_n$ the columns of D):

$$\begin{aligned} &\min_{D \in \mathbb{R}^{d \times n}, C \in \mathbb{R}^{n \times k}} \|Y - DC\|^2 + \lambda \|C\|_0 \\ &\text{subject to } \|b_1\| = \cdots = \|b_n\| = 1. \end{aligned} \tag{2.3}$$

The matrix norm $\|\cdot\|$ is the Frobenius norm, induced by the standard inner product $\langle U, V \rangle = \mathrm{Tr}(U^\top V)$.

Evidently, allowing the dictionary to be overcomplete ($n > d$) helps sparsity. An extreme case is to set $n = k$, in which case an optimal solution consists in letting D be Y with normalized columns. Then, each image can be expressed with a single nonzero coefficient (C is diagonal). This is useless of course, if only because both parties of the communication must have access to the (possibly huge) dictionary, and because this choice may generalize poorly when presented with new images. Interesting scenarios involve n much smaller than k.

The search space in D is a product of several spheres, which is an embedded submanifold of $\mathbb{R}^{d \times n}$ called the *oblique manifold*:

$$\mathrm{OB}(d, n) = (\mathrm{S}^{d-1})^n = \left\{ X \in \mathbb{R}^{d \times n} : \mathrm{diag}(X^\top X) = \mathbf{1} \right\},$$

where $\mathbf{1} \in \mathbb{R}^n$ is the all-ones vector and $\mathrm{diag} \colon \mathbb{R}^{n \times n} \to \mathbb{R}^n$ extracts the diagonal entries of a matrix. The search space in C is the linear manifold $\mathbb{R}^{n \times k}$. Overall, the search space of the dictionary learning optimization problem is

$$\mathrm{OB}(d, n) \times \mathbb{R}^{n \times k},$$

which is an embedded submanifold of $\mathbb{R}^{d \times n} \times \mathbb{R}^{n \times k}$.

We note in closing that the cost function in (2.3) is discontinuous because of the term $\|C\|_0$, making it hard to optimize. A standard reformulation replaces the culprit with $\|C\|_1$: the sum of absolute values of the entries of C. This is continuous but nonsmooth. A possible further step then is to smooth the cost function, for example exploiting that $|x| \approx \sqrt{x^2 + \varepsilon^2}$ or $|x| \approx \varepsilon \log(e^{x/\varepsilon} + e^{-x/\varepsilon})$ for small $\varepsilon > 0$.

Regardless of changes to the cost function, the manifold $\mathrm{OB}(d, n)$ is nonconvex, so that finding a global optimum for dictionary learning as stated above is challenging: see work by Sun et al. [SQW17] for some guarantees.

2.5 Principal component analysis: Stiefel and Grassmann

Let $x_1, \ldots, x_n \in \mathbb{R}^d$ represent a large collection of centered data points in a d-dimensional linear space. We may think of it as a cloud of points. It is often the case that this cloud lies on or near a low-dimensional subspace of $\mathbb{R}^d$, and it may be distributed anisotropically in that subspace, meaning it shows more variation along some directions than others. One of the pillars of data analysis is to determine the main directions of variation of the data, under the name of *principal component analysis* (PCA).

One way to think of a main direction of variation, called a *principal component*, is as a vector $u \in \mathrm{S}^{d-1}$ such that projecting the data points to the one-dimensional subspace spanned by u preserves most of the variance. Specifically, let $X \in \mathbb{R}^{d \times n}$ be the matrix whose columns are the data points and let $uu^\top$ be the orthogonal projector to the span of u. We wish to maximize the following for $u \in \mathrm{S}^{d-1}$:

$$\sum_{i=1}^{n} \|uu^\top x_i\|^2 = \|uu^\top X\|^2 = \langle uu^\top X, uu^\top X \rangle = \langle XX^\top u, u \rangle.$$

We recognize the Rayleigh quotient of $XX^\top$ to be maximized over S^{d-1}. Of course, an optimal solution is given by a dominant eigenvector of $XX^\top$, or equivalently by a dominant left singular vector of X.

Let $u_1 \in \mathrm{S}^{d-1}$ be a principal component. Targeting a second one, we aim to find $u_2 \in \mathrm{S}^{d-1}$, *orthogonal to* u_1, such that projecting the data to the subspace spanned by u_1 and u_2 preserves the most variance. The orthogonal projector to that subspace is $u_1 u_1^\top + u_2 u_2^\top$. We maximize:

$$\|(u_1 u_1^\top + u_2 u_2^\top) X\|^2 = \langle XX^\top u_1, u_1 \rangle + \langle XX^\top u_2, u_2 \rangle,$$

over $u_2 \in \mathrm{S}^{d-1}$ with $u_2^\top u_1 = 0$. This search space for u_2 is an embedded submanifold of $\mathbb{R}^d$.

It is often more convenient to optimize for u_1 and u_2 simultaneously rather than sequentially. Then, since the above cost function is symmetric in u_1 and u_2, as is the constraint $u_2^\top u_1 = 0$, we add weights to the two terms to ensure u_1 captures a principal component and u_2 captures a second principal component:

$$\max_{u_1, u_2 \in \mathrm{S}^{d-1}, u_2^\top u_1 = 0} \alpha_1 \langle XX^\top u_1, u_1 \rangle + \alpha_2 \langle XX^\top u_2, u_2 \rangle,$$

with $\alpha_1 > \alpha_2 > 0$ arbitrary.

More generally, aiming for k principal components, we look for a matrix $U \in \mathbb{R}^{d\times k}$ with k orthonormal columns $u_1, \ldots, u_k \in \mathbb{R}^d$. The set of such matrices is called the *Stiefel manifold*:

$$\mathrm{St}(d, k) = \{U \in \mathbb{R}^{d\times k} : U^\top U = I_k\},$$

where I_k is the identity matrix of size k. It is an embedded submanifold of $\mathbb{R}^{d\times k}$. The orthogonal projector to the subspace spanned by the columns of U is $UU^\top$. Hence, PCA amounts to solving the problem:

$$\max_{U\in\mathrm{St}(d,k)} \sum_{i=1}^{k} \alpha_i \langle XX^\top u_i, u_i\rangle = \max_{U\in\mathrm{St}(d,k)} \langle XX^\top U, UD\rangle, \tag{2.4}$$

where $D \in \mathbb{R}^{k\times k}$ is diagonal with diagonal entries $\alpha_1 > \cdots > \alpha_k > 0$.

It is well known that collecting k top eigenvectors of $XX^\top$ (or, equivalently, k top left singular vectors of X) yields a global optimum of (2.4), meaning this optimization problem can be solved efficiently using tools from numerical linear algebra. Still, the optimization perspective offers significant flexibility that standard linear algebra algorithms cannot match. Specifically, within an optimization framework, it is possible to revisit the variance criterion by changing the cost function. This allows one to promote sparsity or robustness against outliers, for example to develop variants such as sparse PCA [dBEG08, JNRS10] and robust PCA [MT11, GZAL14, MZL19, NNSS20]. There may also be computational advantages, for example in tracking and online models, where the dataset changes or grows with time: it may be cheaper to update a previous good estimator using few optimization steps than to run a complete eigenvalue or singular value decomposition anew.

If the top k principal components are of interest but their ordering is not, then we do not need the weight matrix D. In this scenario, we are seeking an orthonormal basis U for a k dimensional subspace of $\mathbb{R}^d$ such that projecting the data to that subspace preserves as much of the variance as possible. This description makes it clear that the particular basis is irrelevant: only the selected subspace matters. This is apparent in the cost function,

$$f(U) = \langle XX^\top U, U\rangle,$$

which is invariant under orthogonal transformations. Specifically, for all Q in the orthogonal group

$$\mathrm{O}(k) = \{Q \in \mathbb{R}^{k\times k} : Q^\top Q = I_k\}$$

we have $f(UQ) = f(U)$. This induces an *equivalence relation*[1] $\sim$ on the Stiefel

[1] Recall that an equivalence relation $\sim$ on a set M is a reflexive ($a \sim a$), symmetric ($a \sim b \iff b \sim a$) and transitive ($a \sim b$ and $b \sim c \implies a \sim c$) binary relation. The equivalence class $[a]$ is the set of elements of M that are equivalent to a. Each element of M belongs to exactly one equivalence class.

manifold:

$$U \sim V \qquad \iff \qquad V = UQ \text{ for some } Q \in \mathrm{O}(k).$$

This equivalence relation partitions $\mathrm{St}(d,k)$ into *equivalence classes*:

$$[U] = \{V \in \mathrm{St}(d,k) : U \sim V\} = \{UQ : Q \in \mathrm{O}(k)\}.$$

The set of equivalence classes is called the *quotient set*:

$$\mathrm{St}(d,k)/\!\sim\; = \mathrm{St}(d,k)/\mathrm{O}(k) = \{[U] : U \in \mathrm{St}(d,k)\}.$$

Importantly, $U, V \in \mathrm{St}(d,k)$ are equivalent if and only if their columns span the same subspace of $\mathbb{R}^d$. In other words: the quotient set is in one-to-one correspondence with the set of subspaces of dimension k in $\mathbb{R}^d$. With the right geometry, the latter is called the *Grassmann manifold*:

$$\mathrm{Gr}(d,k) = \{ \text{ subspaces of dimension } k \text{ in } \mathbb{R}^d \} \equiv \mathrm{St}(d,k)/\mathrm{O}(k),$$

where the symbol $\equiv$ reads "is equivalent to" (context indicates in what sense). As defined here, the Grassmann manifold is a *quotient manifold*. This type of manifold is more abstract than embedded submanifolds, but we can still develop numerically efficient tools to work with them.

Within our framework, computing the dominant eigenspace of dimension k of the matrix $XX^\top$ can be written as:

$$\max_{[U] \in \mathrm{Gr}(d,k)} \langle XX^\top U, U \rangle.$$

The cost function is well defined over $\mathrm{Gr}(d,k)$ since it depends only on the equivalence class of U, not on U itself.

In analogy with (2.4), for a matrix M in $\mathbb{R}^{m\times n}$, k top left and right singular vectors form a solution of the following problem on a product of Stiefel manifolds:

$$\max_{U \in \mathrm{St}(m,k), V \in \mathrm{St}(n,k)} \langle MV, UD \rangle,$$

where D is a diagonal weight matrix as above.

A book by Trendafilov and Gallo provides more in depth discussion of applications of optimization on manifolds to data analysis [TG21].

2.6 Synchronization of rotations: special orthogonal group

In structure from motion (SfM), the 3D structure of an object is to be reconstructed from several 2D images of it. For example, in the paper *Building Rome in a day* [ASS+09], the authors automatically construct a model of the Colosseum from over 2000 photographs freely available on the Internet. Because the pictures are acquired from an unstructured source, one of the steps in the reconstruction pipeline is to estimate camera locations and *pose*. The pose of a camera is its orientation in space.

In single particle reconstruction through cryo electron microscopy, an electron microscope is used to produce 2D tomographic projections of biological objects such as proteins and viruses. Because shape is a determining factor of function, the goal is to estimate the 3D structure of the object from these projections. Contrary to X-ray crystallography (another fundamental tool of structural biology), the orientations of the objects in the individual projections are unknown. In order to estimate the structure, a useful step is to estimate the individual orientations (though we note that high noise levels do not always allow it, in which case alternative statistical techniques must be used.)

Mathematically, orientations correspond to rotations of $\mathbb{R}^3$. Rotations in $\mathbb{R}^d$ can be represented with orthogonal matrices:

$$\mathrm{SO}(d) = \{R \in \mathbb{R}^{d\times d} : R^\top R = I_d \text{ and } \det(R) = +1\}.$$

The determinant condition excludes reflections of $\mathbb{R}^d$. The set $\mathrm{SO}(d)$ is the *special orthogonal group*: it is both a group (in the mathematical sense of the term) and a manifold (an embedded submanifold of $\mathbb{R}^{d\times d}$)—it is a *Lie group*.

In both applications described above, similar images or projections can be compared to estimate relative orientations. *Synchronization of rotations* is a mathematical abstraction of the ensuing task: it consists in estimating n individual rotation matrices,

$$R_1, \ldots, R_n \in \mathrm{SO}(d),$$

from pairwise relative rotation measurements: for some pairs $(i,j) \in G$, we observe a noisy version of $R_i R_j^{-1}$. Let $H_{ij} \in \mathrm{SO}(d)$ denote such a measurement. Then, one possible formulation of synchronization of rotations is:

$$\min_{\hat{R}_1,\ldots,\hat{R}_n \in \mathrm{SO}(d)} \sum_{(i,j)\in G} \|\hat{R}_i - H_{ij}\hat{R}_j\|^2.$$

This is an optimization problem over $\mathrm{SO}(d)^n$, which is a manifold.

This also comes up in *simultaneous localization and mapping* (SLAM), whereby a robot must simultaneously map its environment and locate itself in it as it moves around [RDTEL21]. An important aspect of SLAM is to keep track of the robot's orientation accurately, by integrating all previously acquired information to correct estimator drift.

2.7 Low-rank matrix completion: fixed-rank manifold

Let $M \in \mathbb{R}^{m\times n}$ be a large matrix of interest. Given some of its entries, our task is to estimate the whole matrix. A commonly cited application for this setup is that of recommender systems, where row i corresponds to a user, column j corresponds to an item (a movie, a book...) and entry M_{ij} indicates how much user i appreciates item j: positive values indicate appreciation, zero is neutral, and negative values indicate dislike. The known entries may be collected from user interactions. Typically, most entries are unobserved. Predicting the missing values may be helpful to automate personalized recommendations.

Of course, without further knowledge about how the entries of the matrix are related, the completion task is ill-posed. Hope comes from the fact that certain users share similar traits, so that what one user likes may be informative about what another, similar user may like. In the same spirit, certain items may be similar enough that whole groups of users may feel similarly about them. One mathematically convenient way to capture this idea is to assume M has (approximately) low rank. The rationale is as follows: if M has rank r, then it can be factored as

$$M = LR^\top,$$

where $L \in \mathbb{R}^{m\times r}$ and $R \in \mathbb{R}^{n\times r}$ are full-rank factor matrices. Row i of L, ℓ_i, attributes r numbers to user i, while the jth row of R, r_j, attributes r numbers to item j. Under the low-rank model, the rating of user i for item j is $M_{ij} = \langle \ell_i, r_j\rangle$. One interpretation is that there are r latent features (these could be movie genres for example): a user has some positive or negative appreciation for each feature, and an item has traits aligned with or in opposition to these features; the rating is obtained as the inner product of the two feature vectors.

Under this model, predicting the user ratings for all items amounts to *low-rank matrix completion*. Let Ω denote the set of pairs (i, j) such that M_{ij} is observed. Allowing for noise in the observations and inaccuracies in the model, we aim to solve

$$\begin{aligned} \min_{X\in\mathbb{R}^{m\times n}} \; & \sum_{(i,j)\in\Omega} (X_{ij} - M_{ij})^2 \\ \text{subject to } \; & \operatorname{rank}(X) = r. \end{aligned}$$

The search space for this optimization problem is the set of matrices of a given size and rank:

$$\mathbb{R}_r^{m\times n} = \{X \in \mathbb{R}^{m\times n} : \operatorname{rank}(X) = r\}.$$

This set is an embedded submanifold of $\mathbb{R}^{m\times n}$, frequently useful in machine learning applications.

Another use for this manifold is solving high-dimensional matrix equations that may come up in systems and control applications: aiming for a low-rank solution may be warranted in certain settings, and exploiting this can lower the computational burden substantially. Yet another context where optimization over low-rank matrices occurs is in completing and denoising approximately separable bivariate functions based on sampled values [Van10, Van13, MV13].

The same set can also be endowed with other geometries, that is, it can be made into a manifold in other ways. For example, exploiting the factored form more directly, note that any matrix in $\mathbb{R}_r^{m\times n}$ admits a factorization as $LR^\top$ with both L and R of full rank r. This correspondence is not one-to-one however, since the pairs (L, R) and $(LJ^{-1}, RJ^\top)$ map to the same matrix in $\mathbb{R}_r^{m\times n}$ for all invertible matrices J: they are equivalent. Similarly to the Grassmann manifold, this leads to a definition of $\mathbb{R}_r^{m\times n}$ as a quotient manifold instead of an embedded submanifold. Many variations on this theme are possible, some of them more useful than others depending on the application [Mey11, Mis14].

The set $\mathbb{R}_r^{m\times n}$ is not closed in $\mathbb{R}^{m\times n}$, which may create difficulties for optimization. The closure of the set corresponds to all matrices of rank at most r (rather than exactly equal to r). That set is not a manifold, but it can be smoothly parameterized by a manifold in several ways [LKB21].

2.8 Gaussian mixture models: positive definite matrices

A common model in machine learning assumes data $x_1, \ldots, x_n \in \mathbb{R}^d$ are sampled independently from a *mixture of K Gaussians*, that is, each data point is sampled from a probability distribution with density of the form

$$f(x) = \sum_{k=1}^{K} w_k \frac{1}{\sqrt{2\pi \det(\Sigma_k)}} e^{-\frac{(x-\mu_k)^\top \Sigma_k^{-1}(x-\mu_k)}{2}},$$

where the centers $\mu_1, \ldots, \mu_K \in \mathbb{R}^d$, covariances $\Sigma_1, \ldots, \Sigma_K \in \mathrm{Sym}(d)^+$ and mixing probabilities $(w_1, \ldots, w_K) \in \Delta_+^{K-1}$ are to be determined. We use the following notation:

$$\mathrm{Sym}(d)^+ = \{X \in \mathbb{R}^{d\times d} : X = X^\top \text{ and } X \succ 0\}$$

for symmetric, positive definite matrices of size d, and

$$\Delta_+^{K-1} = \{w \in \mathbb{R}^K : w_1, \ldots, w_K > 0 \text{ and } w_1 + \cdots + w_K = 1\}$$

for the positive part of the simplex, that is, the set of non-vanishing discrete probability distributions over K objects. In this model, with probability w_k, a point x is sampled from the kth Gaussian, with mean μ_k and covariance Σ_k. The aim is only to estimate the parameters, not to estimate which Gaussian each point x_i was sampled from.

For a given set of observations $x_1, \ldots, x_n$, a maximum likelihood estimator solves:

$$\max_{\substack{\hat\mu_1,\ldots,\hat\mu_K \in \mathbb{R}^d, \\ \hat\Sigma_1,\ldots,\hat\Sigma_K \in \mathrm{Sym}(d)^+, \\ w \in \Delta_+^{K-1}}} \sum_{i=1}^{n} \log\left(\sum_{k=1}^{K} w_k \frac{1}{\sqrt{2\pi \det(\Sigma_k)}} e^{-\frac{(x_i-\mu_k)^\top \Sigma_k^{-1}(x_i-\mu_k)}{2}} \right). \tag{2.5}$$

This is an optimization problem over $\mathbb{R}^{d\times K} \times (\mathrm{Sym}(d)^+)^K \times \Delta_+^{K-1}$, which can be made into a manifold because $\mathrm{Sym}(d)^+$ and Δ_+^{K-1} can be given a manifold structure.

The direct formulation of maximum likelihood estimation for Gaussian mixture models in (2.5) is however not computationally favorable. See [HS15] for a beneficial reformulation, still on a manifold.

2.9 Smooth semidefinite programs

Semidefinite programs (SDPs) are optimization problems of the form

$$\min_{X \in \mathrm{Sym}(n)} \langle C, X \rangle \qquad \text{subject to} \qquad \mathcal{A}(X) = b \text{ and } X \succeq 0, \tag{2.6}$$

where $\mathrm{Sym}(n)$ is the space of real, symmetric matrices of size $n \times n$, $\langle A, B \rangle = \mathrm{Tr}(A^\top B)$ and $\mathcal{A} \colon \mathrm{Sym}(n) \to \mathbb{R}^m$ is a linear map defined by m symmetric matrices $A_1, \ldots, A_m$ as $\mathcal{A}(X)_i = \langle A_i, X \rangle$.

SDPs are convex and they can be solved to global optimality in polynomial time using interior point methods. Still, handling the positive semidefiniteness constraint $X \succeq 0$ and the dimensionality of the problem (namely, the $\frac{n(n+1)}{2}$ variables required to define X) both pose significant computational challenges for large n.

A popular way to address both issues is the Burer–Monteiro approach [BM05], which consists in factorizing X as $X = YY^\top$ with $Y \in \mathbb{R}^{n \times p}$: the number p of columns of Y is a parameter. Notice that X is now automatically positive semidefinite. If $p \geq n$, the SDP can be rewritten equivalently as

$$\min_{Y \in \mathbb{R}^{n \times p}} \langle CY, Y \rangle \qquad \text{subject to} \qquad \mathcal{A}(YY^\top) = b. \tag{2.7}$$

If $p < n$, this corresponds to the SDP with the additional constraint $\mathrm{rank}(X) \leq p$. There is a computational advantage to taking p as small as possible. Interestingly, if the set of matrices X that are feasible for the SDP is compact, then the *Pataki–Barvinok bound* [Pat98, Bar95] provides that at least one of the global optimizers of the SDP has rank r such that $\frac{r(r+1)}{2} \leq m$. In other words: assuming compactness, the Burer–Monteiro formulation is *equivalent* to the original SDP so long as p satisfies $\frac{p(p+1)}{2} \geq m$. This is already the case for $p = O(\sqrt{m})$, which may be significantly smaller than n.

The positive semidefiniteness constraint disappeared, and the dimensionality of the problem went from $O(n^2)$ to np—a potentially appreciable gain. Yet, we lost something important along the way: the Burer–Monteiro problem is not convex. It is not immediately clear how to solve it.

The search space of the Burer–Monteiro problem is the set of feasible points Y:

$$\mathcal{M} = \{Y \in \mathbb{R}^{n \times p} : \mathcal{A}(YY^\top) = b\}. \tag{2.8}$$

Assume the map $h(Y) = \mathcal{A}(YY^\top)$ has the property that its differential at all Y in $\mathcal{M}$ has rank m. Then, $\mathcal{M}$ is a smooth manifold embedded in $\mathbb{R}^{n \times p}$. In this special case, we may try to solve the Burer–Monteiro problem through optimization over that manifold. It turns out that non-convexity is mostly benign in that scenario, in a precise sense [BVB19]:

If $\mathcal{M}$ is compact and $\frac{p(p+1)}{2} > m$, then for a generic cost matrix C the smooth optimization problem $\min_{Y \in \mathcal{M}} \langle CY, Y \rangle$ has no spurious local minima, in the sense *that any point Y which satisfies first- and second-order necessary optimality conditions is a global optimum.*

(Necessary optimality conditions are detailed in Sections 4.2 and 6.1.) Additionally, these global optima map to global optima of the SDP through $X = YY^\top$. This suggests that smooth-and-compact SDPs may be solved to global optimality via optimization on manifolds. The requirement that $\mathcal{M}$ be a regularly defined smooth manifold is not innocuous, but it is satisfied in a number of interesting applications.

There has been a lot of work on this front in recent years, including the early work by Burer and Monteiro [BM03, BM05], the first manifold-inspired perspective by Journée et al. [JBAS10], qualifications of the benign non-convexity at the Pataki–Barvinok threshold [BVB16, BVB19] and below in special cases [BBV16], a proof that p cannot be set much lower than that threshold in general [WW20], smoothed analyses to assess whether points which satisfy necessary optimality conditions approximately are also approximately optimal [BBJN18, PJB18, CM19] and extensions to accommodate scenarios where $\mathcal{M}$ is not a smooth manifold but, more generally, a real algebraic variety [BBJN18, Cif19]. See all these references for applications, including Max-Cut, community detection, the trust-region subproblem, synchronization of rotations and more.

3 Embedded submanifolds: first-order geometry

Our goal is to develop optimization algorithms to solve problems of the form

$$\min_{x \in \mathcal{M}} f(x), \tag{3.1}$$

where $\mathcal{M}$ is a smooth, possibly nonlinear space, and $f: \mathcal{M} \to \mathbb{R}$ is a smooth cost function. In order to do so, our first task is to clarify what we mean by a "smooth space," and a "smooth function" on such a space. Then, we need to develop any tools required to construct optimization algorithms in this setting. Let us start with a bird's-eye view of what this entails.

For smoothness of $\mathcal{M}$, our model space is the unit sphere in $\mathbb{R}^d$:

$$\mathrm{S}^{d-1} = \{x \in \mathbb{R}^d : x^\top x = 1\}. \tag{3.2}$$

Intuitively, we think of S^{d-1} as a smooth nonlinear space in $\mathbb{R}^d$. Our definitions below are compatible with this intuition: we call S^{d-1} an *embedded submanifold* of $\mathbb{R}^d$.

An important element in these definitions is to capture the idea that S^{d-1} can be locally approximated by a linear space around any point x: we call these *tangent spaces*, denoted by $\mathrm{T}_x\mathrm{S}^{d-1}$. This is as opposed to a cube for which no good linearization exists at the edges. More specifically for our example, S^{d-1} is defined by the constraint $x^\top x = 1$. We may expect that differentiating this constraint should yield a suitable linearization, and indeed it does:

$$\mathrm{T}_x\mathrm{S}^{d-1} = \{v \in \mathbb{R}^d : v^\top x + x^\top v = 0\} = \{v \in \mathbb{R}^d : x^\top v = 0\}. \tag{3.3}$$

In the same spirit, it stands to reason that linear spaces and open subsets of linear spaces should also be considered smooth.

Regarding smoothness of functions, we may expect that functions $f: \mathrm{S}^{d-1} \to \mathbb{R}$ obtained by restriction to S^{d-1} of smooth functions on $\mathbb{R}^d$ (smooth in the usual sense for functions from $\mathbb{R}^d$ to $\mathbb{R}$, that is, infinitely differentiable) ought to be considered smooth. We adopt (essentially) this as our definition.

In this early chapter, we give a restricted definition of smoothness, focusing on embedded submanifolds. This allows us to build our initial toolbox more rapidly, and is sufficient to handle many applications. We extend our perspective to the general framework later on, in Chapter 8.

To get started with a list of required tools, it is useful to review briefly the main ingredients of optimization on a *linear* space $\mathcal{E}$:

$$\min_{x \in \mathcal{E}} f(x). \tag{3.4}$$

For example, $\mathcal{E} = \mathbb{R}^d$ or $\mathcal{E} = \mathbb{R}^{n \times p}$. Perhaps the most fundamental algorithm to address this class of problems is *gradient descent*, also known as *steepest descent*. Given an initial guess $x_0 \in \mathcal{E}$, this algorithm produces *iterates* on $\mathcal{E}$ (a sequence of points on $\mathcal{E}$) as follows:[1]

$$x_{k+1} = x_k - \alpha_k \mathrm{grad} f(x_k), \qquad\qquad k = 0, 1, 2, \ldots \tag{3.5}$$

where the $\alpha_k > 0$ are aptly chosen step-sizes and $\mathrm{grad} f \colon \mathcal{E} \to \mathcal{E}$ is the gradient of f. Under mild assumptions, the accumulation points of the sequence $x_0, x_1, x_2, \ldots$ have relevant properties for the optimization problem (3.4). We study these later, in Chapter 4.

From this discussion, we can identify a list of desiderata for a geometric toolbox, meant to solve

$$\min_{x \in \mathrm{S}^{d-1}} f(x) \tag{3.6}$$

with some smooth function f on the sphere. The most pressing point is to find an alternative for the implicit use of linearity in (3.5). Indeed, above, both x_k and $\mathrm{grad} f(x_k)$ are elements of $\mathcal{E}$. Since $\mathcal{E}$ is a linear space, they can be combined with linear operations. Putting aside for now the issue of defining a proper notion of gradient for a function f on S^{d-1}, we must still contend with the issue that S^{d-1} is *not* a linear space: we have no notion of linear combination here.

Alternatively, we can reinterpret iteration (3.5) and say:

To produce $x_{k+1} \in \mathrm{S}^{d-1}$, move away from x_k along the direction $-\mathrm{grad} f(x_k)$ over some distance, while remaining on S^{d-1}.

Surely, if the purpose is to remain on S^{d-1}, it would make little sense to move radially away from the sphere. Rather, using the notion that smooth spaces can be linearized around x by a tangent space $\mathrm{T}_x\mathrm{S}^{d-1}$, we only consider moving along directions in $\mathrm{T}_x\mathrm{S}^{d-1}$. To this end, we introduce the concept of *retraction* at x: a map $\mathrm{R}_x \colon \mathrm{T}_x\mathrm{S}^{d-1} \to \mathrm{S}^{d-1}$ which allows us to move away from x smoothly along a prescribed tangent direction while remaining on the sphere. One suggestion might be as follows, with $\|u\| = \sqrt{u^\top u}$:

$$\mathrm{R}_x(v) = \frac{x + v}{\|x + v\|}. \tag{3.7}$$

In this chapter, we give definitions that allow for this natural proposal.

It remains to make sense of the notion of gradient for a function on a smooth, nonlinear space. Once more, we take inspiration from the linear case. For a smooth function $f \colon \mathcal{E} \to \mathbb{R}$, the gradient is defined with respect to an *inner*

[1] Here, x_k designates an element in a sequence $x_0, x_1, \ldots$ Sometimes, we also use subscript notation such as x_i to select the ith entry of a column vector x. Context tells which is meant.

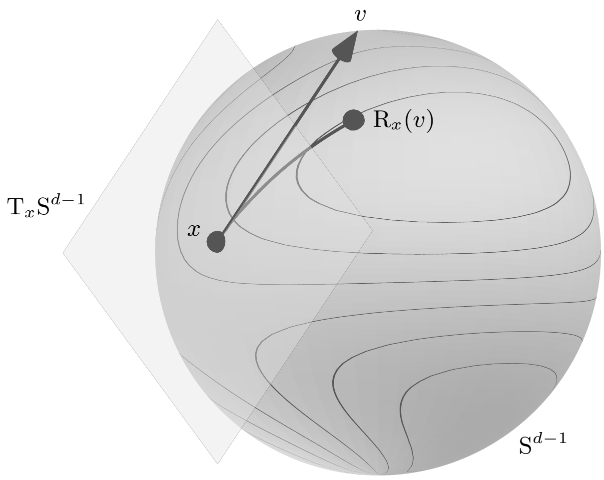


Figure 3.1 Retraction $\mathrm{R}_x(v) = \frac{x+v}{\|x+v\|}$ on the sphere.

product $\langle\cdot,\cdot\rangle : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$ (see Definition 3.1 below for a reminder): $\mathrm{grad} f(x)$ is the unique element of $\mathcal{E}$ such that, for all $v \in \mathcal{E}$,

$$\mathrm{D}f(x)[v] = \langle v, \mathrm{grad} f(x)\rangle\,, \tag{3.8}$$

where $\mathrm{D}f(x) \colon \mathcal{E} \to \mathbb{R}$ is the differential of f at x, that is, it is the linear map defined by:

$$\mathrm{D}f(x)[v] = \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t}. \tag{3.9}$$

Crucially, the gradient of f depends on a choice of inner product (while the differential of f does not).

For example, on $\mathcal{E} = \mathbb{R}^d$ equipped with the standard inner product

$$\langle u, v\rangle = u^\top v \tag{3.10}$$

and the canonical basis $e_1, \ldots, e_d \in \mathbb{R}^d$ (the columns of the identity matrix), the ith coordinate of $\mathrm{grad} f(x) \in \mathbb{R}^d$ is given by

$$\begin{aligned} \mathrm{grad} f(x)_i = \langle e_i, \mathrm{grad} f(x)\rangle &= \mathrm{D}f(x)[e_i] \\ &= \lim_{t\to 0} \frac{f(x+te_i) - f(x)}{t} \triangleq \frac{\partial f}{\partial x_i}(x), \end{aligned} \tag{3.11}$$

that is, the ith partial derivative of f as a function of $x_1, \ldots, x_d \in \mathbb{R}$. This covers a case so common that it is sometimes presented as the definition of the gradient: $\mathrm{grad} f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_d} \end{bmatrix}^\top$.

Turning to our nonlinear example again, in order to define a proper notion of gradient for $f \colon \mathrm{S}^{d-1} \to \mathbb{R}$, we find that we need to (a) provide a meaningful notion of differential $\mathrm{D}f(x)\colon \mathrm{T}_x\mathrm{S}^{d-1} \to \mathbb{R}$, and (b) introduce inner products on the tangent spaces of S^{d-1}. Let us focus on the latter for this outline.

Since $\mathrm{T}_x\mathrm{S}^{d-1}$ is a different linear subspace for various $x \in \mathrm{S}^{d-1}$, we need a different inner product for each point: $\langle \cdot, \cdot \rangle_x$ denotes our choice of inner product on $\mathrm{T}_x\mathrm{S}^{d-1}$. If this choice of inner products varies smoothly with x (in a sense we make precise below), then we call it a *Riemannian metric*, and S^{d-1} equipped with this metric is called a *Riemannian manifold*. This allows us to define the *Riemannian gradient* of f on S^{d-1}: $\mathrm{grad} f(x)$ is the unique tangent vector at x such that, for all $v \in \mathrm{T}_x\mathrm{S}^{d-1}$,

$$\mathrm{D}f(x)[v] = \langle v, \mathrm{grad} f(x) \rangle_x .$$

Thus, first we choose a Riemannian metric, then a notion of gradient ensues.

One arguably natural way of endowing S^{d-1} with a metric is to exploit the fact that each tangent space $\mathrm{T}_x\mathrm{S}^{d-1}$ is a linear subspace of $\mathbb{R}^d$, hence we may define $\langle \cdot, \cdot \rangle_x$ by restricting the inner product of $\mathbb{R}^d$ (3.10) to $\mathrm{T}_x\mathrm{S}^{d-1}$: for all $u, v \in \mathrm{T}_x\mathrm{S}^{d-1}$, $\langle u, v \rangle_x = \langle u, v \rangle$. This is indeed a Riemannian metric, and S^{d-1} endowed with this metric is called a *Riemannian submanifold* of $\mathbb{R}^d$.

For Riemannian submanifolds, the Riemannian gradient is particularly simple to compute. As per our definitions, $f\colon \mathrm{S}^{d-1} \to \mathbb{R}$ is smooth if and only if there exists a function $\bar{f}\colon \mathbb{R}^d \to \mathbb{R}$, smooth in the usual sense, such that f and $\bar{f}$ coincide on S^{d-1}. We will argue that

$$\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad} \bar{f}(x)), \qquad \text{with} \qquad \mathrm{Proj}_x(v) = (I_d - xx^\top)v,$$

where $\mathrm{Proj}_x\colon \mathbb{R}^d \to \mathrm{T}_x\mathrm{S}^{d-1}$ is the orthogonal projector from $\mathbb{R}^d$ to $\mathrm{T}_x\mathrm{S}^{d-1}$ (orthogonal with respect to the inner product on $\mathbb{R}^d$.) The functions f and $\bar{f}$ often have the same analytical expression. For example, $f(x) = x^\top Ax$ (for some matrix $A \in \mathbb{R}^{d\times d}$) is smooth on S^{d-1} because $\bar{f}(x) = x^\top Ax$ is smooth on $\mathbb{R}^d$ and they coincide on S^{d-1}. To summarize:

For Riemannian submanifolds, the Riemannian gradient is the orthogonal projection of the "classical" gradient to the tangent spaces.

With these tools in place, we can justify the following algorithm, an instance of *Riemannian gradient descent* on S^{d-1}: given $x_0 \in \mathrm{S}^{d-1}$,

$$\begin{aligned} x_{k+1} = \mathrm{R}_{x_k}(-\alpha_k \mathrm{grad} f(x_k)), \qquad &\text{with} \qquad \mathrm{grad} f(x) = (I_d - xx^\top)\mathrm{grad} \bar{f}(x), \\ &\text{and} \qquad \mathrm{R}_x(v) = \frac{x+v}{\|x+v\|}, \end{aligned}$$

where $\bar{f}$ is a smooth extension of f to $\mathbb{R}^d$. More importantly, these tools give us a formal framework to design and analyze such algorithms on a large class of smooth, nonlinear spaces.

We now proceed to construct precise definitions, starting with a few reminders of linear algebra and multivariate calculus in linear spaces.

3.1 Reminders of Euclidean space

The letter $\mathcal{E}$ always denotes a *linear space* (or *vector space*) over the reals, that is, a set equipped with (and closed under) vector addition and scalar multiplication by real numbers. Frequent examples include $\mathbb{R}^d$ (column vectors of length d), $\mathbb{R}^{n\times p}$ (matrices of size $n \times p$), $\mathrm{Sym}(n)$ (real, symmetric matrices of size n), $\mathrm{Skew}(n)$ (real, skew-symmetric matrices of size n), and their (linear) subspaces.

We write $\mathrm{span}(u_1, \ldots, u_m)$ to denote the subspace of $\mathcal{E}$ *spanned* by vectors $u_1, \ldots, u_m \in \mathcal{E}$. By extension, $\mathrm{span}(X)$ for a matrix $X \in \mathbb{R}^{n\times m}$ denotes the subspace of $\mathbb{R}^n$ spanned by the columns X.

Given two linear spaces $\mathcal{E}$ and $\mathcal{F}$, a *linear map* or *linear operator* is a map $\mathcal{L} \colon \mathcal{E} \to \mathcal{F}$ such that $\mathcal{L}(au + bv) = a\mathcal{L}(u) + b\mathcal{L}(v)$ for all $u, v \in \mathcal{E}$ and $a, b \in \mathbb{R}$. We let $\mathrm{im}\,\mathcal{L}$ denote the *image* (or the *range*) of $\mathcal{L}$, and we let $\ker \mathcal{L}$ denote the *kernel* (or *null space*) of $\mathcal{L}$.

A *basis* for $\mathcal{E}$ is a set of vectors (elements of $\mathcal{E}$) $e_1, \ldots, e_d$ such that each vector $x \in \mathcal{E}$ can be expressed uniquely as a linear combination $x = a_1 e_1 + \cdots + a_d e_d$ with real coefficients $a_1, \ldots, a_d$. All bases have the same number of elements, called the *dimension* of $\mathcal{E}$ ($\dim \mathcal{E} = d$): it is always finite in our treatment. Each basis induces a one-to-one linear map identifying $\mathcal{E}$ and $\mathbb{R}^d$ to each other: we write $\mathcal{E} \equiv \mathbb{R}^d$.

Topology.
Recall that a *topology* on a set is a collection of subsets called *open* such that (a) the whole set and the empty set are open, (b) any union of opens is open, and (c) the intersection of a finite number of opens is open—more on this in Section 8.2. A subset is *closed* if its complement is open. A subset may be open, closed, both, or neither. We always equip $\mathbb{R}^d$ with its usual topology [Lee12, Ex. A.6]. Each linear space $\mathcal{E}$ of dimension d inherits the topology of $\mathbb{R}^d$ through their identification as above. A *neighborhood* of x in $\mathcal{E}$ is an open subset of $\mathcal{E}$ which contains x. Some authors call such sets *open* neighborhoods. All our neighborhoods are open, hence we omit the qualifier.

Euclidean structure.
It is useful to endow $\mathcal{E}$ with more structure.

Definition 3.1. *An* inner product *on $\mathcal{E}$ is a function $\langle \cdot, \cdot \rangle : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$ with the following properties. For all $u, v, w \in \mathcal{E}$ and $a, b \in \mathbb{R}$, we have:*

1. *Symmetry: $\langle u, v\rangle = \langle v, u\rangle$;*
2. *Linearity: $\langle au + bv, w\rangle = a\langle u, w\rangle + b\langle v, w\rangle$; and*
3. *Positive definiteness: $\langle u, u\rangle \geq 0$ and $\langle u, u\rangle = 0 \iff u = 0$.*

Definition 3.2. *A linear space $\mathcal{E}$ with an inner product $\langle \cdot, \cdot \rangle$ is a* Euclidean space. *An inner product induces a norm on $\mathcal{E}$ called the* Euclidean norm*:*

$$\|u\| = \sqrt{\langle u, u\rangle}. \tag{3.12}$$

Definition 3.3. *A basis $u_1, \ldots, u_d$ of a Euclidean space $\mathcal{E}$ is* orthonormal *if*

$$\forall\, 1 \leq i, j \leq n, \qquad\qquad \langle u_i, u_j \rangle = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

The standard inner product on $\mathbb{R}^d$ and the associated norm are:

$$\langle u, v \rangle = u^\top v = \sum_i u_i v_i, \qquad\qquad \|u\| = \sqrt{\sum_i u_i^2}. \tag{3.13}$$

Similarly, the standard inner product on linear spaces of matrices such as $\mathbb{R}^{n \times p}$ and $\mathrm{Sym}(n)$ is the so-called *Frobenius inner product*, with its associated *Frobenius norm*:

$$\langle U, V \rangle = \mathrm{Tr}(U^\top V) = \sum_{ij} U_{ij} V_{ij}, \qquad\qquad \|U\| = \sqrt{\sum_{ij} U_{ij}^2}, \tag{3.14}$$

where $\mathrm{Tr}(M) = \sum_i M_{ii}$ is the *trace* of a matrix. Summations are over all entries. When we omit to specify it, we mean to use the standard inner product and norm.

We often use the following properties of the above inner product, with matrices U, V, W, A, B of compatible sizes:

$$\begin{aligned} \langle U, V \rangle &= \langle U^\top, V^\top \rangle, & \langle UA, V \rangle &= \langle U, VA^\top \rangle, \\ \langle BU, V \rangle &= \langle U, B^\top V \rangle, & \langle U \odot W, V \rangle &= \langle U, V \odot W \rangle, \end{aligned} \tag{3.15}$$

where $\odot$ denotes entrywise multiplication (*Hadamard product*).

Although we only consider linear spaces over the reals, we can still handle complex matrices. For example, $\mathbb{C}^n$ is a real linear space of dimension $2n$. The standard basis for it is $e_1, \ldots, e_n, ie_1, \ldots, ie_n$, where $e_1, \ldots, e_n$ form the standard basis of $\mathbb{R}^n$ (the columns of the identity matrix of size n), and i is the imaginary unit. Indeed, any vector in $\mathbb{C}^n$ can be written uniquely as a linear combination of those basis vectors using real coefficients. The standard inner product and norm on $\mathbb{C}^n$ as a real linear space are:

$$\langle u, v \rangle = \Re\{u^* v\} = \Re\left\{\sum_k \bar{u}_k v_k\right\}, \qquad\qquad \|u\| = \sqrt{\sum_k |u_k|^2}, \tag{3.16}$$

where u^* is the Hermitian conjugate-transpose of u, $\bar{u}_k$ is the complex conjugate of u_k, $|u_k|$ is its magnitude and $\Re\{a\}$ is the real part of a. This perspective is equivalent to identifying $\mathbb{C}^n$ with $\mathbb{R}^{2n}$, where real and imaginary parts are considered as two vectors in $\mathbb{R}^n$. Likewise, the set of complex matrices $\mathbb{C}^{n \times p}$ is a real linear space of dimension $2np$, with the following standard inner product and norm:

$$\langle U, V \rangle = \Re\{\mathrm{Tr}(U^* V)\} = \Re\left\{\sum_{k\ell} \bar{U}_{k\ell} V_{k\ell}\right\}, \qquad \|U\| = \sqrt{\sum_{k\ell} |U_{k\ell}|^2}. \tag{3.17}$$

The analog of (3.15) in the complex case is:

$$\langle U, V\rangle = \langle U^*, V^*\rangle, \qquad \langle UA, V\rangle = \langle U, VA^*\rangle,$$
$$\langle BU, V\rangle = \langle U, B^*V\rangle, \qquad \langle U \odot W, V\rangle = \langle U, V \odot \bar{W}\rangle. \tag{3.18}$$

Linear maps between Euclidean spaces have *adjoints*, which we now define. These are often useful in deriving gradients of functions—more on this in Section 4.7.

Definition 3.4. *Let $\mathcal{E}$ and $\mathcal{F}$ be two Euclidean spaces, with inner products $\langle \cdot, \cdot\rangle_{\mathcal{E}}$ and $\langle \cdot, \cdot\rangle_{\mathcal{F}}$ respectively. Let $\mathcal{L}\colon \mathcal{E} \to \mathcal{F}$ be a linear map. The* adjoint *of $\mathcal{L}$ is the linear map $\mathcal{L}^*\colon \mathcal{F} \to \mathcal{E}$ defined by the following property:*

$$\forall u \in \mathcal{E}, v \in \mathcal{F}, \qquad \langle \mathcal{L}(u), v\rangle_{\mathcal{F}} = \langle u, \mathcal{L}^*(v)\rangle_{\mathcal{E}}.$$

Definition 3.5. *Let $\mathcal{E}$ be a Euclidean space with inner product $\langle \cdot, \cdot\rangle$. If the linear map $\mathcal{A}\colon \mathcal{E} \to \mathcal{E}$ satisfies*

$$\forall u, v \in \mathcal{E}, \qquad \langle \mathcal{A}(u), v\rangle = \langle u, \mathcal{A}(v)\rangle,$$

that is, if $\mathcal{A} = \mathcal{A}^*$, we say $\mathcal{A}$ is self-adjoint *or* symmetric.

As we can see from (3.15) and (3.18), adjoints and matrix transposes are intimately related: it is an exercise to make this precise.

Self-adjoint linear maps have important spectral properties.

Theorem 3.6 (spectral theorem)**.** *A self-adjoint map $\mathcal{A}$ on a Euclidean space $\mathcal{E}$ admits an orthonormal basis of eigenvectors $v_1, \ldots, v_d \in \mathcal{E}$ associated to real eigenvalues $\lambda_1, \ldots, \lambda_d$ so that $\mathcal{A}(v_i) = \lambda_i v_i$ for $i = 1, \ldots, d$ with $d = \dim \mathcal{E}$.*

Definition 3.7. *A self-adjoint map $\mathcal{A}$ on $\mathcal{E}$ is* positive semidefinite *if, for all $u \in \mathcal{E}$, we have $\langle u, \mathcal{A}(u)\rangle \geq 0$; we write $\mathcal{A} \succeq 0$. Owing to the spectral theorem, this is equivalent to all eigenvalues of $\mathcal{A}$ being nonnegative. Similarly, $\mathcal{A}$ is* positive definite *if $\langle u, \mathcal{A}(u)\rangle > 0$ for all nonzero $u \in \mathcal{E}$; we write $\mathcal{A} \succ 0$. This is equivalent to all eigenvalues of $\mathcal{A}$ being positive.*

Norms on vector spaces induce norms for linear maps.

Definition 3.8. *The* operator norm *of $\mathcal{L}\colon \mathcal{E} \to \mathcal{F}$ is defined as*

$$\|\mathcal{L}\| = \max_{u \in \mathcal{E}, u \neq 0} \frac{\|\mathcal{L}(u)\|_{\mathcal{F}}}{\|u\|_{\mathcal{E}}},$$

where $\|\cdot\|_{\mathcal{E}}$ and $\|\cdot\|_{\mathcal{F}}$ denote the norms on the Euclidean spaces $\mathcal{E}$ and $\mathcal{F}$, respectively.

Equivalently, $\|\mathcal{L}\|$ is the smallest real such that $\|\mathcal{L}(u)\|_{\mathcal{F}} \leq \|\mathcal{L}\|\|u\|_{\mathcal{E}}$ for all $u \in \mathcal{E}$. For a self-adjoint map $\mathcal{A}$ with eigenvalues $\lambda_1, \ldots, \lambda_d$, it is easy to see that $\|\mathcal{A}\| = \max_{1 \leq i \leq d} |\lambda_i|$.

Calculus.
We write $F: A \to B$ to designate a map F whose domain is all of A. If C is a subset of A, we write $F|_C: C \to B$ to designate the *restriction* of F to the domain C, so that $F|_C(x) = F(x)$ for all $x \in C$.

Let U, V be open sets in two linear spaces $\mathcal{E}, \mathcal{F}$. A map $F: U \to V$ is *smooth* if it is infinitely differentiable (class C^∞) on its domain. We also say that F is *smooth at a point* $x \in U$ if there exists a neighborhood U' of x such that $F|_{U'}$ is smooth. Accordingly, F is smooth if it is smooth at all points in its domain.

If $F: U \to V$ is smooth at x, the *differential* of F at x is the linear map $\mathrm{D}F(x): \mathcal{E} \to \mathcal{F}$ defined by

$$\mathrm{D}F(x)[u] = \frac{\mathrm{d}}{\mathrm{d}t} F(x + tu)\bigg|_{t=0} = \lim_{t \to 0} \frac{F(x + tu) - F(x)}{t}. \tag{3.19}$$

For a curve $c: \mathbb{R} \to \mathcal{E}$, we write c' to denote its velocity: $c'(t) = \frac{\mathrm{d}}{\mathrm{d}t} c(t)$.

For a smooth function $f: \mathcal{E} \to \mathbb{R}$ defined on a Euclidean space $\mathcal{E}$, the (Euclidean) *gradient* of f is the map $\mathrm{grad} f: \mathcal{E} \to \mathcal{E}$ defined by the following property:

$$\forall x, v \in \mathcal{E}, \qquad \langle \mathrm{grad} f(x), v \rangle = \mathrm{D}f(x)[v].$$

The (Euclidean) *Hessian* of f at x is the linear map $\mathrm{Hess} f(x): \mathcal{E} \to \mathcal{E}$ defined by

$$\mathrm{Hess} f(x)[v] = \mathrm{D}(\mathrm{grad} f)(x)[v] = \lim_{t \to 0} \frac{\mathrm{grad} f(x + tv) - \mathrm{grad} f(x)}{t}.$$

The *Clairaut–Schwarz theorem* implies that $\mathrm{Hess} f(x)$ is self-adjoint with respect to the inner product of $\mathcal{E}$.

Exercise 3.9 (Adjoint and transpose)**.** *Let $u_1, \ldots, u_n$ form an orthonormal basis of $\mathcal{E}$. Likewise, let $v_1, \ldots, v_m$ form an orthonormal basis of $\mathcal{F}$. Consider a linear map $\mathcal{L}: \mathcal{E} \to \mathcal{F}$. For each $1 \leq i \leq n$, the vector $\mathcal{L}(u_i)$ is an element of $\mathcal{F}$; therefore, it expands uniquely in the basis v as follows:*

$$\mathcal{L}(u_i) = \sum_{j=1}^{m} M_{ji} v_j,$$

where we collect the coefficients into a matrix $M \in \mathbb{R}^{m \times n}$. This matrix represents *$\mathcal{L}$ with respect to the chosen bases. Show that the matrix which represents $\mathcal{L}^*$ with respect to those same bases is $M^\top$: the transpose of M. In particular, a linear map $\mathcal{A}: \mathcal{E} \to \mathcal{E}$ is self-adjoint if and only if the matrix associated to it with respect to the basis $u_1, \ldots, u_n$ is symmetric.*

3.2 Embedded submanifolds of a linear space

We set out to define what it means for a subset $\mathcal{M}$ of a linear space $\mathcal{E}$ to be smooth. Our main angle is to capture the idea that a smooth set can be linearized in some meaningful way around each point. To make sense of what that might mean, consider the sphere S^{d-1}. This is the set of vectors $x \in \mathbb{R}^d$ satisfying

$$h(x) = x^\top x - 1 = 0.$$

As we discussed in the introduction, it can be adequately linearized around each point by the set (3.3). The perspective we used to obtain this linearization is that of differentiating the defining equation. More precisely, consider a truncated Taylor expansion of h:

$$h(x + tv) = h(x) + t\,\mathrm{D}h(x)[v] + O(t^2).$$

If x is in S^{d-1} and v is in $\ker \mathrm{D}h(x)$ (so that $h(x) = 0$ and $\mathrm{D}h(x)[v] = 0$), then $h(x + tv) = O(t^2)$, indicating that $x + tv$ nearly satisfies the defining equation of S^{d-1} for small t. This motivates us to consider the subspace $\ker \mathrm{D}h(x)$ as a linearization of S^{d-1} around x. Since

$$\mathrm{D}h(x)[v] = \lim_{t \to 0} \frac{h(x + tv) - h(x)}{t} = x^\top v + v^\top x = 2x^\top v,$$

the kernel of $\mathrm{D}h(x)$ is the subspace orthogonal to x in $\mathbb{R}^d$ (with respect to the usual inner product). This coincides with (3.3), arguably in line with intuition.

At first, one might think that if a set is defined by an equation of the form $h(x) = 0$ with some smooth function h, then that set is smooth and can be linearized by the kernels of $\mathrm{D}h$. However, this is not the case. Indeed, consider the following example in $\mathbb{R}^2$:

$$\mathcal{X} = \left\{ x \in \mathbb{R}^2 : h(x) = x_1^2 - x_2^2 = 0 \right\}.$$

The defining function h is smooth, yet the set $\mathcal{X}$ is a cross in the plane formed by the union of the lines $x_1 = x_2$ and $x_1 = -x_2$. We want to exclude such sets because of the kink at the origin. If we blindly use the kernel of the differential to linearize $\mathcal{X}$, we first determine

$$\mathrm{D}h(x) = \left[\frac{\partial h}{\partial x_1}(x), \frac{\partial h}{\partial x_2}(x) \right] = \left[2x_1, -2x_2 \right].$$

At $x = 0$, $\mathrm{D}h(0) = [0, 0]$, whose kernel is all of $\mathbb{R}^2$: that does not constitute a reasonable linearization of $\mathcal{X}$ around the origin.

We can gain further insight into the issue at hand by considering additional examples. The zero-sets of the functions $h(x) = x_1^2 - x_2^3$ and $h(x) = x_1^2 - x_2^4$ from $\mathbb{R}^2$ to $\mathbb{R}$ respectively define a cusp and a double parabola, both of which fail our intuitive test of smoothness at the origin. What the cross, cusp and double parabola have in common is that the rank of $\mathrm{D}h(x)$ suddenly drops from one to zero at the origin, whereas for the sphere that rank is maximal on the whole set.

These observations motivate the definition below. Since smoothness is a local notion, the definition is phrased in terms of what the set $\mathcal{M}$ looks like around each point. Since a set $\mathcal{M}$ may be equivalently defined by many different functions h, and since it may not be practical (or even possible: see Section 3.10) to define

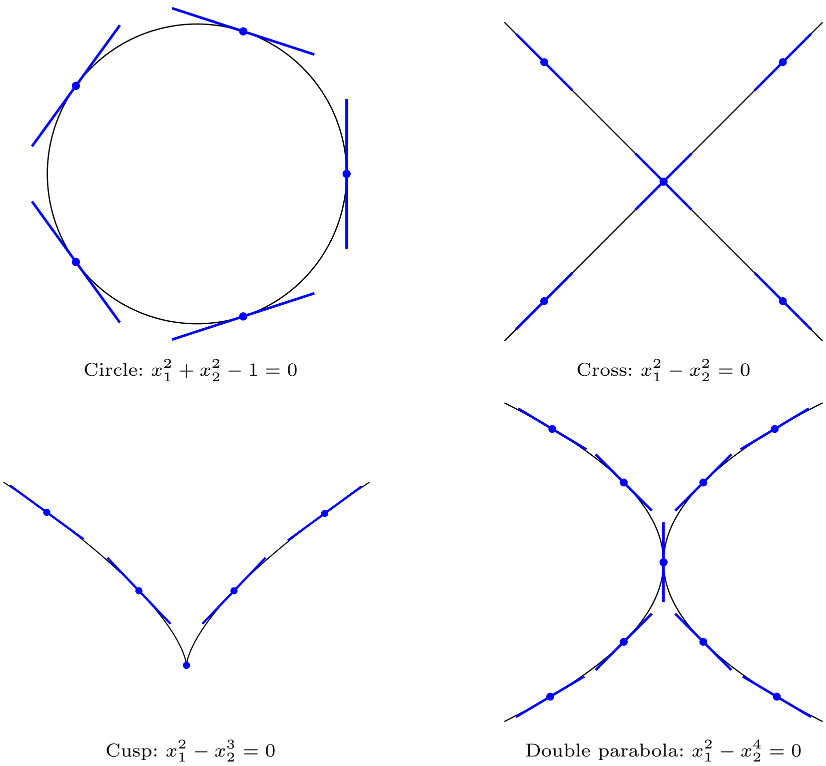


Figure 3.2 Four different sets $\mathcal{S}$ defined as the zero-sets of a smooth function from $\mathbb{R}^2$ to $\mathbb{R}$. For each, the sets $\mathrm{T}_x\mathcal{S}$ (Definition 3.14) are drawn at a few different points. Only the circle (top left) is an embedded submanifold of $\mathbb{R}^2$.

all of $\mathcal{M}$ with a single function h, the definition allows for a different one to be used around each point.

Definition 3.10. *Let $\mathcal{E}$ be a linear space of dimension d. A subset $\mathcal{M}$ of $\mathcal{E}$ is a* (smooth) embedded submanifold of $\mathcal{E}$ of dimension n *if either*

1. *$n = d$ and $\mathcal{M}$ is open in $\mathcal{E}$—we also call this an* open submanifold*; or*
2. *$n = d - k$ for some $k \geq 1$ and, for each $x \in \mathcal{M}$, there exists a neighborhood U of x in $\mathcal{E}$ and a smooth function $h \colon U \to \mathbb{R}^k$ such that*
   (a) If y is in U, then $h(y) = 0$ if and only if $y \in \mathcal{M}$; and
   (b) $\operatorname{rank} \mathrm{D}h(x) = k$.
   Such a function h is called a local defining function *for $\mathcal{M}$ at x.*

If $\mathcal{M}$ is a linear (sub)space, we also call it a linear manifold*.*

Condition 2(a) above can be stated equivalently as:

$$\mathcal{M} \cap U = h^{-1}(0) \triangleq \{y \in U : h(y) = 0\}.$$

It is an exercise to verify that Definition 3.10 excludes various pathological sets such as the cross ($x_1^2 = x_2^2$), cusp ($x_1^2 = x_2^3$) and double parabola ($x_1^2 = x_2^4$).

Differential geometry defines a broader class of smooth sets called *(smooth) manifolds*. We typically omit the word 'smooth' as all of our manifolds are smooth, though bear in mind that in the literature there exist different kinds of manifolds, not all of them smooth. Embedded submanifolds are manifolds. When ⋆ the statements we make hold true for manifolds in general, we use that word to signal it. This is very common throughout Chapters 3 and 5.

The hope is that limiting our initial treatment of manifolds to embedded submanifolds provides a more intuitive entry point to build all the tools we need for optimization. This is all the more relevant considering that many of the manifolds we encounter in applications are in fact embedded submanifolds, presented to us as zero-sets of their local defining functions. All of our optimization algorithms work on general manifolds. The general definitions are in Chapter 8.

To build additional support for our definition of embedded submanifolds, we further argue that small patches of $\mathcal{M}$ can be deformed into linear subspaces in a smooth and smoothly invertible way. This captures an important feature of smoothness, namely: upon zooming close to a point of $\mathcal{M}$, what we see can hardly be distinguished from what we would have seen had $\mathcal{M}$ been a linear subspace of $\mathcal{E}$.

Definition 3.11. *A* diffeomorphism *is a bijective map $F \colon U \to V$ where U, V are open sets and such that both F and F^{-1} are smooth.*

Theorem 3.12. *Let $\mathcal{E}$ be a linear space of dimension d. A subset $\mathcal{M}$ of $\mathcal{E}$ is an embedded submanifold of $\mathcal{E}$ of dimension $n = d - k$ if and only if for each $x \in \mathcal{M}$ there exists a neighborhood U of x in $\mathcal{E}$, an open set V in $\mathbb{R}^d$ and a diffeomorphism $F \colon U \to V$ such that $F(\mathcal{M} \cap U) = E \cap V$ where $E = \{y \in \mathbb{R}^d : y_{n+1} = \cdots = y_d = 0\}$ is a linear subspace of $\mathbb{R}^d$.*

The main tool we need to prove Theorem 3.12 is the standard *inverse function theorem*, stated here without proof [Lee12, Thm. C.34].

Theorem 3.13 (Inverse function theorem)**.** *Suppose $U \subseteq \mathcal{E}$ and $V \subseteq \mathcal{F}$ are open subsets of linear spaces of the same dimension, and $F \colon U \to V$ is smooth. If $\mathrm{D}F(x)$ is invertible at some point $x \in U$, then there exist neighborhoods $U' \subseteq U$ of x and $V' \subseteq V$ of $F(x)$ such that $F|_{U'} \colon U' \to V'$ (the restriction of F to U' and V') is a diffeomorphism.*

Equipped with this tool, we proceed to prove Theorem 3.12.

Proof of Theorem 3.12. We prove one direction of the theorem, namely: we assume $\mathcal{M}$ is an embedded submanifold and construct diffeomorphisms F. The other direction is left as an exercise. For the latter, it is helpful to note that if F is a diffeomorphism with inverse F^{-1}, then $\mathrm{D}F(x)$ is invertible and

$$(\mathrm{D}F(x))^{-1} = \mathrm{D}F^{-1}(F(x)). \tag{3.20}$$

To see this, apply the chain rule to differentiate $F^{-1} \circ F$, noting that this is nothing but the identity map.

If $n = d$ (that is, $\mathcal{M}$ is open in $\mathcal{E}$), the claim is clear: simply let F be any invertible linear map from $\mathcal{E}$ to $\mathbb{R}^d$ (for example, using a basis of $\mathcal{E}$), and restrict its domain and codomain to $U = \mathcal{M}$ and $V = F(U)$.

We now consider the more interesting case where $n = d - k$ with $k \geq 1$. Let $h\colon U \to \mathbb{R}^k$ be any local defining function for $\mathcal{M}$ at x. We work in coordinates on $\mathcal{E}$, which is thus identified with $\mathbb{R}^d$. Then, we can think of $\mathrm{D}h(x)$ as a matrix of size $k \times d$. By assumption, $\mathrm{D}h(x)$ has rank k. This means that it is possible to pick k columns of that matrix which form a $k \times k$ invertible matrix. If needed, permute the chosen coordinates so that the last k columns have that property (this is without loss of generality). Then, we can write $\mathrm{D}h(x)$ in block form so that

$$\mathrm{D}h(x) = \begin{bmatrix} A & B \end{bmatrix},$$

where $B \in \mathbb{R}^{k\times k}$ is invertible and A is in $\mathbb{R}^{k\times(d-k)}$. Now consider the function $F\colon U \to \mathbb{R}^d$ (recall $U \subseteq \mathcal{E}$ is the domain of h) defined by

$$F(y) = (y_1, \ldots, y_{d-k}, h_1(y), \ldots, h_k(y))^\top, \tag{3.21}$$

where $y_1, \ldots, y_d$ denote the coordinates of $y \in \mathcal{E}$. In order to apply the inverse function theorem to F at x, we must verify that F is smooth—this is clear—and that the differential of F at x is invertible. Working this out one row at a time, we get the following expression for that differential:

$$\mathrm{D}F(x) = \begin{bmatrix} I_{d-k} & 0 \\ A & B \end{bmatrix},$$

where I_{d-k} is the identity matrix of size $d-k$, and 0 here denotes a zero matrix of size $(d-k)\times k$. The matrix $\mathrm{D}F(x)$ is invertible, as demonstrated by the following expression for its inverse:

$$(\mathrm{D}F(x))^{-1} = \begin{bmatrix} I_{d-k} & 0 \\ -B^{-1}A & B^{-1} \end{bmatrix}. \tag{3.22}$$

(Indeed, their product is I_d.) Hence, the inverse function theorem asserts that we may reduce U to a possibly smaller neighborhood of x so that F (now restricted to that new neighborhood) is a diffeomorphism from U to $V = F(U)$. The property $F(\mathcal{M} \cap U) = E \cap V$ follows by construction of F from the property $\mathcal{M} \cap U = h^{-1}(0)$. □

In order to understand the local geometry of a set around a point, we aim to describe acceptable directions of movement through that point. This is close in spirit to the tools we look to develop for optimization, as they involve moving away from a point while remaining on the set. Specifically, for a subset $\mathcal{M}$ of a linear space $\mathcal{E}$, consider all the smooth curves of $\mathcal{E}$ which lie entirely on $\mathcal{M}$ and pass through a given point x. Collect their velocities as they do so in a set $\mathrm{T}_x\mathcal{M}$

defined below. In that definition, c is smooth in the usual sense as a map from (an open subset of) $\mathbb{R}$ to $\mathcal{E}$—two linear spaces.

Definition 3.14. *Let $\mathcal{M}$ be a subset of $\mathcal{E}$. For all $x \in \mathcal{M}$, define:*

$$\mathrm{T}_x\mathcal{M} = \{c'(0) \mid c\colon I \to \mathcal{M} \text{ is smooth and } c(0) = x\}, \tag{3.23}$$

where I is any open interval containing $t = 0$. That is, v is in $\mathrm{T}_x\mathcal{M}$ if and only if there exists a smooth curve on $\mathcal{M}$ passing through x with velocity v.

Note that $\mathrm{T}_x\mathcal{M}$ is a subset of $\mathcal{E}$. For the sphere, it is easy to convince oneself that $\mathrm{T}_x\mathcal{M}$ coincides with the subspace in (3.3). We show in the next theorem that this is always the case for embedded submanifolds.

Theorem 3.15. *Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$. Consider $x \in \mathcal{M}$ and the set $\mathrm{T}_x\mathcal{M}$ (3.23). If $\mathcal{M}$ is an open submanifold, then $\mathrm{T}_x\mathcal{M} = \mathcal{E}$. Otherwise, $\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x)$ with h any local defining function at x.*

Proof. For open submanifolds, the claim is clear. By definition, $\mathrm{T}_x\mathcal{M}$ is included in $\mathcal{E}$. The other way around, for any $v \in \mathcal{E}$, consider $c(t) = x + tv$: this is a smooth curve from some non-empty interval around 0 to $\mathcal{M}$ such that $c(0) = x$, hence $v = c'(0)$ is in $\mathrm{T}_x\mathcal{M}$. This shows $\mathcal{E}$ is included in $\mathrm{T}_x\mathcal{M}$, so that the two coincide.

Now consider the case of $\mathcal{M}$ an embedded submanifold of dimension $n = d - k$ with $k \geq 1$. Let $h\colon U \to \mathbb{R}^k$ be a local defining function of $\mathcal{M}$ around x. The proof is in two steps. First, we show that $\mathrm{T}_x\mathcal{M}$ is included in $\ker \mathrm{D}h(x)$. Then, we show that $\mathrm{T}_x\mathcal{M}$ contains a linear subspace of the same dimension as $\ker \mathrm{D}h(x)$. These two facts combined indeed confirm that $\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x)$.

Step 1. If v is in $\mathrm{T}_x\mathcal{M}$, there exists $c\colon I \to \mathcal{M}$, smooth, such that $c(0) = x$ and $c'(0) = v$. Since $c(t)$ is in $\mathcal{M}$, we have $h(c(t)) = 0$ for all $t \in I$ (if need be, restrict the interval I to ensure $c(t)$ remains in the domain of h). Thus, the derivative of $h \circ c$ vanishes at all times:

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} h(c(t)) = \mathrm{D}h(c(t))[c'(t)].$$

In particular, at $t = 0$ this implies $\mathrm{D}h(x)[v] = 0$, that is, $v \in \ker \mathrm{D}h(x)$. This confirms $\mathrm{T}_x\mathcal{M} \subseteq \ker \mathrm{D}h(x)$.

Step 2. To show that $\mathrm{T}_x\mathcal{M}$ contains a subspace of the same dimension as $\ker \mathrm{D}h(x)$ (namely, of dimension $n = d - k$), we must construct smooth curves on $\mathcal{M}$ that pass through x with various velocities. To do so, we call upon Theorem 3.12. The latter provides us with a diffeomorphism $F\colon U \to V$ (where U is a possibly smaller neighborhood of x than the domain of h.) We use F^{-1} to construct smooth curves on $\mathcal{M}$ that pass through x. Specifically, pick an arbitrary $u \in \mathbb{R}^{d-k}$ and let

$$\gamma(t) = F(x) + t \begin{bmatrix} u \\ 0 \end{bmatrix}.$$

(Here, 0 denotes a zero vector of size k.) Notice that γ remains in $E \cap V$ for t

close to zero, where E is the subspace of $\mathbb{R}^d$ consisting of all vectors whose last k entries are zero. Since $F^{-1}(E \cap V) = \mathcal{M} \cap U$, it follows that

$$c(t) = F^{-1}(\gamma(t)) \tag{3.24}$$

resides in $\mathcal{M}$ for t close to zero. Moreover, $c(0) = x$ and c is smooth since F^{-1} is smooth. It follows that c is indeed a smooth curve on $\mathcal{M}$ passing through x. What is the velocity of this curve at x? Applying the chain rule to (3.24), we get

$$c'(t) = \mathrm{D}F^{-1}(\gamma(t))\,[\gamma'(t)].$$

In particular, at $t = 0$ we have

$$c'(0) = \mathrm{D}F^{-1}(F(x)) \begin{bmatrix} u \\ 0 \end{bmatrix}.$$

Since F is a diffeomorphism, we know from (3.20) that $\mathrm{D}F^{-1}(F(x))$ is an invertible linear map, equal to $(\mathrm{D}F(x))^{-1}$. The specific form of $c'(0)$ is unimportant. What matters is that each $c'(0)$ of the form above certainly belongs to $\mathrm{T}_x\mathcal{M}$ (3.23). Since $\mathrm{D}F^{-1}(F(x))$ is invertible and $u \in \mathbb{R}^{d-k}$ is arbitrary, this means that $\mathrm{T}_x\mathcal{M}$ contains a subspace of dimension $d-k$. But we know from the previous step that $\mathrm{T}_x\mathcal{M}$ is included in a subspace of dimension $d-k$, namely, $\ker \mathrm{D}h(x)$. It follows that $\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x)$. Since this holds for all $x \in \mathcal{M}$, the proof is complete. □

Thus, for an embedded submanifold $\mathcal{M}$ of dimension $n = d - k$, for each $x \in \mathcal{M}$, the set $\mathrm{T}_x\mathcal{M}$ is a linear subspace of $\mathcal{E}$ of dimension n. These subspaces are the linearizations of the smooth set $\mathcal{M}$.

Definition 3.16. *We call* $\mathrm{T}_x\mathcal{M}$ *the* tangent space *to* $\mathcal{M}$ *at* x. *Vectors in* $\mathrm{T}_x\mathcal{M}$ *are called* tangent vectors *to* $\mathcal{M}$ *at* x. *The dimension of* $\mathrm{T}_x\mathcal{M}$ *(which is independent of* x*) coincides with the* dimension of $\mathcal{M}$, *denoted by* $\dim \mathcal{M}$.

We consider three brief examples of embedded submanifolds: two obvious by now, and one arguably less obvious. It is good to keep all three in mind when assessing whether a certain proposition concerning embedded submanifolds is likely to be true. Chapter 7 details further examples.

Example 3.17. *The set* $\mathbb{R}^d$ *is a linear manifold of dimension* d *with tangent spaces* $\mathrm{T}_x\mathcal{M} = \mathbb{R}^d$ *for all* $x \in \mathbb{R}^d$. *The affine space* $\{x \in \mathbb{R}^d : Ax = b\}$ *defined by a matrix* $A \in \mathbb{R}^{k \times d}$ *of rank* k *and arbitrary vector* $b \in \mathbb{R}^k$ *is a manifold of dimension* $n = d - k$ *embedded in* $\mathbb{R}^d$.

Example 3.18. *The sphere* $\mathrm{S}^{d-1} = \{x \in \mathbb{R}^d : x^\top x = 1\}$ *is the zero level set of* $h(x) = x^\top x - 1$, *smooth from* $\mathbb{R}^d$ *to* $\mathbb{R}$. *Since* $\mathrm{D}h(x)[v] = 2x^\top v$, *it is clear that* $\operatorname{rank} \mathrm{D}h(x) = 1$ *for all* $x \in \mathrm{S}^{d-1}$. *As a result,* S^{d-1} *is an embedded submanifold of* $\mathbb{R}^d$ *of dimension* $n = d - 1$. *Furthermore, its tangent spaces are given by* $\mathrm{T}_x\mathrm{S}^{d-1} = \ker \mathrm{D}h(x) = \{v \in \mathbb{R}^d : x^\top v = 0\}$.

Example 3.19. *Let* $\mathrm{Sym}(2)_1$ *denote the set of symmetric matrices of size two and rank one, that is,*

$$\mathrm{Sym}(2)_1 = \left\{ X = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} : \operatorname{rank} X = 1 \right\}.$$

This is a subset of $\mathrm{Sym}(2)$*, the linear space of symmetric matrices of size two. The rank function is not a smooth map (it is not even continuous), hence we cannot use it as a local defining function. Nevertheless, we can construct local defining functions for* $\mathrm{Sym}(2)_1$*. Indeed, a matrix of size* 2×2 *has rank one if and only if it is nonzero and its determinant is zero, hence:*

$$\mathrm{Sym}(2)_1 = \left\{ X = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} : x_1 x_3 - x_2^2 = 0 \text{ and } X \neq 0 \right\}.$$

Let $U = \mathrm{Sym}(2) \backslash \{0\}$ *be the open subset of* $\mathrm{Sym}(2)$ *obtained by removing the zero matrix. Consider* $h \colon U \to \mathbb{R}$ *defined by* $h(X) = x_1 x_3 - x_2^2$*. Clearly,* h *is smooth and* $h^{-1}(0) = \mathrm{Sym}(2)_1 \cap U = \mathrm{Sym}(2)_1$*. Furthermore,*

$$\mathrm{D}h(X)[\dot{X}] = \dot{x}_1 x_3 + x_1 \dot{x}_3 - 2 x_2 \dot{x}_2 = \begin{bmatrix} x_3 & -2x_2 & x_1 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix},$$

where $\dot{X}$ *is a matrix in* $\mathrm{Sym}(2)$*: the dot is a visual indication that we should think of* $\dot{X}$ *as a perturbation of* X*. This linear map has rank one provided* $X \neq 0$*, which is always the case in the domain of* h*. Hence,* h *is a defining function for* $\mathrm{Sym}(2)_1$ *around any* X *in* $\mathrm{Sym}(2)_1$*. This confirms that the latter is an embedded submanifold of* $\mathrm{Sym}(2)$ *of dimension* $\dim \mathrm{Sym}(2) - 1 = 3 - 1 = 2$*. Its tangent space at* X *is given by* $\ker \mathrm{D}h(X)$*:*

$$\mathrm{T}_X \mathrm{Sym}(2)_1 = \left\{ \dot{X} = \begin{bmatrix} \dot{x}_1 & \dot{x}_2 \\ \dot{x}_2 & \dot{x}_3 \end{bmatrix} : \dot{x}_1 x_3 + x_1 \dot{x}_3 - 2 x_2 \dot{x}_2 = 0 \right\}.$$

Contrary to the two previous examples, this manifold is neither open nor closed in its embedding space. It is also not connected. Visualized in $\mathbb{R}^3$*, it corresponds to a double, infinite elliptic cone. Indeed,* $X \neq 0$ *is in* $\mathrm{Sym}(2)_1$ *if and only if*

$$(x_1 + x_3)^2 = (2x_2)^2 + (x_1 - x_3)^2.$$

After the linear change of variables $z_1 = x_1 - x_3$*,* $z_2 = 2x_2$ *and* $z_3 = x_1 + x_3$*, the defining equation becomes* $z_1^2 + z_2^2 = z_3^2$*, omitting the origin.*

We can combine manifolds to form new ones. For instance, it is an exercise to show that Cartesian products of manifolds are manifolds.[2]

Proposition 3.20. *Let* $\mathcal{M}, \mathcal{M}'$ *be embedded submanifolds of* $\mathcal{E}, \mathcal{E}'$ *(respectively). Then,* $\mathcal{M} \times \mathcal{M}'$ *is an embedded submanifold of* $\mathcal{E} \times \mathcal{E}'$ *of dimension* $\dim \mathcal{M} + \dim \mathcal{M}'$ *with tangent spaces given by:*

$$\mathrm{T}_{(x,x')}(\mathcal{M} \times \mathcal{M}') = \mathrm{T}_x \mathcal{M} \times \mathrm{T}_{x'} \mathcal{M}'.$$

[2] We collect facts about product manifolds along the way: see Table 7.2 for a summary.

For example, after showing that the sphere S^{d-1} is an embedded submanifold of $\mathbb{R}^d$, it follows that $\mathrm{S}^{d-1} \times \cdots \times \mathrm{S}^{d-1} = (\mathrm{S}^{d-1})^k$ is an embedded submanifold of $(\mathbb{R}^d)^k \equiv \mathbb{R}^{d\times k}$: it is called the *oblique manifold* $\mathrm{OB}(d, k)$.

In closing this section, we equip embedded submanifolds of $\mathcal{E}$ with the topology induced by $\mathcal{E}$.[3] Having a topology notably allows us to define notions such as local optima and convergence of sequences on $\mathcal{M}$. Both are useful when studying iterative optimization algorithms.

Definition 3.21. *A subset $\mathcal{U}$ of $\mathcal{M}$ is* open *(resp.,* closed*) in $\mathcal{M}$ if $\mathcal{U}$ is the intersection of $\mathcal{M}$ with an open (resp., closed) subset of $\mathcal{E}$. This is called the* subspace topology*.*

Echoing the conventions laid out in Section 3.1, our neighborhoods are open.

Definition 3.22. *A* neighborhood *of x in $\mathcal{M}$ is an open subset of $\mathcal{M}$ which contains x. By extension, a neighborhood of a subset of $\mathcal{M}$ is an open set of $\mathcal{M}$ which contains that subset.*

It is an exercise to show that open subsets of a manifold $\mathcal{M}$ are manifolds; we call them *open submanifolds of $\mathcal{M}$*.

Proposition 3.23. *Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$. Any open subset of $\mathcal{M}$ is also an embedded (but not necessarily open) submanifold of $\mathcal{E}$, with same dimension and tangent spaces as $\mathcal{M}$.*

Exercise 3.24. *Complete the proof of Theorem 3.12.*

Exercise 3.25. *Give a proof of Proposition 3.20.*

Exercise 3.26. *Give a proof of Proposition 3.23. In particular, deduce that the relative interior of the simplex,*

$$\Delta_+^{d-1} = \{x \in \mathbb{R}^d : x_1 + \cdots + x_d = 1 \text{ and } x_1, \ldots, x_d > 0\}, \tag{3.25}$$

is an embedded submanifold of $\mathbb{R}^d$. This is useful to represent non-vanishing discrete probability distributions. See also Exercise 3.41.

Exercise 3.27. *Show that the cross $\mathcal{X} = \{x \in \mathbb{R}^2 : x_1^2 = x_2^2\}$ is not an embedded submanifold. It is not sufficient to show that $x \mapsto x_1^2 - x_2^2$ is not a local defining function at the origin: it is necessary to show that no local defining function exists at that point. Hint: proceeding by contradiction, assume there exists a local defining function around the origin and show that its kernel is too large.*

Exercise 3.28. *Show that the cusp $\mathcal{C} = \{x \in \mathbb{R}^2 : x_1^2 = x_2^3\}$ is not an embedded submanifold. Hint: argue that $\mathrm{T}_0\mathcal{C}$ (as defined by (3.23)) is too low-dimensional.*

[3] About terminology: the general definition of submanifolds allows for other topologies. The qualifier 'embedded' (some say 'regular') indicates we use the induced topology. More on this in Section 8.14.

Exercise 3.29. *Show that the double parabola $\mathcal{P} = \{x \in \mathbb{R}^2 : x_1^2 = x_2^4\}$ is not an embedded submanifold, yet $\mathrm{T}_x\mathcal{P}$ as defined by* (3.23) *is a linear subspace of dimension one in $\mathbb{R}^2$ for all $x \in \mathcal{P}$. This example shows that Definition 3.10 is more restrictive than just requiring all sets $\mathrm{T}_x\mathcal{P}$ to be subspaces of the same dimension. Hint: proceeding by contradiction, assume $\mathcal{P}$ is an embedded submanifold and call upon Theorem 3.12 to construct a special diffeomorphism F; then, derive a contradiction from the fact that $\mathcal{P}$ around the origin does not look like a one-dimensional curve. Specifically, you may want to use the fact that it is impossible to have three or more disjoint open intervals of the real line sharing a common accumulation point.*

3.3 Smooth maps on embedded submanifolds

Now that we have a notion of smooth sets, we can introduce the all important notion of smooth maps between smooth sets. It relies heavily on the classical notion of smooth maps between (open subsets of) linear spaces. In optimization, two examples of maps between manifolds are cost functions ($\mathcal{M} \to \mathbb{R}$) and iteration maps ($\mathcal{M} \to \mathcal{M}$); more will come up.

Definition 3.30. *Let $\mathcal{M}$ and $\mathcal{M}'$ be embedded submanifolds of $\mathcal{E}$ and $\mathcal{E}'$ (respectively). A map $F \colon \mathcal{M} \to \mathcal{M}'$ is* smooth at *$x \in \mathcal{M}$ if there exists a function $\bar{F} \colon U \to \mathcal{E}'$ which is smooth on a neighborhood U of x in $\mathcal{E}$ and such that F and $\bar{F}$ coincide on $\mathcal{M} \cap U$, that is, $F(y) = \bar{F}(y)$ for all $y \in \mathcal{M} \cap U$. We call $\bar{F}$ a* (local) smooth extension *of F around x. The map F is* smooth *if it is smooth at all $x \in \mathcal{M}$.*

In the above definition, $\bar{F}$ is a map between open subsets of linear spaces: for it to be smooth means infinitely differentiable on its domain, in the usual sense. By definition, if $\bar{F}$ is any smooth map on $\mathcal{E}$, and $\mathcal{M}$ is an embedded submanifold of $\mathcal{E}$, then the *restriction* $F = \bar{F}|_{\mathcal{M}}$ is smooth on $\mathcal{M}$. This still holds if $\bar{F}$ is only defined on a neighborhood of $\mathcal{M}$ in $\mathcal{E}$, that is, on an open subset of $\mathcal{E}$ which contains $\mathcal{M}$.

Conversely, the following proposition states that a smooth map on $\mathcal{M}$ admits a *smooth extension* to a neighborhood of $\mathcal{M}$: there is no need to pick a different, local smooth extension around each point. While this fact is not needed to establish results hereafter, it is convenient to shorten exposition; so much so that we typically think of it as the definition of a smooth map. See Section 3.10 for a discussion.

Proposition 3.31. *Let $\mathcal{M}$ and $\mathcal{M}'$ be embedded submanifolds of $\mathcal{E}$ and $\mathcal{E}'$. A map $F \colon \mathcal{M} \to \mathcal{M}'$ is smooth if and only if $F = \bar{F}|_{\mathcal{M}}$ where $\bar{F}$ is some smooth map from a neighborhood of $\mathcal{M}$ in $\mathcal{E}$ to $\mathcal{E}'$.*

In particular, a real-valued function $f \colon \mathcal{M} \to \mathbb{R}$ is smooth if and only if there exists a smooth extension $\bar{f}: U \to \mathbb{R}$ defined on a neighborhood U of $\mathcal{M}$ in $\mathcal{E}$ and which coincides with f on $\mathcal{M}$.

Definition 3.32. *A* scalar field *on a manifold $\mathcal{M}$ is a function $f: \mathcal{M} \to \mathbb{R}$. If f is a smooth function, we say it is a* smooth scalar field*. The set of smooth scalar fields on $\mathcal{M}$ is denoted by $\mathfrak{F}(\mathcal{M})$.*

Smoothness is preserved under composition, and also under linear combinations and products of maps when those are defined: see the exercises in the next section.

Exercise 3.33. *Give an example of an embedded submanifold $\mathcal{M}$ in a linear space $\mathcal{E}$ and a smooth function $f: \mathcal{M} \to \mathbb{R}$ for which there does not exist a smooth extension $\bar{f}: \mathcal{E} \to \mathbb{R}$ smooth on all of $\mathcal{E}$. Hint: use Example 3.19, or use Proposition 3.23 and consider removing a point from a simple manifold.*

3.4 The differential of a smooth map

Let $\bar{F}: U \subseteq \mathcal{E} \to \mathcal{E}'$ be a smooth function between two linear spaces, possibly restricted to an open set U. The differential of $\bar{F}$ at $x \in U$ is a linear map $\mathrm{D}\bar{F}(x): \mathcal{E} \to \mathcal{E}'$ defined by:

$$\mathrm{D}\bar{F}(x)[v] = \lim_{t \to 0} \frac{\bar{F}(x + tv) - \bar{F}(x)}{t}. \tag{3.26}$$

This tells us how $\bar{F}(x)$ changes when we push x along v. Applying this definition to a smooth map $F: \mathcal{M} \to \mathcal{M}'$ between two embedded submanifolds is problematic because $x + tv$ generally does not belong to $\mathcal{M}$, even for tiny nonzero values of t: F may not be defined there.

We can propose to resolve this issue in at least two ways:

1. Rely on Definition 3.14: $t \mapsto x + tv$ is nothing but a curve in $\mathcal{E}$ which passes through x with velocity v; we can use curves on $\mathcal{M}$ instead.
2. Rely on Definition 3.30: we can smoothly extend F and differentiate the extension instead.

As it turns out, these two approaches are equivalent. We start with the first one because it is more geometric: it gives the right picture of how things work on general manifolds. The second one is convenient for computation.

For any tangent vector $v \in \mathrm{T}_x\mathcal{M}$, there exists a smooth curve c on $\mathcal{M}$ passing through x with velocity v. Then, $t \mapsto F(c(t))$ itself defines a curve on $\mathcal{M}'$ passing through $F(x)$. That curve is smooth by composition. Thus, it passes through $F(x)$ with a certain velocity. By definition, that velocity is a tangent vector of $\mathcal{M}'$ at $F(x)$. We call this tangent vector the differential of F at x along v, denoted by $\mathrm{D}F(x)[v]$.

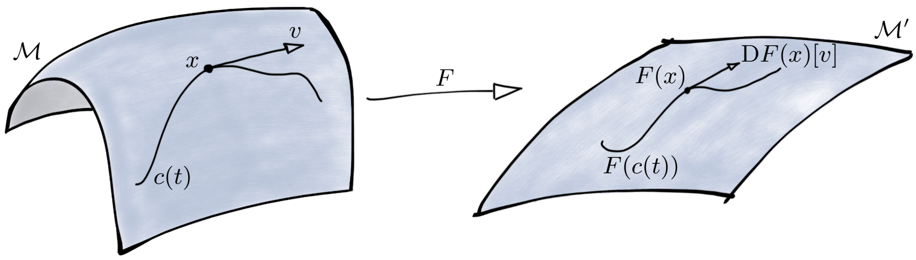


Figure 3.3 A smooth map $F\colon \mathcal{M} \to \mathcal{M}'$ pushes smooth curves c on $\mathcal{M}$ to smooth curves $F \circ c$ on $\mathcal{M}'$.

Definition 3.34. *The* differential *of $F\colon \mathcal{M} \to \mathcal{M}'$ at the point $x \in \mathcal{M}$ is the linear map $\mathrm{D}F(x)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_{F(x)}\mathcal{M}'$ defined by:*

$$\mathrm{D}F(x)[v] = \left.\frac{\mathrm{d}}{\mathrm{d}t}F(c(t))\right|_{t=0} = (F \circ c)'(0), \tag{3.27}$$

where c is a smooth curve on $\mathcal{M}$ passing through x at $t = 0$ with velocity v.

We must clarify two things: (a) that this definition does not depend on the choice of curve c (as many may satisfy the requirements), and (b) that $\mathrm{D}F(x)$ is indeed linear. To do so, we connect with the second approach.

Let $\mathcal{M}$ and $\mathcal{M}'$ be embedded submanifolds of $\mathcal{E}$ and $\mathcal{E}'$. Then, the smooth map $F\colon \mathcal{M} \to \mathcal{M}'$ admits a smooth extension $\bar{F}\colon U \to \mathcal{E}'$, where U is a neighborhood of $\mathcal{M}$ in $\mathcal{E}$ and $F = \bar{F}|_{\mathcal{M}}$. Observe that $F \circ c = \bar{F} \circ c$. The latter is a composition of functions between open subsets of linear spaces, hence the usual chain rule applies:

$$\begin{aligned}\mathrm{D}F(x)[v] &= \left.\frac{\mathrm{d}}{\mathrm{d}t}F(c(t))\right|_{t=0} \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\bar{F}(c(t))\right|_{t=0} = \mathrm{D}\bar{F}(c(0))[c'(0)] = \mathrm{D}\bar{F}(x)[v].\end{aligned} \tag{3.28}$$

This holds for all $v \in \mathrm{T}_x\mathcal{M}$. We summarize as follows.

Proposition 3.35. *With notation as above, $\mathrm{D}F(x) = \mathrm{D}\bar{F}(x)|_{\mathrm{T}_x\mathcal{M}}$.*

This proposition confirms that $\mathrm{D}F(x)$ is linear since $\mathrm{D}\bar{F}(x)$ is linear. It also shows that the definition by eq. (3.27) depends on c only through $c(0)$ and $c'(0)$, as required.

One may now wonder whether eq. (3.28) depends on the choice of smooth extension $\bar{F}$. It does not: that is clear from eq. (3.27). We can also verify it explicitly. Let $\hat{F}$ be another smooth extension of F. Then, for all smooth curves c with $c(0) = x$ and $c'(0) = v$ we have

$$\mathrm{D}\bar{F}(x)[v] = (\bar{F} \circ c)'(0) = (\hat{F} \circ c)'(0) = \mathrm{D}\hat{F}(x)[v].$$

Thus, the choice of smooth extension is inconsequential.

Example 3.36. *Given a real, symmetric matrix $A \in \mathrm{Sym}(d)$, the* Rayleigh quotient *at a nonzero vector $x \in \mathbb{R}^d$ is given by $\frac{x^\top A x}{x^\top x}$. Since this quotient is invariant under scaling of x, we may restrict our attention to unit-norm vectors. This yields a function on the sphere:*

$$f \colon \mathrm{S}^{d-1} \to \mathbb{R} \colon x \mapsto x^\top A x.$$

As we will gradually rediscover, the extreme points (maxima and minima) of f are tightly related to extremal eigenvectors of A. The function f can be smoothly extended to $\mathbb{R}^d$ by $\bar{f}(x) = x^\top A x$, hence f is smooth according to Definition 3.30. Using this smooth extension, we can also obtain an expression for its differential. Indeed, for all $v \in \mathbb{R}^d$,

$$\begin{aligned} \mathrm{D}\bar{f}(x)[v] &= \lim_{t \to 0} \frac{\bar{f}(x+tv) - \bar{f}(x)}{t} \\ &= \lim_{t \to 0} \frac{(x+tv)^\top A (x+tv) - x^\top A x}{t} \\ &= x^\top A v + v^\top A x \\ &= x^\top (A + A^\top) v = 2x^\top A v. \end{aligned}$$

Hence, Proposition 3.35 yields:

$$\mathrm{D}f(x)[v] = \mathrm{D}\bar{f}(x)[v] = 2x^\top A v$$

for all $v \in \mathrm{T}_x\mathrm{S}^{d-1} = \{v \in \mathbb{R}^d : x^\top v = 0\}$. Formally, $\mathrm{D}\bar{f}(x)$ is defined on all of $\mathbb{R}^d$ while $\mathrm{D}f(x)$ is only defined on $\mathrm{T}_x\mathrm{S}^{d-1}$.

Exercise 3.37. *For smooth maps $F_1, F_2 \colon \mathcal{M} \to \mathcal{E}'$ and real numbers a_1, a_2, show that $F \colon x \mapsto a_1 F_1(x) + a_2 F_2(x)$ is smooth and we have linearity:*

$$\mathrm{D}F(x) = a_1 \mathrm{D}F_1(x) + a_2 \mathrm{D}F_2(x).$$

Exercise 3.38. *For smooth maps $f \colon \mathcal{M} \to \mathbb{R}$ and $G \colon \mathcal{M} \to \mathcal{E}'$, show that $fG \colon x \mapsto f(x)G(x)$ is smooth from $\mathcal{M}$ to $\mathcal{E}'$ and we have a product rule:*

$$\mathrm{D}(fG)(x)[v] = \mathrm{D}f(x)[v] \cdot G(x) + f(x) \cdot \mathrm{D}G(x)[v].$$

(The dots $\cdot$ are only used to clarify the factors of the product visually.)

Exercise 3.39. *Let $F \colon \mathcal{M} \to \mathcal{M}'$ and $G \colon \mathcal{M}' \to \mathcal{M}''$ be smooth, where $\mathcal{M}, \mathcal{M}'$ and $\mathcal{M}''$ are embedded submanifolds of $\mathcal{E}, \mathcal{E}'$ and $\mathcal{E}''$ respectively. Show that composition preserves smoothness, that is, $G \circ F \colon x \mapsto G(F(x))$ is smooth. Also show that we have a chain rule:*

$$\mathrm{D}(G \circ F)(x)[v] = \mathrm{D}G(F(x))[\mathrm{D}F(x)[v]]. \tag{3.29}$$

Exercise 3.40. *Let $\mathcal{M}, \mathcal{M}', \mathcal{N}$ be three manifolds, and consider a smooth map $F \colon \mathcal{M} \times \mathcal{M}' \to \mathcal{N}$ (see Proposition 3.20 for the product manifold.) Show that*

$$\mathrm{D}F(x,y)[(u,v)] = \mathrm{D}\big(x \mapsto F(x,y)\big)(x)[u] + \mathrm{D}\big(y \mapsto F(x,y)\big)(y)[v],$$

where $(x,y) \in \mathcal{M} \times \mathcal{M}'$ and $(u,v) \in \mathrm{T}_{(x,y)}(\mathcal{M} \times \mathcal{M}') = \mathrm{T}_x\mathcal{M} \times \mathrm{T}_y\mathcal{M}'$ are *arbitrary. The notation $x \mapsto F(x, y)$ denotes the function from $\mathcal{M}$ to $\mathcal{N}$ obtained by fixing the second input of F to y.*

Exercise 3.41. *Let $\mathcal{M}$ be an embedded submanifold of a linear space $\mathcal{E}$, and let $\mathcal{N}$ be a subset of $\mathcal{M}$ defined by $\mathcal{N} = g^{-1}(0)$, where $g\colon \mathcal{M} \to \mathbb{R}^\ell$ is smooth and* $\operatorname{rank} \mathrm{D}g(x) = \ell$ *for all $x \in \mathcal{N}$. Show that $\mathcal{N}$ is itself an embedded submanifold of $\mathcal{E}$, of dimension* $\dim \mathcal{M} - \ell$*, with tangent spaces $\mathrm{T}_x\mathcal{N} = \ker \mathrm{D}g(x) \subset \mathrm{T}_x\mathcal{M}$. Here, we assume $\ell \geq 1$; see also Exercise 3.26. We call $\mathcal{N}$ an* embedded submanifold of $\mathcal{M}$*; see also Section 8.14.*

3.5 Vector fields and the tangent bundle

A map V which associates to each point $x \in \mathcal{M}$ a tangent vector at x is called a *vector field* on $\mathcal{M}$. For example, the gradient of a function $f\colon \mathcal{M} \to \mathbb{R}$ (still to be defined) is a vector field. In order to define a notion of *smooth* vector field, we need to present V as a map between manifolds. Since the range of V includes tangent vectors from all possible tangent spaces of $\mathcal{M}$, the first step is to introduce the *tangent bundle*: this is the *disjoint union* of all the tangent spaces of $\mathcal{M}$. By "disjoint" we mean that, for each tangent vector $v \in \mathrm{T}_x\mathcal{M}$, we retain the pair (x, v) rather than simply v. This is important to avoid ambiguity because some tangent vectors, seen as vectors in $\mathcal{E}$, may belong to more than one tangent space. For example, the zero vector belongs to all of them.

Definition 3.42. *The* tangent bundle *of a manifold $\mathcal{M}$ is the disjoint union of the tangent spaces of $\mathcal{M}$:*

$$\mathrm{T}\mathcal{M} = \{(x, v) : x \in \mathcal{M} \text{ and } v \in \mathrm{T}_x\mathcal{M}\}. \tag{3.30}$$

With some abuse of notation, for a tangent vector $v \in \mathrm{T}_x\mathcal{M}$, we sometimes conflate the notions of v and (x, v). We may write $(x, v) \in \mathrm{T}_x\mathcal{M}$, or even $v \in$ $\mathrm{T}\mathcal{M}$ *if it is clear from context that the* base *of v is x.*

The tangent bundle is a manifold.

Theorem 3.43. *If $\mathcal{M}$ is an embedded submanifold of $\mathcal{E}$, the tangent bundle $\mathrm{T}\mathcal{M}$ is an embedded submanifold of $\mathcal{E} \times \mathcal{E}$ of dimension $2 \dim \mathcal{M}$.*

Proof. For open submanifolds, the claim is clear: $\mathrm{T}_x\mathcal{M} = \mathcal{E}$ for each $x \in \mathcal{M}$, hence $\mathrm{T}\mathcal{M} = \mathcal{M} \times \mathcal{E}$. This is an open subset of $\mathcal{E} \times \mathcal{E}$, hence it is an open submanifold of that space.

Considering the other case, pick an arbitrary point $\bar{x} \in \mathcal{M}$ and let $h\colon U \to \mathbb{R}^k$ be a local defining function for $\mathcal{M}$ at $\bar{x}$, that is: U is a neighborhood of $\bar{x}$ in $\mathcal{E}$, h is smooth, $\mathcal{M} \cap U = \{x \in U : h(x) = 0\}$, and $\mathrm{D}h(\bar{x})\colon \mathcal{E} \to \mathbb{R}^k$ has rank k.

If needed, restrict the domain U to secure the property $\operatorname{rank} \mathrm{D}h(x) = k$ for all $x \in U$: this is always possible, see Lemma 3.74. Then, we can claim that

$\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x)$ for all $x \in \mathcal{M} \cap U$. Consequently, a pair $(x, v) \in U \times \mathcal{E}$ is in $\mathrm{T}\mathcal{M}$ if and only if it satisfies the following equations:

$$h(x) = 0 \qquad \text{and} \qquad \mathrm{D}h(x)[v] = 0.$$

Accordingly, define the smooth function $H : U \times \mathcal{E} \to \mathbb{R}^{2k}$ as:

$$H(x, v) = \begin{bmatrix} h(x) \\ \mathrm{D}h(x)[v] \end{bmatrix}.$$

The aim is to show that H is a local defining function for $\mathrm{T}\mathcal{M}$. We already have that $\mathrm{T}\mathcal{M} \cap (U \times \mathcal{E}) = H^{-1}(0)$. If we establish that $\mathrm{D}H(x, v)$ has rank $2k$ for all $(x, v) \in \mathrm{T}\mathcal{M} \cap (U \times \mathcal{E})$, we will have shown that $\mathrm{T}\mathcal{M}$ is an embedded submanifold of $\mathcal{E} \times \mathcal{E}$. Let us compute the differential of H (we know it exists since h is smooth):

$$\mathrm{D}H(x, v)[\dot{x}, \dot{v}] = \begin{bmatrix} \mathrm{D}h(x)[\dot{x}] \\ \mathcal{L}(x, v)[\dot{x}] + \mathrm{D}h(x)[\dot{v}] \end{bmatrix} = \begin{bmatrix} \mathrm{D}h(x) & 0 \\ \mathcal{L}(x, v) & \mathrm{D}h(x) \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix},$$

where $\mathcal{L}(x, v) : \mathcal{E} \to \mathcal{E}$ is some linear map which depends on both x and v (it involves the second derivative of h, but its specific form is irrelevant to us.) The block triangular form of $\mathrm{D}H(x, v)$ allows us to conclude that $\operatorname{rank} \mathrm{D}H(x, v) = \operatorname{rank} \mathrm{D}h(x) + \operatorname{rank} \mathrm{D}h(x) = 2k$, as required. Indeed, The rank is at most $2k$ because H maps into $\mathbb{R}^{2k}$, and the rank is at least $2k$ because the two diagonal blocks each have rank k. Since we can build such H on a neighborhood of any point in $\mathrm{T}\mathcal{M}$, we conclude that $\mathrm{T}\mathcal{M}$ is an embedded submanifold of $\mathcal{E} \times \mathcal{E}$. For the dimension, use $\mathrm{T}_{(x,v)}\mathrm{T}\mathcal{M} = \ker \mathrm{D}H(x, v)$ and the rank-nullity theorem to conclude that $\dim \mathrm{T}\mathcal{M} = \dim \mathrm{T}_{(x,v)}\mathrm{T}\mathcal{M} = 2 \dim \mathcal{E} - 2k = 2 \dim \mathcal{M}$. □

The topology we choose for $\mathrm{T}\mathcal{M}$ is the embedded submanifold topology, as in Definition 3.21. This is different from the so-called disjoint union topology, which we never use.

Since $\mathrm{T}\mathcal{M}$ is a manifold, we can now use Definition 3.30 to define smooth vector fields as particular smooth maps from $\mathcal{M}$ to $\mathrm{T}\mathcal{M}$. Be aware that some authors refer to smooth vector fields as vector fields.

Definition 3.44. *A* vector field *on a manifold $\mathcal{M}$ is a map $V : \mathcal{M} \to \mathrm{T}\mathcal{M}$ such that $V(x)$ is in $\mathrm{T}_x\mathcal{M}$ for all $x \in \mathcal{M}$. If V is a smooth map, we say it is a* smooth vector field. *The set of smooth vector fields is denoted by $\mathfrak{X}(\mathcal{M})$.*

A vector field on an embedded submanifold is smooth if and only if it is the restriction of a smooth vector field on a neighborhood of $\mathcal{M}$ in the embedding space.

Proposition 3.45. *For $\mathcal{M}$ an embedded submanifold of $\mathcal{E}$, a vector field V on $\mathcal{M}$ is smooth if and only if there exists a smooth vector field $\bar{V}$ on a neighborhood of $\mathcal{M}$ such that $V = \bar{V}|_{\mathcal{M}}$.*

Proof. Assume $V : \mathcal{M} \to \mathrm{T}\mathcal{M}$ is a smooth vector field on $\mathcal{M}$. Then, since $\mathrm{T}\mathcal{M}$ is an embedded submanifold of $\mathcal{E} \times \mathcal{E}$, by Proposition 3.31, there exists a neighborhood U of $\mathcal{M}$ in $\mathcal{E}$ and a smooth function $\bar{\bar{V}}: U \to \mathcal{E} \times \mathcal{E}$ such that $V = \bar{\bar{V}}|_{\mathcal{M}}$. Denote the two components of $\bar{\bar{V}}$ as $\bar{\bar{V}}(x) = (\bar{\bar{V}}_1(x), \bar{\bar{V}}_2(x))$. Of course, $\bar{\bar{V}}_1, \bar{\bar{V}}_2 : U \to \mathcal{E}$ are smooth. Define $\bar{V}(x) = (x, \bar{\bar{V}}_2(x))$: this is a smooth vector field on U such that $V = \bar{V}|_{\mathcal{M}}$. The other direction is clear. □

In closing, we note a useful identification for the tangent bundle of a product manifold $\mathcal{M} \times \mathcal{M}'$ (Proposition 3.20) which amounts to reordering parameters:

$$\begin{aligned} \mathrm{T}(\mathcal{M} \times \mathcal{M}') &= \{((x, x'), (v, v')) : x \in \mathcal{M}, v \in \mathrm{T}_x\mathcal{M}, x' \in \mathcal{M}', v' \in \mathrm{T}_{x'}\mathcal{M}'\} \\ &\equiv \{((x, v), (x', v')) : x \in \mathcal{M}, v \in \mathrm{T}_x\mathcal{M}, x' \in \mathcal{M}', v' \in \mathrm{T}_{x'}\mathcal{M}'\} \\ &= \mathrm{T}\mathcal{M} \times \mathrm{T}\mathcal{M}'. \end{aligned} \tag{3.31}$$

Exercise 3.46. *For $f \in \mathfrak{F}(\mathcal{M})$ and $V, W \in \mathfrak{X}(\mathcal{M})$, verify that the vector fields fV and $V + W$ are smooth, where we define $(fV)(x) = f(x)V(x)$ and also $(V + W)(x) = V(x) + W(x)$. For pointwise scaling, we purposefully write fV and not Vf. Later, we will give a different meaning to the notation Vf.* ⋆

3.6 Moving on a manifold: retractions

Given a point $x \in \mathcal{M}$ and a tangent vector $v \in \mathrm{T}_x\mathcal{M}$, we often need to move away from x along the direction v while remaining on the manifold: this is the basic operation of a gradient descent algorithm, and of essentially all optimization algorithms on manifolds. We can achieve this by following any smooth curve c on $\mathcal{M}$ such that $c(0) = x$ and $c'(0) = v$, but of course there exist many such curves. A *retraction* picks a particular curve for each possible $(x, v) \in \mathrm{T}\mathcal{M}$. Furthermore, this choice of curve depends smoothly on (x, v), in a sense we make precise using the fact that the tangent bundle $\mathrm{T}\mathcal{M}$ is a manifold.

Definition 3.47. *A* retraction *on a manifold $\mathcal{M}$ is a smooth map*

$$\mathrm{R} \colon \mathrm{T}\mathcal{M} \to \mathcal{M} \colon (x, v) \mapsto \mathrm{R}_x(v)$$

such that each curve $c(t) = \mathrm{R}_x(tv)$ satisfies $c(0) = x$ and $c'(0) = v$.

For $\mathcal{M}$ embedded in $\mathcal{E}$, smoothness of R is understood in the sense of Definition 3.30 for a map from $\mathrm{T}\mathcal{M}$ to $\mathcal{M}$, that is, R is smooth if and only if there exists a smooth map $\overline{\mathrm{R}}$ from a neighborhood of $\mathrm{T}\mathcal{M}$ in $\mathcal{E} \times \mathcal{E}$ into $\mathcal{E}$ such that $\mathrm{R} = \overline{\mathrm{R}}|_{\mathrm{T}\mathcal{M}}$.

Let us illustrate this concept through examples. Chapter 7 has more.

Example 3.48. *On a linear manifold, $\mathrm{R}_x(v) = x + v$ is a retraction.*

Example 3.49. *Let x be a point on the sphere S^{d-1} and let v be tangent at x, that is, $x^\top v = 0$. To move away from x along v while remaining on the sphere,* *one way is to take the step in* $\mathbb{R}^d$ *then to project back to the sphere:*

$$\mathrm{R}_x(v) \triangleq \frac{x+v}{\|x+v\|} = \frac{x+v}{\sqrt{1+\|v\|^2}}. \tag{3.32}$$

Consider the curve $c\colon \mathbb{R} \to \mathrm{S}^{d-1}$ *defined by:*

$$c(t) = \mathrm{R}_x(tv) = \frac{x+tv}{\sqrt{1+t^2\|v\|^2}}.$$

Evidently, $c(0) = x$. *Moreover, one can compute* $c'(0) = v$, *that is: locally around* x, *up to first order, the retraction curve moves along* v. *To verify that* R (3.32) *is smooth, check that* $\overline{\mathrm{R}}_x(v) \triangleq (x+v)/\sqrt{1+\|v\|^2}$ *is a smooth extension to all of* $\mathbb{R}^d \times \mathbb{R}^d$. *This is an example of a retraction based on* metric projection*: we study them in Section 5.12.*

Another reasonable choice is to move away from x *along a great circle:*

$$\mathrm{R}_x(v) \triangleq \cos(\|v\|)x + \frac{\sin(\|v\|)}{\|v\|}v, \tag{3.33}$$

with the usual convention $\sin(0)/0 = 1$. *Indeed, the curve*

$$c(t) = \mathrm{R}_x(tv) = \cos(t\|v\|)x + \frac{\sin(t\|v\|)}{\|v\|}v$$

traces out the great circle on S^{d-1} *passing through* x *at* $t = 0$ *with velocity* $c'(0) = v$. *With the right Riemannian metric (Section 3.7), such curves are* geodesics *(Section 5.8) and the retraction* (3.33) *is the* exponential map *(Section 10.2).*

It is also common to define retractions without referring to curves. To see how, let $\mathrm{R}_x\colon \mathrm{T}_x\mathcal{M} \to \mathcal{M}$ denote the restriction of a smooth map $\mathrm{R}\colon \mathrm{T}\mathcal{M} \to \mathcal{M}$ to the tangent space at x. By the chain rule, each curve $c(t) = \mathrm{R}_x(tv)$ satisfies

$$c(0) = \mathrm{R}_x(0) \qquad \text{and} \qquad c'(0) = \mathrm{DR}_x(0)[v].$$

Thus, R is a retraction exactly if, for all $(x, v) \in \mathrm{T}\mathcal{M}$, we have

1. $\mathrm{R}_x(0) = x$, and
2. $\mathrm{DR}_x(0)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ is the identity map: $\mathrm{DR}_x(0)[v] = v$.

This characterization of retractions is equivalent to Definition 3.47.

Sometimes, it is convenient to relax the definition of retraction to allow maps R that are defined only on an open subset of the tangent bundle, provided all zero vectors belong to its domain. For example, this is the case for the manifold of fixed-rank matrices (Section 7.5).

Exercise 3.50. *Let* $\mathcal{M}, \mathcal{M}'$ *be equipped with retractions* R, R'. *Show that the map* $\mathrm{R}''\colon \mathrm{T}(\mathcal{M} \times \mathcal{M}') \to \mathcal{M} \times \mathcal{M}'$ *defined by* $\mathrm{R}''_{(x,x')}(v, v') = (\mathrm{R}_x(v), \mathrm{R}'_{x'}(v'))$ *is a valid retraction for the product manifold* $\mathcal{M} \times \mathcal{M}'$.

3.7 Riemannian manifolds and submanifolds

It is convenient to equip each tangent space of the manifold $\mathcal{M}$ with an inner product (recall Definition 3.1). This is the key ingredient to define gradients in the next section. Since there are now many inner products (one for each point on the manifold), we distinguish them with a subscript. That said, it is common to omit the subscript when it is clear from context: we do so on occasion in later chapters.

Definition 3.51. *An inner product on* $\mathrm{T}_x\mathcal{M}$ *is a bilinear, symmetric, positive definite function* $\langle \cdot, \cdot \rangle_x : \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M} \to \mathbb{R}$. *It induces a norm for tangent vectors:* $\|u\|_x = \sqrt{\langle u, u \rangle_x}$. *A* metric *on* $\mathcal{M}$ *is a choice of inner product* $\langle \cdot, \cdot \rangle_x$ *for each* $x \in \mathcal{M}$.

Of particular interest are metrics which, in some sense, vary smoothly with x. To give a precise meaning to this requirement, the following definition builds upon the notions of smooth scalar and vector fields.

Definition 3.52. *A metric* $\langle \cdot, \cdot \rangle_x$ *on* $\mathcal{M}$ *is a* Riemannian metric *if it varies smoothly with* x*, in the sense that for all smooth vector fields* V, W *on* $\mathcal{M}$ *the function* $x \mapsto \langle V(x), W(x) \rangle_x$ *is smooth from* $\mathcal{M}$ *to* $\mathbb{R}$.

Definition 3.53. *A manifold with a Riemannian metric is a* Riemannian manifold.

A Euclidean space is a linear space $\mathcal{E}$ with an inner product $\langle \cdot, \cdot \rangle$ (the same at all points)—we call it the *Euclidean metric*. When $\mathcal{M}$ is an embedded submanifold of a Euclidean space $\mathcal{E}$, the tangent spaces of $\mathcal{M}$ are linear subspaces of $\mathcal{E}$. This suggests a particularly convenient way of defining an inner product on each tangent space: simply restrict the inner product of $\mathcal{E}$ to each one. The resulting metric on $\mathcal{M}$ is called the *induced metric*. As we now show, the induced metric is a Riemannian metric, leading to the notion of *Riemannian submanifold*.

Proposition 3.54. *Let* $\mathcal{M}$ *be an embedded submanifold of* $\mathcal{E}$*, and let* $\langle \cdot, \cdot \rangle$ *be the Euclidean metric on* $\mathcal{E}$*. Then, the metric on* $\mathcal{M}$ *defined at each* x *by restriction,* $\langle u, v \rangle_x = \langle u, v \rangle$ *for* $u, v \in \mathrm{T}_x\mathcal{M}$*, is a Riemannian metric.*

Proof. For any two smooth vector fields $V, W \in \mathfrak{X}(\mathcal{M})$, let $\bar{V}, \bar{W}$ be two smooth extensions of V, W to a neighborhood U of $\mathcal{M}$ in $\mathcal{E}$. Then, consider $g(x) = \langle V(x), W(x) \rangle_x$ (a function on $\mathcal{M}$) and let $\bar{g}(x) = \langle \bar{V}(x), \bar{W}(x) \rangle$ (a function on U). Clearly, $\bar{g}$ is smooth and $g = \bar{g}|_{\mathcal{M}}$. Hence, g is smooth. □

Definition 3.55. *Let* $\mathcal{M}$ *be an embedded submanifold of a Euclidean space* $\mathcal{E}$*. Equipped with the Riemannian metric obtained by restriction of the metric of* $\mathcal{E}$*, we call* $\mathcal{M}$ *a* Riemannian submanifold *of* $\mathcal{E}$*.*

This is arguably the most common type of Riemannian manifold in applications. ⋆ Be mindful that a Riemannian submanifold is not merely a submanifold with some Riemannian structure: the words single out a precise choice of metric.

Example 3.56. *Endow $\mathbb{R}^d$ with the standard metric $\langle u, v\rangle = u^\top v$ and consider the sphere $\mathrm{S}^{d-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$, embedded in $\mathbb{R}^d$. With the inherited metric $\langle u, v\rangle_x = \langle u, v\rangle = u^\top v$ on each tangent space $\mathrm{T}_x\mathrm{S}^{d-1}$, the sphere becomes a Riemannian submanifold of $\mathbb{R}^d$.*

Example 3.57. *Let $\mathcal{M}, \mathcal{M}'$ be Riemannian manifolds with metrics $\langle\cdot,\cdot\rangle^{\mathcal{M}}$ and $\langle\cdot,\cdot\rangle^{\mathcal{M}'}$ (respectively). Recall from Proposition 3.20 that the product $\mathcal{M}\times\mathcal{M}'$ is itself a manifold. The* product metric *$\langle\cdot,\cdot\rangle$ on $\mathcal{M}\times\mathcal{M}'$ is defined as follows. For $(u, u'), (v, v')$ in the tangent space $\mathrm{T}_{(x,x')}(\mathcal{M}\times\mathcal{M}')$,*

$$\langle (u,u'),(v,v')\rangle_{(x,x')} = \langle u, v\rangle_x^{\mathcal{M}} + \langle u', v'\rangle_{x'}^{\mathcal{M}'}.$$

It is an exercise to show that this is a Riemannian metric (see Exercise 3.73). With this metric, we call $\mathcal{M}\times\mathcal{M}'$ a Riemannian product manifold.

3.8 Riemannian gradients

Let $\mathcal{M}$ be a Riemannian manifold, that is, a manifold with a Riemannian metric. Given a smooth function $f\colon \mathcal{M}\to\mathbb{R}$, we are finally in a position to define its gradient.

Definition 3.58. *Let $f\colon \mathcal{M}\to\mathbb{R}$ be smooth on a Riemannian manifold $\mathcal{M}$. The* Riemannian gradient *of f is the vector field* $\mathrm{grad}f$ *on $\mathcal{M}$ uniquely defined by the following identities:*

$$\forall (x,v)\in\mathrm{T}\mathcal{M}, \qquad \mathrm{D}f(x)[v] = \langle v, \mathrm{grad}f(x)\rangle_x, \tag{3.34}$$

where $\mathrm{D}f(x)$ is as in Definition 3.34 and $\langle\cdot,\cdot\rangle_x$ is the Riemannian metric.

It is an exercise to show that (3.34) indeed uniquely determines $\mathrm{grad}f(x)$ for each x in $\mathcal{M}$, confirming $\mathrm{grad}f$ is well defined.

To work out the gradient of f, the preferred way is to obtain an expression for $\mathrm{D}f(x)[v]$ and to manipulate it until it takes the form $\langle v, \cdot\rangle_x$, where $\cdot$ is tangent at x. That yields the gradient by uniqueness. We discuss this more in Section 4.7. Alternatively, an indirect approach is through retractions as follows.

Proposition 3.59. *Let $f\colon \mathcal{M}\to\mathbb{R}$ be a smooth function on a Riemannian manifold $\mathcal{M}$ equipped with a retraction* R. *Then, for all $x\in\mathcal{M}$,*

$$\mathrm{grad}f(x) = \mathrm{grad}(f\circ\mathrm{R}_x)(0), \tag{3.35}$$

where $f\circ\mathrm{R}_x\colon \mathrm{T}_x\mathcal{M}\to\mathbb{R}$ is defined on a Euclidean space ($\mathrm{T}_x\mathcal{M}$ with inner product $\langle\cdot,\cdot\rangle_x$), hence its gradient is a "classical" gradient. See also Exercise 10.73 for the gradient of $f\circ\mathrm{R}_x$ away from the origin.

Proof. By the chain rule, for all tangent vectors $v\in\mathrm{T}_x\mathcal{M}$,

$$\mathrm{D}(f\circ\mathrm{R}_x)(0)[v] = \mathrm{D}f(\mathrm{R}_x(0))[\mathrm{DR}_x(0)[v]] = \mathrm{D}f(x)[v],$$

since $\mathrm{R}_x(0) = x$ and $\mathrm{DR}_x(0)$ is the identity map (these are the defining properties of retractions). Using the definition of gradient for both $f \circ \mathrm{R}_x$ and f we conclude that, for all $v \in \mathrm{T}_x\mathcal{M}$,

$$\langle \mathrm{grad}(f \circ \mathrm{R}_x)(0), v\rangle_x = \langle \mathrm{grad}f(x), v\rangle_x \,.$$

The claim follows by uniqueness of the gradient. $\square$

Say $\mathcal{M}$ is embedded in the Euclidean space $\mathcal{E}$ with Euclidean metric $\langle \cdot, \cdot \rangle$ (for now, that metric may or may not be related to the Riemannian metric on $\mathcal{M}$). Since f is smooth, it has a smooth extension $\bar{f}$ defined on a neighborhood of $\mathcal{M}$ in $\mathcal{E}$. The latter has a Euclidean gradient $\mathrm{grad}\bar{f}$. Combining (3.28) with (3.34), we find:

$$\begin{aligned} \forall (x, v) \in \mathrm{T}\mathcal{M}, \qquad \langle v, \mathrm{grad}f(x)\rangle_x &= \mathrm{D}f(x)[v] \\ &= \mathrm{D}\bar{f}(x)[v] = \langle v, \mathrm{grad}\bar{f}(x)\rangle. \end{aligned} \tag{3.36}$$

The core observation is: $\mathrm{T}_x\mathcal{M}$ is a subspace of $\mathcal{E}$, and $\mathrm{grad}\bar{f}(x)$ is a vector in $\mathcal{E}$; as such, the latter can be uniquely decomposed in $\mathcal{E}$ as

$$\mathrm{grad}\bar{f}(x) = \mathrm{grad}\bar{f}(x)_\parallel + \mathrm{grad}\bar{f}(x)_\perp,$$

with one component in $\mathrm{T}_x\mathcal{M}$ and the other orthogonal to $\mathrm{T}_x\mathcal{M}$, orthogonality being judged by the inner product of $\mathcal{E}$. Explicitly, $\mathrm{grad}\bar{f}(x)_\parallel$ is in $\mathrm{T}_x\mathcal{M}$ and

$$\forall v \in \mathrm{T}_x\mathcal{M}, \qquad \left\langle v, \mathrm{grad}\bar{f}(x)_\perp \right\rangle = 0.$$

As a result, we get from (3.36) that, for all (x, v) in $\mathrm{T}\mathcal{M}$,

$$\begin{aligned} \langle v, \mathrm{grad}f(x)\rangle_x &= \left\langle v, \mathrm{grad}\bar{f}(x)\right\rangle \\ &= \left\langle v, \mathrm{grad}\bar{f}(x)_\parallel + \mathrm{grad}\bar{f}(x)_\perp \right\rangle = \left\langle v, \mathrm{grad}\bar{f}(x)_\parallel \right\rangle. \end{aligned} \tag{3.37}$$

This relates the Riemannian gradient of f and the Euclidean gradient of $\bar{f}$.

Now further assume that $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$. Then, since $\mathrm{grad}\bar{f}(x)_\parallel$ is tangent at x and since the Riemannian metric is merely a restriction of the Euclidean metric to the tangent spaces, we find:

$$\forall (x, v) \in \mathrm{T}\mathcal{M}, \qquad \langle v, \mathrm{grad}f(x)\rangle_x = \langle v, \mathrm{grad}\bar{f}(x)_\parallel\rangle_x.$$

By identification, it follows that, for Riemannian submanifolds,

$$\mathrm{grad}f(x) = \mathrm{grad}\bar{f}(x)_\parallel. \tag{3.38}$$

In other words: to determine $\mathrm{grad}f$, first obtain an expression for the (classical) gradient of any smooth extension of f, then orthogonally project to the tangent spaces. This is a practical recipe because we often have access to a smooth extension. It motivates us to introduce orthogonal projectors.

Definition 3.60. *Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$ equipped with a Euclidean metric $\langle \cdot, \cdot \rangle$. The* orthogonal projector *to $\mathrm{T}_x\mathcal{M}$ is the linear map $\mathrm{Proj}_x \colon \mathcal{E} \to \mathcal{E}$ characterized by the following properties:*

1. *Range:* $\mathrm{im}(\mathrm{Proj}_x) = \mathrm{T}_x\mathcal{M}$;
2. *Projector:* $\mathrm{Proj}_x \circ \mathrm{Proj}_x = \mathrm{Proj}_x$;
3. *Orthogonal:* $\langle u - \mathrm{Proj}_x(u), v\rangle = 0$ *for all* $v \in \mathrm{T}_x\mathcal{M}$ *and* $u \in \mathcal{E}$.

For an open submanifold, $\mathrm{T}_x\mathcal{M} = \mathcal{E}$ hence Proj_x is the identity. The useful proposition below summarizes the above discussion.

Proposition 3.61. *Let* $\mathcal{M}$ *be a Riemannian submanifold of* $\mathcal{E}$ *endowed with the metric* $\langle\cdot,\cdot\rangle$ *and let* $f\colon \mathcal{M} \to \mathbb{R}$ *be a smooth function. The Riemannian gradient of* f *is given by*

$$\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad}\bar{f}(x)), \tag{3.39}$$

where $\bar{f}$ *is any smooth extension of* f *to a neighborhood of* $\mathcal{M}$ *in* $\mathcal{E}$.

Example 3.62. *We continue with the Rayleigh quotient from Example 3.36:* $f(x) = x^\top Ax$. *Equip* $\mathbb{R}^d$ *with the standard Euclidean metric* $\langle u, v\rangle = u^\top v$. *Then, using* $A = A^\top$, *for all* $v \in \mathbb{R}^d$,

$$\mathrm{D}\bar{f}(x)[v] = 2x^\top Av = \langle 2Ax, v\rangle .$$

Hence, by identification with Definition 3.58,

$$\mathrm{grad}\bar{f}(x) = 2Ax.$$

To get a notion of gradient for f *on* S^{d-1}, *we need to choose a Riemannian metric for* S^{d-1}. *One convenient choice is to turn* S^{d-1} *into a Riemannian submanifold of* $\mathbb{R}^d$ *by endowing it with the induced Riemannian metric. In that scenario, Proposition 3.61 suggests we should determine the orthogonal projectors of* S^{d-1}. *For the chosen Euclidean metric,*

$$\mathrm{T}_x\mathrm{S}^{d-1} = \{v \in \mathbb{R}^d : x^\top v = 0\} = \{v \in \mathbb{R}^d : \langle x, v\rangle = 0\}$$

is the orthogonal complement of x *in* $\mathbb{R}^d$. *Thus, orthogonal projection from* $\mathbb{R}^d$ *to that tangent space simply removes any component aligned with* x:

$$\mathrm{Proj}_x(u) = u - (x^\top u)x = (I_d - xx^\top)u. \tag{3.40}$$

It follows that the Riemannian gradient of f *on* S^{d-1} *is:*

$$\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad}\bar{f}(x)) = 2\left(Ax - (x^\top Ax)x\right).$$

Notice something quite revealing: for $x \in \mathrm{S}^{d-1}$,

$$\mathrm{grad} f(x) = 0 \qquad \iff \qquad Ax = \underbrace{(x^\top Ax)}_{\text{some scalar}} x.$$

In words: all points where the Riemannian gradient vanishes are eigenvectors of A. *Conversely: the gradient vanishes at all unit-norm eigenvectors of* A. *This basic observation is crucial to understand the behavior of optimization algorithms for* f *on* S^{d-1}.

Orthogonal projectors are self-adjoint (Definition 3.5).

Proposition 3.63. *Let* Proj_x *be the orthogonal projector from* $\mathcal{E}$ *to a linear subspace of* $\mathcal{E}$. *Then,* Proj_x *is self-adjoint. Explicitly:*

$$\forall u, v \in \mathcal{E}, \qquad \langle u, \mathrm{Proj}_x(v) \rangle = \langle \mathrm{Proj}_x(u), v \rangle,$$

where $\langle \cdot, \cdot \rangle$ *is the Euclidean metric on* $\mathcal{E}$.

Proof. From the properties of orthogonal projectors, for all $u, v \in \mathcal{E}$:

$$\begin{aligned} 0 &= \langle u - \mathrm{Proj}_x(u), \mathrm{Proj}_x(v) \rangle \\ &= \langle u, \mathrm{Proj}_x(v) \rangle - \langle \mathrm{Proj}_x(u), \mathrm{Proj}_x(v) \rangle \\ &= \langle u, \mathrm{Proj}_x(v) \rangle - \langle \mathrm{Proj}_x(u), v - (v - \mathrm{Proj}_x(v)) \rangle \\ &= \langle u, \mathrm{Proj}_x(v) \rangle - \langle \mathrm{Proj}_x(u), v \rangle + \underbrace{\langle \mathrm{Proj}_x(u), v - \mathrm{Proj}_x(v) \rangle}_{=0}. \end{aligned}$$

This concludes the proof. □

Exercise 3.64. *Show that* $\mathrm{grad} f(x)$ *is uniquely defined by* (3.34).

Exercise 3.65. *We noted that the relative interior of the simplex,*

$$\mathcal{M} = \Delta_+^{d-1} = \{x \in \mathbb{R}^d : x_1, \ldots, x_d > 0 \text{ and } x_1 + \cdots + x_d = 1\},$$

is an embedded submanifold of $\mathbb{R}^d$ *(Exercise 3.26). Its tangent spaces are:*

$$\mathrm{T}_x\mathcal{M} = \{v \in \mathbb{R}^d : v_1 + \cdots + v_d = 0\}.$$

Show that $\langle u, v \rangle_x = \sum_{i=1}^d \frac{u_i v_i}{x_i}$ *defines a Riemannian metric on* $\mathcal{M}$. *This is called the* Fisher–Rao metric. *Consider a smooth function* $f \colon \mathcal{M} \to \mathbb{R}$ *and a smooth extension* $\bar{f}$ *on a neighborhood of* $\mathcal{M}$ *in* $\mathbb{R}^d$ *(equipped with the canonical Euclidean metric). Give an expression for* $\mathrm{grad} f(x)$ *in terms of* $\mathrm{grad} \bar{f}(x)$.

Note: This exercise provides an example where $\mathrm{grad} f(x)$ *is* not *simply the projection of* $\mathrm{grad} \bar{f}(x)$ *to* $\mathrm{T}_x\mathcal{M}$. *This is because, while* $\mathcal{M}$ *is an embedded submanifold of* $\mathbb{R}^d$ *and it is a Riemannian manifold, it is* not *a Riemannian submanifold of* $\mathbb{R}^d$.

Exercise 3.66. *Let* $\mathcal{E}$ *be a Euclidean space, and let* $\mathcal{L}(\mathcal{E}, \mathcal{E})$ *denote the set of linear maps from* $\mathcal{E}$ *into itself: this is a linear space. If* $\mathcal{E}$ *is identified with* $\mathbb{R}^d$, *then* $\mathcal{L}(\mathcal{E}, \mathcal{E})$ *is identified with* $\mathbb{R}^{d \times d}$. *For* $\mathcal{M}$ *an embedded submanifold of* $\mathcal{E}$, *show that the map (recall Definition 3.60)*

$$\mathrm{Proj} \colon \mathcal{M} \to \mathcal{L}(\mathcal{E}, \mathcal{E}) \colon x \mapsto \mathrm{Proj}_x$$

is smooth. Deduce that if $f \colon \mathcal{M} \to \mathbb{R}$ *is a smooth function on a Riemannian submanifold of* $\mathcal{E}$ *then the Riemannian gradient of* f *is a smooth vector field.*

Note: It is true in general that the Riemannian gradient of a smooth function is smooth, but with the tools we have developed so far the proof is substantially simpler for Riemannian submanifolds. See Section 3.9 for the more general case.

Exercise 3.67. *For a smooth function $f\colon \mathcal{M} \times \mathcal{M}' \to \mathbb{R}$ on a Riemannian product manifold (see Example 3.57), show that*

$$\mathrm{grad} f(x, y) = \Big(\mathrm{grad}\big(x \mapsto f(x, y)\big)(x), \mathrm{grad}\big(y \mapsto f(x, y)\big)(y)\Big),$$

where $x \mapsto f(x, y)$ is the function from $\mathcal{M}$ to $\mathbb{R}$ obtained from f by fixing the second input to y, and similarly for $y \mapsto f(x, y)$.

3.9 Local frames*

This section introduces a technical tool which proves useful in certain proofs. A reader focused on Riemannian submanifolds can safely skip it. We show that embedded submanifolds admit *local frames* (as defined below) around all points. As a first application, we use local frames to show that the gradient of a smooth function is a smooth vector field. Contrast this with Exercise 3.66 which is restricted to Riemannian submanifolds. The section goes on to sketch a proof of a more general result known as the *musical isomorphism*.

Definition 3.68. *Given a point x on a manifold $\mathcal{M}$ of dimension n, a* local frame *around x is a set of smooth vector fields $W_1, \ldots, W_n$ defined on a neighborhood of x in $\mathcal{M}$ such that, for all y in that neighborhood, the vectors $W_1(y), \ldots, W_n(y)$ form a basis for the tangent space $\mathrm{T}_y\mathcal{M}$.*

Proposition 3.69. *Let $\mathcal{M}$ be an embedded submanifold of a linear space $\mathcal{E}$. There exists a local frame around any $x \in \mathcal{M}$.*

Proof. Let $\mathcal{E}$ have dimension d and let $\mathcal{M}$ have dimension $n = d - k$. Theorem 3.12 provides us with a neighborhood U of x in $\mathcal{E}$, an open set V in $\mathbb{R}^d$ and a diffeomorphism $F\colon U \to V$ such that $\mathcal{M} \cap U = F^{-1}(E \cap V)$, where $E = \{y \in \mathbb{R}^d : y_{n+1} = \cdots = y_d = 0\}$. The set $\mathcal{U} = \mathcal{M} \cap U$ is a neighborhood of x on $\mathcal{M}$. We aim to build a local frame on $\mathcal{U}$ using F. The proof echoes that of Theorem 3.15.

Let $e_1, \ldots, e_d$ denote the columns of the identity matrix of size d. Fix an arbitrary $y \in \mathcal{U}$. The point $F(y)$ is in $E \cap V$. For each i in $1, \ldots, n$, consider the curve

$$\gamma_i(t) = F(y) + te_i.$$

Notice that $\gamma_i(t)$ remains in $E \cap V$ for t close to zero. Therefore,

$$c_i(t) = F^{-1}(\gamma_i(t))$$

is a smooth curve which remains in $\mathcal{U}$ for t close to zero. In particular, $c_i(0) = y$. Therefore, $c_i'(0)$ is a tangent vector to $\mathcal{M}$ at y. As a result, we may define vector fields $W_1, \ldots, W_n$ on $\mathcal{U}$ as

$$W_i(y) = c_i'(0) = \big(t \mapsto F^{-1}(F(y) + te_i)\big)'(0).$$

These vector fields are smooth since F and F^{-1} are smooth. It remains to verify that they form a local frame. To this end, use the chain rule to see that

$$W_i(y) = \mathrm{D}F^{-1}(\gamma_i(0))[\gamma_i'(0)] = \mathrm{D}F^{-1}(F(y))[e_i]. \tag{3.41}$$

Since F is a diffeomorphism, we know that $\mathrm{D}F^{-1}(F(y)) = (\mathrm{D}F(y))^{-1}$. In particular, $\mathrm{D}F^{-1}(F(y))$ is invertible. It then follows from linear independence of $e_1, \ldots, e_n$ in $\mathbb{R}^d$ that $W_1(y), \ldots, W_n(y)$ are linearly independent in $\mathrm{T}_y\mathcal{M}$. This holds for all y in $\mathcal{U}$, which is a neighborhood of x in $\mathcal{M}$, concluding the proof.

(We have constructed a rather special local frame called a *coordinate frame*: compare with Sections 8.3 and 8.8.) □

On some manifolds, it is possible to find a *global frame*, that is, a set of smooth vector fields defined on the whole manifold and which provide a basis for all tangent spaces. Such manifolds are called *parallelizable*. For example, $\mathbb{R}^n$ and the circle S^1 are parallelizable. However, the two-dimensional sphere S^2 is not. The obstruction comes from the famous *Hairy Ball Theorem*, which implies that if W is a smooth vector field on S^2 then $W(x) = 0$ for some x. Therefore, any pair of smooth vector fields W_1, W_2 on S^2 fails to provide a basis for at least one of the tangent spaces.

Proposition 3.69 allows us to prove the following statement for embedded submanifolds equipped with a Riemannian metric.

Proposition 3.70. *Let $f\colon \mathcal{M} \to \mathbb{R}$ be a smooth function on a Riemannian manifold $\mathcal{M}$. The gradient vector field* $\mathrm{grad} f$ *is a smooth vector field on $\mathcal{M}$.*

Proof. Pick any point $x \in \mathcal{M}$ and a local frame $W_1, \ldots, W_n$ defined on a neighborhood $\mathcal{U}$ of x in $\mathcal{M}$, where $\dim \mathcal{M} = n$. By the properties of local frames, there exist unique functions $g_1, \ldots, g_n \colon \mathcal{U} \to \mathbb{R}$ such that

$$\mathrm{grad} f(y) = g_1(y)W_1(y) + \cdots + g_n(y)W_n(y)$$

for all $y \in \mathcal{U}$. If $g_1, \ldots, g_n$ are smooth, then $\mathrm{grad} f$ is smooth on $\mathcal{U}$. Since $\mathcal{U}$ is a neighborhood of an arbitrary point, showing so is sufficient to prove that $\mathrm{grad} f$ is smooth. To show that each g_i is indeed smooth, consider the following linear system which defines them: taking the inner product of the above identity with each of the local frame fields against the Riemannian metric yields:

$$\begin{bmatrix} \langle W_1(y), W_1(y)\rangle_y & \cdots & \langle W_n(y), W_1(y)\rangle_y \\ \vdots & & \vdots \\ \langle W_1(y), W_n(y)\rangle_y & \cdots & \langle W_n(y), W_n(y)\rangle_y \end{bmatrix} \begin{bmatrix} g_1(y) \\ \vdots \\ g_n(y) \end{bmatrix}$$
$$= \begin{bmatrix} \langle \mathrm{grad} f(y), W_1(y)\rangle_y \\ \vdots \\ \langle \mathrm{grad} f(y), W_n(y)\rangle_y \end{bmatrix} = \begin{bmatrix} \mathrm{D}f(y)[W_1(y)] \\ \vdots \\ \mathrm{D}f(y)[W_n(y)] \end{bmatrix}.$$

The matrix of this system is invertible for all y in $\mathcal{U}$ and depends smoothly on y. Likewise, the right-hand side depends smoothly on y (consider extensions).

Hence, so does the solution of the system (this can be seen from Cramer's rule for linear systems). $\square$

This last proposition is a corollary of a far more general fact. A *smooth one-form* on a manifold $\mathcal{M}$ is a linear map $X\colon \mathfrak{X}(\mathcal{M}) \to \mathfrak{F}(\mathcal{M})$ which transforms a smooth vector field V into a smooth scalar field $X(V)$ and such that $X(gV) = gX(V)$ for all $g\colon \mathcal{M} \to \mathbb{R}$ (we say X is $\mathfrak{F}(\mathcal{M})$*-linear*). For example, the differential $\mathrm{D}f$ of a smooth function $f\colon \mathcal{M} \to \mathbb{R}$ is a smooth one-form defined by $\mathrm{D}f(V)(x) = \mathrm{D}f(x)[V(x)]$. If $\mathcal{M}$ is a Riemannian manifold with metric $\langle \cdot, \cdot \rangle$, we can create a smooth one-form using any smooth vector field. Indeed, given $U \in \mathfrak{X}(\mathcal{M})$, let $X(V) = \langle U, V \rangle$ where $\langle U, V \rangle$ denotes the function $x \mapsto \langle U(x), V(x) \rangle_x$ on $\mathcal{M}$. In fact, each smooth one-form corresponds to a smooth vector field in this way, and vice versa. This correspondence through the Riemannian metric is called the *musical isomorphism*. Proposition 3.70 is a corollary of that fact: it establishes the correspondence between $\mathrm{D}f$ and $\mathrm{grad} f$.

Proposition 3.71. *Let $\mathcal{M}$ be a Riemannian manifold with metric $\langle \cdot, \cdot \rangle$. If $X\colon \mathfrak{X}(\mathcal{M}) \to \mathfrak{F}(\mathcal{M})$ is a smooth one-form, then there exists a unique smooth vector field $U \in \mathfrak{X}(\mathcal{M})$ such that $X(V) = \langle U, V \rangle$ for all $V \in \mathfrak{X}(\mathcal{M})$.*

Proof sketch. Let X be a smooth one-form. It is somewhat technical to show the following property: given $x \in \mathcal{M}$,

$$V(x) = 0 \qquad \Longrightarrow \qquad X(V)(x) = 0. \tag{3.42}$$

That property is a consequence of $\mathfrak{F}(\mathcal{M})$-linearity, as can be shown using local frames and tools developed later in Section 5.6: see Proposition 5.21 in particular. In the proof sketch here, we simply assume that (3.42) holds. As an example, the property is clear for $X = \mathrm{D}f$.

An important consequence of (3.42) is that $X(V)(x)$ depends on V only through $V(x)$, that is, the dependence of $X(V)$ on V is pointwise. Indeed, for some point $x \in \mathcal{M}$, consider any two smooth vector fields $V_1, V_2 \in \mathfrak{X}(\mathcal{M})$ such that $V_1(x) = V_2(x)$. Then, $(V_1 - V_2)(x) = 0$ and it follows from (3.42) and from linearity of X that $0 = X(V_1 - V_2)(x) = X(V_1)(x) - X(V_2)(x)$. Thus,

$$V_1(x) = V_2(x) \qquad \Longrightarrow \qquad X(V_1)(x) = X(V_2)(x),$$

as claimed.

Therefore, $X(V)(x)$ depends on V only through $V(x)$, linearly. It is then a simple fact about linear functions on the Euclidean space $\mathrm{T}_x\mathcal{M}$ that there exists a unique vector in $\mathrm{T}_x\mathcal{M}$ which we denote by $U(x)$ such that

$$X(V)(x) = \langle U(x), V(x) \rangle_x .$$

As this holds for all $x \in \mathcal{M}$, we deduce that there exists a unique vector field U on $\mathcal{M}$ such that

$$X(V) = \langle U, V \rangle .$$

It remains to show that U is smooth.

To show that U is smooth, we merely need to show that it is smooth around each point. Given $x \in \mathcal{M}$, let $W_1, \ldots, W_n \in \mathfrak{X}(\mathcal{U})$ be a local frame on a neighborhood $\mathcal{U}$ of x, as provided by Proposition 3.69. Again using technical results described in Section 5.6 (see Lemma 5.26 in particular), one can show that, possibly at the expense of replacing $\mathcal{U}$ with a smaller neighborhood of x, there exist smooth vector fields $\tilde{W}_1, \ldots, \tilde{W}_n \in \mathfrak{X}(\mathcal{M})$ such that $\tilde{W}_i|_{\mathcal{U}} = W_i$ for all i. We need this because we do not know how to apply X to W_i but we can apply X to $\tilde{W}_i$; notice that for $X = \mathrm{D}f$ it is clear how to operate on locally defined vector fields, hence this step was not needed to establish Proposition 3.70. By the properties of local frames, there exist unique functions $g_1, \ldots, g_n \colon \mathcal{U} \to \mathbb{R}$ such that

$$U|_{\mathcal{U}} = g_1 W_1 + \cdots + g_n W_n.$$

To show that U is smooth around x, it is sufficient to show that the functions $g_1, \ldots, g_n$ are smooth around x. To this end, observe that

$$\begin{aligned} X(\tilde{W}_j)|_{\mathcal{U}} &= \left\langle U, \tilde{W}_j \right\rangle|_{\mathcal{U}} \\ &= \left\langle U|_{\mathcal{U}}, \tilde{W}_j|_{\mathcal{U}} \right\rangle \\ &= \sum_{i=1}^{n} g_i \left\langle W_i, W_j \right\rangle. \end{aligned}$$

Since X is a smooth one-form, we know that $X(\tilde{W}_j)$ is a smooth function on all of $\mathcal{M}$. Therefore, it is certainly smooth when restricted to $\mathcal{U}$. We can deduce that the functions $g_1, \ldots, g_n$ are smooth with the same reasoning as in the proof of Proposition 3.70. This concludes the proof sketch. □

One-forms are also called *cotangent vector fields*. The musical isomorphism is also called the *tangent–cotangent isomorphism*.

Exercise 3.72. *Let $\mathcal{M}$ be a Riemannian manifold. Show that for all $x \in \mathcal{M}$ there exists an* orthonormal local frame*, that is, a local frame $W_1, \ldots, W_n$ defined on a neighborhood $\mathcal{U}$ of x with the additional property that*

$$\forall y \in \mathcal{U}, \qquad \langle W_i(y), W_j(y) \rangle_y = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Hint: apply the Gram–Schmidt procedure to a local frame and check that its smoothness is preserved.

Exercise 3.73. *Verify that the product metric defined in Example 3.57 is indeed a Riemannian metric.*

3.10 Notes and references

The main sources for this chapter are [AMS08, Lee12, Lee18, O'N83]. More comprehensive reminders of topology, linear algebra and calculus can be found in [Lee12, App. A, B, C].

The treatment given here is restrictive but fully compatible with the general treatment of differential and Riemannian geometry found in those references and also presented in Chapter 8. Explicitly:

1. Every subset $\mathcal{M}$ of a linear space $\mathcal{E}$ which we call an embedded submanifold following Definition 3.10 is a smooth manifold in the sense of Definition 8.21, that is: $\mathcal{M}$ admits a maximal atlas whose associated topology fulfills the usual conditions.
2. There is a unique such maximal atlas whose topology matches the topology in Definition 3.21: we always mean to use that atlas.
3. A subset $\mathcal{M}$ of a linear space $\mathcal{E}$ is an embedded submanifold of $\mathcal{E}$ in the sense of Definition 3.10 if and only if it is an embedded submanifold of $\mathcal{E}$ in the general sense of Definition 8.73.
4. A map between two embedded submanifolds of linear spaces is smooth in the sense of Definition 3.30 if and only if it is smooth in the general sense of Definition 8.5, that is, through charts.
5. The tangent spaces in Definition 3.14 correspond to the general notion of tangent spaces in Definition 8.33 through the standard identification mentioned around eq. (8.26).

See Sections 8.3 and 8.14 for some proofs and references. All other tools constructed here (tangent bundles, vector fields, retractions, Riemannian metrics, etc.) are compatible with their general counterparts constructed in the other sections of Chapter 8.

In Definition 3.10, we require the differential of the local defining function h around x to have rank k at x. The next lemma shows that it would be equivalent to require $\mathrm{D}h$ to have rank k at all points of the domain of h, possibly after reducing that domain [Lee12, Prop. 4.1].

Lemma 3.74. *Let U be a neighborhood of x in $\mathcal{E}$. If $h\colon U \to \mathbb{R}^k$ is smooth and $\operatorname{rank} \mathrm{D}h(x) = k$, one can always restrict the domain U to a possibly smaller neighborhood U' of x such that $\operatorname{rank} \mathrm{D}h(x') = k$ for all x' in U'.*

Proof. The set $U' \subseteq U$ of points x' where $\mathrm{D}h(x')$ has rank k is an open set in $\mathcal{E}$. Indeed, let $A(x') \in \mathbb{R}^{k\times d}$ be the matrix representing $\mathrm{D}h(x')$ in some basis of $\mathcal{E}$, with $d = \dim \mathcal{E}$. Consider the following function on U: $g(x') = \det(A(x')A(x')^\top)$. Notice that $U' = U \backslash g^{-1}(0)$. We know that $g^{-1}(0)$ is closed because it is the preimage of the closed set $\{0\}$ for the continuous function g. Hence, U' is open. By assumption, x is in U'. Thus, it suffices to restrict the domain of h to U'. $\square$

The rank condition in Definition 3.10 is key. Indeed, contrast this with the following fact: *any* closed subset of a linear space $\mathcal{E}$ is the zero-set of a smooth function from $\mathcal{E}$ to $\mathbb{R}$ [Lee12, Thm. 2.29] (and $\mathcal{E}$ can be replaced by a manifold in this statement). For example, there exists a smooth function $h \colon \mathbb{R}^2 \to \mathbb{R}$ such that $h^{-1}(0)$ is a square in $\mathbb{R}^2$. Of course, a square is not smoothly embedded in $\mathbb{R}^2$ due to its corners, so we deduce that $\mathrm{D}h$ must have non-maximal rank at certain points on the square.

Moreover, it is not sufficient for the rank to be constant on the set. Here is an example: consider $h \colon \mathbb{R}^2 \to \mathbb{R}$ defined by $h(x) = (x_1 x_2)^2$. Its zero-set is a cross in $\mathbb{R}^2$. Yet, $\mathrm{D}h(x) = \begin{bmatrix} 2x_1x_2^2, & 2x_1^2x_2 \end{bmatrix}$ has constant rank on the cross: it is zero everywhere. Thus, we see that in order to exclude pathological sets such as this cross it is not sufficient to ask for the rank of $\mathrm{D}h(x)$ to be constant along the zero-set of h. However, it is sufficient to require a constant (possibly not maximal) rank *on a neighborhood* of the zero-set (Proposition 8.77). The above h fails that test.

Some embedded submanifolds of $\mathcal{E}$ with $\dim \mathcal{M} < \dim \mathcal{E}$ cannot be defined with a single defining function. Indeed, if $\mathcal{M}$ is the zero-set of a local defining function $h \colon U \subseteq \mathcal{E} \to \mathbb{R}$, then $\mathcal{M}$ is *orientable* in $\mathcal{E}$ [Lee12, Prop. 15.23]. Yet, one can construct an *open Möbius band* as a non-orientable embedded submanifold of dimension two in $\mathbb{R}^3$. Thus, that manifold cannot be defined using a single defining function.

Proposition 3.31 states that any smooth map between embedded submanifolds of linear spaces can be smoothly extended to a neighborhood of its (co)domain, and vice versa. This follows from the *tubular neighborhood theorem* found for example in [Lee12, Thm. 6.24] and [Lee18, Thm. 5.25], as shown in Propositions 8.79 and 8.80 for the general case of embedded submanifolds of manifolds.

Thus, we could also use Proposition 3.31 as the definition of smooth maps. This is indeed practical in many situations, and this is why we introduced that result early on. However, adopting this as our definition would make it harder to prove, for example, Proposition 3.70. This is because it would require one to exhibit a smooth extension around the whole manifold, as opposed to merely exhibiting a smooth extension around each point of the manifold.

For the special case of a smooth function $f \colon \mathcal{M} \to \mathbb{R}$ where $\mathcal{M}$ is embedded *and closed* in $\mathcal{E}$, it is also possible to smoothly extend f to *all* of $\mathcal{E}$. Indeed, by [Lee12, Prop. 5.5], $\mathcal{M}$ is *properly embedded* in $\mathcal{E}$ if and only if it is closed, and smooth functions on properly embedded submanifolds can be globally smoothly extended [Lee12, Lem. 5.34, Exercise 5-18]: this result (and others referenced above) relies on *partitions of unity* [Lee12, pp40–47]. Importantly, this is *not* generally true for manifolds that are merely embedded (Exercise 3.33).

Definition 3.55 restricts the notion of Riemannian submanifolds of $\mathcal{E}$ to *embedded* submanifolds of $\mathcal{E}$. This is compatible with O'Neill, who reserves the word submanifold for embedded submanifolds [O'N83, pp19, 57]. Certain authors adopt a more general definition, also allowing an *immersed* submanifold (Definition 8.72) to be called a Riemannian submanifold. This is the case of Lee for example [Lee18, p15].

Local frames (Definition 3.68) are discussed in [Lee12, pp177–179]. Likewise, for more on the musical isomorphism (Proposition 3.71), see [Lee12, pp341–343] and [O'N83, Prop. 3.10].

Much of the theory in this book could be constructed with weaker smoothness requirements. Specifically, it is possible to define embedded submanifolds of linear spaces with local defining functions which are differentiable only p times (class C^p instead of C^∞). Likewise, we could work with functions on manifolds which have only limited differentiability properties (see Remark 8.6). One reference among many on this topic is the book by Borwein and Zhu [BZ05, §7].

4 First-order optimization algorithms

In this chapter, we consider a first algorithm to solve problems of the form

$$\min_{x \in \mathcal{M}} f(x), \tag{4.1}$$

where $\mathcal{M}$ is a (smooth) manifold and $f\colon \mathcal{M} \to \mathbb{R}$ is a smooth function called the *cost function* or *objective function*. Discussions in this chapter apply for general manifolds: embedded submanifolds as defined in the previous chapter form one class of examples, and we detail other kinds of manifolds in Chapters 8 and 9.

A *(global) minimizer* or *(global) optimizer* for (4.1) is a point $x \in \mathcal{M}$ such that $f(x) \leq f(y)$ for all $y \in \mathcal{M}$. Defining this notion merely requires $\mathcal{M}$ to be a set and f to be a function: their smoothness is irrelevant. Minimizers may not exist, in which case it would be more appropriate to write (4.1) as $\inf_{x \in \mathcal{M}} f(x)$. Minimizers may also not be unique.

While it is typically our goal to compute a global minimizer, this goal is generally out of reach. A more realistic (though still non-trivial) goal is to aim for a *local minimizer* or *local optimizer*, that is, a point $x \in \mathcal{M}$ such that $f(x) \leq f(y)$ for all y in a neighborhood of x in $\mathcal{M}$. In other words: a local minimizer appears to be optimal when compared only to its immediate surroundings. Likewise, a *strict local minimizer* satisfies $f(x) < f(y)$ for all $y \neq x$ in some neighborhood around x. Recall that a neighborhood of x in $\mathcal{M}$ is an open subset of $\mathcal{M}$ which contains x. Hence, the notion of local minimizer relies on the topology of $\mathcal{M}$. For embedded submanifolds, we defined the topology in Definition 3.21. Just as for global minimizers, it does not rely on smoothness of either $\mathcal{M}$ or f.

Thus, importantly,

The problem we set out to solve is defined independently of the smooth structures we impose.

Yet, smoothness plays a crucial role in helping us solve that problem. As we discuss below, the notions of retraction and gradient afford us efficient means of moving on the manifold while making progress toward our goal. In this sense, the Riemannian geometry we impose on the problem is entirely ours to choose, and an integral part of our responsibilities as algorithm designer.

Since optimization algorithms generate sequences of points on $\mathcal{M}$, it is important to define terms pertaining to *convergence*. These are phrased in terms of the topology on $\mathcal{M}$.

Definition 4.1. *Consider a sequence S of points $x_0, x_1, x_2, \ldots$ on a manifold $\mathcal{M}$. Then,*

1. *A point $x \in \mathcal{M}$ is a* limit *of S if, for every neighborhood $\mathcal{U}$ of x in $\mathcal{M}$, there exists an integer K such that $x_K, x_{K+1}, x_{K+2}, \ldots$ are in $\mathcal{U}$. The topology of a manifold is* Hausdorff *(see Section 8.2), hence a sequence has at most one limit. If x is the limit, we write $\lim_{k\to\infty} x_k = x$ or $x_k \to x$ and we say the sequence* converges *to x.*
2. *A point $x \in \mathcal{M}$ is an* accumulation point *of S if it is the limit of a* subsequence *of S, that is, if every neighborhood $\mathcal{U}$ of x in $\mathcal{M}$ contains an infinite number of elements of S.*

This chapter focuses on Riemannian gradient descent. With just one additional geometric tool, namely, the notion of *vector transport* or *transporter* introduced in Section 10.5, a number of other first-order optimization algorithms can be addressed, including Riemannian versions of nonlinear conjugate gradients and BFGS: see Section 4.9 for pointers.

Exercise 4.2. *Give an example of a sequence that has no limit. Give an example of a sequence that has a single accumulation point yet no limit. Give an example of a sequence that has two distinct accumulation points. Show that if a sequence converges to x, then all of its accumulation points are equal to x. Now consider the particular case of $\mathcal{M}$ an embedded submanifold of a linear space $\mathcal{E}$. Show that a sequence on $\mathcal{M}$ may have a limit in $\mathcal{E}$ yet no limit in $\mathcal{M}$. Argue that this cannot happen if $\mathcal{M}$ is closed in $\mathcal{E}$.*

4.1 A first-order Taylor expansion on curves

Optimization algorithms move from point to point on a manifold by following smooth curves. In order to analyze these algorithms, we need to understand how the cost function varies along those curves. In $\mathbb{R}^n$ for example, we could be interested in how $f(x + tv)$ varies as a function of t close to $t = 0$. The tool of choice for this task is a Taylor expansion. We now apply this concept to manifolds.

Let $c\colon I \to \mathcal{M}$ be a smooth curve on $\mathcal{M}$ with $c(0) = x$ and $c'(0) = v$, where I is an open interval of $\mathbb{R}$ around $t = 0$. Evaluating f along this curve yields a real function:

$$g\colon I \to \mathbb{R}\colon t \mapsto g(t) = f(c(t)).$$

Since $g = f \circ c$ is smooth by composition and maps real numbers to real numbers, it admits a Taylor expansion:

$$g(t) = g(0) + tg'(0) + O(t^2).$$

Clearly, $g(0) = f(x)$. Furthermore, by the chain rule,

$$g'(t) = \mathrm{D}f(c(t))[c'(t)] = \left\langle \mathrm{grad} f(c(t)), c'(t) \right\rangle_{c(t)},$$

so that $g'(0) = \left\langle \mathrm{grad} f(x), v \right\rangle_x$. Overall, we get this Taylor expansion:

$$f(c(t)) = f(x) + t \left\langle \mathrm{grad} f(x), v \right\rangle_x + O(t^2). \tag{4.2}$$

In particular, if the curve is obtained by retraction as $c(t) = \mathrm{R}_x(tv)$,

$$f(\mathrm{R}_x(tv)) = f(x) + t \left\langle \mathrm{grad} f(x), v \right\rangle_x + O(t^2). \tag{4.3}$$

Equivalently, we may eliminate t by introducing $s = tv$:

$$f(\mathrm{R}_x(s)) = f(x) + \left\langle \mathrm{grad} f(x), s \right\rangle_x + O(\|s\|_x^2). \tag{4.4}$$

The latter is a statement about the composition $f \circ \mathrm{R} \colon \mathrm{T}\mathcal{M} \to \mathbb{R}$, called the *pullback* of f to the tangent spaces. It is called this way as it quite literally pulls the cost function from the manifold back to the tangent spaces. In particular, $f \circ \mathrm{R}_x \colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}$ is the pullback of f to the tangent space at x. Importantly, this is a smooth function on a linear space: it has many uses, as we shall soon see.

Later, in Section 5.9, we extend the above reasoning to work out second-order Taylor expansions.

Exercise 4.3. *Given a smooth curve $c \colon [0, 1] \to \mathcal{M}$ with $c(0) = x$ and $c(1) = y$, check that there exists $t \in (0, 1)$ such that*

$$f(y) = f(x) + \left\langle \mathrm{grad} f(c(t)), c'(t) \right\rangle_{c(t)}. \tag{4.5}$$

(See Exercise 5.40 for the next order.)

4.2 First-order optimality conditions

In general, checking whether a point x on $\mathcal{M}$ is a local minimizer for $f \colon \mathcal{M} \to \mathbb{R}$ is difficult. We can however identify certain simple *necessary conditions* for a point x to be a local minimizer. The following definition states such a condition. It is called the *first-order necessary optimality condition*, because it involves first-order derivatives.

Definition 4.4. *A point $x \in \mathcal{M}$ is* critical *(or* stationary*) for a smooth function $f \colon \mathcal{M} \to \mathbb{R}$ if*

$$(f \circ c)'(0) \geq 0$$

for all smooth curves c on $\mathcal{M}$ such that $c(0) = x$.

In words: it is not possible to move away from a critical point x and obtain an initial decrease in the value of f with a linear rate. Notice that it would be equivalent to require $(f \circ c)'(0) = 0$ in the definition: simply consider the curves

$t \mapsto c(t)$ and $t \mapsto c(-t)$ simultaneously. Still equivalently, x is critical for f exactly if $\mathrm{D}f(x) = 0$.

Proposition 4.5. *Any local minimizer of a smooth function $f\colon \mathcal{M} \to \mathbb{R}$ is a critical point of f.*

Proof. Let x be a local minimizer of f: there exists a neighborhood $\mathcal{U}$ of x in $\mathcal{M}$ such that $f(y) \geq f(x)$ for all $y \in \mathcal{U}$. For contradiction, assume there exists a smooth curve $c\colon I \to \mathcal{M}$ with $c(0) = x$ and $(f \circ c)'(0) < 0$. By continuity of $(f \circ c)'$, there exists $\delta > 0$ such that $(f \circ c)'(\tau) < 0$ for all $\tau \in [0, \delta]$. This further implies

$$f(c(t)) = f(c(0)) + \int_0^t (f \circ c)'(\tau) \mathrm{d}\tau < f(x)$$

for all $t \in (0, \delta]$. Yet, since c is continuous, $c^{-1}(\mathcal{U}) = \{t \in I : c(t) \in \mathcal{U}\}$ is open, and it contains 0 because $c(0) = x \in \mathcal{U}$. Hence, $c^{-1}(\mathcal{U}) \cap (0, \delta]$ is non-empty: there exists $t \in (0, \delta]$ (implying $f(c(t)) < f(x)$) such that $c(t)$ is in $\mathcal{U}$ (implying $f(c(t)) \geq f(x)$): a contradiction. $\square$

It is easy to check that *local maximizers* (defined in analogy to local minimizers) are also critical points. The converse of Proposition 4.5 does not hold in general: see Chapter 11 for a special case.

On a Riemannian manifold, the critical points of a function are exactly those points where the Riemannian gradient vanishes.

Proposition 4.6. *Let $f\colon \mathcal{M} \to \mathbb{R}$ be smooth on a Riemannian manifold $\mathcal{M}$. Then, x is a critical point of f if and only if* $\mathrm{grad}f(x) = 0$.

Proof. Let $c\colon I \to \mathcal{M}$ be any smooth curve on $\mathcal{M}$ with $c(0) = x$ and $c'(0) = v$. We know that

$$(f \circ c)'(0) = \mathrm{D}f(x)[v] = \langle \mathrm{grad}f(x), v \rangle_x .$$

If $\mathrm{grad}f(x) = 0$, then x is clearly critical. The other way around, if x is critical, then $\langle \mathrm{grad}f(x), v \rangle_x \geq 0$ for all $v \in \mathrm{T}_x\mathcal{M}$. Considering both v and $-v$, it follows that $\langle \mathrm{grad}f(x), v \rangle_x = 0$ for all $v \in \mathrm{T}_x\mathcal{M}$. Thus, $\mathrm{grad}f(x) = 0$. $\square$

In developing optimization algorithms, one of our more modest goals is to ensure that accumulation points of sequences generated by those algorithms are critical points: we aim for small gradients.

4.3 Riemannian gradient descent

The standard gradient descent algorithm in Euclidean space $\mathcal{E}$ iterates

$$x_{k+1} = x_k - \alpha_k \mathrm{grad}f(x_k), \qquad k = 0, 1, 2, \ldots,$$

starting with some $x_0 \in \mathcal{E}$ and using some step-sizes $\alpha_k > 0$. Inspired by this, the first algorithm we consider for optimization on manifolds is *Riemannian gradient descent* (RGD): given $x_0 \in \mathcal{M}$ and a retraction R on $\mathcal{M}$, iterate

$$x_{k+1} = \mathrm{R}_{x_k}(-\alpha_k \mathrm{grad} f(x_k)), \qquad k = 0, 1, 2, \ldots.$$

See Algorithm 4.1. Importantly, the choice of retraction is part of the algorithm specification.

Algorithm 4.1 RGD: the Riemannian gradient descent method

Input: $x_0 \in \mathcal{M}$
For $k = 0, 1, 2, \ldots$
    Pick a step-size $\alpha_k > 0$
    $x_{k+1} = \mathrm{R}_{x_k}(s_k)$, with step $s_k = -\alpha_k \mathrm{grad} f(x_k)$

To complete the specification of RGD, we need an explicit procedure to pick the step-size α_k at each iteration. This is called the *line-search* phase, and it can be done in various way. Define

$$g(t) = f(\mathrm{R}_{x_k}(-t \mathrm{grad} f(x_k))). \tag{4.6}$$

Line-search is about minimizing g approximately: well enough to make progress, yet bearing in mind that this is only a means to an end; we should not invest too much resources into it. Three common strategies include:

1. Fixed step-size: $\alpha_k = \alpha$ for all k.
2. Optimal step-size: α_k minimizes $g(t)$ exactly; in rare cases, this can be done cheaply.
3. Backtracking: starting with a guess $t_0 > 0$, iteratively reduce it by a factor as $t_i = \tau t_{i-1}$ with $\tau \in (0, 1)$ until t_i is deemed acceptable, and set $\alpha_k = t_i$. There are various techniques to pick t_0.

We discuss this more in Section 4.5. For now, we focus on identifying assumptions that lead to favorable behavior.

Our first assumption about problem (4.1) simply requires that the cost function f be globally lower-bounded. This is normally the case for a well-posed optimization problem.

A4.1. *There exists $f_{\mathrm{low}} \in \mathbb{R}$ such that $f(x) \geq f_{\mathrm{low}}$ for all $x \in \mathcal{M}$.*

We expect that the algorithm may converge (or at least produce interesting points) provided it makes some progress at every iteration. This is the object of the second assumption below. We first confirm that it is sufficient for our purposes, and later show how to fulfill it.

⋆ Going forward in this chapter, we most often write $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ instead of $\langle \cdot, \cdot \rangle_x$ and $\| \cdot \|_x$ when the base point x is clear from context.

A4.2. *At each iteration, the algorithm achieves* sufficient decrease*, in that there exists a constant $c > 0$ such that, for all k,*

$$f(x_k) - f(x_{k+1}) \geq c\|\mathrm{grad} f(x_k)\|^2. \tag{4.7}$$

It is the responsibility of the line-search procedure to ensure this assumption holds. This can be done under some conditions on f and the retraction, as we discuss later. When both assumptions hold, it is straightforward to guarantee that RGD produces points with small gradient. There are no conditions on the initialization $x_0 \in \mathcal{M}$.

Proposition 4.7. *Let f be a smooth function satisfying A4.1 on a Riemannian manifold $\mathcal{M}$. Let $x_0, x_1, x_2, \ldots$ be iterates satisfying A4.2 with constant c. Then,*

$$\lim_{k\to\infty} \|\mathrm{grad} f(x_k)\| = 0.$$

In particular, all accumulation points (if any) are critical points. Furthermore, for all $K \geq 1$, there exists k in $0, \ldots, K-1$ such that

$$\|\mathrm{grad} f(x_k)\| \leq \sqrt{\frac{f(x_0) - f_{\mathrm{low}}}{c}} \frac{1}{\sqrt{K}}.$$

Proof. The proof is based on a standard telescoping sum argument. For all $K \geq 1$, we get the inequality as follows:

$$\begin{aligned} f(x_0) - f_{\mathrm{low}} \overset{\text{A4.1}}{\geq} f(x_0) - f(x_K) &= \sum_{k=0}^{K-1} f(x_k) - f(x_{k+1}) \\ &\overset{\text{A4.2}}{\geq} Kc \min_{k=0,\ldots,K-1} \|\mathrm{grad} f(x_k)\|^2. \end{aligned}$$

To get the limit statement, observe that $f(x_{k+1}) \leq f(x_k)$ for all k by A4.2. Then, taking K to infinity we see that

$$f(x_0) - f_{\mathrm{low}} \geq \sum_{k=0}^{\infty} f(x_k) - f(x_{k+1}),$$

where the right-hand side is a series of nonnegative numbers. The bound implies that the summands converge to zero, thus:

$$0 = \lim_{k\to\infty} f(x_k) - f(x_{k+1}) \geq c \lim_{k\to\infty} \|\mathrm{grad} f(x_k)\|^2,$$

which confirms that $\|\mathrm{grad} f(x_k)\| \to 0$. Now let x be an accumulation point of the sequence of iterates. By definition, there exists a subsequence of iterates $x_{(0)}, x_{(1)}, x_{(2)}, \ldots$ which converges to x. Then, since the norm of the gradient of f is a continuous function, it commutes with the limit and we find:

$$\begin{aligned} 0 = \lim_{k\to\infty} \|\mathrm{grad} f(x_k)\| &= \lim_{k\to\infty} \|\mathrm{grad} f(x_{(k)})\| \\ &= \|\mathrm{grad} f(\lim_{k\to\infty} x_{(k)})\| = \|\mathrm{grad} f(x)\|, \end{aligned}$$

showing all accumulation points are critical points. $\square$

Importantly, the limit statement does *not* say that the sequence of iterates converges to a critical point. It only states that, under the prescribed conditions, the accumulation points of the sequence of iterates (of which there may be one, more than one, or none) are critical points. To preserve conciseness, assuming there exists at least one accumulation point (which is often the case), this property may be summarized as: gradient descent converges to critical *points* (note the plural). See also Section 4.9.

In the next section, we explore regularity conditions to help us guarantee sufficient decrease using simple line-search procedures. The condition we introduce is inspired by the Taylor expansion of f along curves generated by the retraction.

4.4 Regularity conditions and iteration complexity

In order to guarantee sufficient decrease as per A4.2, we need to understand how $f(x_{k+1})$ compares to $f(x_k)$. Recall that $x_{k+1} = \mathrm{R}_{x_k}(s_k)$ with a chosen tangent vector s_k. The Taylor expansion (4.4) thus states:

$$f(x_{k+1}) = f(\mathrm{R}_{x_k}(s_k)) = f(x_k) + \langle \mathrm{grad} f(x_k), s_k \rangle + O(\|s_k\|^2).$$

If the quadratic remainder term stays under control during all iterations, we may deduce a guarantee on the progress $f(x_k) - f(x_{k+1})$. This motivates the following assumption on the pullback $f \circ \mathrm{R}$. We provide further context for this assumption at the end of the section, as well as much later in Corollary 10.54, Lemma 10.57 and Exercise 10.58.

A4.3. *For a given subset S of the tangent bundle $\mathrm{T}\mathcal{M}$, there exists a constant $L > 0$ such that, for all $(x, s) \in S$,*

$$f(\mathrm{R}_x(s)) \leq f(x) + \langle \mathrm{grad} f(x), s \rangle + \frac{L}{2}\|s\|^2. \tag{4.8}$$

Under this assumption (on an appropriate set S to be specified), there exists a range of step-sizes that lead to sufficient decrease.

Proposition 4.8. *Let f be a smooth function on a Riemannian manifold $\mathcal{M}$. For a retraction R, let $f \circ \mathrm{R}$ satisfy A4.3 on a set $S \subseteq \mathrm{T}\mathcal{M}$ with constant L. If the pairs $(x_0, s_0), (x_1, s_1), (x_2, s_2), \ldots$ generated by Algorithm 4.1 with step-sizes*

$$\alpha_k \in [\alpha_{\min}, \alpha_{\max}] \subset (0, 2/L)$$

all lie in S, then the algorithm produces sufficient decrease. Specifically, A4.2 holds with

$$c = \min\left(\alpha_{\min} - \frac{L}{2}\alpha_{\min}^2, \alpha_{\max} - \frac{L}{2}\alpha_{\max}^2\right) > 0.$$

In particular, for $\alpha_k = \frac{1}{L}$ (constant) we have $c = \frac{1}{2L}$.

Proof. By assumption A4.3 on the pullback, for all k,

$$f(x_{k+1}) = f(\mathrm{R}_{x_k}(s_k)) \leq f(x_k) + \langle \mathrm{grad} f(x_k), s_k \rangle + \frac{L}{2}\|s_k\|^2.$$

Reorganizing and using $s_k = -\alpha_k \mathrm{grad} f(x_k)$ reveals

$$f(x_k) - f(x_{k+1}) \geq \left(\alpha_k - \frac{L}{2}\alpha_k^2\right) \|\mathrm{grad} f(x_k)\|^2.$$

The coefficient is quadratic in α_k, positive between its roots at 0 and $2/L$. By assumption on α_k,

$$\alpha_k - \frac{L}{2}\alpha_k^2 \geq \min\left(\alpha_{\min} - \frac{L}{2}\alpha_{\min}^2, \alpha_{\max} - \frac{L}{2}\alpha_{\max}^2\right) > 0,$$

which concludes the proof. □

As a particular case, if a valid constant L is known beforehand, then we get an explicit algorithm and associated guarantee as a corollary of Propositions 4.7 and 4.8.

Corollary 4.9. *Let f be a smooth function satisfying A4.1 on a Riemannian manifold $\mathcal{M}$. For a retraction* R, *let $f \circ \mathrm{R}$ satisfy A4.3 on a set $S \subseteq \mathrm{T}\mathcal{M}$ with constant L. Let $(x_0, s_0), (x_1, s_1), (x_2, s_2), \ldots$ be the pairs generated by Algorithm 4.1 with constant step-size $\alpha_k = 1/L$. If all these pairs are in S, then*

$$\lim_{k\to\infty} \|\mathrm{grad} f(x_k)\| = 0.$$

Furthermore, for all $K \geq 1$, there exists k in $0, \ldots, K-1$ such that

$$\|\mathrm{grad} f(x_k)\| \leq \sqrt{2L(f(x_0) - f_{\mathrm{low}})}\frac{1}{\sqrt{K}}.$$

The conclusion of the above corollary can also be stated as follows: for all $\varepsilon > 0$ there exists k in $0, \ldots, K-1$ such that $\|\mathrm{grad} f(x_k)\| \leq \varepsilon$ provided $K \geq 2L(f(x_0) - f_{\mathrm{low}})\frac{1}{\varepsilon^2}$. Notice that the rate is independent of the dimension of $\mathcal{M}$.

How reasonable is A4.3? Let us contemplate it through the Euclidean lens. Consider f smooth on a Euclidean space $\mathcal{E}$ equipped with the canonical retraction $\mathrm{R}_x(s) = x + s$. If $f \circ \mathrm{R}$ satisfies A4.3 on the whole tangent bundle $\mathrm{T}\mathcal{E} = \mathcal{E} \times \mathcal{E}$, then

$$\forall x, s \in \mathcal{E}, \qquad f(x+s) \leq f(x) + \langle \mathrm{grad} f(x), s \rangle + \frac{L}{2}\|s\|^2. \tag{4.9}$$

This expresses that the difference between f and its first-order Taylor expansion is uniformly upper-bounded by a quadratic. This property holds if (and only if) the gradient of f is *Lipschitz continuous* with constant L, that is, if

$$\forall x, y \in \mathcal{E}, \qquad \|\mathrm{grad} f(y) - \mathrm{grad} f(x)\| \leq L\|y - x\|. \tag{4.10}$$

Indeed, with $c(t) = x + ts$, elementary calculus provides:

$$\begin{aligned} f(x+s) - f(x) &= f(c(1)) - f(c(0)) \\ &= \int_0^1 (f \circ c)'(t) \mathrm{d}t \\ &= \int_0^1 \mathrm{D}f(c(t))[c'(t)] \mathrm{d}t = \int_0^1 \langle \mathrm{grad} f(x+ts), s \rangle \, \mathrm{d}t. \end{aligned}$$

Then, under condition (4.10), by Cauchy–Schwarz we have:

$$\begin{aligned} |f(x+s) - f(x) - \langle \mathrm{grad} f(x), s \rangle| &= \left| \int_0^1 \langle \mathrm{grad} f(x+ts) - \mathrm{grad} f(x), s \rangle \, \mathrm{d}t \right| \\ &\leq \int_0^1 \|\mathrm{grad} f(x+ts) - \mathrm{grad} f(x)\| \|s\| \mathrm{d}t \\ &\leq \|s\| \int_0^1 L \|ts\| \mathrm{d}t \\ &= \frac{L}{2} \|s\|^2. \end{aligned} \tag{4.11}$$

Lipschitz continuity of the gradient (4.10) is a common assumption in Euclidean optimization, valued for the upper-bounds it provides (4.11). When working on manifolds, generalizing (4.10) requires substantial work due to the comparison of gradients at two distinct points (hence of vectors in two distinct tangent spaces) but it can be led to fruition: Sections 10.3, 10.4 and 10.5 provide a detailed discussion involving a special retraction. On the other hand, generalizing (4.11) poses no particular difficulty once a retraction is chosen. This is the reasoning that led to A4.3, which we henceforth call a *Lipschitz-type* assumption. General retractions are covered by Lemma 10.57 under compactness assumptions. In particular, that lemma can be useful to verify regularity assumptions such as A4.3 when the *sublevel sets* of the cost function are compact.

Definition 4.10. *A* sublevel set *of f is a set $\{x \in \mathcal{M} : f(x) \leq \alpha\}$ for some α.*

Exercise 4.11. *For the cost function $f(x) = \frac{1}{2}x^\top Ax$ on the sphere S^{n-1} as a Riemannian submanifold of $\mathbb{R}^n$ equipped with the retraction $\mathrm{R}_x(s) = \frac{x+s}{\|x+s\|}$, determine L such that A4.3 holds over the whole tangent bundle.*

4.5 Backtracking line-search

The simplest result in the previous section is Corollary 4.9, which assumes a constant step-size of $1/L$. In practice however, an appropriate constant L is seldom known. Even when one is available, it may be large due to particular behavior of $f \circ \mathrm{R}$ in a limited part of the domain. That seemingly forces us to take small steps for the whole sequence of iterates, which evidently is not necessary. Indeed, only the local behavior of the cost function around x_k matters to ensure sufficient decrease at iteration k. Thus, we favor line-search algorithms that are *adaptive*.

A common adaptive strategy to pick the step-sizes α_k for RGD is called the *backtracking line-search*: see Algorithm 4.2. For a specified initial step-size $\bar{\alpha}$, this procedure iteratively reduces the tentative step-size by a factor $\tau \in (0,1)$ (often set to 0.8 or 0.5) until the *Armijo–Goldstein* condition is satisfied, namely,

$$f(x) - f(\mathrm{R}_x(-\alpha \mathrm{grad} f(x))) \geq r\alpha \|\mathrm{grad} f(x)\|^2, \tag{4.12}$$

for some constant $r \in (0,1)$ (often set to 10^{-4}).

Algorithm 4.2 Backtracking line-search

Parameters: $\tau, r \in (0,1)$; for example, $\tau = \frac{1}{2}$ and $r = 10^{-4}$
Input: $x \in \mathcal{M}$, $\bar{\alpha} > 0$
Set $\alpha \leftarrow \bar{\alpha}$
While $f(x) - f(\mathrm{R}_x(-\alpha \mathrm{grad} f(x))) < r\alpha \|\mathrm{grad} f(x)\|^2$
    Set $\alpha \leftarrow \tau\alpha$
Output: α

The lemma and corollary below show that, under the regularity condition A4.3, backtracking line-search produces sufficient decrease A4.2, with a constant c which depends on various factors. Importantly, the regularity constant L affects the guarantee but need not be known.

Lemma 4.12. *Let f be a smooth function on a Riemannian manifold $\mathcal{M}$. For a retraction* R*, a point $x \in \mathcal{M}$ and an initial step-size $\bar{\alpha} > 0$, let A4.3 hold for $f \circ \mathrm{R}$ on $\{(x, -\alpha \mathrm{grad} f(x)) : \alpha \in [0, \bar{\alpha}]\}$ with constant L. Then, Algorithm 4.2 with parameters $\tau, r \in (0,1)$ outputs a step-size α such that*

$$f(x) - f(\mathrm{R}_x(-\alpha \mathrm{grad} f(x))) \geq r \min\left(\bar{\alpha}, \frac{2\tau(1-r)}{L}\right) \|\mathrm{grad} f(x)\|^2$$

after computing at most

$$\max\left(1, 2 + \log_{\tau^{-1}}\left(\frac{\bar{\alpha}L}{2(1-r)}\right)\right)$$

retractions and cost function evaluations (assuming $f(x)$ and $\mathrm{grad} f(x)$ *were already computed).*

Proof. Consider $\mathrm{grad} f(x) \neq 0$ (otherwise, the claim is clear). For all step-sizes α considered by Algorithm 4.2, the regularity assumption guarantees

$$f(x) - f(\mathrm{R}_x(-\alpha \mathrm{grad} f(x))) \geq \alpha \|\mathrm{grad} f(x)\|^2 - \frac{L}{2}\alpha^2 \|\mathrm{grad} f(x)\|^2.$$

On the other hand, if the algorithm does not terminate for a certain value α, then

$$f(x) - f(\mathrm{R}_x(-\alpha \mathrm{grad} f(x))) < r\alpha \|\mathrm{grad} f(x)\|^2.$$

If both are true simultaneously, then

$$\alpha > \frac{2(1-r)}{L}.$$

Thus, if α drops below this bound, the line-search algorithm terminates. (Of course, it might also terminate earlier with a longer step-size: we consider the worst case.) This happens either because $\bar{\alpha}$ itself is smaller than $\frac{2(1-r)}{L}$, or as the result of a reduction of α by the factor τ. We conclude that the returned α satisfies:

$$\alpha \geq \min\left(\bar{\alpha}, \frac{2\tau(1-r)}{L}\right).$$

Moreover, the returned α is of the form $\alpha = \bar{\alpha}\tau^{n-1}$ where n is the number of retractions and cost function evaluations issued by Algorithm 4.2. Hence,

$$n = 1 + \log_\tau\left(\frac{\alpha}{\bar{\alpha}}\right) = 1 + \log_{\tau^{-1}}\left(\frac{\bar{\alpha}}{\alpha}\right) \leq 1 + \max\left(0, \log_{\tau^{-1}}\left(\frac{\bar{\alpha}L}{2\tau(1-r)}\right)\right),$$

which concludes the proof. □

When used in conjunction with RGD, one may want to pick the initial step-size $\bar{\alpha}$ dynamically as $\bar{\alpha}_k$ at iteration k. As long as the initializations $\bar{\alpha}_k$ remain bounded away from zero, we retain our convergence result.

Corollary 4.13. *Let f be a smooth function satisfying A4.1 on a Riemannian manifold $\mathcal{M}$. For a retraction R, let $f \circ \mathrm{R}$ satisfy A4.3 on a set $S \subseteq \mathrm{T}\mathcal{M}$ with constant L. Let $x_0, x_1, x_2, \ldots$ be the iterates generated by RGD (Algorithm 4.1) with backtracking line-search (Algorithm 4.2) using fixed parameters $\tau, r \in (0,1)$ and initial step-sizes $\bar{\alpha}_0, \bar{\alpha}_1, \bar{\alpha}_2, \ldots$ If for every k the set $\{(x_k, -\alpha \mathrm{grad} f(x_k)) : \alpha \in [0, \bar{\alpha}_k]\}$ is in S and if $\liminf_{k\to\infty} \bar{\alpha}_k > 0$, then*

$$\lim_{k\to\infty} \|\mathrm{grad} f(x_k)\| = 0.$$

Furthermore, for all $K \geq 1$, there exists k in $0, \ldots, K-1$ such that

$$\|\mathrm{grad} f(x_k)\| \leq \sqrt{\frac{f(x_0) - f_{\mathrm{low}}}{r \min\left(\bar{\alpha}_0, \ldots, \bar{\alpha}_{K-1}, \frac{2\tau(1-r)}{L}\right)}} \frac{1}{\sqrt{K}}.$$

The amount of work per iteration is controlled as in Lemma 4.12.

Proof. By Lemma 4.12, backtracking line-search guarantees decrease in the form

$$f(x) - f(x_{k+1}) \geq c_k \|\mathrm{grad} f(x_k)\|^2, \quad \text{with} \quad c_k = r \min\left(\bar{\alpha}_k, \frac{2\tau(1-r)}{L}\right).$$

Following the same proof as for Proposition 4.7,

$$f(x_0) - f_{\text{low}} \geq \sum_{k=0}^{K-1} f(x_k) - f(x_{k+1}) \geq \sum_{k=0}^{K-1} c_k \|\text{grad} f(x_k)\|^2$$
$$\geq K \cdot \min_{k=0,\ldots,K-1} c_k \cdot \min_{k=0,\ldots,K-1} \|\text{grad} f(x_k)\|^2.$$

This establishes the first claim. For the limit statement, observe that taking $K \to \infty$ on the first line above shows that

$$\lim_{k\to\infty} c_k \|\text{grad} f(x_k)\|^2 = 0.$$

Since $\liminf_{k\to\infty} c_k > 0$, we deduce that $\lim_{k\to\infty} \|\text{grad} f(x_k)\|^2 = 0$. $\square$

As a remark, consider replacing the cost function $f(x)$ by a shifted and positively scaled version of itself, say $g(x) = 8f(x) + 3$. Arguably, the optimization problem did not change, and we might expect a reasonable optimization algorithm initialized at x_0 to produce the same iterates to minimize f or to minimize g. It is easily checked that the combination of Algorithms 4.1 and 4.2 has this invariance property, provided the initial step-sizes $\bar{\alpha}_k$ are chosen in such a way that the first step considered, namely, $-\bar{\alpha}_k \text{grad} f(x_k)$ is invariant under positive scaling of f. For the first iteration, this can be done for example by setting

$$\bar{\alpha}_0 = \frac{\ell_0}{\|\text{grad} f(x_0)\|}$$

with some constant ℓ_0, which is then the length of the first retracted step: it can be set relative to the scale of the search space or to the expected distance between x_0 and a solution (this does not need to be precise). For subsequent iterations, a useful heuristic is (see [NW06, §3.5, eq. (3.60)] for more)

$$\bar{\alpha}_k = 2\frac{f(x_{k-1}) - f(x_k)}{\|\text{grad} f(x_k)\|^2}, \tag{4.13}$$

which also yields the desired invariance. It is common to initialize with a slightly larger value, say, by a factor of $1/\tau$. One may also set $\bar{\alpha}_k$ to be the maximum between the above value and a small reference value, to ensure the first step-size remains bounded away from zero (as required by our convergence theory).

4.6 Local convergence*

In numerical analysis, we study the behavior of sequences generated by iterative algorithms. We distinguish between the *local* and *global* behavior of those sequences. We say that a method enjoys *global convergence* if it generates sequences that converge regardless of their initialization x_0. It is worth noting two common points of confusion here:

1. "Global" convergence is *not* convergence to a global optimizer. It is merely a statement that sequences converge "somewhere".
2. It is common to say that RGD enjoys global convergence, and indeed the results presented in previous sections assume little about the initial point x_0. However, we have not established convergence of RGD. Rather, we have shown that, under some assumptions, all accumulation points of RGD (if any) are critical points.

Still, RGD usually converges to a critical point in practice, hence the habit of calling it globally convergent. Our results in previous sections also qualify how fast the gradient norm converges to zero in the worst case. That rate, however, is underwhelming: it only guarantees a decrease as fast as $1/\sqrt{k}$ where k is the iteration counter. While this result is correct (there exist difficult cost functions even on $\mathbb{R}^n$ that lead to such poor performance), it is common to observe an eventually exponential decrease of the gradient norm. This asymptotic behavior of a sequence is the realm of *local convergence*: the study of how convergent sequences behave once they are close enough to their limit.

The discussions in this section require tools that we have not introduced yet. Specifically, we use:

- The *Riemannian distance* dist, which turns a connected Riemannian manifold into a metric space. See Section 10.1.
- The *exponential map* Exp, which is a special retraction. Of relevance here, it has the property that for $v \in \mathrm{T}_x\mathcal{M}$ small enough and $y = \mathrm{Exp}_x(v)$ we have $\mathrm{dist}(x, y) = \|v\|_x$. See Section 10.2.
- The *Riemannian Hessian* $\mathrm{Hess}f$ of a smooth function $f : \mathcal{M} \to \mathbb{R}$, which is a kind of derivative of the gradient vector field. Of relevance here, (a) $\mathrm{Hess}f(x)$ is a self-adjoint linear map on $\mathrm{T}_x\mathcal{M}$ (hence it has real eigenvalues), and (b) if $\mathrm{grad}f(x) = 0$ and $\mathrm{Hess}f(x) \succ 0$, then x is a strict local minimizer of f. See Sections 5.5 and 6.1.

Local convergence rates are defined in general for sequences in metric spaces. This applies to sequences of real numbers (tracking $f(x_k)$ or $\|\mathrm{grad}f(x_k)\|$) using the absolute value distance on $\mathbb{R}$, and it also applies to sequences of points on $\mathcal{M}$ using the Riemannian distance.

Definition 4.14. *In a metric space with a distance* dist*, a sequence* $a_0, a_1, a_2, \ldots$ converges at least linearly *to* $a_\star$ *if there exist positive reals* $\epsilon_0, \epsilon_1, \epsilon_2, \ldots$ *converging to zero such that* $\mathrm{dist}(a_k, a_\star) \leq \epsilon_k$ *and* $\lim_{k\to\infty} \frac{\epsilon_{k+1}}{\epsilon_k} = \mu$ *for some* $\mu \in (0, 1)$. *The infimum over such* μ *is the* linear convergence factor*. If the latter is zero, the convergence is* superlinear*.*

The above is also called *R-linear convergence*, as opposed to the more restrictive notion of *Q-linear convergence* which forces $\epsilon_k = \mathrm{dist}(a_k, a_\star)$.

Superlinear convergence rates can be further qualified. Anticipating the needs of Chapter 6, we already define *quadratic convergence*. It is straightforward to check that quadratic convergence implies superlinear convergence which itself implies linear convergence.

Definition 4.15. *In a metric space equipped with a distance* dist*, a sequence* $a_0, a_1, a_2, \ldots$ converges at least quadratically *to* $a_\star$ *if there exists a sequence of positive reals* $\epsilon_0, \epsilon_1, \epsilon_2, \ldots$ *converging to zero such that* $\mathrm{dist}(a_k, a_\star) \leq \epsilon_k$ *and* $\lim_{k\to\infty} \frac{\epsilon_{k+1}}{\epsilon_k^2} = \mu$ *for some finite* $\mu \geq 0$*.*

In the remainder of this section, we build up toward a local convergence result for RGD with constant step-size. To this end, we first secure a broader statement called the *local contraction mapping theorem*: it will serve us again to study the Riemannian Newton method in Section 6.2.

An important tool below is the observation that retractions provide local parameterizations of manifolds. More explicitly, R_x provides a diffeomorphism (recall Definition 3.11) between a neighborhood of the origin in $\mathrm{T}_x\mathcal{M}$ (a linear space) and a neighborhood of x in $\mathcal{M}$. This fact is a direct consequence of a generalization of the inverse function theorem, stated now with a couple of relevant corollaries. Note: in the theorem below, $\mathcal{M}$ and $\mathcal{N}$ could also be open subsets of manifolds, as such sets are manifolds too.

Theorem 4.16 (Inverse function theorem on manifolds)**.** *Let* $F\colon \mathcal{M} \to \mathcal{N}$ *be a smooth map between two manifolds. If* $\mathrm{D}F(x)$ *is invertible at some point* $x \in \mathcal{M}$*, then there exist neighborhoods* $U \subseteq \mathcal{M}$ *of* x *and* $V \subseteq \mathcal{N}$ *of* $F(x)$ *such that* $F|_U\colon U \to V$ *is a diffeomorphism.*

Proof sketch. The idea is to reduce the claim to Theorem 3.13. This is best done through charts. For submanifolds, these charts can be built via Theorem 3.12, as in Section 8.3. Details in [Lee12, Thm. 4.5]. □

Corollary 4.17. *Let* R *be a retraction on a manifold* $\mathcal{M}$*. For each* x*, there exists a neighborhood* U *of the origin of* $\mathrm{T}_x\mathcal{M}$ *such that* $\mathrm{R}_x|_U\colon U \to \mathcal{U}$ *is a diffeomorphism, where* $\mathcal{U} = \mathrm{R}_x(U)$ *is a neighborhood of* x *on* $\mathcal{M}$*.*

Proof. The map $\mathrm{R}_x\colon \mathrm{T}_x\mathcal{M} \to \mathcal{M}$ satisfies the assumptions of Theorem 4.16 around the origin of $\mathrm{T}_x\mathcal{M}$ since $\mathrm{DR}_x(0)$ is invertible. □

Corollary 4.18. *Continuing from Corollary 4.17, if* $\mathcal{M}$ *is Riemannian then we can choose* U *as an open ball of some radius* $r > 0$ *around the origin in* $\mathrm{T}_x\mathcal{M}$*, that is,* $U = B(x,r) = \{v \in \mathrm{T}_x\mathcal{M} : \|v\|_x < r\}$*.*

Some simple algorithms come down to the repeated application of a smooth iteration map $F\colon \mathcal{M} \to \mathcal{M}$, so that $x_1 = F(x_0)$, $x_2 = F(x_1) = F(F(x_0))$, etc. The following theorem provides insight into the local convergence of such algorithms near special points.

Theorem 4.19 (Local contraction mapping)**.** *Let* $F\colon \mathcal{M} \to \mathcal{M}$ *be a smooth map to and from a Riemannian manifold* $\mathcal{M}$*. Given* $x_0 \in \mathcal{M}$*, consider the sequence* *defined by*

$$x_{k+1} = F(x_k) \qquad\qquad \text{for } k = 0, 1, 2, \ldots$$

If $x_\star \in \mathcal{M}$ is a fixed point *(that is, $F(x_\star) = x_\star$) and $\|\mathrm{D}F(x_\star)\| < 1$ (that is, all singular values of $\mathrm{D}F(x_\star)$ are strictly smaller than one), then there exists a neighborhood $\mathcal{U}$ of $x_\star$ such that, if the sequence enters $\mathcal{U}$, it stays in $\mathcal{U}$ and converges to $x_\star$ at least linearly with a linear convergence factor which is at most $\|\mathrm{D}F(x_\star)\|$.*

Explicitly, with R an arbitrary retraction on $\mathcal{M}$ we have

$$\lim_{k\to\infty} \frac{\|\xi_{k+1}\|}{\|\xi_k\|} \leq \|\mathrm{D}F(x_\star)\|,$$

where ξ_k is well defined by $\mathrm{R}_{x_\star}(\xi_k) = x_k$ for k large enough. The conclusion follows by letting $\mathrm{R} = \mathrm{Exp}$, in which case $\|\xi_k\| = \mathrm{dist}(x_k, x_\star)$ for large k.

Additionally,

1. *If $\|\mathrm{D}F(x_\star)\| = 0$, the convergence is at least quadratic.*
2. *All claims still hold if F is only smoothly defined in a neighborhood of $x_\star$.*

Proof. Owing to Corollary 4.18, we can choose a radius $r > 0$ such that $\mathrm{R}_{x_\star}$ is a diffeomorphism from $B(x_\star, r)$ to $\mathcal{U} = \mathrm{R}_{x_\star}(B(x_\star, r))$: we now tacitly restrict $\mathrm{R}_{x_\star}$ to those domains. Consider the set $F^{-1}(\mathcal{U})$ (where it is understood that we first restrict F to the neighborhood of $x_\star$ where it is smooth, if need be): it is open (because F is continuous and $\mathcal{U}$ is open) and it contains $x_\star$ (because $F(x_\star) = x_\star$ and $x_\star \in \mathcal{U}$). Thus, $F^{-1}(\mathcal{U})$ is a neighborhood of $x_\star$. It follows that we can select $r' \in (0, r]$ such that

$$\mathcal{U}' \triangleq \mathrm{R}_{x_\star}(B(x_\star, r')) \subseteq F^{-1}(\mathcal{U}) \qquad \text{and} \qquad \mathcal{U}' \subseteq \mathcal{U}.$$

Assume x_k is in $\mathcal{U}'$. Then, $x_{k+1} = F(x_k)$ is in $F(\mathcal{U}')$, which is included in $F(F^{-1}(\mathcal{U}))$, that is: $x_{k+1} \in \mathcal{U}$. Thus, the vectors $\xi_k, \xi_{k+1} \in \mathrm{T}_{x_\star}\mathcal{M}$ are well defined by

$$x_k = \mathrm{R}_{x_\star}(\xi_k), \qquad\qquad x_{k+1} = \mathrm{R}_{x_\star}(\xi_{k+1}).$$

Consider the following map restricted to open sets in $\mathrm{T}_{x_\star}\mathcal{M}$:

$$\tilde{F} \colon B(x_\star, r') \to B(x_\star, r), \qquad\qquad \tilde{F} = \mathrm{R}_{x_\star}^{-1} \circ F \circ \mathrm{R}_{x_\star}.$$

It is defined such that $\xi_{k+1} = \tilde{F}(\xi_k)$. Since $\tilde{F}$ is smooth, we can use a standard Taylor expansion on $\mathrm{T}_{x_\star}\mathcal{M}$ (a Euclidean space) to claim that

$$\tilde{F}(v) = \tilde{F}(0) + \mathrm{D}\tilde{F}(0)[v] + E(v)$$

where $\|E(v)\| \leq c\|v\|^2$ for some constant c, valid for all $v \in B(x_\star, r')$. Notice that $\tilde{F}(0) = 0$. Moreover, $\mathrm{D}\tilde{F}(0) = \mathrm{D}F(x_\star)$: that is due to the chain rule and the fact that $\mathrm{DR}_{x_\star}(0)$ is the identity on $\mathrm{T}_{x_\star}\mathcal{M}$ so that $\mathrm{DR}_{x_\star}^{-1}(x_\star)$ is also the identity. It follows that

$$\xi_{k+1} = \tilde{F}(\xi_k) = \mathrm{D}F(x_\star)[\xi_k] + E(\xi_k).$$

Taking norms on both sides, we find

$$\|\xi_{k+1}\| \leq \|\mathrm{D}F(x_\star)[\xi_k]\| + \|E(\xi_k)\| \leq (\|\mathrm{D}F(x_\star)\| + c\|\xi_k\|)\,\|\xi_k\|. \tag{4.14}$$

Recall that $\|\xi_k\| < r'$. If need be, replace r' by a smaller positive constant such that $\|\mathrm{D}F(x_\star)\| + cr' < 1$: this is possible owing to our assumption that $\|\mathrm{D}F(x_\star)\| < 1$. Doing so, we ensure $\|\xi_{k+1}\| \leq \|\xi_k\|$. In particular, $x_k \in \mathcal{U}' \implies x_{k+1} \in \mathcal{U}'$. By induction, we can now define ξ_K through $x_K = \mathrm{R}_{x_\star}(\xi_K) \in \mathcal{U}'$ for all $K \geq k$. Moreover, we have that $\|\xi_K\|$ converges to zero at least linearly. The linear convergence factor is controlled by:

$$\lim_{K\to\infty} \frac{\|\xi_{K+1}\|}{\|\xi_K\|} \leq \lim_{K\to\infty} \|\mathrm{D}F(x_\star)\| + c\|\xi_K\| = \|\mathrm{D}F(x_\star)\|.$$

Now that convergence is established, we can return to (4.14) and notice that, if $\mathrm{D}F(x_\star) = 0$, then we also have

$$\lim_{K\to\infty} \frac{\|\xi_{K+1}\|}{\|\xi_K\|^2} \leq c.$$

Thus, in that case, the sequence converges at least quadratically. $\square$

We now apply the above theorem to RGD with constant step-size, as this indeed corresponds to the iterative application of a smooth map. More realistically, we would use a backtracking line-search procedure. However, that leads to an iteration map that may lack smoothness as the selected step-size may depend on x discontinuously. With different tools, it is still possible to establish linear convergence of RGD with a line-search, see [AMS08, Thm. 4.5.6]. It is important to note that the retraction used in RGD is unrelated to the retraction used in the proof of the local contraction mapping theorem—For the latter, it makes the most sense to use the exponential retraction.

Theorem 4.20. *Let $\mathcal{M}$ be a Riemannian manifold with a retraction* R. *Let $f\colon \mathcal{M} \to \mathbb{R}$ be a smooth function. Assume $x_\star \in \mathcal{M}$ satisfies*

$$\mathrm{grad} f(x_\star) = 0 \qquad \text{and} \qquad \mathrm{Hess} f(x_\star) \succ 0.$$

Let $0 < \lambda_{\min} \leq \lambda_{\max}$ be the smallest and largest eigenvalues of $\mathrm{Hess} f(x_\star)$, and let $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}}$ denote the condition number of $\mathrm{Hess} f(x_\star)$. Set $L > \frac{1}{2}\lambda_{\max}$. Given $x_0 \in \mathcal{M}$, constant step-size Riemannian gradient descent iterates

$$x_{k+1} = F(x_k), \qquad \text{with} \qquad F(x) = \mathrm{R}_x\left(-\frac{1}{L}\mathrm{grad} f(x)\right).$$

There exists a neighborhood of $x_\star$ such that, if the above sequence enters the neighborhood, then it stays in that neighborhood and it converges to $x_\star$ at least linearly. If $L = \lambda_{\max}$, the linear convergence factor is at most $1 - 1/\kappa$.

Proof. Let us check the assumptions of Theorem 4.19. First, it is clear that

$F(x_\star) = x_\star$. Second, let us investigate $\mathrm{D}F(x_\star)\colon \mathrm{T}_{x_\star}\mathcal{M} \to \mathrm{T}_{x_\star}\mathcal{M}$. In particular, we need to differentiate through $\mathrm{R}\colon \mathrm{T}\mathcal{M} \to \mathcal{M}$:

$$\mathrm{D}F(x)[v] = \mathrm{D}(x \mapsto \mathrm{R}(x, G(x)))(x)[v] = \mathrm{DR}(x, G(x))[(v, \mathrm{D}G(x)[v])],$$

with $G(x) = -\frac{1}{L}\mathrm{grad}f(x)$. This simplifies at $x = x_\star$ since $G(x_\star) = 0$. Indeed, Lemma 4.21 below justifies the following:

$$\mathrm{D}F(x_\star)[v] = \mathrm{DR}(x_\star, 0)[(v, \mathrm{D}G(x_\star)[v])] = v + \mathrm{D}G(x_\star)[v].$$

Anticipating concepts from Chapter 5, the property $G(x_\star) = 0$ also implies via Proposition 5.3 that

$$\mathrm{D}G(x_\star)[v] = \nabla_v G = -\frac{1}{L}\mathrm{Hess}f(x_\star)[v] \tag{4.15}$$

for all $v \in \mathrm{T}_{x_\star}\mathcal{M}$, where ∇ is the Riemannian connection on $\mathcal{M}$. Thus,

$$\mathrm{D}F(x_\star) = \mathrm{Id} - \frac{1}{L}\mathrm{Hess}f(x_\star), \tag{4.16}$$

where Id is the identity map on $\mathrm{T}_{x_\star}\mathcal{M}$. Since both Id and $\mathrm{Hess}f(x_\star)$ are self-adjoint (Proposition 5.15), we find that $\mathrm{D}F(x_\star)$ is self-adjoint. Its eigenvalues (all real) are given by

$$1 - \frac{\lambda_1}{L} \leq \cdots \leq 1 - \frac{\lambda_n}{L},$$

where $\lambda_1 \geq \cdots \geq \lambda_n > 0$ are the eigenvalues of $\mathrm{Hess}f(x_\star)$. It follows that the operator norm of $\mathrm{D}F(x_\star)$ is

$$\|\mathrm{D}F(x_\star)\| = \max\left(\left|1 - \frac{\lambda_1}{L}\right|, \left|1 - \frac{\lambda_n}{L}\right|\right). \tag{4.17}$$

Under our assumption on L, it is easy to check that $\|\mathrm{D}F(x_\star)\| < 1$. All conclusions now follow from Theorem 4.19. $\square$

Lemma 4.21. *For each point x on a manifold $\mathcal{M}$, we have*

$$\mathrm{T}_{(x,0)}\mathrm{T}\mathcal{M} = \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M}. \tag{4.18}$$

Let $\mathrm{R}\colon \mathrm{T}\mathcal{M} \to \mathcal{M}$ *be a retraction on $\mathcal{M}$. Given $x \in \mathcal{M}$, the differential of* R *at $(x, 0)$ is a linear map* $\mathrm{DR}(x, 0)\colon \mathrm{T}_{(x,0)}\mathrm{T}\mathcal{M} \to \mathrm{T}_{\mathrm{R}(x,0)}\mathcal{M}$. *Equivalently, it is a linear map* $\mathrm{DR}(x, 0)\colon \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$. *For all $u, v \in \mathrm{T}_x\mathcal{M}$, it holds*

$$\mathrm{DR}(x, 0)[(u, v)] = u + v. \tag{4.19}$$

Proof. To verify (4.18), note the following:

1. For each $u \in \mathrm{T}_x\mathcal{M}$, we can pick a smooth curve c on $\mathcal{M}$ such that $c(0) = x$ and $c'(0) = u$; then, $\gamma(t) = (c(t), 0)$ is a smooth curve on $\mathrm{T}\mathcal{M}$ and $\gamma'(0) = (u, 0)$ is tangent to $\mathrm{T}\mathcal{M}$ at $\gamma(0) = (x, 0)$.
2. For each $v \in \mathrm{T}_x\mathcal{M}$, the curve $\gamma(t) = (x, tv)$ is smooth on $\mathrm{T}\mathcal{M}$ and $\gamma'(0) = (0, v)$ is tangent to $\mathrm{T}\mathcal{M}$ at $\gamma(0) = (x, 0)$.
3. By linearity, $\mathrm{T}_{(x,0)}\mathrm{T}\mathcal{M}$ contains all pairs $(u, v) \in \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M}$.

4. The two sets are in fact the same since they are linear spaces of the same dimension.

Let us establish (4.19). By linearity of differentials, we have

$$\mathrm{DR}(x,0)[(u,v)] = \mathrm{DR}(x,0)[(u,0)] + \mathrm{DR}(x,0)[(0,v)].$$

For the first term, pick a smooth curve c on $\mathcal{M}$ such that $c(0) = x$ and $c'(0) = u$. Then,

$$\mathrm{DR}(x,0)[(u,0)] = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{R}(c(t),0)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}c(t)\bigg|_{t=0} = c'(0) = u.$$

For the second term, we have

$$\mathrm{DR}(x,0)[(0,v)] = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{R}(x,tv)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{R}_x(tv)\bigg|_{t=0} = v.$$

This concludes the proof. We used both defining properties of R. □

The following lemma connects the regularity assumption A4.3 from Section 4.4 to the Hessian of f at $x_\star$ as needed in Theorem 4.20.

Lemma 4.22. *Let* R *be a retraction on a Riemannian manifold* $\mathcal{M}$. *If A4.3 holds for* $f\colon \mathcal{M} \to \mathbb{R}$ *with constant* L *on* $\mathrm{T}\mathcal{M}$ *and* $x_\star \in \mathcal{M}$ *is critical, then*

$$\forall v \in \mathrm{T}_{x_\star}\mathcal{M}, \qquad\qquad \langle v, \mathrm{Hess} f(x_\star)[v]\rangle_{x_\star} \leq L\|v\|_{x_\star}^2.$$

In particular, L *is valid for Theorem 4.20 since* $L \geq \lambda_{\max}(\mathrm{Hess} f(x_\star))$.

Proof. We only need A4.3 for all pairs $(x_\star, v) \in \mathrm{T}\mathcal{M}$; that provides:

$$f(\mathrm{R}_{x_\star}(tv)) \leq f(x_\star) + L\frac{t^2}{2}\|v\|_{x_\star}^2.$$

Anticipating results from Chapter 5, we deduce from (5.28) that

$$f(\mathrm{R}_{x_\star}(tv)) = f(x_\star) + \frac{t^2}{2}\langle v, \mathrm{Hess} f(x_\star)[v]\rangle_{x_\star} + O(t^3).$$

The two combine to yield $\langle v, \mathrm{Hess} f(x_\star)[v]\rangle_{x_\star} \leq L\|v\|_{x_\star}^2 + O(t)$. Take $t \to 0$ to conclude. □

Theorem 4.20 above provides circumstances for iterates x_k of RGD to converge to a local minimizer $x_\star$ at least linearly. We remark in closing that, using tools from Section 10.4, it is easily argued that the gradient norm $\|\mathrm{grad} f(x_k)\|$ also converges to zero, and likewise the cost function value $f(x_k)$ converges to $f(x_\star)$, both at least linearly. Specifically, this is done with Corollaries 10.48 and 10.54 by arguing that the gradient of f is Lipschitz continuous in a neighborhood of $x_\star$.

4.7 Computing gradients*

This section provides some guidance on how to obtain an expression for the gradient of a function. It can be skipped safely. The reader may find it helpful to return to this section when working on particular applications.

The gradient of a function $f \colon \mathcal{M} \to \mathbb{R}$ on a Riemannian manifold (recall Definition 3.58 or 8.57) is defined in full generality as the unique vector field $\mathrm{grad} f$ on $\mathcal{M}$ such that, for all points $x \in \mathcal{M}$ and all tangent vectors $v \in \mathrm{T}_x\mathcal{M}$,

$$\mathrm{D}f(x)[v] = \langle \mathrm{grad} f(x), v \rangle_x \,, \tag{4.20}$$

where $\langle \cdot, \cdot \rangle_x$ is the inner product on $\mathrm{T}_x\mathcal{M}$ (the Riemannian metric at x). This suggests a general strategy to obtain a formula for $\mathrm{grad} f(x)$:

1. Determine an expression for the directional derivative $\mathrm{D}f(x)[v]$, and
2. Re-arrange it until it is of the form $\langle g, v \rangle_x$, with some $g \in \mathrm{T}_x\mathcal{M}$.

At this point, we get the gradient by identification: $\mathrm{grad} f(x) = g$. This requires essentially two steps: first, to write out $\mathrm{D}f(x)[v]$ somewhat explicitly as an inner product between two quantities; second, to use the notion of adjoint of a linear map (recall Definition 3.4) to isolate v.

In working out directional derivatives, three rules get most of the work done (it is an exercise to verify them):

1. The chain rule: as for (3.29), let $F \colon \mathcal{M} \to \mathcal{M}'$ and $G \colon \mathcal{M}' \to \mathcal{M}''$ be smooth maps between manifolds $\mathcal{M}, \mathcal{M}', \mathcal{M}''$. The composition $H = G \circ F$ defined by $H(x) = G(F(x))$ is smooth with differential:
$$\mathrm{D}H(x)[v] = \mathrm{D}G(F(x))[\mathrm{D}F(x)[v]]. \tag{4.21}$$
2. The product rule: let F, G be two smooth maps from a manifold $\mathcal{M}$ to matrix spaces such that $F(x)$ and $G(x)$ can be matrix-multiplied to form the product map $H = FG$ defined by $H(x) = F(x)G(x)$. For example, F maps $\mathcal{M}$ to $\mathbb{R}^{n\times k}$ and G maps $\mathcal{M}$ to $\mathbb{R}^{k\times d}$. Then, H is smooth with differential:
$$\mathrm{D}H(x)[v] = \mathrm{D}F(x)[v]G(x) + F(x)\mathrm{D}G(x)[v]. \tag{4.22}$$
This rule holds for any type of product. For example, with the entrywise product $H(x) = F(x) \odot G(x)$, we have
$$\mathrm{D}H(x)[v] = \mathrm{D}F(x)[v] \odot G(x) + F(x) \odot \mathrm{D}G(x)[v]. \tag{4.23}$$
Likewise, with the Kronecker product $H(x) = F(x) \otimes G(x)$,
$$\mathrm{D}H(x)[v] = \mathrm{D}F(x)[v] \otimes G(x) + F(x) \otimes \mathrm{D}G(x)[v]. \tag{4.24}$$
3. Inner product rule: let $F, G \colon \mathcal{M} \to \mathcal{E}$ be two smooth maps from a manifold $\mathcal{M}$ to a linear space $\mathcal{E}$ equipped with an inner product $\langle \cdot, \cdot \rangle$. Then, the scalar function $h(x) = \langle F(x), G(x) \rangle$ is smooth with differential:
$$\mathrm{D}h(x)[v] = \langle \mathrm{D}F(x)[v], G(x) \rangle + \langle F(x), \mathrm{D}G(x)[v] \rangle \,. \tag{4.25}$$

(To differentiate inner products of two smooth vector fields on a manifold, we need more tools: see Sections 5.4 and 5.7, specifically Theorems 5.6 and 5.29.)

Then, to see how the notion of adjoint comes up, consider the common example of a cost function $f\colon \mathcal{M} \to \mathbb{R}$ with

$$f(x) = \|F(x)\|_{\mathcal{E}}^2 = \langle F(x), F(x)\rangle_{\mathcal{E}}\,, \tag{4.26}$$

where $F\colon \mathcal{M} \to \mathcal{E}$ is a map from a Riemannian manifold to a Euclidean space (e.g., a matrix space) endowed with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{E}}$ and associated norm $\|\cdot\|_{\mathcal{E}}$. From (4.25), we know that

$$\mathrm{D}f(x)[v] = \langle \mathrm{D}F(x)[v], F(x)\rangle_{\mathcal{E}} + \langle F(x), \mathrm{D}F(x)[v]\rangle_{\mathcal{E}} = 2\,\langle F(x), \mathrm{D}F(x)[v]\rangle_{\mathcal{E}}\,.$$

The linear map $\mathrm{D}F(x)\colon \mathrm{T}_x\mathcal{M} \to \mathcal{E}$ has an adjoint with respect to the inner products on $\mathrm{T}_x\mathcal{M}$ and $\mathcal{E}$; we denote it by $\mathrm{D}F(x)^*\colon \mathcal{E} \to \mathrm{T}_x\mathcal{M}$. It follows by definition that

$$\mathrm{D}f(x)[v] = 2\,\langle \mathrm{D}F(x)^*[F(x)], v\rangle_x\,.$$

This holds for all $v \in \mathrm{T}_x\mathcal{M}$, thus by identification with (4.20) we find:

$$\mathrm{grad} f(x) = 2\mathrm{D}F(x)^*[F(x)]. \tag{4.27}$$

This highlights the importance of computing adjoints of linear maps in obtaining gradients. Formulas (3.15) and (3.18) are particularly helpful in this respect. We further illustrate the computation of adjoints in examples below.

In many cases, it is sufficient to work out the gradient of a function defined on a Euclidean space, then to use a rule to convert it to a Riemannian gradient. For example, Proposition 3.61 shows how to obtain the Riemannian gradient of a function f defined on a Riemannian submanifold of a Euclidean space $\mathcal{E}$ by orthogonal projection to tangent spaces. Thus, below we focus on the Euclidean case.

Example 4.23. *Consider $F\colon \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ defined by $F(X) = X^k$ for some positive integer k. Using the product rule repeatedly, it is easy to see that*

$$\begin{aligned}\mathrm{D}F(X)[U] &= UX^{k-1} + XUX^{k-2} + X^2UX^{k-3} + \cdots + X^{k-2}UX + X^{k-1}U \\ &= \sum_{\ell=1}^{k} X^{\ell-1}UX^{k-\ell}.\end{aligned} \tag{4.28}$$

Equipping $\mathbb{R}^{n\times n}$ with the usual trace inner product, we find that the adjoint is simply $\mathrm{D}F(X)^* = \mathrm{D}F(X^\top)$. Indeed, for all $U, V \in \mathbb{R}^{n\times n}$, using (3.15),

$$\begin{aligned}\langle \mathrm{D}F(X)[U], V\rangle &= \sum_{\ell=1}^{k} \left\langle X^{\ell-1}UX^{k-\ell}, V\right\rangle \\ &= \sum_{\ell=1}^{k} \left\langle U, (X^\top)^{\ell-1}V(X^\top)^{k-\ell}\right\rangle \\ &= \left\langle U, \mathrm{D}F(X^\top)[V]\right\rangle.\end{aligned} \tag{4.29}$$

Similarly, for $F(X) = X^k$ defined on $\mathbb{C}^{n\times n}$ equipped with the usual inner product (3.17)*, the expression for* $\mathrm{D}F(X)$ *is unchanged, and* $\mathrm{D}F(X)^* = \mathrm{D}F(X^*)$.

Example 4.24. *Consider $F(X) = X^{-1}$ defined on the (open) set of invertible matrices (real or complex). One can show that F is smooth on that domain. By definition,*

$$F(X)X = I$$

for all X in the domain of F. Differentiating that identity at X along U on both sides, the product rule yields

$$\mathrm{D}F(X)[U]X + F(X)U = 0.$$

Hence, we get the following useful expression:

$$\mathrm{D}F(X)[U] = -X^{-1}UX^{-1}. \tag{4.30}$$

Equipping $\mathbb{R}^{n\times n}$ or $\mathbb{C}^{n\times n}$ with its usual inner product, the adjoint $\mathrm{D}F(X)^*$ *is* $\mathrm{D}F(X^\top)$ *or* $\mathrm{D}F(X^*)$*, as in the previous example.*

Example 4.25. *Consider a differentiable scalar function $g\colon \mathbb{R} \to \mathbb{R}$, and let $\tilde{g}\colon \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ denote its entrywise application to matrices so that $\tilde{g}(X)_{ij} = g(X_{ij})$. Then, with g' the derivative of g,*

$$\forall i, j, \qquad (\mathrm{D}\tilde{g}(X)[U])_{ij} = \mathrm{D}g(X_{ij})[U_{ij}] = g'(X_{ij})U_{ij}.$$

Letting $\tilde{g}'\colon \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ denote the entrywise application of g' to matrices, we can summarize this as

$$\mathrm{D}\tilde{g}(X)[U] = \tilde{g}'(X) \odot U. \tag{4.31}$$

This differential is self-adjoint with respect to the usual inner product, that is, $\mathrm{D}\tilde{g}(X)^* = \mathrm{D}\tilde{g}(X)$*, since for all $U, V \in \mathbb{R}^{n\times m}$, using* (3.15)*, we have*

$$\langle \mathrm{D}\tilde{g}(X)[U], V\rangle = \langle \tilde{g}'(X) \odot U, V\rangle = \langle U, \tilde{g}'(X) \odot V\rangle = \langle U, \mathrm{D}\tilde{g}(X)[V]\rangle.$$

There does not always exist a complex equivalent because for $g\colon \mathbb{C} \to \mathbb{C}$, even if $\mathrm{D}g(x)[u]$ *is well defined, there may not exist a function $g'\colon \mathbb{C} \to \mathbb{C}$ such that* $\mathrm{D}g(x)[u] = g'(x)u$*, i.e., g is not necessarily* complex differentiable. *Fortunately, this is not an obstacle to computing directional derivatives: it merely means there may not exist as simple an expression as above.*

Example 4.26. *Consider a function g from $\mathbb{R}$ to $\mathbb{R}$ or from $\mathbb{C}$ to $\mathbb{C}$ with a convergent Taylor series, that is,*

$$g(x) = \sum_{k=0}^{\infty} a_k x^k$$

for some coefficients $a_0, a_1, \ldots$, with x possibly restricted to a particular domain.

Such functions can be extended to matrix functions*, that is, to functions from* $\mathbb{R}^{n\times n}$ *to* $\mathbb{R}^{n\times n}$ *or from* $\mathbb{C}^{n\times n}$ *to* $\mathbb{C}^{n\times n}$ *simply by defining*

$$G(X) = \sum_{k=0}^{\infty} a_k X^k. \tag{4.32}$$

We can gain insight into this definition by considering the ubiquitous special case where X *is diagonalizable, that is,* $X = VDV^{-1}$ *for some diagonal matrix* $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ *containing the eigenvalues of* X *and some invertible matrix* V *containing its eigenvectors. Indeed, in this case,*

$$\begin{aligned} G(X) &= \sum_{k=0}^{\infty} a_k (VDV^{-1})^k \\ &= V\left(\sum_{k=0}^{\infty} a_k D^k\right) V^{-1} = V \mathrm{diag}(g(\lambda_1), \ldots, g(\lambda_n)) V^{-1}. \end{aligned}$$

Thus, G *is well defined at* X *provided the eigenvalues of* X *belong to the domain of definition of* g*. In this case, the matrix function* G *transforms the eigenvalues through* g*.*

Important examples include the matrix exponential, matrix logarithm and matrix square root functions. In Matlab, these are available as `expm`*,* `logm` *and* `sqrtm` *respectively. In Manopt, the differentials are available as* `dexpm`*,* `dlogm` *and* `dsqrtm`*.*

This view of matrix functions is sufficient for our discussion but it has its limitations. In particular, the Taylor series expansion does not make it immediately clear why the matrix logarithm and matrix square root can be defined for all matrices whose real eigenvalues (if any) are positive. For a more formal discussion of matrix functions—including definitions that allow us to go beyond Taylor series and diagonalizable matrices—as well as details regarding domains of definition and numerical computation, see [Hig08]. Generalized matrix functions (which apply to non-square matrices) and their differentials are discussed in [Nof17].

Provided one can compute the matrix function, a theorem by Mathias offers a convenient way to compute its directional derivatives (also called Gâteaux *and, under stronger conditions,* Fréchet derivative*) [Mat96], [Hig08, §3.2, Thm. 3.6, 3.8, eq. (3.16)]: if* g *is* $2n-1$ *times continuously differentiable on some open domain in* $\mathbb{R}$ *or* $\mathbb{C}$ *and the eigenvalues of* X *belong to this domain, then,*

$$G\left(\begin{bmatrix} X & U \\ 0 & X \end{bmatrix}\right) = \begin{bmatrix} G(X) & \mathrm{D}G(X)[U] \\ 0 & G(X) \end{bmatrix}. \tag{4.33}$$

Thus, for the cost of one matrix function computation on a matrix of size $2n\times 2n$*, we get* $G(X)$ *and* $\mathrm{D}G(X)[U]$*. This is useful, though we should bear in mind that computing matrix functions is usually easier for symmetric or Hermitian matrices: here, even if* X *is favorable in that regard, the structure is lost by forming the block matrix. If the matrices are large or poorly conditioned, it may* *help to explore alternatives [AMH09], [Hig08, §10.6, §11.8]. If an eigenvalue decomposition of* X *is available, there exists an explicit expression for* $\mathrm{D}G(X)[U]$ *involving* Loewner matrices *[Hig08, Cor. 3.12].*

We can gain insight into the adjoint of the directional derivative of a matrix function through (4.32) *and Example 4.23. Indeed,*

$$\begin{aligned}\mathrm{D}G(X)[U] &= \sum_{k=0}^{\infty} a_k \mathrm{D}\big(X \mapsto X^k\big)\,(X)[U] \\ &= \sum_{k=0}^{\infty} a_k \sum_{\ell=1}^{k} X^{\ell-1} U X^{k-\ell}. \end{aligned} \tag{4.34}$$

Assume the Taylor expansion coefficients a_k *are real: This holds for the matrix exponential, logarithm and square root. It is then straightforward to see that the adjoint with respect to the usual inner product obeys*

$$\mathrm{D}G(X)^* = \mathrm{D}G(X^*). \tag{4.35}$$

Indeed,

$$\begin{aligned}\langle \mathrm{D}G(X)[U], V\rangle &= \left\langle \sum_{k=0}^{\infty} a_k \sum_{\ell=1}^{k} X^{\ell-1} U X^{k-\ell}, V \right\rangle \\ &= \sum_{k=0}^{\infty} a_k \sum_{\ell=1}^{k} \left\langle X^{\ell-1} U X^{k-\ell}, V\right\rangle \\ &= \sum_{k=0}^{\infty} a_k \sum_{\ell=1}^{k} \left\langle U, (X^*)^{\ell-1} V (X^*)^{k-\ell}\right\rangle = \langle U, \mathrm{D}G(X^*)[V]\rangle\,. \end{aligned} \tag{4.36}$$

Of course, $X^* = X^\top$ *in the real case.*

Example 4.27. *Formulas for the directional derivatives of factors of certain* matrix factorizations *are known, including QR, LU, Cholesky, polar factorization, eigenvalue decomposition and SVD. See [Deh95, §3.1], [DMV99, DE99, Fep17, FL19, BZA20] and [AMS08, Ex. 8.1.5] among others.*

Example 4.28. *The directional derivative of* $g(X) = \log(\det(X))$ *is given by* $\mathrm{D}g(X)[U] = \mathrm{Tr}(X^{-1}U)$*, provided that* $\det(X)$ *is positive if it is real. Indeed, using* $\det(AB) = \det(A)\det(B)$ *and* $\log(ab) = \log(a) + \log(b)$*,*

$$\begin{aligned}\log(\det(X + tU)) &= \log(\det(X(I_n + tX^{-1}U))) \\ &= \log(\det(X)) + \log(\det(I_n + tX^{-1}U)) \\ &= \log(\det(X)) + t\,\mathrm{Tr}(X^{-1}U) + O(t^2), \end{aligned} \tag{4.37}$$

where we used $\det(I_n + tA) = 1 + t\,\mathrm{Tr}(A) + O(t^2)$ *then* $\log(1 + \lambda t) = \lambda t + O(t^2)$*. For the former claim, check that if* $\lambda_1, \ldots, \lambda_n$ *denote the eigenvalues of* A *then* $\det(I_n + tA) = \prod_{i=1}^{n}(1 + \lambda_i t)$*. In particular, if we restrict* g *to the set of positive* *definite matrices, then g is real valued and we conclude that* $\mathrm{grad}g(X) = X^{-1}$ *with respect to the usual inner product. The related function*

$$h(X) = \log(\det(X^{-1})) = \log(1/\det(X)) = -\log(\det(X)) = -g(X)$$

has derivatives $\mathrm{D}h(X)[U] = -\operatorname{Tr}(X^{-1}U)$ *and* $\mathrm{grad}h(X) = -X^{-1}$.

Example 4.29. *We now work out a gradient as a full example. Consider the following function which maps a pair of square matrices (X, Y) to a real number, with A and B two given real matrices (A, B, X, Y are all in $\mathbb{R}^{n\times n}$):*

$$f(X,Y) = \frac{1}{2}\|A \odot \exp(X^{-1}B)Y\|^2.$$

Here, $\exp$ *denotes the matrix exponential and* $\odot$ *denotes entrywise multiplication. See Exercise 3.67 for pointers regarding gradients on a product manifold. Define* $Q(X,Y) = A \odot \left[\exp(X^{-1}B)Y\right]$*, so that*

$$f(X,Y) = \frac{1}{2}\|Q(X,Y)\|^2 = \frac{1}{2}\left\langle Q(X,Y), Q(X,Y)\right\rangle.$$

Then, using the product rule on the inner product $\langle \cdot, \cdot \rangle$, we get the directional derivative of f at (X,Y) along the direction $(\dot{X}, \dot{Y})$ (a pair of matrices of the same size as (X,Y)):

$$\mathrm{D}f(X,Y)[\dot{X},\dot{Y}] = \left\langle \mathrm{D}Q(X,Y)[\dot{X},\dot{Y}], Q(X,Y)\right\rangle.$$

We focus on the differential of Q for now. Using that A is constant, the product rule on $\exp(\cdot)Y$ *and the chain rule on* $\exp$*, we get:*

$$\mathrm{D}Q(X,Y)[\dot{X},\dot{Y}] = A \odot \left[\mathrm{Dexp}(X^{-1}B)[U]Y + \exp(X^{-1}B)\dot{Y}\right],$$

where Dexp *is the differential of the matrix exponential, and U is the differential of $(X,Y) \mapsto X^{-1}B$ at (X,Y) along $(\dot{X},\dot{Y})$, that is, $U = -X^{-1}\dot{X}X^{-1}B$. Combining and using $W = X^{-1}B$ for short, we find*

$$\mathrm{D}f(X,Y)[\dot{X},\dot{Y}] = \left\langle A \odot \left[\mathrm{Dexp}(W)[-X^{-1}\dot{X}W]Y + \exp(W)\dot{Y}\right], Q(X,Y)\right\rangle.$$

We re-arrange the terms in this expression to reach the form $\langle \dot{X}, \cdot\rangle + \langle \dot{Y}, \cdot\rangle$. This mostly requires using the notion of adjoint of linear maps: recall Section 3.1. First using the adjoint of entrywise multiplication with respect to the usual inner product as in (3.15)*, then linearity of the inner product:*

$$\begin{aligned}\mathrm{D}f(X,Y)[\dot{X},\dot{Y}] = &\left\langle \mathrm{Dexp}(W)[-X^{-1}\dot{X}W]Y, A \odot Q(X,Y)\right\rangle \\ &+ \left\langle \exp(W)\dot{Y}, A \odot Q(X,Y)\right\rangle.\end{aligned}$$

Let $Z = A \odot Q(X,Y)$ for short; using the adjoint of matrix multiplication for both terms:

$$\mathrm{D}f(X,Y)[\dot{X},\dot{Y}] = \left\langle \mathrm{Dexp}(W)[-X^{-1}\dot{X}W], ZY^\top\right\rangle + \left\langle \dot{Y}, \exp(W)^\top Z\right\rangle.$$

The gradient with respect to Y is readily apparent from the second term. We focus on the first term. Using the adjoint of the differential of the matrix exponential at $X^{-1}B$ (denoted by a star), we get:

$$\begin{aligned}\left\langle \mathrm{Dexp}(W)[-X^{-1}\dot{X}W], ZY^\top \right\rangle &= \left\langle -X^{-1}\dot{X}W, \mathrm{Dexp}(W)^*[ZY^\top] \right\rangle \\ &= \left\langle \dot{X}, -(X^{-1})^\top \mathrm{Dexp}(W)^*[ZY^\top]W^\top \right\rangle.\end{aligned}$$

This reveals the gradient of f with respect to X. We can go one step further using the fact that $\mathrm{Dexp}(W)^* = \mathrm{Dexp}(W^\top)$. *To summarize:*

$$\mathrm{grad}f(X,Y) = \Big(-(X^{-1})^\top \mathrm{Dexp}(W^\top)[ZY^\top]W^\top,\ \exp(W)^\top Z \Big).$$

Considering that $W, \exp(W)$ and Q must be computed in order to evaluate f, it is clear that computing $\mathrm{grad}f$ *is not significantly more expensive, and much of the computations can be reused.*

If A, B, X, Y are in $\mathbb{C}^{n\times n}$ and we use the real inner product over complex matrices (3.17) *as in Section 3.1,* $\mathrm{grad}f$ *takes on the same expression except all transposes are replaced by conjugate-transposes, and $Z = \overline{A} \odot Q(X,Y)$. See also Example 4.30.*

Example 4.30. *Consider the function $f\colon \mathbb{C}^n \to \mathbb{R}$ defined by*

$$f(x) = \sum_{i=1}^m |x^* A_i x - b_i|^2,$$

where $A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$ and $b \in \mathbb{C}^m$ are given. This is a real-valued function of a complex vector. As in Section 3.1, we consider $\mathbb{C}^n$ to be a real vector space *of dimension $2n$. We further equip $\mathbb{C}^n$ with the (real) inner product $\langle u, v\rangle = \Re\{u^*v\}$ as in* (3.16). *To work out the gradient of f, as usual, we first work out its directional derivatives. Define $z = z(x)$ in $\mathbb{C}^m$ with $z_i(x) = x^*A_ix - b_i$; observe that $f(x) = \sum_{i=1}^m |z_i|^2$ and:*

$$\mathrm{D}z_i(x)[u] = u^*A_ix + x^*A_iu.$$

Above, we have used the convenient fact that complex conjugation is a linear map on the real vector space $\mathbb{C}^n$; as such, the differential of the conjugate is the conjugate of the differential. With the identities $|a|^2 = \overline{a}a$ and $\overline{a}b + \overline{b}a = 2\Re\{\overline{a}b\}$, *we find:*

$$\begin{aligned}
\mathrm{D}f(x)[u] &= \sum_{i=1}^{m} \overline{z_i} \cdot \mathrm{D}z_i(x)[u] + \overline{\mathrm{D}z_i(x)[u]} \cdot z_i \\
&= 2\sum_{i=1}^{m} \Re\{\overline{z_i} \cdot (u^* A_i x + x^* A_i u)\} \\
&= 2\Re\left\{\sum_{i=1}^{m} \overline{z_i} \cdot (u^* A_i x)\right\} + 2\Re\left\{\sum_{i=1}^{m} \overline{z_i} \cdot (x^* A_i u)\right\} \\
&= 2\left\langle u, \sum_{i=1}^{m} \overline{z_i} \cdot A_i x \right\rangle + 2\left\langle \sum_{i=1}^{m} z_i \cdot A_i^* x, u \right\rangle.
\end{aligned}$$

By identification in the definition $\langle \mathrm{grad} f(x), u\rangle = \mathrm{D}f(x)[u]$, *we deduce:*

$$\mathrm{grad} f(x) = 2\left(\sum_{i=1}^{m} \overline{z_i} A_i + z_i A_i^*\right) x.$$

In the particular case where $A_i = A_i^*$ *and (as one might expect in that case) where* b *is real, we also have that* z *is real and the gradient further simplifies to* $\mathrm{grad} f(x) = 4\left(\sum_{i=1}^{m} z_i A_i\right) x$.

The *cheap gradient principle* [GW08, p88] asserts that, for a wide class of functions f, computing the gradient of f at a point requires no more than a multiple (often five or less) of the number of arithmetic operations required to evaluate f itself at that point. Furthermore, much of the computations required to evaluate the cost function can be reused to evaluate its gradient at the same point. Thus, if it appears that computing the gradient takes inordinately more time than it takes to evaluate the cost, chances are the code can be improved. Anticipating the introduction of Hessians, we note that a similar fact holds for Hessian-vector products [Pea94].

The latter principle is at the heart of *automatic differentiation* (AD): algorithms that automatically compute the derivatives of a function, based simply on code to compute that function. AD can significantly speed up development time. Packages such as Manopt offer AD for optimization on manifolds.

Exercise 4.31. *Prove rules* (4.21), (4.22), (4.23), (4.24) *and* (4.25).

Exercise 4.32. *The (principal) matrix square root function* $F(X) = X^{1/2}$ *is well defined provided real eigenvalues of* X *are positive [Hig08, Thm. 1.29]. Show that* $\mathrm{D}F(X)[U] = E$, *where* E *is the matrix which solves the Sylvester equation* $EX^{1/2} + X^{1/2}E = U$. *Hint: consider* $G(X) = X^2$ *and* $F = G^{-1}$.

Exercise 4.33. *For a matrix function whose Taylor expansion coefficients are real, show that* $\mathrm{D}F(X)[U^*] = (\mathrm{D}F(X^*)[U])^*$. *Combining with* (4.36), *this yields:* $\mathrm{D}F(X)^*[U] = \mathrm{D}F(X^*)[U] = (\mathrm{D}F(X)[U^*])^*$.

4.8 Numerically checking a gradient*

After writing code to evaluate a cost function $f(x)$ and its Riemannian gradient $\mathrm{grad} f(x)$, it is often helpful to run numerical tests to catch possible mistakes early. This section describes such tests. In the Matlab toolbox Manopt, they are implemented as `checkgradient`.

The first thing to test is that $\mathrm{grad} f(x)$ is indeed in the tangent space at x. This being secured, consider the Taylor expansion (4.3):

$$f(\mathrm{R}_x(tv)) = f(x) + t \langle \mathrm{grad} f(x), v \rangle_x + O(t^2).$$

This says that, for all $x \in \mathcal{M}$ and $v \in \mathrm{T}_x\mathcal{M}$, with all retractions,

$$E(t) \triangleq |f(\mathrm{R}_x(tv)) - f(x) - t \langle \mathrm{grad} f(x), v \rangle_x| = O(t^2). \tag{4.38}$$

Taking the logarithm on both sides, we find that $\log(E(t))$ must grow approximately linearly in $\log(t)$, with a slope of two (or more[1]) when t is small:

$$\log(E(t)) \approx 2\log(t) + \text{constant}.$$

This suggests a procedure to check the gradient numerically:

1. Generate a random point $x \in \mathcal{M}$;
2. Generate a random tangent vector $v \in \mathrm{T}_x\mathcal{M}$ with $\|v\|_x = 1$;
3. Compute $f(x)$ and $\mathrm{grad} f(x)$. Check that $\mathrm{grad} f(x)$ is in $\mathrm{T}_x\mathcal{M}$, and compute $\langle \mathrm{grad} f(x), v \rangle_x$;
4. Compute $E(t)$ for several values of t logarithmically spaced on the interval $[10^{-8}, 10^0]$;
5. Plot $E(t)$ as a function of t, in a log–log plot;
6. Check that the plot exhibits a slope of two (or more) over several orders of magnitude.

We do not expect to see a slope of two over the whole range: On the one hand, for large t, the Taylor approximation may be poor; On the other hand, for small t, floating-point arithmetic strongly affects the computation of $E(t)$ (see also the discussion in Section 6.4.6). Still, we do expect to see a range of values of t for which the numerical computation is accurate and the Taylor expansion is valid. If the curve does not exhibit a slope of two over at least one or two orders of magnitude, this is a strong sign that there is a mistake in the computation of the gradient (or the cost function, or the retraction, or the inner product).

Example 4.34. *With some symmetric matrix A and size n, recall the cost function $f(X) = -\frac{1}{2}\mathrm{Tr}(X^\top AX)$ defined on the Stiefel manifold* $\mathrm{St}(n,p)$. *Its gradient is the orthogonal projection of $-AX$ to the tangent space at X. Figure 4.1 plots the numerical gradient check described above, obtained first with an incorrect gradient (the minus sign was forgotten), then with the correct gradient. Notice how*

[1] If the Taylor remainder happens to be $O(t^k)$ with $k > 2$, we should get a slope of k. This is good but rare.

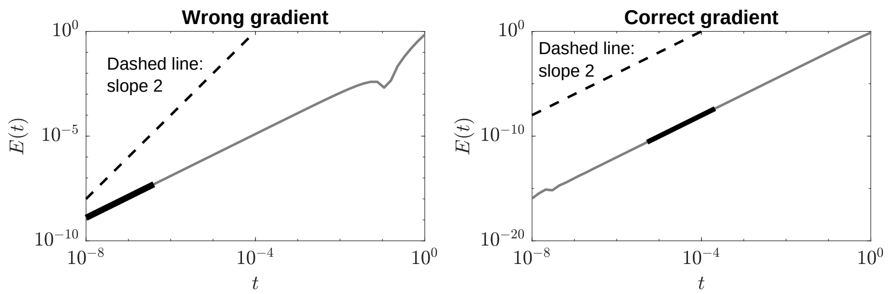


Figure 4.1 Example 4.34 illustrates a numerical procedure to check gradient computation code. The dashed lines have a slope of two: this serves as a visual reference. The solid curves represent the function $E(t)$ (4.38) in a log–log plot. Part of each solid curve is overlaid with a thicker line. The (average) slopes of those thick lines are nearly one (left) and two (right), strongly suggesting the left gradient is incorrect, and suggesting the right gradient is correct (as is indeed the case).

for the incorrect gradient the blue curve has (mostly) a slope of one, whereas for the correct gradient it has (mostly) a slope of two. This figure is obtained with the following Matlab code, using Manopt.

```
n = 50;
A = randn(n, n);
A = A + A';

inner = @(U, V) U(:)'*V(:); % = trace(U'*V)
St = stiefelfactory(n, 3);
problem.M = St;
problem.cost = @(X) -0.5*inner(X, A*X);

problem.grad = @(X) St.proj(X, A*X);  % Oops, forgot -
checkgradient(problem); % First panel

problem.grad = @(X) St.proj(X, -A*X); % This is better
checkgradient(problem); % Second panel

X = steepestdescent(problem); % Call to RGD
X = trustregions(problem);    % Call to RTR (Chapter 6)
```

4.9 Notes and references

Absil et al. give a thorough treatment and history of Riemannian gradient descent in [AMS08, §4], with references going back to [Lue72, Gab82, Smi94, Udr94, HM96, Rap97, EAS98]. Gabay [Gab82] details the important work of Lichnewsky [Lic79], who generalized Luenberger's pioneering paper [Lue72] from Riemannian submanifolds of Euclidean space to general manifolds, and designed a Riemannian nonlinear conjugate gradients method for nonlinear eigenvalue problems.

The first iteration complexity analyses in the Riemannian setting appear about the same time on public repositories in [ZS16, BAC18, BFM17], under various related models. The analysis presented here is largely based on [BAC18]. Before that, analyses with similar ingredients including Lipschitz-type assumptions (but phrased as asymptotic convergence results) appear notably in [dCNdLO98, Thm. 5.1].

Several first-order optimization algorithms on Riemannian manifolds are available, including nonlinear conjugate gradients [AMS08, SI15, Sat16] (pioneered by Lichnewsky), BFGS [BM06, QGA10b, RW12, HGA15, HAG16] (pioneered by Gabay) and (variance reduced) stochastic gradients [Bon13, ZRS16, KSM18, SKM19]. See also the book by Sato [Sat21] which provides a general introduction to Riemannian optimization and an in depth treatment of the nonlinear conjugate gradients method. Quadratic convergence results for the latter appear in [Smi94, §5]. Regarding stochastic methods, Hosseini and Sra propose a survey in [HS20].

There is also recent work focused on nonsmooth cost functions on smooth manifolds, including proximal point methods and subgradient methods [BFM17, CMSZ20], gradient sampling [HU17] and ADMM-type algorithms [KGB16].

Many of these more advanced algorithms require *transporters* or *vector transports*, which we cover in Section 10.5: these are tools to transport tangent vectors and linear maps from one tangent space to another.

In (4.11), we considered the standard proof that Lipschitz continuity of the gradient of f (in the Euclidean case) implies uniform bounds on the truncation error of first-order Taylor expansions of f. If f is twice continuously differentiable, it is not difficult to show that the converse also holds because the gradient is Lipschitz-continuous if and only if the Hessian is bounded. See [BAJN20, Cor. 5.1] for a more general discussion assuming Hölder continuity of the gradient.

See [BH19] for first-order necessary optimality conditions when x, in addition to living on a manifold $\mathcal{M}$, may be further restricted by equality and inequality constraints. Second-order optimality conditions are also investigated in Section 6.1 and in [YZS14].

After Proposition 4.7, we observed that (under the stated conditions) the statement guarantees all accumulation points of RGD are critical points, but it does not guarantee convergence of the iterates: there could be more than one accumulation point. This type of behavior is undesirable, and by all accounts uncommon. The local convergence results outlined in Section 4.6 exclude such pathological cases near critical points where the Riemannian Hessian (introduced in Section 5.5) is positive definite. For more general conditions based on analyticity, see notably [AK06]. See [Lag07] and [BH15, Thm. 4.1, Cor. 4.2] for connections to optimization on manifolds. At their core, these results rely on the *Kurdyka–Łojasiewicz inequality* for real analytic functions [Łoj65]. See also Lemma 11.28 and Theorem 11.29 for the geodesically strongly convex case via the *Polyak–Łojasiewicz inequality*.

The superlinear convergence claim in Theorem 4.19 also appears with a distinction regarding the degree of differentiability of the iteration map in [AMS08, Thm. 4.5.3].

5 Embedded submanifolds: second-order geometry

In previous chapters, we developed a notion of gradient for smooth functions on manifolds. We explored how this notion is useful both to analyze optimization problems and to design algorithms for them. In particular, we found that local minimizers are critical points, that is, the gradient vanishes there. Furthermore, we showed under a regularity condition that following the negative gradient allows us to find critical points.

A tool of choice in those developments has been a type of first-order Taylor expansion of the cost function along a curve. Concretely, for a smooth curve c on $\mathcal{M}$ passing through x at $t = 0$ with velocity $c'(0) = v$, we considered the composition $g = f \circ c$ (a smooth function from reals to reals), and its truncated Taylor expansion

$$g(t) = g(0) + tg'(0) + O(t^2) = f(x) + t \left\langle \mathrm{grad} f(x), v \right\rangle_x + O(t^2).$$

To gain further control over g, it is natural to ask what happens if we truncate the expansion one term later, that is, if we write

$$g(t) = f(x) + t \left\langle \mathrm{grad} f(x), v \right\rangle_x + \frac{t^2}{2} g''(0) + O(t^3).$$

In the Euclidean case, with the straight curve $c(t) = x + tv$, we would find the well-known formula

$$g(t) = f(x + tv) = f(x) + t \left\langle \mathrm{grad} f(x), v \right\rangle + \frac{t^2}{2} \left\langle \mathrm{Hess} f(x)[v], v \right\rangle + O(t^3),$$

where $\mathrm{Hess} f(x)$ is the *Hessian* of f at x.

This leads us to ponder: can we define an equivalent of the Hessian of a function on a Riemannian manifold? To make progress on this question, we first review how Hessians are defined on Euclidean spaces.

Recall from Section 3.1 the definition of the gradient and Hessian of a smooth function $f : \mathcal{E} \to \mathbb{R}$ on a Euclidean space $\mathcal{E}$ with inner product $\langle \cdot, \cdot \rangle$. The gradient of f is the map $\mathrm{grad} f : \mathcal{E} \to \mathcal{E}$ which satisfies $\langle \mathrm{grad} f(x), v \rangle = \mathrm{D} f(x)[v]$ for all $x, v \in \mathcal{E}$. The Hessian of f at x is the linear map $\mathrm{Hess} f(x) : \mathcal{E} \to \mathcal{E}$ defined by

$$\mathrm{Hess} f(x)[v] = \mathrm{D}(\mathrm{grad} f)(x)[v] = \lim_{t \to 0} \frac{\mathrm{grad} f(x + tv) - \mathrm{grad} f(x)}{t}.$$

Thus, $\mathrm{Hess} f(x)[v]$ tells us how much the gradient changes if x is perturbed along v, up to first order.

For the special case where $\mathcal{E} = \mathbb{R}^n$ is equipped with the standard inner product $\langle u, v \rangle = u^\top v$, we already reasoned that the gradient is the vector of partial derivatives of f. In that case, we also recover a familiar form of the Hessian as the symmetric matrix of second-order partial derivatives:

$$\mathrm{grad} f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}, \qquad \mathrm{Hess} f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{bmatrix}.$$

Indeed, we can confirm this by working out the directional derivative of the gradient vector field $\mathrm{grad} f$ at x along $v \in \mathbb{R}^n$:

$$\begin{aligned} \mathrm{D}(\mathrm{grad} f)(x)[v] &= \begin{bmatrix} \mathrm{D}\left(\frac{\partial f}{\partial x_1}\right)(x)[v] \\ \vdots \\ \mathrm{D}\left(\frac{\partial f}{\partial x_n}\right)(x)[v] \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) v_1 + \cdots + \frac{\partial^2 f}{\partial x_n \partial x_1}(x) v_n \\ \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) v_1 + \cdots + \frac{\partial^2 f}{\partial x_n \partial x_n}(x) v_n \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \mathrm{Hess} f(x)[v]. \end{aligned}$$

While this special case is important, the general definition of the Hessian as the derivative of the gradient vector field is more relevant: it leads the way forward.

Accordingly, to extend the concept of Hessian to Riemannian manifolds, we need a good notion of derivative of vector fields. As we shall see, the derivative we already have (Definition 3.34) is not appropriate. To overcome this, we introduce a new notion of derivative for vector fields called a *connection* or *covariant derivative*. That naturally leads to a notion of covariant derivative of a vector field along a curve. With a particularly apt choice of connection called the *Riemannian connection* or *Levi-Civita connection*, we will be able to complete the Taylor expansion above as follows:

$$\begin{aligned} f(c(t)) = f(x) + t \langle \mathrm{grad} f(x), v \rangle_x &+ \frac{t^2}{2} \langle \mathrm{Hess} f(x)[v], v \rangle_x \\ &+ \frac{t^2}{2} \langle \mathrm{grad} f(x), c''(0) \rangle_x + O(t^3), \end{aligned}$$

where $\mathrm{Hess} f(x)$ is the *Riemannian Hessian* of f at x we are about to define, and $c''(t)$ is the *acceleration* along c: the covariant derivative of its velocity vector field $c'(t)$. Importantly, $\mathrm{Hess} f(x)$ retains familiar properties. For example, it is symmetric as a linear map from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_x\mathcal{M}$; that is: it is self-adjoint with respect to the Riemannian metric $\langle\cdot,\cdot\rangle_x$. These considerations occupy us for most of this chapter.

We already encountered the Riemannian Hessian in studying the local convergence behavior of gradient descent in Section 4.6. In Chapter 6, we use the notion of Riemannian Hessian and the extended Taylor expansion to develop so-called *second-order optimization algorithms*.

5.1 The case for another derivative of vector fields

Consider the unit sphere S^{d-1} as a Riemannian submanifold of $\mathbb{R}^d$ with the canonical inner product $\langle u, v\rangle = u^\top v$. For a given symmetric matrix A of size d, let $f(x) = \frac{1}{2}x^\top Ax$ be defined on S^{d-1}. We know from Example 3.62 that the Riemannian gradient of f is the following smooth vector field on S^{d-1}:

$$V(x) = \mathrm{grad} f(x) = Ax - (x^\top Ax)x.$$

Since V is a smooth map from S^{d-1} to its tangent bundle (two manifolds), we already have a notion of differential for V provided by Definition 3.34. We can compute the latter via (3.28). Explicitly, with the smooth extension

$$\bar{V}(x) = Ax - (x^\top Ax)x$$

defined on all of $\mathbb{R}^d$, we have for all tangent vectors $u \in \mathrm{T}_x\mathrm{S}^{d-1}$:

$$\begin{aligned} \mathrm{D}V(x)[u] = \mathrm{D}\bar{V}(x)[u] &= Au - (x^\top Ax)u - (u^\top Ax + x^\top Au)x \\ &= \mathrm{Proj}_x(Au) - (x^\top Ax)u - (u^\top Ax)x, \end{aligned} \tag{5.1}$$

where $\mathrm{Proj}_x(v) = v - (x^\top v)x$ is the orthogonal projector from $\mathbb{R}^d$ to $\mathrm{T}_x\mathrm{S}^{d-1}$. Evidently, $\mathrm{D}V(x)[u]$ is *not* always tangent to S^{d-1} at x: the first two terms in (5.1) are tangent, but the third one is not whenever $u^\top Ax \neq 0$. Thus, if we were to use that notion of derivative of gradient vector fields to define Hessians, we would find ourselves in the uncomfortable situation where $\mathrm{Hess} f(x)[u]$, defined as $\mathrm{D}(\mathrm{grad} f)(x)[u]$, might not be a tangent vector at x. As a result, $\mathrm{Hess} f(x)$ would not be a linear map to and from $\mathrm{T}_x\mathrm{S}^{d-1}$, and terms such as $\langle \mathrm{Hess} f(x)[u], u\rangle_x$ would make no sense. We need a new derivative for vector fields.

5.2 Another look at differentials of vector fields in linear spaces

We aim to define a new derivative for vector fields on manifolds. In so doing, we follow the axiomatic approach, that is: we prescribe properties we would like that derivative to have, and later we show there exists a unique operator that satisfies them. Of course, the classical derivative of vector fields on linear spaces should qualify: let us have a look at some of its elementary properties for inspiration.

Let $\mathcal{E}$ be a linear space. Recall that the differential of a smooth vector field $V \in \mathfrak{X}(\mathcal{E})$ at a point x along u is given by:

$$\mathrm{D}V(x)[u] = \lim_{t \to 0} \frac{V(x+tu) - V(x)}{t}. \tag{5.2}$$

Given three smooth vector fields $U, V, W \in \mathfrak{X}(\mathcal{E})$, two vectors $u, w \in \mathcal{E}$ (we think of them as being "tangent at x"), two real numbers $a, b \in \mathbb{R}$ and a smooth function $f \in \mathfrak{F}(\mathcal{E})$, we know from classical calculus that the following properties hold:

1. $\mathrm{D}V(x)[au + bw] = a\mathrm{D}V(x)[u] + b\mathrm{D}V(x)[w]$;
2. $\mathrm{D}(aV + bW)(x)[u] = a\mathrm{D}V(x)[u] + b\mathrm{D}W(x)[u]$; and
3. $\mathrm{D}(fV)(x)[u] = \mathrm{D}f(x)[u] \cdot V(x) + f(x)\mathrm{D}V(x)[u]$.

Furthermore, the map $x \mapsto \mathrm{D}V(x)[U(x)]$ is smooth since U and V are smooth, and it defines a vector field on $\mathcal{E}$. This constitutes a first set of properties we look to preserve on manifolds.

5.3 Differentiating vector fields on manifolds: connections

Our new notion of derivative for vector fields on manifolds is called a *connection* (or *affine connection*), traditionally denoted by ∇ (read: "nabla"). Given a tangent vector $u \in \mathrm{T}_x\mathcal{M}$ and a vector field V, we think of $\nabla_u V$ as a derivative of V at x along u. Formally, we should write $\nabla_{(x,u)} V$, but the base point x is typically clear from context. Note that we do not need a Riemannian metric yet.

Definition 5.1. *A* connection *on a manifold $\mathcal{M}$ is an operator*

$$\nabla \colon \mathrm{T}\mathcal{M} \times \mathfrak{X}(\mathcal{M}) \to \mathrm{T}\mathcal{M} \colon (u, V) \mapsto \nabla_u V$$

such that $\nabla_u V$ is in $\mathrm{T}_x\mathcal{M}$ whenever u is in $\mathrm{T}_x\mathcal{M}$ and which satisfies four properties for all $U, V, W \in \mathfrak{X}(\mathcal{M})$, $u, w \in \mathrm{T}_x\mathcal{M}$ and $a, b \in \mathbb{R}$:

0. *Smoothness: $(\nabla_U V)(x) \triangleq \nabla_{U(x)} V$ defines a smooth vector field $\nabla_U V$;*
1. *Linearity in u: $\nabla_{au+bw} V = a\nabla_u V + b\nabla_w V$;*
2. *Linearity in V: $\nabla_u(aV + bW) = a\nabla_u V + b\nabla_u W$; and*
3. *Leibniz rule: $\nabla_u(fV) = \mathrm{D}f(x)[u] \cdot V(x) + f(x)\nabla_u V$.*

The field $\nabla_U V$ is the covariant derivative *of V along U with respect to ∇.*

See Section 5.6 for a more common (and equivalent) definition.

There exist many connections. For example, on a linear space $\mathcal{E}$,

$$\nabla_u V = \mathrm{D}V(x)[u] \tag{5.3}$$

is a connection by design. More interestingly, there exist connections on manifolds. Here is an example for $\mathcal{M}$ embedded in a Euclidean space $\mathcal{E}$: based on the discussion in Section 5.1, one may surmise that a possible fix for the standard notion of derivative of vector fields is to project the result to tangent spaces. This can be done as follows:

$$\nabla_u V = \mathrm{Proj}_x\big(\mathrm{D}\bar{V}(x)[u]\big), \tag{5.4}$$

where Proj_x is the projector from $\mathcal{E}$ to $\mathrm{T}_x\mathcal{M}$—orthogonal with respect to the Euclidean metric on $\mathcal{E}$—and $\bar{V}$ is any smooth extension of V. That is indeed a valid connection.

Theorem 5.2. *Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$. The operator ∇ defined by* (5.4) *is a connection on $\mathcal{M}$.*

Proof. It is helpful to denote the connection (5.3) on $\mathcal{E}$ by $\bar{\nabla}$. Then,

$$\nabla_u V = \mathrm{Proj}_x\big(\bar{\nabla}_u \bar{V}\big). \tag{5.5}$$

If $\mathcal{M}$ is an open submanifold of $\mathcal{E}$, the claim is clear since Proj_x is identity and we may take $\bar{V} = V$. We now handle $\mathcal{M}$ not open in $\mathcal{E}$. Consider $U, V, W \in \mathfrak{X}(\mathcal{M})$ together with smooth extensions $\bar{U}, \bar{V}, \bar{W} \in \mathfrak{X}(O)$ defined on a neighborhood O of $\mathcal{M}$ in $\mathcal{E}$. As we just argued, $\bar{\nabla}$ is a connection on O since O is an open submanifold of $\mathcal{E}$. Also consider $a, b \in \mathbb{R}$ and $u, w \in \mathrm{T}_x\mathcal{M}$. Using consecutively that $\bar{\nabla}$ is a connection and that Proj_x is linear, it is straightforward to verify linearity in the first argument:

$$\begin{aligned}\nabla_{au+bw} V &= \mathrm{Proj}_x\big(\bar{\nabla}_{au+bw}\bar{V}\big)\\ &= \mathrm{Proj}_x\big(a\bar{\nabla}_u\bar{V} + b\bar{\nabla}_w\bar{V}\big)\\ &= a\nabla_u V + b\nabla_w V.\end{aligned}$$

Likewise, linearity in the second argument holds since:

$$\begin{aligned}\nabla_u(aV + bW) &= \mathrm{Proj}_x\big(\bar{\nabla}_u(a\bar{V} + b\bar{W})\big)\\ &= \mathrm{Proj}_x\big(a\bar{\nabla}_u\bar{V} + b\bar{\nabla}_u\bar{W}\big)\\ &= a\nabla_u V + b\nabla_u W.\end{aligned}$$

To verify the Leibniz rule, consider an arbitrary $f \in \mathfrak{F}(\mathcal{M})$ and smooth extension $\bar{f} \in \mathfrak{F}(O)$. Then, using that $\bar{f}\bar{V}$ is a smooth extension for fV on O it follows that

$$\begin{aligned}\nabla_u(fV) &= \mathrm{Proj}_x\big(\bar{\nabla}_u(\bar{f}\bar{V})\big)\\ &= \mathrm{Proj}_x\big(\mathrm{D}\bar{f}(x)[u]\cdot\bar{V}(x) + \bar{f}(x)\bar{\nabla}_u\bar{V}\big)\\ &= \mathrm{D}f(x)[u]\cdot V(x) + f(x)\nabla_u V,\end{aligned}$$

as desired. Finally, we see that $\nabla_U V$ is smooth as per Exercise 3.66. $\square$

Not only do connections exist, but actually: there exist infinitely many of them on any manifold $\mathcal{M}$. For instance, we can consider (5.4) with other projectors.

As a result, the connection (5.4) may seem arbitrary. In the next section, we show that the number of connections can be reduced to one if we require further properties to hold with respect to a Riemannian structure on $\mathcal{M}$. As it turns out, the connection (5.4) satisfies those additional properties if $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$. If we endow $\mathcal{M}$ with a different Riemannian metric, there still exists a preferred connection for $\mathcal{M}$ but it may differ from (5.4).

We close this section with an observation: all connections coincide at critical points of a vector field. For optimization, we will see manifestations of this fact when applied to the gradient vector field at a critical point of a cost function f. The proof relies on local frames: it can safely be skipped.

Proposition 5.3. *Let $\mathcal{M}$ be a manifold with arbitrary connection ∇. Given a smooth vector field $V \in \mathfrak{X}(\mathcal{M})$ and a point $x \in \mathcal{M}$, if $V(x) = 0$ then*

$$\nabla_u V = \mathrm{D}V(x)[u]$$

for all $u \in \mathrm{T}_x\mathcal{M}$. In particular, $\mathrm{D}V(x)[u]$ is tangent at x.

Proof. In a neighborhood $\mathcal{U}$ of x on $\mathcal{M}$, we can expand V in a local frame $W_1, \ldots, W_n \in \mathfrak{X}(\mathcal{U})$ (Proposition 3.69):

$$V|_{\mathcal{U}} = g_1 W_1 + \cdots + g_n W_n,$$

where $g_1, \ldots, g_n \colon \mathcal{U} \to \mathbb{R}$ are smooth. Given $u \in \mathrm{T}_x\mathcal{M}$, the properties of connections allow us to write the following (see Section 5.6 for a technical point about why it makes sense to say $\nabla_u V = \nabla_u(V|_{\mathcal{U}})$):

$$\nabla_u V = \sum_i \nabla_u(g_i W_i) = \sum_i \mathrm{D}g_i(x)[u] \cdot W_i(x) + g_i(x)\nabla_u W_i.$$

Moreover, a direct computation reveals

$$\mathrm{D}V(x)[u] = \sum_i \mathrm{D}(g_i W_i)(x)[u] = \sum_i \mathrm{D}g_i(x)[u] \cdot W_i(x) + g_i(x)\mathrm{D}W_i(x)[u].$$

Since $V(x) = 0$, we know $g_i(x) = 0$ for all i, hence

$$\nabla_u V = \sum_i \mathrm{D}g_i(x)[u] \cdot W_i(x) = \mathrm{D}V(x)[u].$$

This concludes the proof. $\square$

Exercise 5.4. *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two manifolds, respectively equipped with connections $\nabla^{(1)}$ and $\nabla^{(2)}$. Consider the product manifold $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$. Show that the map $\nabla \colon \mathrm{T}\mathcal{M} \times \mathfrak{X}(\mathcal{M}) \to \mathrm{T}\mathcal{M}$ defined by*

$$\begin{aligned}\nabla_{(u_1,u_2)}(V_1, V_2) = \Big(&\nabla^{(1)}_{u_1} V_1(\cdot, x_2) + \mathrm{D}V_1(x_1, \cdot)(x_2)[u_2], \\ &\nabla^{(2)}_{u_2} V_2(x_1, \cdot) + \mathrm{D}V_2(\cdot, x_2)(x_1)[u_1]\Big)\end{aligned} \tag{5.6}$$

for all (u_1, u_2) tangent to $\mathcal{M}$ at (x_1, x_2) is a connection on $\mathcal{M}$—we call it the product connection. *Notation such as $V_1(\cdot, x_2)$ represents the map obtained from $V_1 \colon \mathcal{M}_1 \times \mathcal{M}_2 \to \mathrm{T}\mathcal{M}_1$ by fixing the second input to x_2. In particular, $V_1(\cdot, x_2)$ is a vector field on $\mathcal{M}_1$, while $V_1(x_1, \cdot)$ is a map from $\mathcal{M}_2$ to the linear space $\mathrm{T}_{x_1}\mathcal{M}_1$.*

5.4 Riemannian connections

There exist many connections on a manifold, which means we have leeway to be more demanding. Upon equipping the manifold with a Riemannian metric, we require two further properties so that the connection and the metric interact nicely. This is the object of our next theorem, called the *fundamental theorem of Riemannian geometry*. In particular, the two additional properties ensure the Hessian as defined later in this chapter is a self-adjoint map on each tangent space.

In order to state the desired properties, we need to introduce a few notational
⋆ definitions. Mind the difference between Uf and fU.

Definition 5.5. *For $U, V \in \mathfrak{X}(\mathcal{M})$ and $f \in \mathfrak{F}(\mathcal{U})$ with $\mathcal{U}$ open in $\mathcal{M}$, define:*

- $Uf \in \mathfrak{F}(\mathcal{U})$ *such that* $(Uf)(x) = \mathrm{D}f(x)[U(x)]$*;*
- $[U, V] \colon \mathfrak{F}(\mathcal{U}) \to \mathfrak{F}(\mathcal{U})$ *such that* $[U, V]f = U(Vf) - V(Uf)$*; and*
- $\langle U, V \rangle \in \mathfrak{F}(\mathcal{M})$ *such that* $\langle U, V \rangle(x) = \langle U(x), V(x) \rangle_x$.

The notation Uf captures the *action* of a smooth vector field U on a smooth function f through *derivation*, transforming f into another smooth function. The *commutator* $[U, V]$ of such action is called the *Lie bracket*. Even in linear spaces $[U, V]f$ is nonzero in general.[1] Notice that

$$Uf = \langle \mathrm{grad} f, U \rangle, \tag{5.7}$$

owing to the definitions of Uf, $\langle V, U \rangle$ and $\mathrm{grad} f$.

Theorem 5.6. *On a Riemannian manifold $\mathcal{M}$, there exists a unique connection ∇ which satisfies two additional properties for all $U, V, W \in \mathfrak{X}(\mathcal{M})$:*

4. *Symmetry:* $[U, V]f = (\nabla_U V - \nabla_V U)f$ *for all* $f \in \mathfrak{F}(\mathcal{M})$*; and*
5. *Compatibility with the metric:* $U\langle V, W \rangle = \langle \nabla_U V, W \rangle + \langle V, \nabla_U W \rangle$.

This connection is called the Levi-Civita *or* Riemannian connection*.*

A connection which satisfies the symmetry property is a *symmetric connection* (also called *torsion-free*)—this is defined independently of the Riemannian structure. Compatibility with the Riemannian metric is a type of product rule for differentiation through inner products. Unless otherwise stated, we always equip a Riemannian manifold with its Riemannian connection.

Before we prove Theorem 5.6, let us check its statement against the connections we know. As expected, the Riemannian connection on Euclidean spaces is nothing but classical vector field differentiation (5.3).

Theorem 5.7. *The Riemannian connection on a Euclidean space $\mathcal{E}$ with any Euclidean metric $\langle \cdot, \cdot \rangle$ is $\nabla_u V = \mathrm{D}V(x)[u]$: the* canonical Euclidean connection*.*

[1] In $\mathbb{R}^2$, consider $U(x) = (1, 0)$, $V(x) = (0, x_1 x_2)$ and $f(x) = x_2$. Then, $[U, V]f = f$.

Proof. We first establish compatibility with the metric, as it will be useful to prove symmetry. To this end, we go back to the definition of derivatives as limits. Consider three vector fields $U, V, W \in \mathfrak{X}(\mathcal{E})$. Owing to smoothness of the latter and to the definition of ∇,

$$\begin{aligned} V(x + tU(x)) &= V(x) + t\mathrm{D}V(x)[U(x)] + O(t^2) \\ &= V(x) + t(\nabla_U V)(x) + O(t^2). \end{aligned}$$

Define the function $f = \langle V, W \rangle$. Using bilinearity of the metric,

$$\begin{aligned} (Uf)(x) &= \mathrm{D}f(x)[U(x)] \\ &= \lim_{t \to 0} \frac{\langle V(x + tU(x)), W(x + tU(x)) \rangle - \langle V(x), W(x) \rangle}{t} \\ &= \lim_{t \to 0} \frac{\langle V(x) + t(\nabla_U V)(x), W(x) + t(\nabla_U W)(x) \rangle - \langle V(x), W(x) \rangle}{t} \\ &= \big(\langle \nabla_U V, W \rangle + \langle V, \nabla_U W \rangle \big)(x) \end{aligned}$$

for all x, as desired.

To establish symmetry, we develop the left-hand side first. Recall the definition of Lie bracket: $[U, V]f = U(Vf) - V(Uf)$. Focusing on the first term, note that

$$(Vf)(x) = \mathrm{D}f(x)[V(x)] = \langle \mathrm{grad} f(x), V(x) \rangle_x .$$

We can now use compatibility with the metric:

$$U(Vf) = U \langle \mathrm{grad} f, V \rangle = \langle \nabla_U (\mathrm{grad} f), V \rangle + \langle \mathrm{grad} f, \nabla_U V \rangle .$$

Consider the term $\nabla_U(\mathrm{grad} f)$: this is the derivative of the gradient vector field of f along U. By definition, this is the (Euclidean) Hessian of f along U. We write $\nabla_U(\mathrm{grad} f) = \mathrm{Hess} f[U]$, with the understanding that $(\mathrm{Hess} f[U])(x) = \mathrm{Hess} f(x)[U(x)] = \nabla_{U(x)}(\mathrm{grad} f)$. Overall,

$$U(Vf) = \langle \mathrm{Hess} f[U], V \rangle + \langle \mathrm{grad} f, \nabla_U V \rangle .$$

Likewise for the other term,

$$V(Uf) = \langle \mathrm{Hess} f[V], U \rangle + \langle \mathrm{grad} f, \nabla_V U \rangle .$$

It is a standard fact from multivariate calculus that the Euclidean Hessian is self-adjoint, that is: $\langle \mathrm{Hess} f[U], V \rangle = \langle \mathrm{Hess} f[V], U \rangle$. (This is the Clairaut–Schwarz theorem, which you may remember as the fact that partial derivatives in $\mathbb{R}^n$ commute.) Hence,

$$\begin{aligned} [U, V]f &= U(Vf) - V(Uf) \\ &= \langle \mathrm{grad} f, \nabla_U V - \nabla_V U \rangle \\ &= (\nabla_U V - \nabla_V U)f, \end{aligned}$$

concluding the proof. □

For $\mathcal{M}$ an embedded submanifold of a Euclidean space $\mathcal{E}$, the connection ∇ (5.4) we defined by projection to tangent spaces is always symmetric, regardless of any Riemannian structure on $\mathcal{M}$. To show this, it is convenient to introduce notation in analogy with (5.5):

$$\nabla_U V = \mathrm{Proj}\big(\bar{\nabla}_{\bar{U}} \bar{V}\big), \tag{5.8}$$

where $\bar{\nabla}$ is the canonical Euclidean connection on $\mathcal{E}$; $\bar{U}, \bar{V}$ are smooth extensions of U, V; and Proj takes as input a smooth vector field on a neighborhood of $\mathcal{M}$ in $\mathcal{E}$ and returns a smooth vector field on $\mathcal{M}$ obtained by orthogonal projection at each point. Thus,

$$\nabla_{U(x)} V = (\nabla_U V)(x) = \mathrm{Proj}_x\big((\bar{\nabla}_{\bar{U}} \bar{V})(x)\big) = \mathrm{Proj}_x(\bar{\nabla}_{\bar{U}(x)} \bar{V})$$

are all equivalent notations.

Theorem 5.8. *Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$. The connection ∇ defined by* (5.4) *is symmetric on $\mathcal{M}$.*

Proof. Let $\bar{\nabla}$ denote the canonical Euclidean connection on $\mathcal{E}$. If $\mathcal{M}$ is an open submanifold of $\mathcal{E}$, the claim is clear since ∇ is then nothing but $\bar{\nabla}$ with restricted domains. We now consider $\mathcal{M}$ not open in $\mathcal{E}$. To establish symmetry of ∇, we rely heavily on the fact that $\bar{\nabla}$ is itself symmetric on (any open subset of) the embedding space $\mathcal{E}$.

Consider $U, V \in \mathfrak{X}(\mathcal{M})$ and $f, g \in \mathfrak{F}(\mathcal{M})$ together with smooth extensions $\bar{U}, \bar{V} \in \mathfrak{X}(O)$ and $\bar{f}, \bar{g} \in \mathfrak{F}(O)$ to a neighborhood O of $\mathcal{M}$ in $\mathcal{E}$. We use the identity $Uf = (\bar{U}\bar{f})|_{\mathcal{M}}$ repeatedly, then the fact that $\bar{\nabla}$ is symmetric on O:

$$\begin{aligned}
[U, V]f &= U(Vf) - V(Uf) \\
&= U\big((\bar{V}\bar{f})|_{\mathcal{M}}\big) - V\big((\bar{U}\bar{f})|_{\mathcal{M}}\big) \\
&= \big(\bar{U}(\bar{V}\bar{f})\big)|_{\mathcal{M}} - \big(\bar{V}(\bar{U}\bar{f})\big)|_{\mathcal{M}} \\
&= \big([\bar{U}, \bar{V}]\bar{f}\big)|_{\mathcal{M}} \\
&= \big((\bar{\nabla}_{\bar{U}}\bar{V} - \bar{\nabla}_{\bar{V}}\bar{U})\bar{f}\big)|_{\mathcal{M}} \\
&= (\bar{W}\bar{f})|_{\mathcal{M}},
\end{aligned} \tag{5.9}$$

where we defined $\bar{W} = \bar{\nabla}_{\bar{U}}\bar{V} - \bar{\nabla}_{\bar{V}}\bar{U} \in \mathfrak{X}(O)$. We know from Section 5.1 that the individual vector fields $\bar{\nabla}_{\bar{U}}\bar{V}$ and $\bar{\nabla}_{\bar{V}}\bar{U}$ need not be tangent along $\mathcal{M}$. Yet, we are about to show that their difference is. Assume this for now, that is, assume $\bar{W}$ is a smooth extension of a vector field W on $\mathcal{M}$. Then,

$$W = \bar{W}|_{\mathcal{M}} = \mathrm{Proj}(\bar{W}) = \mathrm{Proj}\big(\bar{\nabla}_{\bar{U}}\bar{V} - \bar{\nabla}_{\bar{V}}\bar{U}\big) = \nabla_U V - \nabla_V U.$$

Furthermore, $(\bar{W}\bar{f})|_{\mathcal{M}} = Wf$, so that continuing from (5.9) we find:

$$[U, V]f = (\bar{W}\bar{f})|_{\mathcal{M}} = Wf = (\nabla_U V - \nabla_V U)f,$$

which is exactly what we want. Thus, it only remains to show that $\bar{W}(x)$ is indeed tangent to $\mathcal{M}$ for all $x \in \mathcal{M}$.

To this end, let $x \in \mathcal{M}$ be arbitrary and let $\bar{h}\colon O' \to \mathbb{R}^k$ be a local defining function for $\mathcal{M}$ around x so that $\mathcal{M} \cap O' = \bar{h}^{-1}(0)$, and we ensure $O' \subseteq O$. Consider the restriction $h = \bar{h}|_{\mathcal{M}\cap O'}$: of course, h is nothing but the zero function. Applying (5.9) to h, we find:

$$0 = [U, V]h = (\bar{W}\bar{h})|_{\mathcal{M}\cap O'}.$$

Evaluate this at x:

$$0 = (\bar{W}\bar{h})(x) = \mathrm{D}\bar{h}(x)\big[\bar{W}(x)\big]\,.$$

In words: $\bar{W}(x)$ is in the kernel of $\mathrm{D}\bar{h}(x)$, meaning it is in the tangent space at x. This concludes the proof. ☐

In the special case where $\mathcal{M}$ inherits the metric from its embedding Euclidean space, ∇ as defined above is *the* Riemannian connection.

Theorem 5.9. *Let $\mathcal{M}$ be a Riemannian submanifold of a Euclidean space. The connection ∇ defined by* (5.4) *is the Riemannian connection on $\mathcal{M}$.*

Proof. In light of Theorem 5.8, it remains to check compatibility with the metric, that is, property 5 in Theorem 5.6. Consider $U, V, W \in \mathfrak{X}(\mathcal{M})$ together with smooth extensions $\bar{U}, \bar{V}, \bar{W} \in \mathfrak{X}(O)$ defined on a neighborhood O of $\mathcal{M}$ in $\mathcal{E}$. Let $\langle\cdot,\cdot\rangle$ denote the metric on the embedding space $\mathcal{E}$ (which $\mathcal{M}$ inherits). Since $\langle V, W\rangle = \langle \bar{V}, \bar{W}\rangle|_{\mathcal{M}}$ and $Uf = (\bar{U}\bar{f})|_{\mathcal{M}}$, setting $f = \langle V, W\rangle$ and $\bar{f} = \langle \bar{V}, \bar{W}\rangle$ we find that $U\langle V, W\rangle = (\bar{U}\langle \bar{V}, \bar{W}\rangle)|_{\mathcal{M}}$. Using compatibility of $\bar{\nabla}$ with the metric:

$$U\langle V, W\rangle = \left(\bar{U}\langle \bar{V}, \bar{W}\rangle\right)|_{\mathcal{M}} = \Big(\big\langle \bar{\nabla}_{\bar{U}}\bar{V}, \bar{W}\big\rangle + \big\langle \bar{V}, \bar{\nabla}_{\bar{U}}\bar{W}\big\rangle\Big)\Big|_{\mathcal{M}}. \tag{5.10}$$

Pick $x \in \mathcal{M}$. Then, $\bar{W}(x) = W(x) = \mathrm{Proj}_x(W(x))$. Recall that Proj_x is self-adjoint (Proposition 3.63), that is, $\langle u, \mathrm{Proj}_x(v)\rangle = \langle \mathrm{Proj}_x(u), v\rangle$ for all $u, v \in \mathcal{E}$. Consequently,

$$\begin{aligned}\big\langle \bar{\nabla}_{\bar{U}}\bar{V}, \bar{W}\big\rangle(x) &= \big\langle (\bar{\nabla}_{\bar{U}}\bar{V})(x), \mathrm{Proj}_x(W(x))\big\rangle \\ &= \big\langle \mathrm{Proj}_x\big((\bar{\nabla}_{\bar{U}}\bar{V})(x)\big), W(x)\big\rangle_x \\ &= \langle \nabla_U V, W\rangle(x).\end{aligned}$$

Combining twice with (5.10), we find indeed that

$$U\langle V, W\rangle = \langle \nabla_U V, W\rangle + \langle V, \nabla_U W\rangle\,.$$

This concludes the proof. ☐

The previous theorem gives a conveniently clear picture of how to differentiate vector fields on a Riemannian submanifold $\mathcal{M}$ embedded in a Euclidean space: first differentiate the vector field in the linear space (a classical derivative), then orthogonally project the result to the tangent spaces of $\mathcal{M}$. More generally, if $\mathcal{M}$ is not a Riemannian submanifold, then this procedure still defines a symmetric connection, but it may not be the Riemannian connection.

We now return to Theorem 5.6. To provide the missing proof, we need a technical observation: a Lie bracket “is” a smooth vector field.

Proposition 5.10. *Let U, V be two smooth vector fields on a manifold $\mathcal{M}$. There exists a unique smooth vector field W on $\mathcal{M}$ such that $[U, V]f = Wf$ for all $f \in \mathfrak{F}(\mathcal{M})$. Therefore, we identify $[U, V]$ with that smooth vector field. Explicitly, if ∇ is any symmetric connection, then $[U, V] = \nabla_U V - \nabla_V U$.*

Proof. Say $\mathcal{M}$ is embedded in the Euclidean space $\mathcal{E}$. In Theorem 5.8 we have shown that ∇ as defined by (5.4) is a symmetric connection for $\mathcal{M}$. Thus, $[U, V]f = Wf$ for all $f \in \mathfrak{F}(\mathcal{M})$ with $W = \nabla_U V - \nabla_V U$. That vector field is unique because two vector fields $W_1, W_2 \in \mathfrak{X}(\mathcal{M})$ such that $W_1 f = W_2 f$ for all $f \in \mathfrak{F}(\mathcal{M})$ are necessarily equal. Indeed, for contradiction, assume $W_1 f = W_2 f$ for all $f \in \mathfrak{F}(\mathcal{M})$ yet $W_3 = W_1 - W_2 \neq 0$: there exists $\tilde{x} \in \mathcal{M}$ such that $W_3(\tilde{x}) \neq 0$. Consider the linear function $\bar{f}(x) = \langle x, W_3(\tilde{x})\rangle$ and its restriction $f = \bar{f}|_{\mathcal{M}}$; we have

$$(W_1 f)(\tilde{x}) - (W_2 f)(\tilde{x}) = (W_3 f)(\tilde{x}) = \mathrm{D}f(\tilde{x})[W_3(\tilde{x})] = \|W_3(\tilde{x})\|^2 \neq 0,$$

which is a contradiction. (Here, $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ come from $\mathcal{E}$.)

A comment is in order. Note that $[U, V]$ is defined irrespective of any connection. The above proof shows that $[U, V]$ is equivalent to $\nabla_U V - \nabla_V U$ for *any* symmetric connection ∇, and relies on Theorem 5.8 for the existence of a symmetric connection. Because of that, the proof here is limited to manifolds embedded in a Euclidean space. In Section 8.10, we see a proof that holds for manifolds in general. □

Proof sketch of Theorem 5.6. It is easy to verify uniqueness. Indeed, assume ∇ is a symmetric connection which is also compatible with the metric. For all $U, V, W \in \mathfrak{X}(\mathcal{M})$, compatibility with the metric implies

$$\begin{aligned} U\langle V, W\rangle &= \langle \nabla_U V, W\rangle + \langle V, \nabla_U W\rangle, \\ V\langle W, U\rangle &= \langle \nabla_V W, U\rangle + \langle W, \nabla_V U\rangle, \\ W\langle U, V\rangle &= \langle \nabla_W U, V\rangle + \langle U, \nabla_W V\rangle. \end{aligned}$$

Add the first two lines and subtract the third: owing to Proposition 5.10 and symmetry of ∇, we find after some reorganizing that

$$\begin{aligned} 2\,\langle \nabla_U V, W\rangle = U\langle V, W\rangle + V\langle W, U\rangle - W\langle U, V\rangle \\ - \langle U, [V, W]\rangle + \langle V, [W, U]\rangle + \langle W, [U, V]\rangle. \end{aligned} \tag{5.11}$$

This is the *Koszul formula*. Notice that the right-hand side is independent of ∇. For fixed U, V, the fact that this identity holds for all W implies that $\nabla_U V$ is uniquely determined. To see this, consider for each $x \in \mathcal{M}$ a set of vector fields $W_1, \ldots, W_n \in \mathfrak{X}(\mathcal{M})$ such that $W_1(x), \ldots, W_n(x)$ form a basis of $\mathrm{T}_x\mathcal{M}$: this uniquely determines $(\nabla_U V)(x)$. Thus, there can be at most one Riemannian connection.

To prove existence, we also rely on the Koszul formula but we need more advanced tools. Let $U, V \in \mathfrak{X}(\mathcal{M})$ be arbitrary (fixed). We can verify that the right-hand side of (5.11) defines a smooth one-form $X: \mathfrak{X}(\mathcal{M}) \to \mathfrak{F}(\mathcal{M})$ (recall the definition in Section 3.9). Indeed,

$$X(W) = U\langle V, W\rangle + V\langle W, U\rangle - W\langle U, V\rangle \\ - \langle U, [V, W]\rangle + \langle V, [W, U]\rangle + \langle W, [U, V]\rangle$$

linearly maps a smooth vector field W to a smooth scalar field $X(W)$ with the property that $X(fW) = fX(W)$ for all $W \in \mathfrak{X}(\mathcal{M})$ and $f \in \mathfrak{F}(\mathcal{M})$. This can be verified through direct computation using $[U, fV] = f[U, V] + (Uf) \cdot V$ as follows from Proposition 5.10. Then, the musical isomorphism (Proposition 3.71) implies that there exists a unique smooth vector field $Z \in \mathfrak{X}(\mathcal{M})$ such that $X(W) = \langle Z, W\rangle$. We use this to define an operator $\nabla: \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$ as $\nabla_U V = \frac{1}{2}Z$. It then remains to verify that ∇ is a symmetric connection which is compatible with the metric—some details require Section 5.6.

A comment is in order. In Section 3.9 we only sketched the proof of the musical isomorphism. The sketched parts were clear or unnecessary for one-forms such as $\mathrm{D}f$. Likewise, one could verify this for the one-form X defined above, though this can be lengthy. It is similarly technical but more instructive to study the missing details outlined later in Section 5.6. For readers who are only interested in Riemannian submanifolds, the situation is rather simpler: we already proved existence of the Riemannian connection as a pointwise operator (constructively) in Proposition 5.9. □

Exercise 5.11. *A* derivation *on $\mathcal{M}$ is a map $\mathcal{D}: \mathfrak{F}(\mathcal{M}) \to \mathfrak{F}(\mathcal{M})$ such that, for all $f, g \in \mathfrak{F}(\mathcal{M})$ and $a, b \in \mathbb{R}$, we have:*

1. *Linearity: $\mathcal{D}(af + bg) = a\mathcal{D}(f) + b\mathcal{D}(g)$, and*
2. *Leibniz rule: $\mathcal{D}(fg) = g\mathcal{D}(f) + f\mathcal{D}(g)$.*

Show that the action of a smooth vector field on a smooth function (as per Definition 5.5) is a derivation. (See Section 5.13 for context.)

Exercise 5.12. *Show that the Lie bracket $[U, V]$ of two smooth vector fields $U, V \in \mathfrak{X}(\mathcal{M})$ is a derivation, as per the definition in the previous exercise. It is instructive to do so without using connections or Proposition 5.10.*

Exercise 5.13. *Continuing from Exercise 5.4, show that if $\nabla^{(1)}, \nabla^{(2)}$ are the Riemannian connections on $\mathcal{M}_1, \mathcal{M}_2$ (respectively), then the product connection defined by* (5.6) *is the Riemannian connection on the Riemannian product manifold $\mathcal{M}_1 \times \mathcal{M}_2$ whose metric is defined in Example 3.57. (Concepts from later sections may help; specifically, Proposition 5.15 and Theorem 5.29.)*

5.5 Riemannian Hessians

The Riemannian Hessian of a function is defined as the covariant derivative of its gradient vector field with respect to the Riemannian connection ∇, which we defined in Theorem 5.6. At any point x on the manifold $\mathcal{M}$, the Hessian defines a linear map from the tangent space $\mathrm{T}_x\mathcal{M}$ into itself.

Definition 5.14. *Let $\mathcal{M}$ be a Riemannian manifold with its Riemannian connection ∇. The* Riemannian Hessian *of $f \in \mathfrak{F}(\mathcal{M})$ at $x \in \mathcal{M}$ is the linear map* $\mathrm{Hess}f(x)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ *defined as follows:*

$$\mathrm{Hess}f(x)[u] = \nabla_u \mathrm{grad}f.$$

Equivalently, $\mathrm{Hess}f$ *maps $\mathfrak{X}(\mathcal{M})$ to $\mathfrak{X}(\mathcal{M})$ as* $\mathrm{Hess}f[U] = \nabla_U \mathrm{grad}f$.

The two special properties of the Riemannian connection together lead to symmetry of the Hessian. By the spectral theorem, this implies that the $\dim \mathcal{M}$ eigenvalues of $\mathrm{Hess}f(x)$ are real, and that corresponding eigenvectors may be chosen to form a basis of $\mathrm{T}_x\mathcal{M}$ orthonormal with respect to $\langle \cdot, \cdot \rangle_x$ (see Theorem 3.6).

Proposition 5.15. *The Riemannian Hessian is self-adjoint with respect to the Riemannian metric. That is, for all $x \in \mathcal{M}$ and $u, v \in \mathrm{T}_x\mathcal{M}$,*

$$\langle \mathrm{Hess}f(x)[u], v \rangle_x = \langle u, \mathrm{Hess}f(x)[v] \rangle_x .$$

Proof. Pick any two vector fields $U, V \in \mathfrak{X}(\mathcal{M})$ such that $U(x) = u$ and $V(x) = v$. Recalling the notation for vector fields acting on functions as derivations (Definition 5.5) and using compatibility of the Riemannian connection with the Riemannian metric, we find:

$$\begin{aligned}\langle \mathrm{Hess}f[U], V \rangle &= \langle \nabla_U \mathrm{grad}f, V \rangle \\ &= U\langle \mathrm{grad}f, V \rangle - \langle \mathrm{grad}f, \nabla_U V \rangle \\ &= U(Vf) - (\nabla_U V)f.\end{aligned}$$

Similarly,

$$\langle U, \mathrm{Hess}f[V] \rangle = V(Uf) - (\nabla_V U)f.$$

Thus, recalling the definition of Lie bracket, we get

$$\begin{aligned}\langle \mathrm{Hess}f[U], V \rangle - \langle U, \mathrm{Hess}f[V] \rangle &= U(Vf) - V(Uf) - (\nabla_U V)f + (\nabla_V U)f \\ &= [U, V]f - (\nabla_U V - \nabla_V U)f \\ &= 0,\end{aligned}$$

where we could conclude owing to symmetry of the connection. □

To compute the Riemannian Hessian, we must compute the Riemannian connection. For the particular case of a Riemannian submanifold of a Euclidean space, we know how to do this from Theorem 5.9. In practical terms, we simply need to consider a smooth extension of the Riemannian gradient vector field, differentiate it in the classical sense, then orthogonally project the result to the tangent spaces.

Corollary 5.16. *Let $\mathcal{M}$ be a Riemannian submanifold of a Euclidean space. Consider a smooth function $f\colon \mathcal{M} \to \mathbb{R}$. Let $\bar{G}$ be a smooth extension of* $\mathrm{grad} f$ *— that is, $\bar{G}$ is any smooth vector field defined on a neighborhood of $\mathcal{M}$ in the embedding space such that $\bar{G}(x) = \mathrm{grad} f(x)$ for all $x \in \mathcal{M}$. Then,*

$$\mathrm{Hess} f(x)[u] = \mathrm{Proj}_x\big(\mathrm{D}\bar{G}(x)[u]\big).$$

More can be said about the important special case of Riemannian submanifolds: see Section 5.11. The following example illustrates how to use Corollary 5.16 in practice.

Example 5.17. *Consider the cost function $\bar{f}(x) = \frac{1}{2}x^\top Ax$ for some symmetric matrix $A \in \mathbb{R}^{d\times d}$ and its restriction $f = \bar{f}|_{\mathrm{S}^{d-1}}$ to the sphere S^{d-1} as a Riemannian submanifold of $\mathbb{R}^d$. We already determined the Euclidean and Riemannian gradients of $\bar{f}$ and f, respectively:*

$$\begin{aligned} &\mathrm{grad}\bar{f}(x) = Ax, \\ &\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad}\bar{f}(x)) = (I_d - xx^\top)Ax = Ax - (x^\top Ax)x. \end{aligned}$$

To obtain the Riemannian Hessian of f, we must differentiate a smooth extension of $\mathrm{grad} f$ *in $\mathbb{R}^d$ and project the result to the tangent spaces of S^{d-1}. As is typical, the analytic expression of* $\mathrm{grad} f$ *provides a natural candidate for a smooth extension; we simply pick:*

$$\bar{G}(x) = Ax - (x^\top Ax)x.$$

The differential of $\bar{G}$ follows from the product rule (see also Section 4.7):

$$\mathrm{D}\bar{G}(x)[u] = Au - (u^\top Ax + x^\top Au)x - (x^\top Ax)u.$$

Orthogonally project to the tangent space at x to reveal the Hessian:

$$\begin{aligned} \mathrm{Hess} f(x)[u] = \mathrm{Proj}_x\big(\mathrm{D}\bar{G}(x)[u]\big) &= \mathrm{Proj}_x(Au) - (x^\top Ax)u \\ &= Au - (x^\top Au)x - (x^\top Ax)u. \end{aligned}$$

This linear map is formally defined only on $\mathrm{T}_x\mathrm{S}^{d-1}$ (not on all of $\mathbb{R}^d$).

Exercise 5.18. *Continuing Example 5.17, show that if* $\mathrm{grad} f(x)$ *is zero and* $\mathrm{Hess} f(x)$ *is positive semidefinite (i.e., $\langle u, \mathrm{Hess} f(x)[u]\rangle_x \geq 0$ for all $u \in \mathrm{T}_x\mathrm{S}^{d-1}$), then x is a global minimizer of f, that is, x is an eigenvector of A associated to its smallest (left-most) eigenvalue. This is an unusual property: we do not normally expect to be able to certify global optimality based on local conditions alone. See also Exercise 9.51.*

Example 5.19. *Let us derive an expression for the Riemannian Hessian of a smooth function $f\colon \mathcal{M}_1 \times \mathcal{M}_2 \to \mathbb{R}$ on a Riemannian product manifold. Fix a point $x = (x_1, x_2) \in \mathcal{M}_1 \times \mathcal{M}_2$. With $f(\cdot, x_2)\colon \mathcal{M}_1 \to \mathbb{R}$, we denote the function obtained from f by fixing its second input to x_2; likewise for $f(x_1, \cdot)\colon \mathcal{M}_2 \to \mathbb{R}$.*

From Exercise 3.67, the Riemannian gradient of f at x is given by

$$\mathrm{grad}f(x_1, x_2) = \big(G_1(x_1, x_2), G_2(x_1, x_2)\big) \quad \textit{with} \quad \begin{aligned} G_1(x_1, x_2) &= \mathrm{grad}f(\cdot, x_2)(x_1), \\ G_2(x_1, x_2) &= \mathrm{grad}f(x_1, \cdot)(x_2). \end{aligned}$$

Then, the Riemannian Hessian of f at x along any tangent vector $u = (u_1, u_2)$ at x follows from Exercise 5.13 as:

$$\begin{aligned} \mathrm{Hess}f(x_1, x_2)[u_1, u_2] = \big(&\mathrm{Hess}f(\cdot, x_2)(x_1)[u_1] + \mathrm{D}G_1(x_1, \cdot)(x_2)[u_2], \\ &\mathrm{Hess}f(x_1, \cdot)(x_2)[u_2] + \mathrm{D}G_2(\cdot, x_2)(x_1)[u_1]\big). \end{aligned} \tag{5.12}$$

Above, $G_1(x_1, \cdot)\colon \mathcal{M}_2 \to \mathrm{T}_{x_1}\mathcal{M}_1$ denotes the map obtained from G_1 by fixing its first input to x_1, and likewise for $G_2(\cdot, x_2)\colon \mathcal{M}_1 \to \mathrm{T}_{x_2}\mathcal{M}_2$.

As a side note, if $\mathcal{M}_1, \mathcal{M}_2$ are both Riemannian submanifolds of respective embedding spaces $\mathcal{E}_1, \mathcal{E}_2$, then the product manifold $\mathcal{M}_1 \times \mathcal{M}_2$ is a Riemannian submanifold of $\mathcal{E}_1 \times \mathcal{E}_2$. In that case, Corollary 5.16 and Section 5.11 provide other ways to get to the Hessian of f.

5.6 Connections as pointwise derivatives*

Definition 5.1 is not standard: the standard definition follows. In this section, we argue that they are equivalent. A reader focused on Riemannian submanifolds can safely skip this section.

Definition 5.20. *A* connection *on a manifold $\mathcal{M}$ is an operator*

$$\nabla\colon \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})\colon (U, V) \mapsto \nabla_U V$$

which has three properties for all $U, V, W \in \mathfrak{X}(\mathcal{M})$, $f, g \in \mathfrak{F}(\mathcal{M})$ and $a, b \in \mathbb{R}$:

1. *$\mathfrak{F}(\mathcal{M})$-linearity in U: $\nabla_{fU+gW} V = f\nabla_U V + g\nabla_W V$;*
2. *$\mathbb{R}$-linearity in V: $\nabla_U(aV + bW) = a\nabla_U V + b\nabla_U W$; and*
3. *Leibniz rule: $\nabla_U(fV) = (Uf)V + f\nabla_U V$.*

The field $\nabla_U V$ is the covariant derivative *of V along U with respect to ∇.*

It is clear that if ∇ is a connection as per Definition 5.1 then it is also a connection as per Definition 5.20, with $(\nabla_U V)(x) \triangleq \nabla_{U(x)} V$. The other way around is less clear.

Specifically, we must show that a connection in the sense of Definition 5.20 acts pointwise with respect to U, that is, $(\nabla_U V)(x)$ depends on U only through $U(x)$. This gives meaning to the notation $\nabla_u V$ as being equal to $(\nabla_U V)(x)$ for arbitrary $U \in \mathfrak{X}(\mathcal{M})$ such that $U(x) = u$.

That is the object of the following proposition. It is a consequence of $\mathfrak{F}(\mathcal{M})$-linearity in U. Note that dependence on V is through more than just $V(x)$ (and indeed, connections are not $\mathfrak{F}(\mathcal{M})$-linear in V). The main tool of the proof is the existence of local frames, as introduced in Section 3.9. Furthermore, a technical point requires some extra work, which we defer until after the proof.

In the remainder of this section, the word 'connection' refers to Definition 5.20. By the end of the section, we will have established that this is equivalent to Definition 5.1.

Proposition 5.21. *For any connection ∇ and smooth vector fields U, V on a manifold $\mathcal{M}$, the vector field $\nabla_U V$ at x depends on U only through $U(x)$.*

Proof. It is sufficient to show that if $U(x) = 0$ then $(\nabla_U V)(x) = 0$. Indeed, let $U_1, U_2 \in \mathfrak{X}(\mathcal{M})$ be two vector fields with $U_1(x) = U_2(x)$. Then, using the claim,

$$(\nabla_{U_1} V)(x) - (\nabla_{U_2} V)(x) = (\nabla_{U_1} V - \nabla_{U_2} V)(x) = (\nabla_{U_1 - U_2} V)(x) = 0.$$

To prove the claim, consider a local frame $W_1, \ldots, W_n$ on a neighborhood $\mathcal{U}$ of x in $\mathcal{M}$ (Proposition 3.69). Given a vector field $U \in \mathfrak{X}(\mathcal{M})$ with $U(x) = 0$, there exist unique smooth functions $g_1, \ldots, g_n \in \mathfrak{F}(\mathcal{U})$ such that

$$U|_{\mathcal{U}} = g_1 W_1 + \cdots + g_n W_n.$$

Clearly, $U(x) = 0$ implies $g_1(x) = \cdots = g_n(x) = 0$. By a technical lemma given hereafter (Lemma 5.27), it is legitimate to write:

$$\begin{aligned}(\nabla_U V)(x) &= (\nabla_{g_1 W_1 + \cdots + g_n W_n} V)(x) \\ &= g_1(x)(\nabla_{W_1} V)(x) + \cdots + g_n(x)(\nabla_{W_n} V)(x) = 0,\end{aligned} \tag{5.13}$$

which concludes the proof. $\square$

In the proof above, it is not immediately clear why (5.13) holds, because $\nabla_{U|_{\mathcal{U}}} V$ is not formally defined: normally, ∇ is fed two smooth vector fields on all of $\mathcal{M}$. To support this notation and the claim that $\nabla_U V$ and $\nabla_{U|_{\mathcal{U}}} V$ coincide at x, we work through a number of lemmas. The first one concerns the existence of *bump functions* in linear spaces. It is an exercise in analysis to build such functions [Lee12, Lem. 2.22].

Lemma 5.22. *Given any real numbers $0 < r_1 < r_2$ and any point x in a Euclidean space $\mathcal{E}$ with norm $\|\cdot\|$, there exists a smooth function $b \colon \mathcal{E} \to \mathbb{R}$ such that $b(y) = 1$ if $\|y - x\| \le r_1$, $b(y) = 0$ if $\|y - x\| \ge r_2$, and $b(y) \in (0, 1)$ if $\|y - x\| \in (r_1, r_2)$.*

Using bump functions, we can show that $(\nabla_U V)(x)$ depends on U and V only through their values in a neighborhood around x. This is the object of the two following lemmas.

Lemma 5.23. *Let V_1, V_2 be smooth vector fields on a manifold $\mathcal{M}$ equipped with a connection ∇. If $V_1|_{\mathcal{U}} = V_2|_{\mathcal{U}}$ on some open set $\mathcal{U}$ of $\mathcal{M}$, then $(\nabla_U V_1)|_{\mathcal{U}} = (\nabla_U V_2)|_{\mathcal{U}}$ for all $U \in \mathfrak{X}(\mathcal{M})$.*

Proof. For $\mathcal{M}$ an embedded submanifold of a Euclidean space $\mathcal{E}$, there exists an open set O in $\mathcal{E}$ such that $\mathcal{U} = \mathcal{M} \cap O$. Furthermore, there exist $0 < r_1 < r_2$ such that $\bar{B}(x, r_2)$—the closed ball of radius r_2 around x in $\mathcal{E}$—is included in O. Hence, by Lemma 5.22 there exists a smooth function $\bar{b} \in \mathfrak{F}(\mathcal{E})$ which is constantly equal to 1 on $\bar{B}(x, r_1)$ and constantly equal to 0 outside of $\bar{B}(x, r_2)$. With $b = \bar{b}|_{\mathcal{M}} \in \mathfrak{F}(\mathcal{M})$, it follows that the vector field $V = b \cdot (V_1 - V_2)$ is the zero vector field on $\mathcal{M}$. Hence, $\nabla_U V = 0$. Using $\mathbb{R}$-linearity of ∇ in V and the Leibniz rule:

$$0 = \nabla_U V = \nabla_U (b\,(V_1 - V_2)) = (Ub)(V_1 - V_2) + b\,(\nabla_U V_1 - \nabla_U V_2)\,.$$

Evaluating this at x and using $V_1(x) = V_2(x)$ and $b(x) = 1$, we find $(\nabla_U V_1)(x) = (\nabla_U V_2)(x)$. Repeat for all $x \in \mathcal{U}$. $\square$

Lemma 5.24. *Let U_1, U_2 be smooth vector fields on a manifold $\mathcal{M}$ equipped with a connection ∇. If $U_1|_{\mathcal{U}} = U_2|_{\mathcal{U}}$ on some open set $\mathcal{U}$ of $\mathcal{M}$, then $(\nabla_{U_1} V)|_{\mathcal{U}} = (\nabla_{U_2} V)|_{\mathcal{U}}$ for all $V \in \mathfrak{X}(\mathcal{M})$.*

Proof. Construct $b \in \mathfrak{F}(\mathcal{M})$ as in the proof of Lemma 5.23. Then, $U = b \cdot (U_1 - U_2)$ is the zero vector field on $\mathcal{M}$. By $\mathfrak{F}(\mathcal{M})$-linearity of ∇ in U,

$$0 = \nabla_U V = \nabla_{b\,(U_1 - U_2)} V = b \cdot (\nabla_{U_1} V - \nabla_{U_2} V).$$

Evaluating this at x and using $b(x) = 1$ yields the result. $\square$

We now use bump functions to show that a smooth function defined on a neighborhood of a point x on a manifold can always be extended into a smooth function defined on the whole manifold, in such a way that its value at and around x is unaffected. This is a weak version of a result known as the *extension lemma* [Lee12, Lem. 2.26].

Lemma 5.25. *Let $\mathcal{U}$ be a neighborhood of a point x on a manifold $\mathcal{M}$. Given a smooth function $f \in \mathfrak{F}(\mathcal{U})$, there exists a smooth function $g \in \mathfrak{F}(\mathcal{M})$ and a neighborhood $\mathcal{U}' \subseteq \mathcal{U}$ of x such that $g|_{\mathcal{U}'} = f|_{\mathcal{U}'}$.*

Proof. For $\mathcal{M}$ an embedded submanifold of a Euclidean space $\mathcal{E}$, we know from Proposition 3.23 that $\mathcal{U}$ itself is an embedded submanifold of $\mathcal{E}$. Hence, there exists a smooth extension $\bar{f}$ of f defined on a neighborhood O of x in $\mathcal{E}$. For this O, construct $\bar{b} \in \mathfrak{F}(\mathcal{E})$ as in the proof of Lemma 5.23, with $0 < r_1 < r_2$ such that $\bar{B}(x, r_2) \subset O$. Consider $g \colon \mathcal{E} \to \mathbb{R}$ defined by

$$\bar{g}(y) = \begin{cases} \bar{b}(y)\bar{f}(y) & \text{if } \|y - x\| \leq r_2, \\ 0 & \text{otherwise.} \end{cases}$$

It is an exercise in real analysis to verify that $\bar{g}$ is smooth in $\mathcal{E}$; hence, $g = \bar{g}|_{\mathcal{M}}$ is smooth on $\mathcal{M}$. Furthermore, $\bar{g}$ is equal to $\bar{f}$ on $\bar{B}(x, r_1)$. Set $\mathcal{U}' = \mathcal{U} \cap B(x, r_1)$, where $B(x, r_1)$ is the open ball of radius r_1 around x in $\mathcal{E}$. This is a neighborhood of x on $\mathcal{M}$ such that $g|_{\mathcal{U}'} = f|_{\mathcal{U}'}$. $\square$

Likewise, there is a smooth extension lemma for vector fields, and we state a weak version of it here. The proof is essentially the same as for the previous lemma [Lee12, Lem. 8.6].

Lemma 5.26. *Let $\mathcal{U}$ be a neighborhood of a point x on a manifold $\mathcal{M}$. Given a smooth vector field $U \in \mathfrak{X}(\mathcal{U})$, there exists a smooth vector field $V \in \mathfrak{X}(\mathcal{M})$ and a neighborhood $\mathcal{U}' \subseteq \mathcal{U}$ of x such that $V|_{\mathcal{U}'} = U|_{\mathcal{U}'}$.*

Equipped with the last three lemmas, we can finally state the technical result necessary to support the proof of Proposition 5.21.

Lemma 5.27. *Let U, V be two smooth vector fields on a manifold $\mathcal{M}$ equipped with a connection ∇. Further let $\mathcal{U}$ be a neighborhood of $x \in \mathcal{M}$ such that $U|_{\mathcal{U}} = g_1W_1 + \cdots + g_nW_n$ for some $g_1, \ldots, g_n \in \mathfrak{F}(\mathcal{U})$ and $W_1, \ldots, W_n \in \mathfrak{X}(\mathcal{U})$. Then,*

$$(\nabla_U V)(x) = g_1(x)(\nabla_{W_1} V)(x) + \cdots + g_n(x)(\nabla_{W_n} V)(x),$$

where each vector $(\nabla_{W_i} V)(x)$ is understood to mean $(\nabla_{\tilde{W}_i} V)(x)$ with $\tilde{W}_i$ any smooth extension of W_i to $\mathcal{M}$ around x.

Proof. Combining Lemmas 5.25 and 5.26, we know there exist smooth extensions $\tilde{g}_1, \ldots, \tilde{g}_n \in \mathfrak{F}(\mathcal{M})$ and $\tilde{W}_1, \ldots, \tilde{W}_n \in \mathfrak{X}(\mathcal{M})$ that coincide with $g_1, \ldots, g_n$ and $W_1, \ldots, W_n$ on a neighborhood $\mathcal{U}' \subseteq \mathcal{U}$ of x, so that $\tilde{U} = \tilde{g}_1\tilde{W}_1 + \cdots + \tilde{g}_n\tilde{W}_n$ is a smooth vector field on $\mathcal{M}$ which agrees with U locally: $U|_{\mathcal{U}'} = \tilde{U}|_{\mathcal{U}'}$. Thus, by Lemma 5.24,

$$\begin{aligned}
(\nabla_U V)(x) &= (\nabla_{\tilde{U}} V)(x) \\
&= (\nabla_{\tilde{g}_1\tilde{W}_1 + \cdots + \tilde{g}_n\tilde{W}_n} V)(x) \\
&= \tilde{g}_1(x)(\nabla_{\tilde{W}_1} V)(x) + \cdots + \tilde{g}_n(x)(\nabla_{\tilde{W}_n} V)(x) \\
&= g_1(x)(\nabla_{W_1} V)(x) + \cdots + g_n(x)(\nabla_{W_n} V)(x).
\end{aligned}$$

The stated definition of $(\nabla_{W_i} V)(x)$ is independent of the choice of smooth extension owing to Lemma 5.24. $\square$

In the proofs above, the most important feature of $(U, V) \mapsto \nabla_U V$ we have used is that it is $\mathfrak{F}(\mathcal{M})$-linear in U. With that in mind, it is easy to revisit those proofs and fill in the missing parts for the proof of Proposition 3.71.

Anticipating our needs for Section 5.7, we note that Lemmas 5.23, 5.25 and 5.26 also allow us to make sense of the notation

$$(\nabla_u(gW))(x) = \mathrm{D}g(x)[u] \cdot W(x) + g(x) \cdot (\nabla_u W)(x), \tag{5.14}$$

where $g \in \mathfrak{F}(\mathcal{U})$ and $W \in \mathfrak{X}(\mathcal{U})$ are merely defined on a neighborhood $\mathcal{U}$ of x. Specifically, $(\nabla_u W)(x)$ represents $(\nabla_u \tilde{W})(x)$ where $\tilde{W} \in \mathfrak{X}(\mathcal{M})$ is any smooth extension of W around x, as justified by Lemmas 5.23 and 5.26.

5.7 Differentiating vector fields on curves

Recall that one of our goals in this chapter is to develop second-order Taylor expansions for $g = f \circ c$ with a smooth cost function $f \colon \mathcal{M} \to \mathbb{R}$ evaluated along a smooth curve $c\colon I \to \mathcal{M}$ defined on some interval I. We already determined in Section 4.1 that the first derivative of g is

$$g'(t) = \langle \mathrm{grad} f(c(t)), c'(t) \rangle_{c(t)}.$$

To obtain a second-order expansion of g, we must differentiate g'.

A connection ∇ does not, in a direct way, tell us how to compute this derivative, since $(\mathrm{grad} f) \circ c$ and c' are not vector fields on $\mathcal{M}$. Rather, for all t in the domain of c, these maps each provide a tangent vector at $c(t)$, smoothly varying with t: they are called *smooth vector fields on c*. We expect that g'' should involve a kind of derivative of these vector fields on c, through a kind of product rule. In short: we need a derivative for vector fields on c.

Fortunately, a connection ∇ on a manifold $\mathcal{M}$ *induces* a notion of derivative of vector fields along curves, with natural properties. The proof of that statement below involves local frames, which we discussed in Section 3.9. Readers who skipped that section may want to consider Proposition 5.31 instead: that one is limited to Riemannian submanifolds but its proof does not require local frames.

Definition 5.28. *Let $c\colon I \to \mathcal{M}$ be a smooth curve on $\mathcal{M}$ defined on an open interval I. A map $Z\colon I \to \mathrm{T}\mathcal{M}$ is a* vector field on c *if $Z(t)$ is in $\mathrm{T}_{c(t)}\mathcal{M}$ for all $t \in I$, and Z is a* smooth vector field on c *if it is also smooth as a map from I to $\mathrm{T}\mathcal{M}$. The set of smooth vector fields on c is denoted by $\mathfrak{X}(c)$.*

Theorem 5.29. *Let $c\colon I \to \mathcal{M}$ be a smooth curve on a manifold equipped with a connection ∇. There exists a unique operator $\frac{\mathrm{D}}{\mathrm{d}t}\colon \mathfrak{X}(c) \to \mathfrak{X}(c)$ which satisfies the following properties for all $Y, Z \in \mathfrak{X}(c)$, $U \in \mathfrak{X}(\mathcal{M})$, $g \in \mathfrak{F}(I)$, and $a, b \in \mathbb{R}$:*

1. *$\mathbb{R}$-linearity: $\frac{\mathrm{D}}{\mathrm{d}t}(aY + bZ) = a\frac{\mathrm{D}}{\mathrm{d}t}Y + b\frac{\mathrm{D}}{\mathrm{d}t}Z$;*
2. *Leibniz rule: $\frac{\mathrm{D}}{\mathrm{d}t}(gZ) = g'Z + g\frac{\mathrm{D}}{\mathrm{d}t}Z$;*
3. *Chain rule: $\left(\frac{\mathrm{D}}{\mathrm{d}t}(U \circ c)\right)(t) = \nabla_{c'(t)}U$ for all $t \in I$.*

We call $\frac{\mathrm{D}}{\mathrm{d}t}$ the induced covariant derivative *(induced by ∇). If moreover $\mathcal{M}$ is a Riemannian manifold and ∇ is compatible with its metric $\langle \cdot, \cdot \rangle$ (e.g., if ∇ is the Riemannian connection), then the induced covariant derivative also satisfies:*

4. *Product rule: $\frac{\mathrm{d}}{\mathrm{d}t}\langle Y, Z \rangle = \left\langle \frac{\mathrm{D}}{\mathrm{d}t}Y, Z \right\rangle + \left\langle Y, \frac{\mathrm{D}}{\mathrm{d}t}Z \right\rangle$,*

where $\langle Y, Z \rangle \in \mathfrak{F}(I)$ is defined by $\langle Y, Z \rangle(t) = \langle Y(t), Z(t) \rangle_{c(t)}$.

Before moving on to the proof, a comment is in order. In light of the chain rule (property 3), one may wonder why we need to define $\frac{\mathrm{D}}{\mathrm{d}t}$ at all: can it not always be computed through an application of ∇? The key is that not all vector fields $Z \in \mathfrak{X}(c)$ are of the form $U \circ c$ for some $U \in \mathfrak{X}(\mathcal{M})$. Indeed, consider a smooth curve c such that $c(t_1) = c(t_2) = x$ (it crosses itself). It could well be that $Z(t_1) \neq Z(t_2)$. Then, we would not know how to define $U(x)$: should it be equal to $Z(t_1)$ or $Z(t_2)$? For that reason, we really do need to introduce $\frac{\mathrm{D}}{\mathrm{d}t}$ as a separate concept.

Proof of Theorem 5.29. We first prove uniqueness under properties 1–3. Pick an arbitrary $\bar{t} \in I$. There exists a local frame $W_1, \ldots, W_n \in \mathfrak{X}(\mathcal{U})$ defined on a neighborhood $\mathcal{U}$ of $c(\bar{t})$ in $\mathcal{M}$ (see Proposition 3.69). Since c is continuous, $J = c^{-1}(\mathcal{U})$ is an open subset of I which contains $\bar{t}$. Furthermore, by the properties of local frames, there exist unique smooth functions $g_1, \ldots, g_n \colon J \to \mathbb{R}$ such that

$$\forall t \in J, \qquad Z(t) = g_1(t)W_1(c(t)) + \cdots + g_n(t)W_n(c(t)).$$

Using the first two properties of the covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$, we get

$$\forall t \in J, \qquad \frac{\mathrm{D}}{\mathrm{d}t}Z(t) = \sum_{i=1}^{n} g_i'(t)W_i(c(t)) + g_i(t)\frac{\mathrm{D}}{\mathrm{d}t}(W_i \circ c)(t).$$

Now using the third property, we find

$$\forall t \in J, \qquad \frac{\mathrm{D}}{\mathrm{d}t}Z(t) = \sum_{i=1}^{n} g_i'(t)W_i(c(t)) + g_i(t)\nabla_{c'(t)}W_i. \tag{5.15}$$

(As a technicality, see the discussion around eq. (5.14) for how to interpret $\nabla_{c'(t)}W_i$, considering W_i is only defined locally around $c(t)$.) Expression (5.15) is fully determined by the connection ∇. Since this argument can be repeated on a neighborhood of each $\bar{t}$ in I, it follows that $\frac{\mathrm{D}}{\mathrm{d}t}$ is uniquely determined by the connection ∇ and the three stated properties.

To prove existence, simply consider (5.15) as the definition of an operator $\frac{\mathrm{D}}{\mathrm{d}t}$ on a neighborhood of each $\bar{t}$. It is an exercise to verify that this definition satisfies properties 1–3. Since we have uniqueness, it is clear that definitions obtained on overlapping domains J and J' are compatible, so that (5.15) prescribes a smooth vector field on all of c.

Now consider the case where $\mathcal{M}$ is a Riemannian manifold and ∇ is compatible with the Riemannian metric. We prove the 4th property. To this end, expand Y in the local frame:

$$\forall t \in J, \qquad Y(t) = f_1(t)W_1(c(t)) + \cdots + f_n(t)W_n(c(t)).$$

Using also the expansion of Z, we have the following identity on J:

$$\langle Y, Z \rangle = \sum_{i,j=1}^{n} f_i g_j \langle W_i \circ c, W_j \circ c \rangle .$$

Differentiate this with respect to t:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle Y, Z \rangle = \sum_{i,j=1}^{n} (f_i' g_j + f_i g_j') \langle W_i \circ c, W_j \circ c \rangle + f_i g_j \frac{\mathrm{d}}{\mathrm{d}t}\langle W_i \circ c, W_j \circ c \rangle . \tag{5.16}$$

On the other hand, by uniqueness we know that (5.15) is a valid expression for $\frac{\mathrm{D}}{\mathrm{d}t}Z$ so that

$$\left\langle Y, \frac{\mathrm{D}}{\mathrm{d}t}Z \right\rangle = \sum_{i,j=1}^{n} f_i g_j' \langle W_i \circ c, W_j \circ c \rangle + f_i g_j \langle W_i \circ c, \nabla_{c'} W_j \rangle .$$

Similarly,

$$\left\langle \frac{\mathrm{D}}{\mathrm{d}t} Y, Z \right\rangle = \sum_{i,j=1}^{n} f_i' g_j \left\langle W_i \circ c, W_j \circ c \right\rangle + f_i g_j \left\langle \nabla_{c'} W_i, W_j \circ c \right\rangle.$$

Summing up these identities and comparing to (5.16), we find that property 4 holds if

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle W_i \circ c, W_j \circ c \rangle = \langle \nabla_{c'} W_i, W_j \circ c \rangle + \langle W_i \circ c, \nabla_{c'} W_j \rangle.$$

This is indeed the case owing to compatibility of ∇ with the metric, since

$$\frac{\mathrm{d}}{\mathrm{d}t} \big(\langle W_i, W_j \rangle \circ c \big)(t)$$

is the directional derivative of $\langle W_i, W_j \rangle$ at $c(t)$ along $c'(t)$. $\square$

Example 5.30. *Let f be a smooth function on a Riemannian manifold $\mathcal{M}$ equipped with the Riemannian connection ∇ and induced covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$. Applying the chain rule property of Theorem 5.29 to Definition 5.14 for the Riemannian Hessian, we get the following expression:*

$$\operatorname{Hess} f(x)[u] = \nabla_u \operatorname{grad} f = \frac{\mathrm{D}}{\mathrm{d}t} \operatorname{grad} f(c(t)) \bigg|_{t=0}, \tag{5.17}$$

where $c \colon I \to \mathcal{M}$ is any smooth curve such that $c(0) = x$ and $c'(0) = u$. This is true in particular with $c(t) = \mathrm{R}_x(tu)$ for any retraction R on $\mathcal{M}$.

For the special case where ∇ is the connection defined by (5.4) on a manifold $\mathcal{M}$ embedded in a Euclidean space $\mathcal{E}$, the induced covariant derivative admits a particularly nice expression. Consider a smooth curve $c \colon I \to \mathcal{M}$. We can also think of it as a smooth curve $c \colon I \to \mathcal{E}$. Thus, a vector field Z along c on $\mathcal{M}$ is smooth exactly if it is smooth as a vector field along c in $\mathcal{E}$. As a result, it makes sense to write $\frac{\mathrm{d}}{\mathrm{d}t} Z$ to denote the classical (or extrinsic) derivative of Z in the embedding space $\mathcal{E}$. We are about to show that the operator $\frac{\mathrm{D}}{\mathrm{d}t} \colon \mathfrak{X}(c) \to \mathfrak{X}(c)$ defined by

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \mathrm{Proj}_{c(t)} \left(\frac{\mathrm{d}}{\mathrm{d}t} Z(t) \right) \tag{5.18}$$

is the covariant derivative induced by ∇. Thus, similarly to ∇ (5.4), it suffices to take a classical derivative in the embedding space, followed by an orthogonal projection to the tangent spaces. In particular, if $\mathcal{M}$ is (an open subset of) a linear space, then $\frac{\mathrm{D}}{\mathrm{d}t} Z = \frac{\mathrm{d}}{\mathrm{d}t} Z$, as expected.

Proposition 5.31. *Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$ with connection ∇ as in (5.4). The operator $\frac{\mathrm{D}}{\mathrm{d}t}$ defined by (5.18) is the induced covariant derivative, that is, it satisfies properties 1–3 in Theorem 5.29. If $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$, then $\frac{\mathrm{D}}{\mathrm{d}t}$ also satisfies property 4 in that same theorem.*

Proof. Properties 1 and 2 follow directly from linearity of projectors. For the chain rule, consider $U \in \mathfrak{X}(\mathcal{M})$ with smooth extension $\bar{U}$, and $Z(t) = U(c(t)) = \bar{U}(c(t))$. Then, $\frac{\mathrm{d}}{\mathrm{d}t}Z(t) = \mathrm{D}\bar{U}(c(t))[c'(t)] = \bar{\nabla}_{c'(t)}\bar{U}$ with $\bar{\nabla}$ the Riemannian connection on $\mathcal{E}$. It follows from (5.5) that

$$\frac{\mathrm{D}}{\mathrm{d}t}Z(t) = \mathrm{Proj}_{c(t)}\big(\bar{\nabla}_{c'(t)}\bar{U}\big) = \nabla_{c'(t)}U,$$

as desired for property 3.

Property 4 follows as a consequence of Theorem 5.29, but we verify it explicitly anyway because this is an important special case and because the proof below does not require local frames. Let $\langle\cdot,\cdot\rangle$ be the Euclidean metric. Consider two vector fields $Y, Z \in \mathfrak{X}(c)$. Differentiate the function $t \mapsto \langle Y(t), Z(t)\rangle$ treating Y, Z as vector fields along c in $\mathcal{E}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle Y, Z\rangle = \left\langle \frac{\mathrm{d}}{\mathrm{d}t}Y, Z\right\rangle + \left\langle Y, \frac{\mathrm{d}}{\mathrm{d}t}Z\right\rangle.$$

Since Z is tangent to $\mathcal{M}$, $Z = \mathrm{Proj}_c Z$ (and similarly for Y). Now using that Proj is self-adjoint, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle Y, Z\rangle &= \left\langle \frac{\mathrm{d}}{\mathrm{d}t}Y, \mathrm{Proj}_c Z\right\rangle + \left\langle \mathrm{Proj}_c Y, \frac{\mathrm{d}}{\mathrm{d}t}Z\right\rangle \\ &= \left\langle \mathrm{Proj}_c \frac{\mathrm{d}}{\mathrm{d}t}Y, Z\right\rangle + \left\langle Y, \mathrm{Proj}_c \frac{\mathrm{d}}{\mathrm{d}t}Z\right\rangle = \left\langle \frac{\mathrm{D}}{\mathrm{d}t}Y, Z\right\rangle + \left\langle Y, \frac{\mathrm{D}}{\mathrm{d}t}Z\right\rangle.\end{aligned}$$

Conclude using that $\langle\cdot,\cdot\rangle$ is the metric both in the embedding space and on $\mathcal{M}$ since $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$. ☐

Example 5.32. *For a smooth function f on a Riemannian submanifold $\mathcal{M}$ of a Euclidean space $\mathcal{E}$, we can apply* (5.18) *to* (5.17) *to find*

$$\begin{aligned}\mathrm{Hess}f(x)[u] &= \mathrm{Proj}_x\left(\lim_{t\to 0}\frac{\mathrm{grad}f(c(t)) - \mathrm{grad}f(c(0))}{t}\right) \\ &= \lim_{t\to 0}\frac{\mathrm{Proj}_x(\mathrm{grad}f(c(t))) - \mathrm{grad}f(x)}{t},\end{aligned} \tag{5.19}$$

where the subtraction makes sense because $\mathrm{grad}f(c(t))$ *is an element of the linear embedding space $\mathcal{E}$ for all t. This holds for any smooth curve c such that $c(0) = x$ and $c'(0) = u$. Picking a retraction curve for example, this justifies the claim that, for some aptly chosen $\bar{t} > 0$,*

$$\mathrm{Hess}f(x)[u] \approx \frac{\mathrm{Proj}_x(\mathrm{grad}f(\mathrm{R}_x(\bar{t}u))) - \mathrm{grad}f(x)}{\bar{t}}. \tag{5.20}$$

This is a finite difference approximation *of the Hessian. Assuming* $\mathrm{grad}f(x)$ *is readily available, it affords us a straightforward way to approximate* $\mathrm{Hess}f(x)[u]$ *for the computational cost of one retraction, one gradient evaluation, and one projection. The parameter $\bar{t}$ should be small enough for the mathematical approximation to be accurate, yet large enough to avoid catastrophic numerical errors. We revisit this concept in more generality in Section 10.6.*

Exercise 5.33. *In the proof of Theorem 5.29, show that the operator* (5.15) *satisfies properties 1–3.*

Exercise 5.34. *Continuing from Exercise 5.4, show that if* $\nabla^{(1)}, \nabla^{(2)}$ *are connections on* $\mathcal{M}_1, \mathcal{M}_2$ *with induced covariant derivatives* $\frac{\mathrm{D}}{\mathrm{d}t}^{(1)}, \frac{\mathrm{D}}{\mathrm{d}t}^{(2)}$ *(respectively), then the covariant derivative* $\frac{\mathrm{D}}{\mathrm{d}t}$ *induced on* $\mathcal{M}_1 \times \mathcal{M}_2$ *by the product connection* ∇ (5.6) *is given simply by:*

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \left(\frac{\mathrm{D}}{\mathrm{d}t}^{(1)} Z_1(t), \frac{\mathrm{D}}{\mathrm{d}t}^{(2)} Z_2(t) \right), \tag{5.21}$$

where $Z = (Z_1, Z_2)$ *is a smooth vector field along a curve* $c = (c_1, c_2)$ *on* $\mathcal{M}_1 \times \mathcal{M}_2$*, so that* Z_1, Z_2 *are smooth vector fields along* c_1, c_2 *on* $\mathcal{M}_1, \mathcal{M}_2$*, respectively. In order to verify the chain rule property with a smooth vector field* $U = (U_1, U_2)$ *on* $\mathcal{M}_1 \times \mathcal{M}_2$*, it is helpful first to establish that*

$$\frac{\mathrm{D}}{\mathrm{d}t}^{(1)} (U_1 \circ c)(t) = \nabla^{(1)}_{c_1'(t)} U_1(\cdot, c_2(t)) + \mathrm{D}U_1(c_1(t), \cdot)(c_2(t))[c_2'(t)]. \tag{5.22}$$

(Likewise for U_2*.) Hint: expand* U_1 *in a local frame and use Exercise 3.40.*

Exercise 5.35. *For* $Z \in \mathfrak{X}(c)$*, show that* $\frac{\mathrm{D}}{\mathrm{d}t}(Z \circ \phi)(t) = \phi'(t) \left(\frac{\mathrm{D}}{\mathrm{d}t} Z\right)(\phi(t))$ *where* $\phi \colon \mathbb{R} \to I$ *is any smooth reparameterization of* $c \colon I \to \mathcal{M}$.

5.8 Acceleration and geodesics

If the manifold $\mathcal{M}$ is equipped with a covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$, we can use it to define the notion of acceleration along a curve on $\mathcal{M}$.

Definition 5.36. *Let* $c \colon I \to \mathcal{M}$ *be a smooth curve. Its* velocity *is the vector field* $c' \in \mathfrak{X}(c)$*. The* acceleration of c *is the smooth vector field* $c'' \in \mathfrak{X}(c)$ *defined by:*

$$c'' = \frac{\mathrm{D}}{\mathrm{d}t} c'.$$

We also call c'' *the* intrinsic acceleration *of* c.

When $\mathcal{M}$ is embedded in a linear space $\mathcal{E}$, a curve c on $\mathcal{M}$ is also a curve in $\mathcal{E}$. It is then convenient to distinguish notationally between the acceleration of c on the manifold (as defined above) and the classical acceleration of c in the embedding space. We write

$$\ddot{c} = \frac{\mathrm{d}^2}{\mathrm{d}t^2} c$$

for the classical or *extrinsic* acceleration. In that spirit, we use notations c' and $\dot{c}$ interchangeably for velocity since the two notions coincide.

When $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$ (with the associated Riemannian connection), the induced covariant derivative takes on a convenient form (5.18), so that

$$c''(t) = \mathrm{Proj}_{c(t)}(\ddot{c}(t)). \tag{5.23}$$

We also state this as $c'' = \mathrm{Proj}_c(\ddot{c})$ for short. In words: on a Riemannian submanifold, the acceleration of a curve is the tangential part of its extrinsic acceleration in the embedding space. See Section 5.11 for a discussion of the normal part

Example 5.37. *Consider the sphere* $\mathrm{S}^{d-1} = \{x \in \mathbb{R}^d : x^\top x = 1\}$ *equipped with the Riemannian submanifold geometry of* $\mathbb{R}^d$ *with the canonical metric. For a given* $x \in \mathrm{S}^{d-1}$ *and* $v \in \mathrm{T}_x\mathrm{S}^{d-1}$ *(nonzero), consider the curve*

$$c(t) = \cos(t\|v\|)x + \frac{\sin(t\|v\|)}{\|v\|}v,$$

which traces a so-called great circle on the sphere. The (extrinsic) velocity and acceleration of c *in* $\mathbb{R}^d$ *are easily derived:*

$$\begin{aligned}\dot{c}(t) &= -\|v\|\sin(t\|v\|)x + \cos(t\|v\|)v,\\ \ddot{c}(t) &= -\|v\|^2\cos(t\|v\|)x - \|v\|\sin(t\|v\|)v = -\|v\|^2 c(t).\end{aligned}$$

The velocity $c'(t)$ *matches* $\dot{c}(t)$*. Owing to* (5.23)*, to get the (intrinsic) acceleration of* c *on* S^{d-1}*, we project:*

$$c''(t) = \mathrm{Proj}_{c(t)}\ddot{c}(t) = (I_d - c(t)c(t)^\top)\ddot{c}(t) = 0.$$

Thus, c *is a curve with zero acceleration on the sphere (even though its acceleration in* $\mathbb{R}^d$ *is nonzero.)*

Curves with zero acceleration play a particular role in geometry, as they provide a natural generalization of the concept of straight lines $t \mapsto x + tv$ from linear spaces to manifolds. Reading the definition below, recall that by default we equip a Riemannian manifold with its Riemannian connection ∇, which induces a covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$: it is with the latter that c'' is to be interpreted.

Definition 5.38. *On a Riemannian manifold* $\mathcal{M}$*, a* geodesic *is a smooth curve* $c\colon I \to \mathcal{M}$ *such that* $c''(t) = 0$ *for all* $t \in I$*, where* I *is an open interval of* $\mathbb{R}$*.*

Owing to (5.23), a curve c on a Riemannian submanifold $\mathcal{M}$ is a geodesic if and only if its extrinsic acceleration $\ddot{c}$ is everywhere normal to $\mathcal{M}$. Geodesics are further discussed in Section 10.2—they play a minor role in practical optimization algorithms.

Exercise 5.39. *Let* $c(t) = (c_1(t), c_2(t))$ *be a smooth curve on the product manifold* $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$*. Its velocity is given by* $c'(t) = (c_1'(t), c_2'(t))$*. Equip* $\mathcal{M}_1$ *and* $\mathcal{M}_2$ *with Riemannian structures, and let* $\mathcal{M}$ *be their Riemannian product as in Example 3.57. Argue that* $c''(t) = (c_1''(t), c_2''(t))$*, where accelerations are defined with respect to the Riemannian connections. Deduce that* c *is a geodesic on* $\mathcal{M}_1 \times \mathcal{M}_2$ *if and only if* c_1, c_2 *are geodesics on* $\mathcal{M}_1, \mathcal{M}_2$*, respectively.*

5.9 A second-order Taylor expansion on curves

On a Riemannian manifold $\mathcal{M}$, consider a smooth function $f : \mathcal{M} \to \mathbb{R}$ and a smooth curve $c : I \to \mathcal{M}$ passing through x at $t = 0$ with velocity v. In this section, we build a second-order Taylor expansion for the function $g = f \circ c$, as announced in the introduction of this chapter.

Since g is a smooth function from $I \subseteq \mathbb{R}$ to $\mathbb{R}$, it has a Taylor expansion:

$$f(c(t)) = g(t) = g(0) + tg'(0) + \frac{t^2}{2} g''(0) + O(t^3).$$

We have the tools necessary to investigate the derivatives of g. Indeed,

$$g'(t) = \mathrm{D}f(c(t))[c'(t)] = \langle \mathrm{grad} f(c(t)), c'(t) \rangle_{c(t)},$$

so that

$$(f \circ c)'(0) = g'(0) = \langle \mathrm{grad} f(x), v \rangle_x . \tag{5.24}$$

Moreover, using in turn properties 4 and 3 of Theorem 5.29 regarding the covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$ induced by the Riemannian connection ∇, followed by Definition 5.14 for the Hessian, we compute:

$$\begin{aligned}
g''(t) &= \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathrm{grad} f(c(t)), c'(t) \rangle_{c(t)} \\
\text{(property 4)} \quad &= \left\langle \frac{\mathrm{D}}{\mathrm{d}t} (\mathrm{grad} f \circ c)(t), c'(t) \right\rangle_{c(t)} + \left\langle \mathrm{grad} f(c(t)), \frac{\mathrm{D}}{\mathrm{d}t} c'(t) \right\rangle_{c(t)} \\
\text{(property 3)} \quad &= \left\langle \nabla_{c'(t)} \mathrm{grad} f, c'(t) \right\rangle_{c(t)} + \langle \mathrm{grad} f(c(t)), c''(t) \rangle_{c(t)} \\
&= \langle \mathrm{Hess} f(c(t))[c'(t)], c'(t) \rangle_{c(t)} + \langle \mathrm{grad} f(c(t)), c''(t) \rangle_{c(t)} .
\end{aligned}$$

Evaluating $g''(t)$ at $t = 0$ yields:

$$(f \circ c)''(0) = g''(0) = \langle \mathrm{Hess} f(x)[v], v \rangle_x + \langle \mathrm{grad} f(x), c''(0) \rangle_x . \tag{5.25}$$

These all combine to form:

$$\begin{aligned}
f(c(t)) = f(x) + t \langle \mathrm{grad} f(x), v \rangle_x &+ \frac{t^2}{2} \langle \mathrm{Hess} f(x)[v], v \rangle_x \\
&+ \frac{t^2}{2} \langle \mathrm{grad} f(x), c''(0) \rangle_x + O(t^3).
\end{aligned} \tag{5.26}$$

To be clear, formula (5.26) holds for all smooth curves c satisfying $c(0) = x$ and $c'(0) = v$. Of particular interest for optimization is the Taylor expansion of f along a retraction curve. That is the topic of Section 5.10.

Exercise 5.40. *For a smooth curve $c : [0, 1] \to \mathcal{M}$ on a Riemannian manifold $\mathcal{M}$ with $c(0) = x$ and $c(1) = y$, show that there exists $t \in (0, 1)$ such that*

$$\begin{aligned}
f(y) = f(x) + \langle \mathrm{grad} f(x), c'(0) \rangle_x &+ \frac{1}{2} \langle \mathrm{Hess} f(c(t))[c'(t)], c'(t) \rangle_{c(t)} \\
&+ \frac{1}{2} \langle \mathrm{grad} f(c(t)), c''(t) \rangle_{c(t)} .
\end{aligned} \tag{5.27}$$

(Hint: use the mean value theorem.) Show that the speed $\|c'(t)\|_{c(t)}$ *of the curve* c *is constant if* c *is a geodesic. Deduce a proof of Lemma 5.41 below.*

Lemma 5.41. *Let* $c(t)$ *be a geodesic connecting* $x = c(0)$ *to* $y = c(1)$, *and assume* $\mathrm{Hess}f(c(t)) \succeq \mu \, \mathrm{Id}$ *for some* $\mu \in \mathbb{R}$ *and all* $t \in [0, 1]$. *Then,*

$$f(y) \geq f(x) + \langle \mathrm{grad}f(x), v \rangle_x + \frac{\mu}{2} \|v\|_x^2.$$

Here is some context for Lemma 5.41. If $x \in \mathcal{M}$ is such that $\mathrm{grad}f(x) = 0$ and $\mathrm{Hess}f(x) \succeq \mu' \, \mathrm{Id}$ for some $\mu' > 0$, then by continuity of eigenvalues there exists a neighborhood $\mathcal{U}$ of x in which $\mathrm{Hess}f(z) \succeq \mu \, \mathrm{Id}$ for all $z \in \mathcal{U}$ and some $\mu > 0$. If $\mathcal{U}$ is appropriately chosen, the lemma implies that x is the unique critical point in $\mathcal{U}$, and it is the global minimizer in that set (hence an isolated local minimizer for f on all of $\mathcal{M}$). This is relevant in connection with Chapter 11 about *geodesic convexity*: the neighborhood can be chosen to be a geodesically convex geodesic ball, and f restricted to that ball is μ-strongly convex, in a geodesic sense. This can ease the study of the local convergence behavior of optimization algorithms near isolated local minimizers by paralleling Section 11.5.

5.10 Second-order retractions

Continuing from the Taylor expansion (5.26) established above, we consider the important case where c is a retraction curve, that is,

$$c(t) = \mathrm{R}_x(tv)$$

for a point $x \in \mathcal{M}$ and a vector $v \in \mathrm{T}_x\mathcal{M}$. A direct application of (5.26) yields

$$\begin{aligned} f(\mathrm{R}_x(tv)) = f(x) + t \langle \mathrm{grad}f(x), v \rangle_x &+ \frac{t^2}{2} \langle \mathrm{Hess}f(x)[v], v \rangle_x \\ &+ \frac{t^2}{2} \langle \mathrm{grad}f(x), c''(0) \rangle_x + O(t^3). \end{aligned} \tag{5.28}$$

The last term involving the acceleration of c at $t = 0$ is undesirable, as it is of order t^2 and depends on the retraction. Fortunately, it vanishes if $\mathrm{grad}f(x) = 0$ or $c''(0) = 0$. The latter happens in particular if c is a geodesic. Retractions whose curves are geodesics are studied later in Section 10.2: they are called *exponential maps*. More generally though, notice that we only need the acceleration to vanish at $t = 0$. This suggests the following definition.

Definition 5.42. *A* second-order retraction R *on a Riemannian manifold* $\mathcal{M}$ *is a retraction such that, for all* $x \in \mathcal{M}$ *and all* $v \in \mathrm{T}_x\mathcal{M}$, *the curve* $c(t) = \mathrm{R}_x(tv)$ *has zero acceleration at* $t = 0$, *that is,* $c''(0) = 0$.

Second-order retractions are not hard to come by: see Section 5.12 for a common construction that works on Riemannian submanifolds. The following example illustrates that construction on the sphere.

Example 5.43. *Consider the following retraction on the sphere* S^{d-1}*:*

$$\mathrm{R}_x(v) = \frac{x+v}{\|x+v\|}.$$

That retraction is second order. Indeed, with $c(t) = \mathrm{R}_x(tv)$*:*

$$\begin{aligned}
c(t) &= \frac{x+tv}{\sqrt{1+t^2\|v\|^2}} = \left(1 - \frac{1}{2}\|v\|^2t^2 + O(t^4)\right)(x+tv), \\
\dot{c}(t) &= -\|v\|^2 t(x+tv) + \left(1 - \frac{1}{2}\|v\|^2t^2\right)v + O(t^3), \\
\ddot{c}(t) &= -\|v\|^2(x+tv) - \|v\|^2tv - \|v\|^2tv + O(t^2) \\
&= -\|v\|^2(x+3tv) + O(t^2).
\end{aligned}$$

Of course, $c'(0) = \dot{c}(0) = v$*. As for acceleration,* $\ddot{c}(0) = -\|v\|^2x$*, so that:*

$$c''(0) = \mathrm{Proj}_x(\ddot{c}(0)) = 0,$$

as announced.

We summarize two important particular cases of the Taylor expansion (5.28) into a useful statement regarding the pullback $f \circ \mathrm{R}_x$.

Proposition 5.44. *Consider a Riemannian manifold* $\mathcal{M}$ *equipped with any retraction* R*, and a smooth function* $f\colon \mathcal{M} \to \mathbb{R}$*. If* x *is a critical point of* f *(that is, if* $\mathrm{grad}f(x) = 0$*), then*

$$f(\mathrm{R}_x(s)) = f(x) + \frac{1}{2}\left\langle \mathrm{Hess}f(x)[s], s\right\rangle_x + O(\|s\|_x^3). \tag{5.29}$$

If R *is a second-order retraction, then for all points* $x \in \mathcal{M}$ *we have*

$$f(\mathrm{R}_x(s)) = f(x) + \left\langle \mathrm{grad}f(x), s\right\rangle_x + \frac{1}{2}\left\langle \mathrm{Hess}f(x)[s], s\right\rangle_x + O(\|s\|_x^3). \tag{5.30}$$

Proof. Simply rewrite (5.28) with $s = tv$. ☐

That the first identity holds for all retractions is useful to study the behavior of optimization algorithms at or close to critical points.

Proposition 5.44 suggests an alternative way to compute the Riemannian Hessian. Indeed, the direct way is to use the definition as we did in Example 5.17. This requires computing with the Riemannian connection, which may not be straightforward for general manifolds. If a second-order retraction is on hand or if we are only interested in the Hessian at critical points, an alternative is to use the next result. In practical terms, it suggests to compose f with R_x (which yields a smooth function from a linear space to the reals), then to compute the Hessian of the latter in the usual way. This echoes Proposition 3.59 stating that $\mathrm{grad}f(x) = \mathrm{grad}(f \circ \mathrm{R}_x)(0)$.

Proposition 5.45. *If the retraction is second order or if* $\mathrm{grad}f(x) = 0$*, then*

$$\mathrm{Hess}f(x) = \mathrm{Hess}(f \circ \mathrm{R}_x)(0),$$

where the right-hand side is the Hessian of $f \circ \mathrm{R}_x \colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}$ *at* $0 \in \mathrm{T}_x\mathcal{M}$. *The latter is a "classical" Hessian since* $\mathrm{T}_x\mathcal{M}$ *is a Euclidean space. See also Exercise 10.73 for the Hessian of* $f \circ \mathrm{R}_x$ *away from the origin.*

Proof. If R is second order, expand $\hat{f}_x(s) = f(\mathrm{R}_x(s))$ using (5.30):

$$\hat{f}_x(s) = f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{1}{2} \langle \mathrm{Hess} f(x)[s], s \rangle_x + O(\|s\|_x^3).$$

The gradient and Hessian of $\hat{f}_x \colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}$ with respect to s follow easily, using the fact that $\mathrm{Hess} f(x)$ is self-adjoint:

$$\begin{aligned} \mathrm{grad} \hat{f}_x(s) &= \mathrm{grad} f(x) + \mathrm{Hess} f(x)[s] + O(\|s\|_x^2), \text{ and} \\ \mathrm{Hess} \hat{f}_x(s)[\dot{s}] &= \mathrm{Hess} f(x)[\dot{s}] + O(\|s\|_x \|\dot{s}\|_x). \end{aligned}$$

Evaluating at $s = 0$ yields $\mathrm{Hess} \hat{f}_x(0) = \mathrm{Hess} f(x)$, as announced. The proof is similar if x is a critical point, starting with (5.29). □

We close this section with a remark. Recall that critical points of a function and (first-order) retractions are defined on a manifold $\mathcal{M}$ independently of any Riemannian structure on $\mathcal{M}$. Proposition 5.45 further tells us that, *at a critical point* x, the Riemannian Hessian of a function on $\mathcal{M}$ depends on the Riemannian metric only through the inner product on $\mathrm{T}_x\mathcal{M}$. In particular, the *signature* of the Hessian at x, that is, the number of positive, zero and negative eigenvalues, is independent on the Riemannian structure.

Exercise 5.46. *Let* R *be a retraction on a manifold* $\mathcal{M}$*, and fix* $x \in \mathcal{M}$. *Consider a smooth curve* $w \colon I \to \mathrm{T}_x\mathcal{M}$ *in the tangent space at* x *such that* $w(0) = 0$. *This induces a smooth curve* $c(t) = \mathrm{R}_x(w(t))$ *on* $\mathcal{M}$. *Of course,* $c(0) = x$. *It is also easy to confirm that* $c'(0) = w'(0)$.

Show that if $\mathcal{M}$ *is Riemannian and* R *is a second-order retraction, then we also have* $c''(0) = w''(0)$. *Hint: expand the differential* $\mathrm{DR}_x(v)$ *for* v *close to the origin using a local frame around* x*, e.g., as provided by Exercise 3.72. Then use the properties of* $\frac{\mathrm{D}}{\mathrm{d}t}$ *and* ∇ *to compute* $c''(t)$ *and work out* $c''(0)$.

Here is one take-away: for all $u, v \in \mathrm{T}_x\mathcal{M}$ *we can create a curve* c *on* $\mathcal{M}$ *such that* $c'(0) = u$ *and* $c''(0) = v$ *as* $c(t) = \mathrm{R}_x\Big(tu + \frac{t^2}{2}v\Big)$. *This is always doable since the exponential map is a second-order retraction (Section 10.2).*

5.11 Special case: Riemannian submanifolds*

The special case where $\mathcal{M}$ is a Riemannian submanifold of a Euclidean space $\mathcal{E}$ merits further attention. Consider a smooth function $f \colon \mathcal{M} \to \mathbb{R}$ and a point $x \in \mathcal{M}$ together with a tangent vector $u \in \mathrm{T}_x\mathcal{M}$. Let $\bar{f}$ be a smooth extension of f to a neighborhood of $\mathcal{M}$ in $\mathcal{E}$, and let c be any smooth curve on $\mathcal{M}$ satisfying $c(0) = x$ and $c'(0) = u$. We know from Proposition 3.61 that

$$\mathrm{grad} f(c(t)) = \mathrm{Proj}_{c(t)}(\mathrm{grad} \bar{f}(c(t))), \tag{5.31}$$

where Proj_x is the orthogonal projector from $\mathcal{E}$ to $\mathrm{T}_x\mathcal{M}$. We know from Exercise 3.66 that $x \mapsto \mathrm{Proj}_x$ is smooth. Starting from here, formulas (5.17) and (5.18) combine to yield:

$$\begin{aligned}\mathrm{Hess} f(x)[u] &= \frac{\mathrm{D}}{\mathrm{d}t}\mathrm{grad} f(c(t))\bigg|_{t=0} \\ &= \mathrm{Proj}_x\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Proj}_{c(t)}(\mathrm{grad}\bar{f}(c(t)))\bigg|_{t=0}\right) \\ &= \mathrm{Proj}_x\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Proj}_{c(t)}\bigg|_{t=0}(\mathrm{grad}\bar{f}(x))\right) \\ &\quad + \mathrm{Proj}_x\left(\mathrm{Proj}_x\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{grad}\bar{f}(c(t))\bigg|_{t=0}\right)\right). \end{aligned} \tag{5.32}$$

This simplifies noting that $\mathrm{Proj}_x \circ \mathrm{Proj}_x = \mathrm{Proj}_x$. Let us introduce notation for the differentials of the projector, based on Definition 3.34:

$$\mathcal{P}_u \triangleq \mathrm{D}(x \mapsto \mathrm{Proj}_x)(x)[u] = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Proj}_{c(t)}\bigg|_{t=0}. \tag{5.33}$$

Intuitively, that differential measures how the tangent spaces of $\mathcal{M}$ vary, that is, how $\mathcal{M}$ "bends" in its embedding space. Plugging this notation in (5.32), we can write the Riemannian Hessian as follows:

$$\mathrm{Hess} f(x)[u] = \mathrm{Proj}_x\big(\mathcal{P}_u(\mathrm{grad}\bar{f}(x))\big) + \mathrm{Proj}_x\big(\mathrm{Hess}\bar{f}(x)[u]\big). \tag{5.34}$$

It is instructive to investigate $\mathcal{P}_u$ more closely. To this end, let $P(t) = \mathrm{Proj}_{c(t)}$. In particular, $P(0) = \mathrm{Proj}_x$ and $P'(0) = \mathcal{P}_u$. By definition of projectors, $P(t)P(t) = P(t)$ for all t. Differentiate with respect to t to find that $P'(t)P(t) + P(t)P'(t) = P'(t)$ for all t. At $t = 0$, this reveals a useful identity:

$$\mathcal{P}_u = \mathcal{P}_u \circ \mathrm{Proj}_x + \mathrm{Proj}_x \circ \mathcal{P}_u. \tag{5.35}$$

Let $\mathrm{Proj}_x^\perp = \mathrm{Id} - \mathrm{Proj}_x$ denote the orthogonal projector to the *normal space* $\mathrm{N}_x\mathcal{M}$, that is, the orthogonal complement of $\mathrm{T}_x\mathcal{M}$ in $\mathcal{E}$. Then, the identity above can be reorganized in two ways to find:

$$\mathcal{P}_u \circ \mathrm{Proj}_x^\perp = \mathrm{Proj}_x \circ \mathcal{P}_u \qquad \text{and} \qquad \mathcal{P}_u \circ \mathrm{Proj}_x = \mathrm{Proj}_x^\perp \circ \mathcal{P}_u. \tag{5.36}$$

Combining (5.34) and (5.36) warrants the following statement.

Corollary 5.47. *Let $\mathcal{M}$ be a Riemannian submanifold of a Euclidean space $\mathcal{E}$. For a smooth function $f \colon \mathcal{M} \to \mathbb{R}$ with smooth extension $\bar{f}$ to a neighborhood of $\mathcal{M}$ in $\mathcal{E}$, the Riemannian Hessian of f is given by:*

$$\mathrm{Hess} f(x)[u] = \mathrm{Proj}_x(\mathrm{Hess}\bar{f}(x)[u]) + \mathcal{P}_u(\mathrm{Proj}_x^\perp(\mathrm{grad}\bar{f}(x))),$$

where $\mathcal{P}_u$ is the differential of $x \mapsto \mathrm{Proj}_x$ at x along u and $\mathrm{Proj}_x^\perp = \mathrm{Id} - \mathrm{Proj}_x$.

This casts the Riemannian Hessian as the projected Euclidean Hessian of a smooth extension, plus a correction term which depends on that extension only through the normal part of its Euclidean gradient.

Evidently, the differential of the projector captures important aspects of the geometry of $\mathcal{M}$ as judged by the embedding space $\mathcal{E}$. We introduce two new objects that contain the relevant information.

From (5.36), we find that if $v \in \mathrm{T}_x\mathcal{M}$ is a tangent vector at x, then $v = \mathrm{Proj}_x(v)$ so that

$$\mathcal{P}_u(v) = \mathcal{P}_u(\mathrm{Proj}_x(v)) = \mathrm{Proj}_x^{\perp}(\mathcal{P}_u(v)).$$

Notice how if v is tangent at x then $\mathcal{P}_u(v)$ is normal at x. Likewise, if $w \in \mathrm{N}_x\mathcal{M}$ is a normal vector at x, then $w = \mathrm{Proj}_x^{\perp}(w)$ so that

$$\mathcal{P}_u(w) = \mathcal{P}_u(\mathrm{Proj}_x^{\perp}(w)) = \mathrm{Proj}_x(\mathcal{P}_u(w)).$$

The output is necessarily a tangent vector at x. These considerations motivate us to define two special bilinear maps.

Definition 5.48. *Let $\mathcal{M}$ be a Riemannian submanifold of a Euclidean space $\mathcal{E}$. At a point $x \in \mathcal{M}$, the normal space $\mathrm{N}_x\mathcal{M}$ is the orthogonal complement of $\mathrm{T}_x\mathcal{M}$ in $\mathcal{E}$. The* second fundamental form *at x is the map:*

$$\mathrm{II} \colon \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M} \to \mathrm{N}_x\mathcal{M} \colon (u, v) \mapsto \mathrm{II}(u, v) = \mathcal{P}_u(v). \tag{5.37}$$

(Read "two" for II.*) The* Weingarten map *at x is the map:*

$$\mathcal{W} \colon \mathrm{T}_x\mathcal{M} \times \mathrm{N}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M} \colon (u, w) \mapsto \mathcal{W}(u, w) = \mathcal{P}_u(w). \tag{5.38}$$

For both, the map $\mathcal{P}_u \colon \mathcal{E} \to \mathcal{E}$ is defined by (5.33).

These two objects describe $\mathcal{P}_u$ fully. Indeed, for all $z \in \mathcal{E}$ we can decompose $\mathcal{P}_u(z)$ as follows:

$$\mathcal{P}_u(z) = \mathrm{II}(u, \mathrm{Proj}_x(z)) + \mathcal{W}(u, \mathrm{Proj}_x^{\perp}(z)). \tag{5.39}$$

The maps II and $\mathcal{W}$ are further related through the inner product $\langle \cdot, \cdot \rangle$, which denotes both the Euclidean inner product and the Riemannian metric at x since $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$. Indeed, for all $u, v \in \mathrm{T}_x\mathcal{M}$ and $w \in \mathrm{N}_x\mathcal{M}$ it holds that

$$\langle \mathrm{II}(u, v), w \rangle = \langle \mathcal{P}_u(v), w \rangle = \langle v, \mathcal{P}_u(w) \rangle = \langle v, \mathcal{W}(u, w) \rangle. \tag{5.40}$$

The middle equality above uses that $\mathcal{P}_u$ is self-adjoint on $\mathcal{E}$.

Going back to Corollary 5.47, we get an identity for the Hessian:

$$\mathrm{Hess} f(x)[u] = \mathrm{Proj}_x(\mathrm{Hess} \bar{f}(x)[u]) + \mathcal{W}(u, \mathrm{Proj}_x^{\perp}(\mathrm{grad} \bar{f}(x))). \tag{5.41}$$

In bilinear form on $\mathrm{T}_x\mathcal{M}$, combining with (5.40) we also have

$$\begin{aligned} \langle v, \mathrm{Hess} f(x)[u] \rangle &= \left\langle v, \mathrm{Hess} \bar{f}(x)[u] \right\rangle + \langle v, \mathcal{W}(u, \mathrm{Proj}_x^{\perp}(\mathrm{grad} \bar{f}(x))) \rangle \\ &= \left\langle v, \mathrm{Hess} \bar{f}(x)[u] \right\rangle + \langle \mathrm{II}(u, v), \mathrm{grad} \bar{f}(x) \rangle. \end{aligned} \tag{5.42}$$

In this last identity, it is still clear that only the normal part of $\mathrm{grad} \bar{f}(x)$ plays a role since $\mathrm{II}(u, v)$ is normal at x.

While it is not obvious from the definition (5.37), we may surmise from (5.42) that II is symmetric in its inputs. This is indeed the case, as we show through a couple of lemmas.

Lemma 5.49. *Let U, V be two smooth vector fields on a manifold $\mathcal{M}$ embedded in $\mathcal{E}$. Let $\bar{U}, \bar{V}$ be smooth extensions of U, V to a neighborhood O of $\mathcal{M}$. Then, $[\bar{U}, \bar{V}]$ is a smooth extension of $[U, V]$ to the same neighborhood O.*

Proof. This is a by-product of the proof of Theorem 5.8, with the hindsight that the Lie bracket of two vector fields is a vector field. □

Lemma 5.50 (Gauss formula)**.** *Let $\mathcal{M}$ be a Riemannian submanifold of $\mathcal{E}$, respectively endowed with their Riemannian connections ∇ and $\bar{\nabla}$. Let V be a smooth vector field on $\mathcal{M}$ with smooth extension $\bar{V}$. For all $(x, u) \in \mathrm{T}\mathcal{M}$, the vector $\bar{\nabla}_u \bar{V}$ splits in a tangent and a normal part to $\mathcal{M}$ at x as:*

$$\bar{\nabla}_u \bar{V} = \mathrm{Proj}_x(\bar{\nabla}_u \bar{V}) + \mathrm{Proj}_x^\perp(\bar{\nabla}_u \bar{V}) = \nabla_u V + \mathrm{II}(u, v),$$

where $v = V(x) = \bar{V}(x)$.

Proof. We already know that $\mathrm{Proj}_x(\bar{\nabla}_u \bar{V}) = \nabla_u V$ (Theorem 5.9). It remains to show that $\mathrm{Proj}_x^\perp(\bar{\nabla}_u \bar{V}) = \mathrm{II}(u, v)$. This is clear from the following computation, where we let $P(t) = \mathrm{Proj}_{c(t)}$ along a smooth curve c on $\mathcal{M}$ with $c(0) = x$ and $c'(0) = u$:

$$\begin{aligned}
\mathrm{Proj}_x^\perp(\bar{\nabla}_u \bar{V}) &= \mathrm{Proj}_x^\perp\left(\frac{\mathrm{d}}{\mathrm{d}t} \bar{V}(c(t)) \bigg|_{t=0} \right) \\
&= \mathrm{Proj}_x^\perp\left(\frac{\mathrm{d}}{\mathrm{d}t} P(t)(\bar{V}(c(t))) \bigg|_{t=0} \right) \\
&= \mathrm{Proj}_x^\perp\big(\mathcal{P}_u(v) + \mathrm{Proj}_x(\bar{\nabla}_u \bar{V})\big) \\
&= \mathcal{P}_u(\mathrm{Proj}_x(v))) \\
&= \mathrm{II}(u, v).
\end{aligned}$$

In the second to last step, we used (5.36) and $\mathrm{Proj}_x^\perp \circ \mathrm{Proj}_x = 0$. □

Proposition 5.51. *For all $u, v \in \mathrm{T}_x\mathcal{M}$ it holds that* $\mathrm{II}(u, v) = \mathrm{II}(v, u)$.

Proof. Pick $U, V \in \mathfrak{X}(\mathcal{M})$ such that $U(x) = u$ and $V(x) = v$; for example, let $U(y) = \mathrm{Proj}_y(u)$ and $V(y) = \mathrm{Proj}_y(v)$. Let $\bar{U}, \bar{V}$ be two smooth extensions for them. Lemma 5.50 yields

$$\mathrm{II}(u, v) - \mathrm{II}(v, u) = \mathrm{Proj}_x^\perp(\bar{\nabla}_u \bar{V} - \bar{\nabla}_v \bar{U}) = \mathrm{Proj}_x^\perp([\bar{U}, \bar{V}](x)).$$

Conclude with Lemma 5.49 which tells us $[\bar{U}, \bar{V}](x)$ is tangent at x. □

Lemma 5.50 has a counter-part for the covariant derivatives of a smooth vector field Z along a curve c on $\mathcal{M}$. The derivative of Z in the embedding space splits in a tangent and a normal part:

$$\frac{\mathrm{d}}{\mathrm{d}t} Z(t) = \frac{\mathrm{D}}{\mathrm{d}t} Z(t) + \mathrm{II}(c'(t), Z(t)). \tag{5.43}$$

Particularized to $Z = c'$, it follows that the extrinsic acceleration of c (denoted by $\ddot{c}$) and the intrinsic acceleration of c (denoted by c'') satisfy

$$\ddot{c}(t) = c''(t) + \mathrm{II}(c'(t), c'(t)). \tag{5.44}$$

This splits $\ddot{c}$ into tangent and normal parts. In particular, it makes clear the fact that c is a geodesic (that is, c'' is identically zero) if and only if $\ddot{c}$ is normal at all times. For a geodesic γ on $\mathcal{M}$ satisfying $\gamma(0) = x$ and $\gamma'(0) = u$ we know that $\gamma''(0) = 0$ so that $\ddot{\gamma}(0) = \mathrm{II}(u, u)$. This gives meaning to $\mathrm{II}(u, u)$ as the *extrinsic* acceleration of the geodesic γ in the embedding space. This informs us regarding *extrinsic curvature* of $\mathcal{M}$ in its embedding space, and may be useful to interpret (5.42).

Exercise 5.52. *Give a proof for formula* (5.43).

5.12 Special case: metric projection retractions*

Let $\mathcal{E}$ be a Euclidean space with the inner product $\langle \cdot, \cdot \rangle$ and associated norm $\|\cdot\|$. For an embedded submanifold $\mathcal{M}$, it is natural to consider the following as a tentative retraction, with $(x, v) \in \mathrm{T}\mathcal{M}$:

$$\mathrm{R}_x(v) = \operatorname*{arg\,min}_{x' \in \mathcal{M}} \|x' - (x + v)\|. \tag{5.45}$$

In fact, several of the retractions discussed in Chapter 7 are of that form. In this section, we argue that (5.45) indeed defines a retraction (albeit not necessarily on the whole tangent bundle) and that this retraction is second order if $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$.

Let $\mathrm{dist}_{\mathcal{M}} \colon \mathcal{E} \to \mathbb{R}$ denote the distance from a point of $\mathcal{E}$ to $\mathcal{M}$:

$$\mathrm{dist}_{\mathcal{M}}(y) = \inf_{x \in \mathcal{M}} \|x - y\|. \tag{5.46}$$

For a given $y \in \mathcal{E}$ the set

$$P_{\mathcal{M}}(y) = \{x \in \mathcal{M} : \|x - y\| = \mathrm{dist}_{\mathcal{M}}(y)\}$$

is the *metric projection* or *nonlinear orthogonal projection* of y to $\mathcal{M}$. It may be empty, or it may contain one or more points.

Let $A \subseteq \mathcal{E}$ be the set of points $y \in \mathcal{E}$ for which $P_{\mathcal{M}}(y)$ is a singleton, that is, for which there exists a unique point $x \in \mathcal{M}$ which is closest to y. It is an exercise to show that A may be neither open nor closed. However, strong properties hold on the interior of A, that is, on the largest subset of A which is open in $\mathcal{E}$. We state the following theorem without proof: see Section 5.13 for references.

Theorem 5.53. *Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$. Let $A \subseteq \mathcal{E}$ be the domain where $P_{\mathcal{M}}$ is single-valued, and let Ω denote the interior of A.*

1. *For $y \in A$ and $x = P_{\mathcal{M}}(y)$, we have that $y - x$ is orthogonal to $\mathrm{T}_x\mathcal{M}$ and $\{x + t(y - x) : t \in [0, 1)\} \subset \Omega$. In particular, $\mathcal{M} \subset \Omega$.*

2. *Ω is dense in A; if $\mathcal{M}$ is closed in $\mathcal{E}$, then the closure of Ω equals $\mathcal{E}$.*
3. *The restriction $P_{\mathcal{M}}\colon \Omega \to \mathcal{M}$ is smooth, and $\mathrm{D}P_{\mathcal{M}}(x) = \mathrm{Proj}_x$ for all $x \in \mathcal{M}$ (the orthogonal projector from $\mathcal{E}$ to $\mathrm{T}_x\mathcal{M}$).*

Now consider the following subset of the tangent bundle of $\mathcal{M}$:

$$\mathcal{O} = \{(x, v) \in \mathrm{T}\mathcal{M} : x + v \in \Omega\}. \tag{5.47}$$

It is open since the map $(x, v) \mapsto x + v$ is continuous (in fact, smooth) from $\mathrm{T}\mathcal{M}$ to $\mathcal{E}$. Moreover, $\mathcal{O}$ contains all pairs $(x, 0) \in \mathrm{T}\mathcal{M}$ since $\mathcal{M}$ is included in Ω. Formula (5.45) defines a retraction on $\mathcal{O}$.

Proposition 5.54. *On an embedded submanifold $\mathcal{M}$ of a Euclidean space $\mathcal{E}$ with norm $\|\cdot\|$, metric projection induces a retraction* (5.45) *as:*

$$\mathrm{R}\colon \mathcal{O} \to \mathcal{M}\colon (x, v) \mapsto \mathrm{R}(x, v) = \mathrm{R}_x(v) = P_{\mathcal{M}}(x + v).$$

This is called the metric projection retraction*.*

Proof. Clearly, $\mathrm{R}_x(0) = P_{\mathcal{M}}(x) = x$ for all $x \in \mathcal{M}$. From Theorem 5.53, we see that R is smooth on its domain by composition. By the same theorem, for all $(x, v) \in \mathrm{T}\mathcal{M}$ it holds that

$$\mathrm{D}\mathrm{R}_x(0)[v] = \mathrm{D}P_{\mathcal{M}}(x)[v] = \mathrm{Proj}_x(v) = v.$$

This confirms that $\mathrm{D}\mathrm{R}_x(0)$ is the identity on $\mathrm{T}_x\mathcal{M}$, as needed. □

Absil and Malick show that this retraction is part of a large family of second-order retractions (recall Definition 5.42) [AM12, Ex. 23]. For the case at hand, Razvan-Octavian Radu shared the short proof below.

Proposition 5.55. *If $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$, the retraction in Proposition 5.54 is second order.*

Proof. For an arbitrary $(x, v) \in \mathrm{T}\mathcal{M}$, consider the retraction curve $c(t) = \mathrm{R}_x(tv) = P_{\mathcal{M}}(x + tv)$. From Theorem 5.53, we know that $x + tv - c(t)$ is orthogonal to $\mathrm{T}_{c(t)}\mathcal{M}$ for all t. (We could also see this by noting that $c(t)$ is a critical point of $x' \mapsto \|x' - (x + tv)\|^2$ on $\mathcal{M}$.) This is all we need for our purpose.

Let $P(t) = \mathrm{Proj}_{c(t)}$ denote orthogonal projection to $\mathrm{T}_{c(t)}\mathcal{M}$: this is smooth in t (see Exercise 3.66). Since $x + tv - c(t)$ is orthogonal to $\mathrm{T}_{c(t)}\mathcal{M}$ for all t, we have that

$$g(t) = P(t)(x + tv - c(t))$$

is identically zero as a function from $I \subseteq \mathbb{R}$ (the domain of c) to $\mathcal{E}$. Thus, the (classical) derivative $g'(t)$ is also identically zero from I to $\mathcal{E}$:

$$g'(t) = P'(t)(x + tv - c(t)) + P(t)(v - c'(t)) \equiv 0.$$

At $t = 0$, we can use $c(0) = x$ to see that $0 = g'(0) = \mathrm{Proj}_x(v - c'(0))$. Since v and $c'(0)$ are both tangent vectors at x, this simply recovers the fact that $c'(0) = v$.

Differentiating once more, we have that $g''(t)$ is also identically zero from I to $\mathcal{E}$:

$$g''(t) = P''(t)(x + tv - c(t)) + 2P'(t)(v - c'(t)) - P(t)\frac{\mathrm{d}}{\mathrm{d}t}c'(t) \equiv 0.$$

At $t = 0$, we use $c(0) = x$ and $c'(0) = v$ to see that

$$0 = -g''(0) = \mathrm{Proj}_x\left(\frac{\mathrm{d}}{\mathrm{d}t}c'(0)\right) = \frac{\mathrm{D}}{\mathrm{d}t}c'(0) = c''(0),$$

where the last two equalities follow (5.23): this is where we use that $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$. □

The domain of R_x is the open subset $\mathcal{O}_x = \{v \in \mathrm{T}_x\mathcal{M} : (x, v) \in \mathcal{O}\}$. Clearly, $\mathcal{O}_x$ contains the origin, hence it also contains an open ball around the origin. However, $\mathcal{O}_x$ itself is not necessarily *star-shaped* with respect to the origin, that is: it is not true that $v \in \mathcal{O}_x$ implies $tv \in \mathcal{O}_x$ for all $t \in [0, 1]$. Indeed, consider metric projection to the set of matrices of fixed rank $\mathbb{R}_r^{m\times n}$ as defined by (7.49). Given $X \in \mathbb{R}_r^{m\times n}$, let $\dot{X} = -X$: this is a tangent vector to $\mathbb{R}_r^{m\times n}$ at X. Consider the line $t \mapsto X + t\dot{X}$: projection of $X + t\dot{X}$ to $\mathbb{R}_r^{m\times n}$ is well defined for all t except $t = 1$. It is an exercise to show that the same issue can arise with an embedded submanifold which is a closed set as well.

For an embedded submanifold $\mathcal{M}$ in $\mathcal{E}$, the domain A of $P_{\mathcal{M}}$ is all of $\mathcal{E}$ if and only if $\mathcal{M}$ is an affine subspace of $\mathcal{E}$. However, even if A (and a fortiori Ω) is not all of $\mathcal{E}$, it can be the case that $\mathcal{O} = \mathrm{T}\mathcal{M}$. This happens in particular when $\mathcal{M}$ is the boundary of a non-empty, closed, convex set, as then the sets of the form $x + \mathrm{T}_x\mathcal{M}$ are supporting hyperplanes of the convex hull of $\mathcal{M}$: projecting an element of $x + \mathrm{T}_x\mathcal{M}$ to $\mathcal{M}$ is the same as projecting to the convex hull of $\mathcal{M}$, which is globally defined. One example of this is metric projection onto the sphere (7.9), which is the boundary of the unit Euclidean ball. The metric projection retraction (7.24) for the Stiefel manifold $\mathrm{St}(n, p)$ with $p < n$ is also defined on the whole tangent bundle. This extends to $\mathrm{SO}(n)$ and (with some care regarding its two components) to $\mathrm{O}(n)$ (Section 7.4).

In Section 10.7, we define third-order retractions. Metric projection retractions are not third order in general: see Exercise 10.88.

Exercise 5.56. *Show that A may be neither open nor closed. Hint: consider $\mathcal{M} = \{(t, t^2) : t \in \mathbb{R}\}$ in $\mathbb{R}^2$ and show that $A = \mathbb{R}^2\backslash\{(0, t) : t > 1/2\}$.*

Exercise 5.57. *Show that $\mathcal{O}_x$ (the domain of the metric projection retraction restricted to $\mathrm{T}_x\mathcal{M}$) can fail to be star-shaped even if $\mathcal{M}$ is closed in $\mathcal{E}$. Hint: consider $\mathcal{M} = \{(t, \cos(t)) : t \in \mathbb{R}\} \subset \mathbb{R}^2$.*

5.13 Notes and references

Definition 5.1 for connections is not standard. As explained in Section 5.6, the usual approach is to define ∇ as an operator mapping two smooth vector fields to a smooth vector field, then to prove that this operator acts pointwise in its first argument. The latter point confirms that the two definitions are equivalent, but it is technical. Leading with Definition 5.1 makes it possible to skip these technicalities at first.

The pointwise dependence of ∇ in its first argument is a consequence of $\mathfrak{F}(\mathcal{M})$-linearity, and holds more generally for all tensor fields (see Section 10.7): most references give the proof at that level of generality. See for example [Lee12, Lem. 12.24], [Lee18, Prop. 4.5] or the remark after Def. 3.9 as well as Prop. 2.2 and Cor. 2.3 in [O'N83]. In the same vein, $\nabla_u V$ depends on V only locally through the values of V in a neighborhood of x (as in Lemma 5.23) or along any smooth curve passing through x with velocity u (owing to the chain rule property in Theorem 5.29)—see also [Lee18, Prop. 4.26].

Existence and uniqueness of the Riemannian connection is proved in most Riemannian geometry textbooks, e.g., [Lee18, Thm. 5.10] and [O'N83, Thm. 3.11]. Likewise, for existence and uniqueness of the covariant derivative of vector fields along curves, see [Lee18, Thm. 4.24 and Prop. 5.5] and [O'N83, Prop. 3.18].

We showed that the Riemannian connection for a Euclidean space corresponds to the usual directional derivative, and that the Riemannian connection on a Riemannian submanifold is obtained through orthogonal projection of the Riemannian connection in the embedding space [Lee18, Prop. 5.12], [AMS08, Prop. 5.3.2]. As part of that proof, we show symmetry in Theorem 5.8. This involves showing that if $\bar{U}, \bar{V}$ (smooth vector fields in the embedding space) are tangent to a submanifold, then their Lie bracket is also tangent to that submanifold: a similar statement appears as [Lee12, Cor. 8.32].

In the proof of Theorem 5.6, we use the fact that the Lie bracket $[U, V]$ of two smooth vector fields $U, V \in \mathfrak{X}(\mathcal{M})$ is itself a smooth vector field (Proposition 5.10). Our proof is non-standard and restricted to embedded submanifolds. We provide a general proof in Section 8.10. In the meantime, we get some insight along these lines: Exercise 5.11 introduces *derivations*, and claims smooth vector fields are derivations. In fact, the converse is true as well: smooth vector fields are one-to-one with derivations [Lee12, Prop. 8.15]. Then, Exercise 5.12 claims Lie brackets are derivations, so that Lie brackets are indeed smooth vector fields. The proof in Section 8.10 follows yet another path.

We follow the definition of Riemannian Hessian in Absil et al. [AMS08, §5.5]. The definition of second-order retractions and Proposition 5.45 follow that reference too. Absil et al. readily stress the importance of the fact that, at a critical point, it does not matter whether the retraction is second order. A broader discussion of various types of Hessians and second covariant derivatives of smooth functions is presented in [AMS08, §5.6]. See also Section 10.7.

The extension lemmas (Lemmas 5.25 and 5.26) hold for general manifolds. They are stated here to provide extensions in a neighborhood around a single point. More generally, these hold to obtain extensions around any closed set. This can be shown using partitions of unity [Lee12, Lem. 2.26, 8.6]. On this topic, bump functions on Euclidean spaces (Lemma 5.22) can be used to construct partitions of unity, which in turn can be used to construct bump functions on any manifold [Lee12, Lem. 2.22, Thm. 2.23, Prop. 2.25].

Definitions (5.37) and (5.38) for the second fundamental form and the Weingarten map are not standard but they are equivalent to the standard ones given in [Lee18, pp225–230]. Moreover, both maps (and their properties as laid out in Section 5.11) extend as is to the more general situation of a Riemannian submanifold of a Riemannian manifold, as defined in Section 8.14. The Hessian formula (5.41) involving the Weingarten map—and its construction—appear first in [AMT13].

Here are a few additional references for Section 5.11: The Gauss formula (Lemma 5.50) is discussed in [Lee18, Thm. 8.2], symmetry of II (Proposition 5.51) is stated in [Lee18, Prop. 8.1], the way covariant derivatives split on submanifolds appears in [O'N83, Prop. 4.8], the implications of the latter for geodesics on submanifolds is spelled out in [Lee18, Cor. 5.2] and [O'N83, Cor. 4.10], and the relation to extrinsic curvature is pointed out in [Lee18, Prop. 8.10].

Absil and Malick study metric projection retractions under the name *projective retraction*: they prove Propositions 5.54 and 5.55, and extend the discussion to broad classes of retractions [AM12, §3.1,§4.3]. Our proof of Proposition 5.55 is different. Moreover, the statements here are more specific regarding the domain of definition of the retractions, building on Theorem 5.53 which is a particular case of results presented by Dudek and Holly [DH94, Thms 3.8 and 3.13, Cor. 3.14, Thm. 4.1]. The discussion of when the domain of the metric projection retraction is the whole tangent bundle relies on certain basic facts which appear in [DH94, Thm. 5.1, 5.3, 6.4]. The two exercises of Section 5.12 parallel [DH94, Ex. 6.1, 6.4].

From Theorem 5.53 it is fairly direct to build a so-called *tubular neighborhood* for $\mathcal{M}$ in $\mathcal{E}$ [Lee18, Thm. 5.25]. The other way around, the proof of Proposition 5.55 generalizes easily to show that retractions built from tubular neighborhoods in a natural way are second order.

Breiding and Vannieuwenhoven study the sensitivity of metric projection to Riemannian submanifolds of Euclidean space in terms of extrinsic curvature, via the Weingarten map [BV21].

6 Second-order optimization algorithms

In Chapter 4, we used the Riemannian gradient of a function to develop Riemannian gradient descent: a first-order optimization algorithm. Now that we have developed the concept of Riemannian Hessian, we are in a good position to develop second-order optimization algorithms.

We first consider a Riemannian version of Newton's method: a pillar of both optimization and numerical analysis. When initialized close to certain local minimizers, this algorithm enjoys a quadratic local convergence rate. This is significantly faster than gradient descent which converges only linearly to such points, but the speedup comes at a cost:

1. Each iteration of Newton's method involves solving a linear system of equations in a tangent space: this is more expensive than computing a gradient step.
2. The global convergence behavior of Newton's method is erratic: it can easily diverge, whereas gradient descent usually converges.

To combine the best of both gradient descent and Newton's method, we turn to the Riemannian trust-regions method. That algorithm occupies us for most of the chapter. It preserves the favorable global behavior of gradient descent and keeps the per-iteration computational cost under control, while also preserving superlinear local converge rates to favorable local minimizers. This is arguably the most robust algorithm for smooth optimization on manifolds to date.

We open the chapter with a discussion of second-order optimality conditions. After describing Newton's method, we look into the conjugate gradients method: a matrix-free algorithm to solve the type of linear systems that arise in the computation of Newton steps. That algorithm resurfaces later in the chapter in a truncated form that is more suitable for the trust-region method.

6.1 Second-order optimality conditions

Before we move on to discuss second-order optimization algorithms, we secure second-order necessary optimality conditions: this is in the same spirit as the first-order conditions developed in Section 4.2.

Definition 6.1. *A point $x \in \mathcal{M}$ is* second-order critical *(or* second-order stationary*) for a smooth function $f : \mathcal{M} \to \mathbb{R}$ if*

$$(f \circ c)'(0) = 0 \qquad \text{and} \qquad (f \circ c)''(0) \geq 0$$

for all smooth curves c on $\mathcal{M}$ such that $c(0) = x$.

In words: it is not possible to move away from a second-order critical point x and obtain an initial decrease in the value of f with linear or even quadratic rate. In particular, second-order critical points are critical points.

Proposition 6.2. *Any local minimizer of a smooth function $f : \mathcal{M} \to \mathbb{R}$ is a second-order critical point of f.*

Proof. Let x be a local minimizer of f. We know from Proposition 4.5 that x is critical. For contradiction, assume x is not second-order critical. Thus, there exists a smooth curve $c : I \to \mathcal{M}$ with $c(0) = x$, $(f \circ c)'(0) = 0$ and $(f \circ c)''(0) < 0$. By continuity of $(f \circ c)''$, there exists $\delta > 0$ such that $(f \circ c)''(\tau) < 0$ for all $\tau \in [0, \delta]$. Taylor's theorem on $f \circ c$ implies that for each $t \in [0, \delta]$ there exists $\tau \in [0, \delta]$ such that

$$f(c(t)) = f(c(0)) + t \cdot (f \circ c)'(0) + \frac{t^2}{2} \cdot (f \circ c)''(\tau).$$

Thus, $f(c(t)) < f(x)$ for all $t \in (0, \delta]$: a contradiction. □

On a Riemannian manifold, second-order critical points are characterized by gradients and Hessians. (Recall Definition 3.7 for the basic spectral properties of self-adjoint maps.)

Proposition 6.3. *Let $f : \mathcal{M} \to \mathbb{R}$ be smooth on a Riemannian manifold $\mathcal{M}$. Then, x is a second-order critical point of f if and only if $\mathrm{grad} f(x) = 0$ and $\mathrm{Hess} f(x) \succeq 0$.*

Proof. Let $c : I \to \mathcal{M}$ be any smooth curve on $\mathcal{M}$ with $c(0) = x$, and let $v = c'(0)$, $u = c''(0)$. We know from Section 5.9 that

$$\begin{aligned}(f \circ c)'(0) &= \langle \mathrm{grad} f(x), v \rangle_x \quad \text{and} \\ (f \circ c)''(0) &= \langle \mathrm{grad} f(x), u \rangle_x + \langle \mathrm{Hess} f(x)[v], v \rangle_x .\end{aligned}$$

If $\mathrm{grad} f(x) = 0$ and $\mathrm{Hess} f(x) \succeq 0$, then x is second-order critical. The other way around, assume x is second-order critical. Since the above hold for all $v \in \mathrm{T}_x\mathcal{M}$, we first find that $\mathrm{grad} f(x) = 0$, and subsequently also that $\mathrm{Hess} f(x) \succeq 0$. □

It is also possible to establish *sufficient* conditions for local optimality by strengthening the second-order requirements.

Definition 6.4. *A point $x \in \mathcal{M}$ is* strictly second-order critical *(or* strictly second-order stationary*) for a smooth function $f : \mathcal{M} \to \mathbb{R}$ if*

$$(f \circ c)'(0) = 0 \qquad \text{and} \qquad (f \circ c)''(0) > 0$$

for all smooth curves c such that $c(0) = x$ and $c'(0) \neq 0$.

The proof of the next proposition relies on retractions to provide local parameterizations of a manifold. We could also use charts directly, as in Section 8.1.

Proposition 6.5. *If x is a strict second-order critical point for $f\colon \mathcal{M} \to \mathbb{R}$, then x is a strict local minimizer of f.*

Proof. Assume we have a retraction R on $\mathcal{M}$—This is not restrictive: see Section 5.12 for embedded submanifolds or Section 10.2 for the general case. Since $\mathrm{DR}_x(0)$ is invertible (it is the identity map), the inverse function theorem implies that R_x provides a diffeomorphism between a neighborhood of the origin of $\mathrm{T}_x\mathcal{M}$ and a neighborhood of x in $\mathcal{M}$: see Corollary 4.17. As a result, x is a strict local minimizer for f on $\mathcal{M}$ if and only if the origin is a strict local minimizer for $\hat{f}_x = f \circ \mathrm{R}_x$ on $\mathrm{T}_x\mathcal{M}$. Since $\mathrm{T}_x\mathcal{M}$ is a linear space, we can endow it with an inner product $\langle \cdot, \cdot \rangle$, so that $\hat{f}_x$ has a (Euclidean) gradient and Hessian. For some nonzero $v \in \mathrm{T}_x\mathcal{M}$, let $c(t) = \mathrm{R}_x(tv)$. Then, $\hat{f}_x(tv) = f(c(t))$ and therefore:

$$\langle \mathrm{grad}\hat{f}_x(0), v \rangle = \frac{\mathrm{d}}{\mathrm{d}t}\hat{f}_x(tv)\bigg|_{t=0} = (f \circ c)'(0) = 0,$$
$$\langle \mathrm{Hess}\hat{f}_x(0)[v], v \rangle = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\hat{f}_x(tv)\bigg|_{t=0} = (f \circ c)''(0) > 0.$$

The above hold for all $v \neq 0$, so that $\mathrm{grad}\hat{f}_x(0) = 0$ and $\mathrm{Hess}\hat{f}_x(0) \succ 0$. The claim now follows from the Euclidean case [NW06, Thm. 2.4]. $\square$

The following proposition is clear: its proof is a slight modification of that for Proposition 6.3.

Proposition 6.6. *Let $f\colon \mathcal{M} \to \mathbb{R}$ be smooth on a Riemannian manifold $\mathcal{M}$. Then, x is a strict second-order critical point of f if and only if* $\mathrm{grad}f(x) = 0$ *and* $\mathrm{Hess}f(x) \succ 0$.

6.2 Riemannian Newton's method

All optimization algorithms we consider are retraction based, in the sense that they iterate

$$x_{k+1} = \mathrm{R}_{x_k}(s_k)$$

for some step s_k. Thus, the change in cost function value from one iterate to the next can be understood through the pullbacks $\hat{f}_x = f \circ \mathrm{R}_x$:

$$f(x_{k+1}) = f(\mathrm{R}_{x_k}(s_k)) = \hat{f}_{x_k}(s_k).$$

Accordingly, a strategy to design algorithms is to pick a model $m_{x_k}\colon \mathrm{T}_{x_k}\mathcal{M} \to \mathbb{R}$ which suitably approximates $\hat{f}_{x_k}$, and to choose s_k as an (approximate) minimizer of m_{x_k}. Given our work building Taylor expansions (recall equation (5.28)), we know that close to critical points it holds that

$$\hat{f}_x(s) \approx m_x(s) \triangleq f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{1}{2} \langle \mathrm{Hess} f(x)[s], s \rangle_x .$$

The model m_x is a quadratic function of s on the linear space $\mathrm{T}_x\mathcal{M}$. A minimizer of m_x, if one exists, must be a critical point of m_x. To determine the gradient of m_x, we use the fact that $\mathrm{Hess} f(x)$ is self-adjoint to compute:

$$\langle \mathrm{grad}\, m_x(s), u \rangle_x = \mathrm{D} m_x(s)[u] = \langle \mathrm{grad} f(x), u \rangle_x + \langle \mathrm{Hess} f(x)[s], u \rangle_x .$$

The above holds for all $u \in \mathrm{T}_x\mathcal{M}$, hence by identification we find:

$$\mathrm{grad}\, m_x(s) = \mathrm{grad} f(x) + \mathrm{Hess} f(x)[s].$$

Thus, a tangent vector $s \in \mathrm{T}_x\mathcal{M}$ is a critical point of m_x if and only if

$$\mathrm{Hess} f(x)[s] = -\mathrm{grad} f(x). \tag{6.1}$$

This defines a linear system of equations called the *Newton equations* for the unknown $s \in \mathrm{T}_x\mathcal{M}$. So long as $\mathrm{Hess} f(x)$ is invertible, there exists a unique solution called the *Newton step*: we use it to define Algorithm 6.1. It is an easy exercise to show that if $\mathrm{Hess} f(x)$ is positive definite then the Newton step is the minimizer of m_x. In contrast, if the Hessian is invertible but not positive definite, then the Newton step does *not* correspond to a minimizer of m_x: following that step may lead us astray.

Algorithm 6.1 Riemannian Newton's method

Input: $x_0 \in \mathcal{M}$
For $k = 0, 1, 2, \ldots$
  Solve $\mathrm{Hess} f(x_k)[s_k] = -\mathrm{grad} f(x_k)$ for $s_k \in \mathrm{T}_{x_k}\mathcal{M}$
  $x_{k+1} = \mathrm{R}_{x_k}(s_k)$

Recall the various notions of local convergence rates introduced in Section 4.6. As we now show, Newton's method may converge locally quadratically (Definition 4.15). This is much faster than the typical linear local convergence of gradient descent, though we should bear in mind that (a) Newton steps are more expensive to compute, and (b) the global convergence behavior of Newton's method can be unwieldy. The proof below relies on the local contraction mapping theorem from Section 4.6.

Theorem 6.7. *Let $f \colon \mathcal{M} \to \mathbb{R}$ be smooth on a Riemannian manifold $\mathcal{M}$. If $x_\star \in \mathcal{M}$ is such that $\mathrm{grad} f(x_\star) = 0$ and $\mathrm{Hess} f(x_\star)$ is invertible, then there exists a neighborhood of $x_\star$ on $\mathcal{M}$ such that, for all x_0 in that neighborhood, Newton's method (Algorithm 6.1) generates an infinite sequence of iterates $x_0, x_1, x_2, \ldots$ which converges at least quadratically to $x_\star$.*

Proof. Let $\mathcal{U} = \{x \in \mathcal{M} \colon \det(\mathrm{Hess} f(x)) \neq 0\}$ be the subset of $\mathcal{M}$ where the Riemannian Hessian of f is invertible. This is a neighborhood of $x_\star$. Indeed, it contains $x_\star$ by assumption, and it is open because its complement is closed (the determinant of the Hessian is a continuous function). Newton's method iterates the map $F \colon \mathcal{U} \to \mathcal{M}$ given by

$$F(x) = \mathrm{R}_x(-V(x)) \qquad \text{with} \qquad V(x) = \mathrm{Hess} f(x)^{-1}[\mathrm{grad} f(x)].$$

At a critical point $x_\star$, we have $V(x_\star) = 0$ hence Lemma 4.21 provides the following for all $u \in \mathrm{T}_{x_\star}\mathcal{M}$:

$$\mathrm{D}F(x_\star)[u] = \mathrm{DR}(x_\star, 0)[u, -\mathrm{D}V(x_\star)[u]] = u - \mathrm{D}V(x_\star)[u]. \tag{6.2}$$

Moreover, Proposition 5.3 provides $\mathrm{D}V(x_\star)[u] = \nabla_u V$ with the Riemannian connection ∇, and

$$\begin{aligned} \nabla_u V &= -\left(\mathrm{Hess} f(x_\star)^{-1} \circ \nabla_u \mathrm{Hess} f \circ \mathrm{Hess} f(x_\star)^{-1}\right)[\mathrm{grad} f(x_\star)] \\ &\qquad + \mathrm{Hess} f(x_\star)^{-1}[\nabla_u \mathrm{grad} f] \\ &= u. \end{aligned} \tag{6.3}$$

In the intermediate expression above, the first term involves the covariant derivative of the Hessian tensor field: see Section 10.7. Its precise definition does not matter in the end since that linear operator is applied to the vector $\mathrm{grad} f(x_\star)$ which is zero. The second term evaluates to u since $\nabla_u \mathrm{grad} f = \mathrm{Hess} f(x_\star)[u]$ by definition. The above results combined establish that $\mathrm{D}F(x_\star) = 0$. All claims now follow from the local contraction mapping theorem (Theorem 4.19). □

Notice that the above theorem does *not* require the retraction to be second order. Essentially, this is due to Proposition 5.44 and the fact that $x_\star$ is a critical point.

From an optimization perspective, Theorem 6.7 is only beneficial if $\mathrm{Hess} f(x_\star)$ is positive definite. Indeed, by Proposition 6.2, critical points with an invertible Hessian which is not positive definite are certainly not local minimizers (in fact, they could be local *maximizers*). Yet, this theorem tells us Newton's method may converge to such points.

Partly because of this, given an initialization x_0, it is hard to predict where Newton's method may converge (if it converges at all). After all, the neighborhood in Theorem 6.7 may be arbitrarily small. To compensate for such issues, we add safeguards and other enhancements to this bare algorithm in Section 6.4.

Still, owing to its fast local convergence, Newton's method is relevant for favorable problems, or more generally to refine approximate solutions (for example, obtained through gradient descent). Thus, before moving on entirely, we discuss a practical algorithm to compute the Newton step s_k in the next section. This proves useful for the safeguarded algorithm as well.

Exercise 6.8. *Let $g(v) = \frac{1}{2}\langle v, Hv\rangle_x - \langle b, v\rangle_x$ be a quadratic function defined on a tangent space $\mathrm{T}_x\mathcal{M}$ of a Riemannian manifold. Assume H is positive definite* *on $\mathrm{T}_x\mathcal{M}$. Show that g has a unique minimizer which coincides with its unique critical point, that is, the solution of the linear system $Hv = b$. More generally, show that v minimizes g over a linear subspace of $\mathrm{T}_x\mathcal{M}$ if and only if* $\mathrm{grad}g(v)$ *is orthogonal to that subspace, and that this minimizer exists and is unique.*

Exercise 6.9. *Theorem 6.7 controls the local convergence of Newton's method to a zero of a gradient vector field. More generally, let $U \in \mathfrak{X}(\mathcal{M})$ be a smooth vector field (not necessarily the gradient of some function) on a manifold equipped with a connection ∇. Let $J_x : \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ denote the* Jacobian *of U at x, defined by $J_x(u) = \nabla_u U$ for all $u \in \mathrm{T}_x\mathcal{M}$. Newton's method for the vector field U iterates $x_{k+1} = F(x_k)$ with $F(x) = \mathrm{R}_x(-(J_x)^{-1}[U(x)])$, using some retraction* R*. Highlight the small changes needed in the proof of Theorem 6.7 to see that if $U(x_\star) = 0$ and $J_{x_\star}$ is invertible for some $x_\star \in \mathcal{M}$ then Newton's method initialized in a sufficiently small neighborhood of $x_\star$ converges to $x_\star$ at least quadratically.*

6.3 Computing Newton steps: conjugate gradients

Consider a cost function $f : \mathcal{M} \to \mathbb{R}$. Let $x \in \mathcal{M}$ be such that the Riemannian Hessian of f at x is positive definite. Then, to compute Newton's step at x we minimize a quadratic approximation of f lifted to the tangent space at x. Explicitly, we seek $v \in \mathrm{T}_x\mathcal{M}$ to minimize

$$g(v) = \frac{1}{2} \langle v, Hv \rangle_x - \langle b, v \rangle_x , \tag{6.4}$$

where we let $H = \mathrm{Hess}f(x)$ and $b = -\mathrm{grad}f(x)$ for short. Since H is positive definite by assumption, g has a unique minimizer which coincides with its unique critical point (Exercise 6.8). As

$$\mathrm{grad}g(v) = Hv - b, \tag{6.5}$$

that minimizer is the unique solution $s \in \mathrm{T}_x\mathcal{M}$ of the linear system $Hs = b$. Below, we assume $b \neq 0$ as otherwise the task is trivial.

Since H is a linear map on the linear space $\mathrm{T}_x\mathcal{M}$, we could in principle do the following: choose a basis for $\mathrm{T}_x\mathcal{M}$, represent H as a matrix and b as a vector with respect to that basis (see Exercise 3.9), and solve the resulting linear system in matrix form using any standard solver (e.g., based on LU, QR or Cholesky decomposition). However, that would be impractical because we seldom have access to a preferred basis of a tangent space (we would need to generate one), and computing the representation of H in that basis would be expensive.

It is far more fruitful to resort to a *matrix-free solver*, that is, an algorithm which only requires access to H as a linear map $v \mapsto Hv$. This is indeed what we usually have at our disposal in applications. Such solvers do not require access to H in matrix form.

The most famous matrix-free solver for systems with a positive definite map is the *conjugate gradients method* (CG): see Algorithm 6.2. Using exactly one computation of the form $v \mapsto Hv$ per iteration, this method generates three finite sequences of tangent vectors in $\mathrm{T}_x\mathcal{M}$:

1. $p_0, p_1, p_2, \ldots$ are linearly independent: they span a subspace of $\mathrm{T}_x\mathcal{M}$ of increasing dimension;
2. $v_0, v_1, v_2, \ldots$ are increasingly better approximations of the minimizer of g; and
3. $r_0, r_1, r_2, \ldots$ are the residues: the smaller r_n, the better v_n approximates the sought solution.

Algorithm 6.2 CG: Conjugate gradients on a tangent space

Input: positive definite map H on $\mathrm{T}_x\mathcal{M}$ and $b \in \mathrm{T}_x\mathcal{M}$, $b \neq 0$
Set $v_0 = 0, r_0 = b, p_0 = r_0$
For $n = 1, 2, \ldots$
  Compute Hp_{n-1} (this is the only call to H)
  $\alpha_n = \frac{\|r_{n-1}\|_x^2}{\langle p_{n-1}, Hp_{n-1}\rangle_x}$
  $v_n = v_{n-1} + \alpha_n p_{n-1}$
  $r_n = r_{n-1} - \alpha_n H p_{n-1}$
  If $r_n = 0$, **output** $s = v_n$: the solution of $Hs = b$
  $\beta_n = \frac{\|r_n\|_x^2}{\|r_{n-1}\|_x^2}$
  $p_n = r_n + \beta_n p_{n-1}$

We begin with a simple fact clarifying how the residue r_n informs us about the quality of v_n as a candidate minimizer for g.

Lemma 6.10. *If Algorithm 6.2 generates the vectors $v_0, \ldots, v_n$ and $r_0, \ldots, r_n$ before termination, then*

$$r_n = -\mathrm{grad}g(v_n) = b - Hv_n. \tag{6.6}$$

Thus, the algorithm terminates with v_n if and only if v_n minimizes g.

Proof. The proof is by induction. Clearly, $r_0 = b = -\mathrm{grad}g(v_0)$ since $v_0 = 0$. Assume $r_{n-1} = -\mathrm{grad}g(v_{n-1}) = b - Hv_{n-1}$. Then, by construction in Algorithm 6.2, we have $r_n = r_{n-1} - \alpha_n Hp_{n-1} = b - Hv_{n-1} - \alpha_n Hp_{n-1} = b - H(v_{n-1} + \alpha_n p_{n-1}) = b - Hv_n = -\mathrm{grad}g(v_n)$. The last part holds since the algorithm terminates with v_n if and only if $r_n = 0$. □

The key fact about the CG algorithm is that the vectors $p_0, p_1, \ldots$ are orthogonal with respect to a special inner product. The standard proof is by induction to show simultaneously Lemmas 6.11, 6.12 and 6.13 below: see for example [TB97, Thm. 38.1]. For exposition, we state Lemma 6.11 without proof, then we use it to prove the two subsequent lemmas. The intent is to clarify how the properties of $p_0, p_1, \ldots$ unlock all the other important features of CG.

Lemma 6.11. *If Algorithm 6.2 generates the vectors $p_0, p_1, \ldots, p_{n-1}$ before termination, then they are H*-conjugate*, that is, they are nonzero and*

$$\forall i \neq j, \qquad\qquad \langle p_i, Hp_j \rangle_x = 0.$$

In particular, $p_0, \ldots, p_{n-1}$ are linearly independent.

The above lemma states that the vectors $p_0, p_1, \ldots$ form an orthogonal basis with respect to the special inner product $\langle u, v \rangle_H = \langle u, Hv \rangle_x$. The fact that each p_{n-1} is nonzero also confirms that α_n is well defined since $\langle p_{n-1}, Hp_{n-1} \rangle_x$ is positive by positive definiteness of H.

The remarkable feature of H-conjugacy is that it makes minimizing g trivial in that basis. The CG method exploits this to build the sequence $v_0, v_1, \ldots$, as stated in the next lemma. A more constructive proof would start with the fact that any vector v in $\mathrm{span}(p_0, \ldots, p_{n-1})$ expands as $v = y_1 p_0 + \cdots + y_n p_{n-1}$ for some coefficients $y_1, \ldots, y_n$, then observing that $g(v)$ is a quadratic function of those coefficients with a *diagonal* Hessian matrix owing to H-conjugacy.

Lemma 6.12. *If Algorithm 6.2 generates the vectors $v_0, \ldots, v_n$ and $p_0, \ldots, p_{n-1}$ before termination, then*

$$v_n = \operatorname*{arg\,min}_{v \in \mathrm{span}(p_0, \ldots, p_{n-1})} g(v). \tag{6.7}$$

In particular, $\mathrm{grad}g(v_n)$ *is orthogonal to* $\mathrm{span}(p_0, \ldots, p_{n-1})$.

Proof. Unrolling the recursion for v_n in Algorithm 6.2 with $v_0 = 0$, we see that

$$v_n = v_{n-1} + \alpha_n p_{n-1} = \cdots = \alpha_1 p_0 + \cdots + \alpha_n p_{n-1}. \tag{6.8}$$

Thus, it is clear that v_n is in the span of $p_0, \ldots, p_{n-1}$. To show that v_n minimizes g in that span, we proceed by induction. For $n = 0$, we see that $v_0 = 0$ is valid since that is the only vector in the trivial span. Assume v_{n-1} minimizes g in the span of $p_0, \ldots, p_{n-2}$. Equivalently, $\mathrm{grad}g(v_{n-1})$ is orthogonal to $p_0, \ldots, p_{n-2}$ (Exercise 6.8). By Lemma 6.10, this means r_{n-1} is orthogonal to $p_0, \ldots, p_{n-2}$. For the same reason, to show that v_n minimizes g in the span of $p_0, \ldots, p_{n-1}$ we must show that r_n is orthogonal to those vectors. Consider the following for $i = 0, \ldots, i-1$:

$$\langle r_n, p_i \rangle_x = \langle r_{n-1} - \alpha_n H p_{n-1}, p_i \rangle_x = \langle r_{n-1}, p_i \rangle_x - \alpha_n \langle p_{n-1}, Hp_i \rangle_x .$$

For $i \leq n-2$, both terms on the right-most side are zero by induction hypothesis (for the firm term) and by H-conjugacy (for the second term, see Lemma 6.11). For $i = n-1$, the right-most side is zero by definition of α_n. Indeed,

$$\langle r_{n-1}, p_{n-1} \rangle_x = \langle r_{n-1}, r_{n-1} + \beta_{n-1} p_{n-2} \rangle_x = \langle r_{n-1}, r_{n-1} \rangle_x ,$$

where the last equality holds by orthogonality of r_{n-1} and p_{n-2}. $\square$

From the above lemma, we can infer the following:

1. We get steady progress in the sense that $g(v_n) \leq g(v_{n-1}) \leq \cdots \leq g(v_0)$ because each v_i is obtained by minimizing g over a subspace that contains all previous subspaces.
2. The algorithm terminates with the minimizer of g after at most $\dim \mathcal{M}$ iterations because if n reaches $\dim \mathcal{M}$ then $p_0, \ldots, p_{n-1}$ span the whole space $\mathrm{T}_x\mathcal{M}$ hence $r_n = -\mathrm{grad} g(v_n) = 0$.

To quantify how much progress we can hope to achieve as iterations progress, it is instructive to study the sequence of subspaces spanned by the H-conjugate directions. The following lemma gives a convenient characterization.

Lemma 6.13. *If Algorithm 6.2 generates the vectors* $p_0, \ldots, p_{n-1}$ *before termination, then*

$$\mathrm{span}(p_0, \ldots, p_{n-1}) = \mathrm{span}(b, Hb, H^2 b, \ldots, H^{n-1} b). \tag{6.9}$$

We call the right-hand side the Krylov subspace $\mathcal{K}^n$.

Proof. The proof is by induction on n. The identity (6.9) certainly holds for $n = 1$ since $p_0 = b$. Now assume (6.9) holds. Under that induction hypothesis, we aim to show that if the algorithm generates p_n then

$$\mathrm{span}(p_0, \ldots, p_n) = \mathcal{K}^{n+1}.$$

To show this equality, it is sufficient to show both of the following:

1. $\dim \mathcal{K}^{n+1} \leq \dim \mathrm{span}(p_0, \ldots, p_n)$, and
2. $\mathrm{span}(p_0, \ldots, p_n) \subseteq \mathcal{K}^{n+1}$.

We know from Lemma 6.11 that $p_0, \ldots, p_n$ are linearly independent, so that $\dim \mathrm{span}(p_0, \ldots, p_n) = n + 1$. Moreover, $\dim \mathcal{K}^{n+1} \leq n + 1$ because $\mathcal{K}^{n+1}$ is generated by $n + 1$ vectors. Thus, the first part is clear. For the second part, we already know by induction hypothesis that $p_0, \ldots, p_{n-1}$ are included in $\mathcal{K}^{n+1}$. It remains to show the same for p_n. To this end, consider the following where we use (6.6) for r_n:

$$p_n = r_n + \beta_n p_{n-1} = b - Hv_n + \beta_n p_{n-1}.$$

Of course, b is in $\mathcal{K}^{n+1}$. From (6.8) we know v_n is in $\mathrm{span}(p_0, \ldots, p_{n-1})$. By induction hypothesis, this means $v_n \in \mathcal{K}^n$. By definition, for all $u \in \mathcal{K}^n$ it holds that Hu is in $\mathcal{K}^{n+1}$. Thus, $Hv_n \in \mathcal{K}^{n+1}$. The induction hypothesis also provides $p_{n-1} \in \mathcal{K}^{n+1}$. Hence, p_n is in $\mathcal{K}^{n+1}$. □

Let $s \in \mathrm{T}_x\mathcal{M}$ be our target, that is, the unique solution to $Hs = b$. We wish to assess the size of the error vector $v_n - s$ at iteration n. We could do so in the norm $\|\cdot\|_x$ we already have on $\mathrm{T}_x\mathcal{M}$, but we choose to use the norm associated to the inner product $\langle u, v \rangle_H = \langle u, Hv \rangle_x$ instead, namely, the norm

$$\|u\|_H = \sqrt{\langle u, u \rangle_H} = \sqrt{\langle u, Hu \rangle_x}. \tag{6.10}$$

Explicitly, we aim to bound $\|v_n - s\|_H$. To this end, notice that the approximation error $\|v - s\|_H$ for a vector $v \in \mathrm{T}_x\mathcal{M}$ obeys:

$$\begin{aligned}\|v - s\|_H^2 &= \langle v - s, H(v - s)\rangle_x \\ &= \langle v, Hv\rangle_x - \langle s, Hv\rangle_x - \langle v, Hs\rangle_x + \langle s, Hs\rangle_x \\ &= \langle v, Hv\rangle_x - 2\langle v, b\rangle_x + \langle s, Hs\rangle_x \\ &= 2g(v) + \langle s, Hs\rangle_x, \end{aligned} \tag{6.11}$$

where we used that H is self-adjoint and $Hs = b$. Since the last term on the right-hand side is independent of v, we conclude that minimizing $\|v - s\|_H$ over a subset of $\mathrm{T}_x\mathcal{M}$ is *equivalent* to minimizing $g(v)$ over that same subset. Therefore, Lemma 6.12 tells us that v_n is the vector in $\mathrm{span}(p_0, \ldots, p_{n-1})$ which minimizes $\|v_n - s\|_H$: this is why the H-norm is particularly relevant. Further combining with Lemma 6.13 reveals the following key fact about the CG algorithm:

$$v_n = \operatorname*{arg\,min}_{v \in \mathcal{K}^n} \|v - s\|_H, \tag{6.12}$$

where $\mathcal{K}^n$ is the Krylov subspace. Let us reformulate this once more: Lemmas 6.12 and 6.13 combined with the observation (6.11) reveal that

$$\begin{aligned} v_n &= a_0 b + a_1 Hb + a_2 H^2 b + \cdots + a_{n-1} H^{n-1} b \\ &= \left(a_0 I + a_1 H + \cdots + a_{n-1} H^{n-1}\right) b \end{aligned} \tag{6.13}$$

with coefficients $a_0, \ldots, a_{n-1} \in \mathbb{R}$ such that $\|v_n - s\|_H$ is minimized. Substituting Hs for b in (6.13), we deduce that the error vector at iteration n is

$$\begin{aligned} v_n - s &= \left(a_0 I + a_1 H + \cdots + a_{n-1} H^{n-1}\right) Hs - s \\ &= \left(a_0 H + a_1 H^2 + \cdots + a_{n-1} H^n - I\right) s. \end{aligned} \tag{6.14}$$

The parenthesized expression on the right-hand side is a polynomial in H. Specifically, it is $q_n(H)$ with the polynomial

$$q_n(z) = -1 + a_0 z + a_1 z^2 + \cdots + a_{n-1} z^n. \tag{6.15}$$

Thus,

$$v_n - s = q_n(H)s. \tag{6.16}$$

The polynomial q_n has degree at most n and satisfies $q_n(0) = -1$. Let $\mathcal{Q}_n$ denote the set of such polynomials. Since every polynomial in $\mathcal{Q}_n$ can be written in the form (6.15) for some choice of coefficients $a_0, \ldots, a_{n-1}$, and since the CG method generates v_n such that $\|v_n - s\|_H$ is minimized, it follows that the CG method guarantees

$$\|v_n - s\|_H = \min_{q \in \mathcal{Q}_n} \|q(H)s\|_H. \tag{6.17}$$

To turn this conclusion into an interpretable bound on the error after n CG iterations, we now investigate the effect of applying a polynomial to the map H. To this end, let $u_1, \ldots, u_d$ be a basis of eigenvectors of H, orthonormal with respect to $\langle \cdot, \cdot \rangle_x$ (we write $d = \dim \mathcal{M}$ for short). These exist since H is self-adjoint. Moreover, let $\lambda_1, \ldots, \lambda_d$ be associated eigenvalues: $Hu_i = \lambda_i u_i$. The unknown vector s expands as

$$s = \sum_{i=1}^{d} \langle u_i, s \rangle_x u_i.$$

Hence, t applications of H to this vector yield:

$$H^t s = \sum_{i=1}^{d} \lambda_i^t \langle u_i, s \rangle_x u_i.$$

More generally, applying $q(H)$ to s for some polynomial q yields:

$$q(H)s = \sum_{i=1}^{d} q(\lambda_i) \langle u_i, s \rangle_x u_i.$$

We conclude that, for any polynomial q,

$$\frac{\|q(H)s\|_H^2}{\|s\|_H^2} = \frac{\langle q(H)s, Hq(H)s \rangle_x}{\langle s, Hs \rangle_x} = \frac{\sum_{i=1}^{d} q(\lambda_i)^2 \lambda_i \langle u_i, s \rangle_x^2}{\sum_{i=1}^{d} \lambda_i \langle u_i, s \rangle_x^2} \leq \max_{1 \leq i \leq d} q(\lambda_i)^2,$$

where the inequality is due to positivity of the eigenvalues. Combined with (6.17), it follows that

$$\|v_n - s\|_H \leq \|s\|_H \cdot \min_{q \in \mathcal{Q}_n} \max_{1 \leq i \leq d} |q(\lambda_i)|. \tag{6.18}$$

In words: the relative error after n iterations, in the H-norm, is controlled by the existence of a polynomial q in $\mathcal{Q}_n$ with small absolute value when evaluated at each of the eigenvalues of H.

Based on these considerations, it follows easily that if H has only $k \leq d$ distinct eigenvalues then CG terminates in k iterations. To verify this, it suffices to construct a polynomial q of degree k with single roots at the distinct eigenvalues and such that $q(0) = -1$. More generally, if $\lambda_{\min}$ and $\lambda_{\max}$ denote the smallest and largest eigenvalues of H, then $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}}$ is the *condition number* of H, and it can be shown that

$$\|v_n - s\|_H \leq \|s\|_H \cdot 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^n \leq \|s\|_H \cdot 2e^{-n/\sqrt{\kappa}}, \tag{6.19}$$

so that the error decreases exponentially fast as CG iterates (hence linear convergence as per Definition 4.14). This is shown by exhibiting an appropriate polynomial q with small absolute value over the whole interval $[\lambda_{\min}, \lambda_{\max}]$: see [TB97, Thm. 38.5] for a classical construction based on Chebyshev polynomials, as illustrated in Figure 6.1.

We close with a few comments.

1. That CG terminates in at most $\dim \mathcal{M}$ iterations is of little practical relevance, in part because numerical round-off errors typically prevent this (specifically,

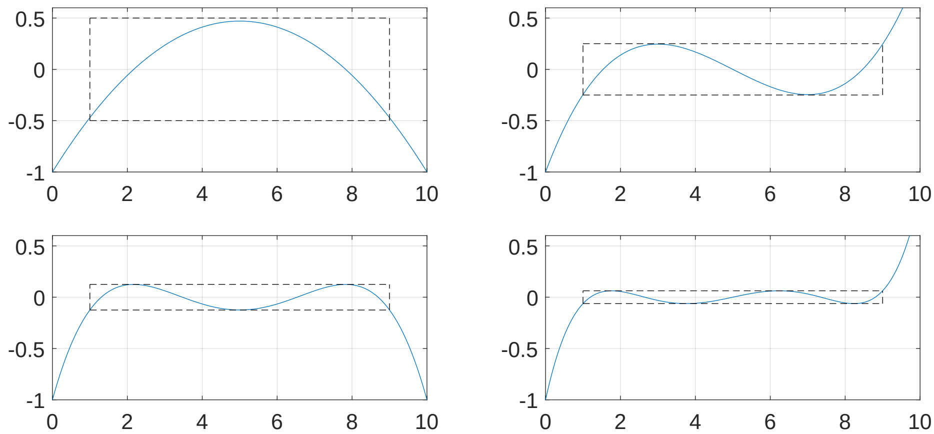


Figure 6.1 The convergence rate of CG is governed by the existence of special polynomials, as shown by (6.18). For illustration of (6.19), let $[\lambda_{\min}, \lambda_{\max}] = [1, 9]$ and $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}} = 9$. For $n = 2, 3, 4, 5$, the plots show the polynomial $q_n(x) = -\frac{T_n(\ell(x))}{T_n(\ell(0))}$, where $\ell(x) = \frac{2x - \lambda_{\max} - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}$ maps $[\lambda_{\min}, \lambda_{\max}]$ to $[-1, 1]$ and $T_n(x) = \cos(n \arccos(x))$ defines the Chebyshev polynomial of the first kind and of degree n on $[-1, 1]$. One can check that q_n is in $\mathcal{Q}_n$ (that is, q_n is a polynomial of degree n with $q_n(0) = -1$) and, as depicted, that $|q_n(x)| \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n$ for all $x \in [\lambda_{\min}, \lambda_{\max}]$.

because numerically the vectors p_i are not exactly H-conjugate). However, the progressive improvement of the iterates v_n as predicted by (6.19) is borne out empirically, and the role of the condition number κ is indeed critical. In practice, CG is terminated after a set number of iterations, or when a target relative tolerance is met. For example, we may replace the stopping criterion $r_n = 0$ with $\|r_n\|_x \leq \varepsilon_{\text{tolerance}} \|b\|_x$.

2. Reconsidering the bigger picture, we want to keep in mind that the goal is to minimize $f(x)$: solving the linear system which arises in Newton's method is only a means to an end. Since CG can produce adequate approximate solutions to the linear system in few iterations, it is often beneficial to terminate CG early and proceed with an approximate Newton step: this is at the heart of the developments regarding the trust-region method in the next section.
3. In practice, $\mathrm{Hess} f(x)$ may not be positive definite. If such is the case, we ought to be able to detect it. For example, the inner product $\langle p_{n-1}, H p_{n-1} \rangle_x$ may turn out to be negative. In the trust-region method, such events are monitored and appropriate actions are taken.
4. Regarding numerical errors again, in Algorithm 6.2, the vectors p_i may slowly build-up a non-tangent component (even though this cannot happen mathematically). Experience shows that it is sometimes beneficial to ensure p_{n-1} is tangent (up to machine precision) before computing $H p_{n-1}$. For embedded submanifolds, this can be done through orthogonal projection for example.

Doing this at every iteration appears to be sufficient to ensure the other sequences (namely, r_i and v_i) also remain numerically tangent.

Exercise 6.14. *An alternative to CG is to run gradient descent on $g(v)$ (6.4) in the tangent space. Since g is a quadratic, it is easy to check that it has L-Lipschitz continuous gradient with $L = \lambda_{\max}(H)$. Show that running $v_{n+1} = v_n - \frac{1}{L}\mathrm{grad}g(v_n)$ with $v_0 = 0$ leads to $\|v_n - s\|_x \leq e^{-n/\kappa}\|s\|_x$ where $\kappa = \frac{\lambda_{\max}(H)}{\lambda_{\min}(H)}$. Contrast this with the role of κ in (6.19) for CG.*

6.4 Riemannian trust regions

The trust-region method addresses the fundamental shortcomings of Newton's method, while preserving its fast local convergence properties under favorable circumstances. The premise is the same: around a point x, we approximate the pullback $f \circ \mathrm{R}_x$ with a simpler model in the tangent space:

$$f(\mathrm{R}_x(s)) \approx m_x(s) = f(x) + \langle \mathrm{grad}f(x), s \rangle_x + \frac{1}{2} \langle H_x(s), s \rangle_x .$$

Here, H_x is allowed to be *any* self-adjoint linear map on $\mathrm{T}_x\mathcal{M}$ (in fact, we will relax this even further). Of course, the model is a better match for $f \circ \mathrm{R}_x$ if H_x is chosen to be the Hessian of $f \circ \mathrm{R}_x$. From Proposition 5.44, we also know that, close to critical points, this is essentially the same as $\mathrm{Hess}f(x)$ (exactly the same for second-order retractions).

In a key departure from Newton's method however, we do not select the step by blindly jumping to the critical point of the model (which might not even exist). Rather, we insist on reducing the value of m_x. Moreover, since the model is only a local approximation of the pullback, we only *trust* it in a ball around the origin in the tangent space: the *trust region*. Specifically, at the iterate x_k, we define the model

$$m_k(s) = f(x_k) + \langle \mathrm{grad}f(x_k), s \rangle_{x_k} + \frac{1}{2} \langle H_k(s), s \rangle_{x_k} \tag{6.20}$$

for some map $H_k \colon \mathrm{T}_{x_k}\mathcal{M} \to \mathrm{T}_{x_k}\mathcal{M}$ to be specified, and we pick the tentative next iterate x_k^+ as $\mathrm{R}_{x_k}(s_k)$ such that the step s_k approximately solves the *trust-region subproblem*:

$$\min_{s \in \mathrm{T}_{x_k}\mathcal{M}} m_k(s) \quad \text{subject to} \quad \|s\|_{x_k} \leq \Delta_k, \tag{6.21}$$

where Δ_k is the radius of the trust region at iteration k. Specific requirements are discussed later, but at the very least $m_k(s_k)$ should be smaller than $m_k(0)$. The step is accepted ($x_{k+1} = x_k^+$) or rejected ($x_{k+1} = x_k$) based on the performance of x_k^+ as judged by the actual cost function f, compared to the expected improvement as predicted by the model. Depending on how the two compare, the trust-region radius may also be adapted. See Algorithm 6.3 for details: it is called the *Riemannian trust-region method* (RTR).

Algorithm 6.3 RTR: the Riemannian trust-region method

Parameters: maximum radius $\bar{\Delta} > 0$, threshold $\rho' \in (0, 1/4)$
Input: $x_0 \in \mathcal{M}$, $\Delta_0 \in (0, \bar{\Delta}]$
For $k = 0, 1, 2, \ldots$

Pick a map $H_k \colon \mathrm{T}_{x_k}\mathcal{M} \to \mathrm{T}_{x_k}\mathcal{M}$ to define m_k (6.20).
Approximately solve the subproblem (6.21), yielding s_k.
The tentative next iterate is $x_k^+ = \mathrm{R}_{x_k}(s_k)$.
Compute the ratio of actual to model improvement:

$$\rho_k = \frac{f(x_k) - f(x_k^+)}{m_k(0) - m_k(s_k)}. \tag{6.22}$$

Accept or reject the tentative next iterate:

$$x_{k+1} = \begin{cases} x_k^+ & \text{if } \rho_k > \rho' \text{ (accept)}, \\ x_k & \text{otherwise (reject)}. \end{cases} \tag{6.23}$$

Update the trust-region radius:

$$\Delta_{k+1} = \begin{cases} \frac{1}{4}\Delta_k & \text{if } \rho_k < \frac{1}{4}, \\ \min(2\Delta_k, \bar{\Delta}) & \text{if } \rho_k > \frac{3}{4} \text{ and } \|s_k\|_{x_k} = \Delta_k, \\ \Delta_k & \text{otherwise.} \end{cases} \tag{6.24}$$

Running RTR, we hope to find a minimizer of f, but that is too much to ask in general. More realistically, we hope to find a point $x = x_k$ such that

$$\|\mathrm{grad} f(x)\|_x \leq \varepsilon_g \qquad \text{and} \qquad \mathrm{Hess} f(x) \succeq -\varepsilon_H \,\mathrm{Id}, \tag{6.25}$$

where Id is the identity map on $\mathrm{T}_x\mathcal{M}$, and ε_H may be infinite if we only care about first-order optimality conditions. One of the main goals of this chapter is to show that, regardless of initialization, under suitable assumptions, RTR provides such a point in a bounded number of iterations.

Of course, to provide such guarantees we must specify conditions on the maps H_k, requirements on how well the trust-region subproblems are to be solved, and regularity conditions on the pullbacks $f \circ \mathrm{R}_{x_k}$. We do this in the subsections below. In Section 6.7, we discuss more restrictive settings which make it possible to verify all assumptions discussed below in a straightforward manner.

6.4.1 Conditions on the model

The model m_{x_k} is determined by a choice of map H_k from $\mathrm{T}_{x_k}\mathcal{M}$ to itself. The simpler this map, the easier it may be to solve the trust-region subproblem (6.21). In choosing H_k, we aim to strike a balance between model accuracy, computational efficiency, and convenience. With the goal (6.25) determined by $\varepsilon_g, \varepsilon_H > 0$ in mind, we introduce the following requirements.

For iterations with large gradient, the conditions are particularly mild. In essence, this is because for such iterations the main focus is on reducing the gradient norm, which can be done with any first-order accurate model.

A6.1. *For all iterations k such that $\|\mathrm{grad} f(x_k)\|_{x_k} > \varepsilon_g$, we require that:*

1. *H_k is radially linear, that is:*

$$\forall s \in \mathrm{T}_{x_k}\mathcal{M}, \alpha \geq 0, \qquad H_k(\alpha s) = \alpha H_k(s); \qquad \textit{and} \tag{6.26}$$

2. *H_k is uniformly bounded, that is, there exists $c_0 \geq 0$, independent of k, such that*

$$\forall s \in \mathrm{T}_{x_k}\mathcal{M}, \qquad \left|\langle s, H_k(s)\rangle_{x_k}\right| \leq c_0\|s\|_{x_k}^2. \tag{6.27}$$

(We can gain insight into the latter through Corollary 10.47.)

An extreme case consists in selecting $H_k = L \cdot \mathrm{Id}$ for some $L > 0$. This is convenient, computationally inexpensive, and allows us to solve the subproblem (6.21) in closed form: RTR then takes gradient steps. However, the model does not capture second-order information at all, which may slow down convergence. Alternatively, a convenient radially linear (but not linear) map H_k can be obtained from finite difference approximations of the Hessian using gradients, see Section 10.6. Naturally, if it is practical to use the Hessian of f (or that of $f \circ \mathrm{R}_{x_k}$) itself for H_k, then the enhanced accuracy of the model is a strong incentive to do so.

For iterations with small gradient, if there is a desire to reach approximate satisfaction of second-order necessary optimality conditions ($\varepsilon_H < \infty$), we need the model to be (at least approximately) second-order accurate.

A6.2. *For all iterations k such that $\|\mathrm{grad} f(x_k)\|_{x_k} \leq \varepsilon_g$, we require H_k to be linear and self-adjoint. Furthermore, there must exist $c_1 \geq 0$ independent of k such that*

$$\|\mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0) - H_k\| \leq \frac{c_1\Delta_k}{3}, \tag{6.28}$$

where $\|\cdot\|$ denotes the operator norm of a self-adjoint map, that is, the largest magnitude of any of its eigenvalues.

The convergence results below guarantee H_k is, eventually, almost positive semidefinite. This is only meaningful if H_k is close to $\mathrm{Hess} f(x_k)$ in operator norm. In turn, $\mathrm{Hess} f(x_k)$ is equal to $\mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0)$ if the retraction is second order (and for a general retraction they are close if x_k is nearly critical): see Propositions 5.44 and 5.45 (and Exercise 10.73 for first-order retractions). Overall, the conceptually simplest situation is that for which we use a second-order retraction and a quadratically-accurate model, in which case:

$$H_k = \mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0) = \mathrm{Hess} f(x_k). \tag{6.29}$$

Then, A6.2 holds with $c_1 = 0$.

6.4.2 Requirements on solving the subproblem

Once a model is selected through a choice of map H_k, the key (and typically most computationally expensive) part of an iteration of RTR is to solve the trust-region subproblem (6.21) approximately, producing a step s_k. Numerous efficient algorithms have been proposed over the past few decades: we detail one that is particularly well suited to the Riemannian case in Section 6.5. For now, we merely specify minimum requirements on how well the task ought to be solved.

We require sufficient decrease in the value of the *model*, similar to but different from the analysis of Riemannian gradient descent in Section 4.3 which required sufficient decrease in the value of the actual cost function. So long as first-order criticality has not been approximately attained, sufficient decrease is defined with respect to the gradient norm. The subproblem solver we discuss in Section 6.5 satisfies the assumption below, see Exercise 6.26.

A6.3. *There exists $c_2 > 0$ such that, for all k with $\|\mathrm{grad} f(x_k)\|_{x_k} > \varepsilon_g$, the step s_k satisfies*

$$m_k(0) - m_k(s_k) \geq c_2 \min\left(\Delta_k, \frac{\|\mathrm{grad} f(x_k)\|_{x_k}}{c_0}\right) \|\mathrm{grad} f(x_k)\|_{x_k}, \tag{6.30}$$

where c_0 is the constant in A6.1.

This condition is easily satisfied by computing the so-called *Cauchy point*: the minimizer of the subproblem when restricted to the negative gradient direction. Given the gradient at x_k, it can be computed with one call to H_k. It is an exercise to establish the following lemma.

Lemma 6.15. *Let $g_k = \mathrm{grad} f(x_k)$ for convenience. The Cauchy point is the tangent vector $s_k^C = -t\mathrm{grad} f(x_k)$ with $t \geq 0$ such that $m_k(s_k^C)$ is minimal under the constraint $\|s_k^C\|_{x_k} \leq \Delta_k$. Under A6.1, we can compute the corresponding optimal t explicitly as:*

$$t = \begin{cases} \min\left(\frac{\|g_k\|_{x_k}^2}{\langle g_k, H_k(g_k)\rangle_{x_k}}, \frac{\Delta_k}{\|g_k\|_{x_k}}\right) & \textit{if } \langle g_k, H_k(g_k)\rangle_{x_k} > 0, \\ \frac{\Delta_k}{\|g_k\|_{x_k}} & \textit{otherwise.} \end{cases}$$

Furthermore, setting $s_k = s_k^C$ in RTR satisfies A6.3 with $c_2 = \frac{1}{2}$.

Once the gradient is small and if $\varepsilon_H < \infty$, it becomes necessary to focus on second-order optimality conditions.

A6.4. *There exists $c_3 > 0$ such that, for all k with $\|\mathrm{grad} f(x_k)\|_{x_k} \leq \varepsilon_g$ and $\lambda_{\min}(H_k) < -\varepsilon_H$, the step s_k satisfies*

$$m_k(0) - m_k(s_k) \geq c_3 \Delta_k^2 \varepsilon_H. \tag{6.31}$$

(Note that H_k has real eigenvalues owing to A6.2.)

This condition too can be satisfied with explicit, finite procedures by computing *eigensteps*: moving up to the boundary of the trust region along a direction which certifies that the smallest eigenvalue of H_k is strictly smaller than $-\varepsilon_H$. Proving the next lemma is an exercise.

Lemma 6.16. *Under A6.2, if $\lambda_{\min}(H_k) < -\varepsilon_H$ then there exists a tangent vector $u \in \mathrm{T}_{x_k}\mathcal{M}$ satisfying*

$$\|u\|_{x_k} = 1, \qquad \langle \mathrm{grad} f(x_k), u \rangle_{x_k} \leq 0, \qquad \text{and} \qquad \langle u, H_k(u) \rangle_{x_k} < -\varepsilon_H.$$

Setting $s_k = \Delta_k u$ (called an eigenstep) in RTR satisfies A6.4 with $c_3 = \frac{1}{2}$.

Eigensteps are rarely (if ever) computed in practice. More pragmatically, the existence of eigensteps serves to show that a global minimizer of the subproblem also satisfies A6.4.

Corollary 6.17. *If H_k is linear and self-adjoint for every iteration k, then setting s_k to be a global minimizer of the subproblem* (6.21) *at every iteration satisfies both A6.3 and A6.4 with $c_2 = c_3 = \frac{1}{2}$. Likewise, setting s_k to achieve at least a fraction $\alpha \in (0, 1]$ of the optimal model decrease satisfies the assumptions with $c_2 = c_3 = \frac{\alpha}{2}$.*

Exercise 6.18. *Give a proof of Lemma 6.15.*

Exercise 6.19. *Give a proof of Lemma 6.16.*

6.4.3 Regularity conditions

As we did when analyzing the Riemannian gradient method, we require that the cost function be lower-bounded.

A6.5. *There exists $f_{\mathrm{low}} \in \mathbb{R}$ such that $f(x_k) \geq f_{\mathrm{low}}$ for all iterates $x_0, x_1, \ldots$*

Likewise, we still require a first-order, Lipschitz-type condition on the pullbacks of f for the given retraction R. The set S_g is specified later on.

A6.6. *For a given subset S_g of the tangent bundle $\mathrm{T}\mathcal{M}$, there exists a constant $L_g > 0$ such that, for all $(x, s) \in S_g$,*

$$f(\mathrm{R}_x(s)) \leq f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{L_g}{2} \|s\|_x^2.$$

In addition to these, we now also include a second-order Lipschitz-type condition. When $\mathcal{M}$ is a Euclidean space and $\mathrm{R}_x(s) = x+s$, this one holds in particular if $\mathrm{Hess} f$ is Lipschitz continuous with constant L_H. The set S_H is specified later on; it is empty if $\varepsilon_H = \infty$.

A6.7. *For a given subset S_H of the tangent bundle $\mathrm{T}\mathcal{M}$, there exists a constant $L_H > 0$ such that, for all $(x, s) \in S_H$,*

$$f(\mathrm{R}_x(s)) \leq f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{1}{2} \langle s, \mathrm{Hess}(f \circ \mathrm{R}_x)(0)[s] \rangle_x + \frac{L_H}{6} \|s\|_x^3.$$

We note that, in particular, the sets S_g and S_H will not be required to contain any tangent vectors of norm larger than $\bar{\Delta}$, since this is the largest trust-region radius ever considered. This is useful notably when the retraction is not globally defined (or well behaved). In addition, all root points of elements in S_g and S_H are iterates $x_0, x_1, x_2, \ldots$ generated by RTR. This can be helpful when the iterates are easily shown to lie in a compact subset of $\mathcal{M}$, for example if the sublevel sets of f are compact, as then A6.6 and A6.7 hold by Lemma 10.57: see Section 6.7.

We gain further insight into the regularity assumptions from Corollary 10.54 and Exercise 10.58 (for A6.6) and from Corollary 10.56 and Exercise 10.87 (for A6.7).

6.4.4 Iteration complexity

Given tolerances $\varepsilon_g > 0$ and $\varepsilon_H > 0$, we show that RTR produces an iterate x_k which satisfies the following termination conditions in a bounded number of iterations:

$$\|\mathrm{grad} f(x_k)\|_{x_k} \leq \varepsilon_g \qquad \text{and} \qquad \lambda_{\min}(H_k) \geq -\varepsilon_H. \tag{6.32}$$

We stress that ε_H may be set to infinity if only first-order optimality conditions are targeted. Accordingly, we separate the theorem statement in two scenarios. See the discussion around eq. (6.29) to relate the guarantees on H_k to the eigenvalues of $\mathrm{Hess} f(x_k)$.

Following the standard proofs for trust regions in Euclidean space, the analysis is based on three supporting lemmas which we state and prove below. In a nutshell, they show that:

1. The trust-region radius cannot become arbitrarily small. Essentially, this is because regularity of the cost function ensures the model m_k is sufficiently accurate for small steps, which ultimately ensures step acceptance. This prevents trust-region radius reductions beyond a certain point.
2. Combining the latter with our sufficient decrease assumptions, successful steps initiated from iterates with large gradient produce large decrease in the cost function value (and similarly at iterates where H_k has a "large" negative eigenvalue). Yet, the total amount of cost decrease is bounded by $f(x_0) - f_{\mathrm{low}}$, so that there cannot be arbitrarily many successful steps.
3. The number of successful steps as above is at least a fraction of the total number of iterations, because a large number of consecutive failures would eventually violate the fact that the trust-region radius is lower-bounded: every so often, there must be a successful step.

We state the main theorem—the proof comes later in this section.

Theorem 6.20. *Let $S = \{(x_0, s_0), (x_1, s_1), \ldots\}$ be the pairs of iterates and tentative steps generated by RTR under A6.1, A6.2, A6.3, A6.4 and A6.5. Further* *assume A6.6 and A6.7 hold with constants L_g and L_H on the sets*

$$S_g = \{(x_k, s_k) \in S : \|\mathrm{grad} f(x_k)\|_{x_k} > \varepsilon_g\}, \text{ and}$$
$$S_H = \{(x_k, s_k) \in S : \|\mathrm{grad} f(x_k)\|_{x_k} \le \varepsilon_g \text{ and } \lambda_{\min}(H_k) < -\varepsilon_H\}.$$

Define

$$\lambda_g = \frac{1}{4}\min\left(\frac{1}{c_0}, \frac{c_2}{L_g + c_0}\right) \qquad \text{and} \qquad \lambda_H = \frac{3}{4}\frac{c_3}{L_H + c_1}. \tag{6.33}$$

We consider two scenarios, depending on whether second-order optimality conditions are targeted or not:

1. *If $\varepsilon_g \le \frac{\Delta_0}{\lambda_g}$ and $\varepsilon_H = \infty$, there exists t with $\|\mathrm{grad} f(x_t)\|_{x_t} \le \varepsilon_g$ and*
$$t \le \frac{3}{2}\frac{f(x_0) - f_{\mathrm{low}}}{\rho' c_2 \lambda_g}\frac{1}{\varepsilon_g^2} + \frac{1}{2}\log_2\left(\frac{\Delta_0}{\lambda_g \varepsilon_g}\right) = O\left(\frac{1}{\varepsilon_g^2}\right). \tag{6.34}$$
(In this scenario, A6.2, A6.4 and A6.7 are irrelevant.)
2. *If $\varepsilon_g \le \frac{\Delta_0}{\lambda_g}$, $\varepsilon_g \le \frac{c_2}{c_3}\frac{\lambda_H}{\lambda_g^2}$ and $\varepsilon_H < \frac{c_2}{c_3}\frac{1}{\lambda_g}$, there exists $t' \ge t$ such that $\|\mathrm{grad} f(x_{t'})\|_{x_{t'}} \le \varepsilon_g$ and $\lambda_{\min}(H_{t'}) \ge -\varepsilon_H$ with*
$$t' \le \frac{3}{2}\frac{f(x_0) - f_{\mathrm{low}}}{\rho' c_3 \lambda^2}\frac{1}{\varepsilon^2 \varepsilon_H} + \frac{1}{2}\log_2\left(\frac{\Delta_0}{\lambda\varepsilon}\right) = O\left(\frac{1}{\varepsilon^2\varepsilon_H}\right), \tag{6.35}$$
where $(\lambda, \varepsilon) = (\lambda_g, \varepsilon_g)$ if $\lambda_g\varepsilon_g \le \lambda_H\varepsilon_H$, and $(\lambda, \varepsilon) = (\lambda_H, \varepsilon_H)$ otherwise.

Since the algorithm is a descent method, $f(x_{t'}) \le f(x_t) \le f(x_0)$.

To build a proof of the theorem above, we work through a sequence of three lemmas. This first one lower-bounds the trust-region radius.

Lemma 6.21. *Under the assumptions of Theorem 6.20, let $x_0, \ldots, x_n$ be iterates generated by RTR. If none of them satisfy the termination conditions* (6.32)*, then*

$$\Delta_k \ge \min(\Delta_0, \lambda_g\varepsilon_g, \lambda_H\varepsilon_H) \tag{6.36}$$

for $k = 0, \ldots, n$.

Proof. Our goal is to show that if Δ_k is small, then ρ_k must be large. By the mechanism of RTR (specifically, eq. (6.24)), this guarantees Δ_k cannot decrease further. By definition of ρ_k (6.22), using $m_k(0) = f(x_k)$,

$$1 - \rho_k = 1 - \frac{f(x_k) - f(\mathrm{R}_{x_k}(s_k))}{m_k(0) - m_k(s_k)} = \frac{f(\mathrm{R}_{x_k}(s_k)) - m_k(s_k)}{m_k(0) - m_k(s_k)}.$$

Consider an iteration k such that $\|\mathrm{grad} f(x_k)\|_{x_k} > \varepsilon_g$. Then, the numerator is upper-bounded owing to A6.6 and A6.1:

$$\begin{aligned} &f(\mathrm{R}_{x_k}(s_k)) - m_k(s_k) \\ &\quad = f(\mathrm{R}_{x_k}(s_k)) - f(x_k) - \langle \mathrm{grad} f(x_k), s_k\rangle_{x_k} - \frac{1}{2}\langle H_k(s_k), s_k\rangle_{x_k} \\ &\quad \le \frac{L_g + c_0}{2}\|s_k\|_{x_k}^2. \end{aligned}$$

Furthermore, the denominator is lower-bounded by A6.3:

$$m_k(0) - m_k(s_k) \geq c_2 \min\left(\Delta_k, \frac{\varepsilon_g}{c_0}\right)\varepsilon_g.$$

Hence, using $\|s_k\|_{x_k} \leq \Delta_k$, we have

$$1 - \rho_k \leq \frac{1}{2}\frac{L_g + c_0}{c_2\varepsilon_g}\frac{\Delta_k^2}{\min\left(\Delta_k, \frac{\varepsilon_g}{c_0}\right)}.$$

If $\Delta_k \leq \frac{\varepsilon_g}{c_0}$, the last factor is equal to Δ_k. If additionally $\Delta_k \leq \frac{c_2\varepsilon_g}{L_g+c_0}$, then $1 - \rho_k \leq \frac{1}{2}$. Using (6.33), we summarize this as: if $\Delta_k \leq 4\lambda_g\varepsilon_g$, then $\rho_k \geq \frac{1}{2}$ and the mechanism of RTR implies $\Delta_{k+1} \geq \Delta_k$.

Now, consider k such that $\|\mathrm{grad}f(x_k)\|_{x_k} \leq \varepsilon_g$ and $\lambda_{\min}(H_k) < -\varepsilon_H$. Then, the numerator is upper-bounded by A6.7, A6.2 and $\|s_k\|_{x_k} \leq \Delta_k$:

$$\begin{aligned} & f(\mathrm{R}_{x_k}(s_k)) - m_k(s_k) \\ &\quad = f(\mathrm{R}_{x_k}(s_k)) - f(x_k) - \langle \mathrm{grad}f(x_k), s_k\rangle_{x_k} - \frac{1}{2}\langle \mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0)[s_k], s_k\rangle_{x_k} \\ &\qquad + \frac{1}{2}\langle (\mathrm{Hess}(f \circ \mathrm{R}_{x_k}(0) - H_k)[s_k], s_k\rangle_{x_k} \\ &\quad \leq \frac{L_H + c_1}{6}\Delta_k^3, \end{aligned}$$

and the denominator is lower-bounded by A6.4:

$$m_k(0) - m_k(s_k) \geq c_3\Delta_k^2\varepsilon_H. \tag{6.37}$$

Combining, we get

$$1 - \rho_k \leq \frac{L_H + c_1}{6c_3\varepsilon_H}\Delta_k.$$

Again, considering (6.33), we find that if $\Delta_k \leq 4\lambda_H\varepsilon_H$, then $\rho_k \geq \frac{1}{2}$ and as a result $\Delta_{k+1} \geq \Delta_k$.

We have established that if $\Delta_k \leq 4\min(\lambda_g\varepsilon_g, \lambda_H\varepsilon_H)$ then $\Delta_{k+1} \geq \Delta_k$. Since RTR does not reduce the radius by more than a factor four per iteration, the claim follows. □

The second lemma upper-bounds the total number of successful (that is, accepted) steps before termination conditions are met.

Lemma 6.22. *Under the assumptions of Theorem 6.20, let $x_0, \ldots, x_n$ be iterates generated by RTR. If none of them satisfy the termination conditions* (6.32), *define the set of* successful steps *among those as*

$$S_n = \left\{k \in \{0, \ldots, n\} : \rho_k > \rho'\right\},$$

and let U_n designate the unsuccessful steps, *so that S_n and U_n form a partition* *of* $\{0, \ldots, n\}$*. In the first scenario of Theorem 6.20, the number of successful steps is bounded as*

$$|S_n| \leq \frac{f(x_0) - f_{\text{low}}}{\rho' c_2} \frac{1}{\lambda_g \varepsilon_g^2}. \tag{6.38}$$

Similarly, in the second scenario we have

$$|S_n| \leq \frac{f(x_0) - f_{\text{low}}}{\rho' c_3} \frac{1}{\min(\lambda_g \varepsilon_g, \lambda_H \varepsilon_H)^2 \varepsilon_H}. \tag{6.39}$$

Proof. Clearly, if $k \in U_n$, then $f(x_k) = f(x_{k+1})$. On the other hand, if $k \in S_n$, then the definition of ρ_k (6.22) combined with A6.3 and A6.4 ensures:

$$\begin{aligned} f(x_k) - f(x_{k+1}) &= \rho_k \big(m_k(0) - m_k(s_k)\big) \\ &\geq \rho' \min\left(c_2 \min\left(\Delta_k, \frac{\varepsilon_g}{c_0}\right) \varepsilon_g \ , \ c_3 \Delta_k^2 \varepsilon_H \right). \end{aligned}$$

By Lemma 6.21 and the assumption $\lambda_g \varepsilon_g \leq \Delta_0$, it holds that

$$\Delta_k \geq \min\left(\lambda_g \varepsilon_g, \lambda_H \varepsilon_H\right).$$

Furthermore, using $\lambda_g \leq 1/c_0$ reveals that

$$\min(\Delta_k, \varepsilon_g / c_0) \geq \min(\Delta_k, \lambda_g \varepsilon_g) \geq \min\left(\lambda_g \varepsilon_g, \lambda_H \varepsilon_H\right).$$

Hence,

$$f(x_k) - f(x_{k+1}) \geq \rho' \min\big(c_2 \lambda_g \varepsilon_g^2, c_2 \lambda_H \varepsilon_g \varepsilon_H, c_3 \lambda_g^2 \varepsilon_g^2 \varepsilon_H, c_3 \lambda_H^2 \varepsilon_H^3\big). \tag{6.40}$$

In the first scenario, $\varepsilon_H = \infty$ and the above simplifies to:

$$f(x_k) - f(x_{k+1}) \geq \rho' c_2 \lambda_g \varepsilon_g^2.$$

Sum over iterations up to n and use A6.5 (lower-bounded f):

$$f(x_0) - f_{\text{low}} \geq f(x_0) - f(x_{n+1}) = \sum_{k \in S_n} f(x_k) - f(x_{k+1}) \geq |S_n| \rho' c_2 \lambda_g \varepsilon_g^2.$$

Hence,

$$|S_n| \leq \frac{f(x_0) - f_{\text{low}}}{\rho' c_2 \lambda_g} \frac{1}{\varepsilon_g^2}.$$

Similarly, in the second scenario, starting over from (6.40) and assuming both $c_3 \lambda_g^2 \varepsilon_g^2 \varepsilon_H \leq c_2 \lambda_H \varepsilon_g \varepsilon_H$ and $c_3 \lambda_g^2 \varepsilon_g^2 \varepsilon_H \leq c_2 \lambda_g \varepsilon_g^2$ (which is equivalent to $\varepsilon_g \leq c_2 \lambda_H / c_3 \lambda_g^2$ and $\varepsilon_H \leq c_2 / c_3 \lambda_g$), the same telescoping sum yields

$$f(x_0) - f_{\text{low}} \geq |S_n| \rho' c_3 \min(\lambda_g \varepsilon_g, \lambda_H \varepsilon_H)^2 \varepsilon_H.$$

Reorganize this as a bound on $|S_n|$ to conclude. $\square$

Our third and last lemma lower-bounds the number of successful steps before termination as a fraction of the total number of iterations before termination. It captures the fact that we cannot have arbitrarily long strings of rejections.

Lemma 6.23. *Under the assumptions of Theorem 6.20, let $x_0, \ldots, x_n$ be iterates generated by RTR. If none of them satisfy the termination conditions* (6.32)*, using the notation S_n and U_n of Lemma 6.22, it holds that*

$$|S_n| \geq \frac{2}{3}(n+1) - \frac{1}{3}\max\left(0, \log_2\left(\frac{\Delta_0}{\lambda_g \varepsilon_g}\right), \log_2\left(\frac{\Delta_0}{\lambda_H \varepsilon_H}\right)\right). \tag{6.41}$$

Proof. The proof rests on the lower-bound for Δ_k from Lemma 6.21. For all $k \in S_n$, it holds that $\Delta_{k+1} \leq 2\Delta_k$. For all $k \in U_k$, it holds that $\Delta_{k+1} \leq 2^{-2}\Delta_k$. Hence,

$$\Delta_n \leq 2^{|S_n|} 2^{-2|U_n|} \Delta_0.$$

On the other hand, Lemma 6.21 gives

$$\Delta_n \geq \min(\Delta_0, \lambda_g \varepsilon_g, \lambda_H \varepsilon_H).$$

Combine, divide by Δ_0 and take the log in base 2 to see that:

$$|S_n| - 2|U_n| \geq \min\left(0, \log_2\left(\frac{\lambda_g \varepsilon_g}{\Delta_0}\right), \log_2\left(\frac{\lambda_H \varepsilon_H}{\Delta_0}\right)\right).$$

Use $|S_n| + |U_n| = n + 1$ to conclude. □

With these lemmas available, the main theorem follows easily.

Proof of Theorem 6.20. For each scenario, Lemmas 6.22 and 6.23 provide an upper-bound and a lower-bound on $|S_n|$, and it suffices to combine them to produce an upper-bound on n. For example, in the first scenario, if n is such that none of the iterates $x_0, \ldots, x_n$ have gradient smaller than ε_g, then

$$n \leq \frac{3}{2}\frac{f(x_0) - f_{\mathrm{low}}}{\rho' c_2}\frac{1}{\lambda_g \varepsilon_g^2} + \frac{1}{2}\log_2\left(\frac{\Delta_0}{\lambda_g \varepsilon_g}\right) - 1.$$

Thus, by contraposition, after a number of iterations larger than the right-hand side, an iterate with sufficiently small gradient must have been found. The same argument applies in the second scenario. □

6.4.5 Critical accumulation points

Building on Theorem 6.20 above, it is also possible to show that all accumulation points of RTR are critical points. We start with a straightforward corollary of this theorem that ensures RTR keeps generating points with small gradient, then we strengthen that corollary with an additional assumption.

Corollary 6.24. *Let $S = \{(x_0, s_0), (x_1, s_1), \ldots\}$ be the pairs of iterates and tentative steps generated by RTR under A6.1, A6.3 and A6.5 with $\varepsilon_g = 0$ (we aim for first-order criticality). Further assume A6.6 holds on S. Then,*

$$\liminf_{k \to \infty} \|\mathrm{grad} f(x_k)\|_{x_k} = 0, \tag{6.42}$$

that is: for all $\varepsilon > 0$ and K there exists $k \geq K$ such that $\|\mathrm{grad} f(x_k)\|_{x_k} \leq \varepsilon$.

Proof. Let $\varepsilon > 0$ and $K \in \{0, 1, 2, \ldots\}$ be arbitrary. Our assumptions imply that $S_K = \{(x_K, s_K), (x_{K+1}, s_{K+1}), \ldots\}$ is a sequence of pairs of iterates and tentative steps generated by RTR under A6.1, A6.3 and A6.5 with $\varepsilon_g = \varepsilon$ and $\varepsilon_H = \infty$, and that A6.6 holds on S_K. Thus, Theorem 6.20 guarantees that there exists $k \geq K$ such that $\|\mathrm{grad} f(x_k)\|_{x_k} \leq \varepsilon$. □

To strengthen the above corollary, we introduce a new assumption. If the function $x \mapsto \|\mathrm{grad} f(x)\|_x$ is Lipschitz continuous (see Section 10.4; this notably holds if $\mathrm{Hess} f$ is continuous and bounded), then that assumption is satisfied in particular if the retraction does not unduly distort distances, that is, if the Riemannian distance between x and $\mathrm{R}_x(s)$ is bounded by some constant times $\|s\|_x$ (see also A6.9 below). The latter holds for the exponential retraction (Section 10.2). The assumption below also holds if S is contained in a compact set, see Proposition 6.31.

A6.8. *For a given subset S of the tangent bundle $\mathrm{T}\mathcal{M}$, there exists a constant $L_{gn} > 0$ such that, for all $(x, s) \in S$,*

$$\big|\|\mathrm{grad} f(\mathrm{R}_x(s))\|_{\mathrm{R}_x(s)} - \|\mathrm{grad} f(x)\|_x\big| \leq L_{gn}\|s\|_x.$$

Under that assumption it is possible to show that all accumulation points of RTR are critical points. The statement below is similar in spirit to [AMS08, Thm. 7.4.4], though the precise assumptions are different, hence the proof is also adapted.

Proposition 6.25. *Let $S = \{(x_0, s_0), (x_1, s_1), \ldots\}$ be the pairs of iterates and tentative steps generated by RTR under A6.1, A6.3 and A6.5 with $\varepsilon_g = 0$ (we aim for first-order criticality). Further assume A6.6 and A6.8 hold on S. Then,*

$$\lim_{k \to \infty} \|\mathrm{grad} f(x_k)\|_{x_k} = 0. \tag{6.43}$$

In particular, all accumulation points of $x_0, x_1, \ldots$ (if any) are critical points.

Proof. If iteration k is unsuccessful ($\rho_k \leq \rho'$), then $x_{k+1} = x_k$. If iteration k is successful, then A6.8 guarantees

$$\|\mathrm{grad} f(x_{k+1})\|_{x_{k+1}} \geq \|\mathrm{grad} f(x_k)\|_{x_k} - L_{gn}\|s_k\|_{x_k}.$$

Fix an arbitrary index m such that $\mathrm{grad} f(x_m) \neq 0$. For all $\ell \geq m$, we have

$$\|\mathrm{grad} f(x_{\ell+1})\|_{x_{\ell+1}} \geq \|\mathrm{grad} f(x_m)\|_{x_m} - L_{gn} \sum_{\substack{m \leq k \leq \ell \\ \rho_k > \rho'}} \|s_k\|_{x_k}.$$

Pick the smallest $\ell \geq m$ such that $\|\mathrm{grad} f(x_{\ell+1})\|_{x_{\ell+1}} \leq \frac{1}{2}\|\mathrm{grad} f(x_m)\|_{x_m}$: we know such ℓ exists owing to Corollary 6.24. Then,

$$\|\mathrm{grad} f(x_m)\|_{x_m} \leq 2L_{gn} \sum_{\substack{m \leq k \leq \ell \\ \rho_k > \rho'}} \|s_k\|_{x_k} \leq 2L_{gn} \sum_{\substack{m \leq k \leq \ell \\ \rho_k > \rho'}} \Delta_k. \tag{6.44}$$

Given our choice of ℓ, we have $\|\mathrm{grad} f(x_k)\|_{x_k} > \frac{1}{2}\|\mathrm{grad} f(x_m)\|_{x_m}$ for $k =$

$m, \ldots, \ell$. Also, $x_{k+1} = \mathrm{R}_{x_k}(s_k)$ for all k such that $\rho_k > \rho'$. It thus follows from A6.3 and from the definition of ρ_k (6.22) that

$$\begin{aligned} f(x_m) - f(x_{\ell+1}) &= \sum_{\substack{m \leq k \leq \ell \\ \rho_k > \rho'}} f(x_k) - f(x_{k+1}) \\ &\geq \sum_{\substack{m \leq k \leq \ell \\ \rho_k > \rho'}} \frac{\rho' c_2}{2} \min\left(\Delta_k, \frac{\|\mathrm{grad} f(x_m)\|_{x_m}}{2c_0}\right) \|\mathrm{grad} f(x_m)\|_{x_m}. \end{aligned} \tag{6.45}$$

There are two scenarios to consider. Either $\frac{\|\mathrm{grad} f(x_m)\|_{x_m}}{2c_0} \leq \Delta_k$ for some k in the summation range, in which case we use the corresponding term to lower-bound the sum:

$$f(x_m) - f(x_{\ell+1}) \geq \frac{\rho' c_2}{4c_0} \|\mathrm{grad} f(x_m)\|_{x_m}^2. \tag{6.46}$$

Or $\frac{\|\mathrm{grad} f(x_m)\|_{x_m}}{2c_0} > \Delta_k$ for all k in the summation range, in which case we use both (6.45) and (6.44) to see that

$$\begin{aligned} f(x_m) - f(x_{\ell+1}) &\geq \frac{\rho' c_2}{2} \|\mathrm{grad} f(x_m)\|_{x_m} \sum_{\substack{m \leq k \leq \ell \\ \rho_k > \rho'}} \Delta_k \\ &\geq \frac{\rho' c_2}{4L_{gn}} \|\mathrm{grad} f(x_m)\|_{x_m}^2. \end{aligned} \tag{6.47}$$

The sequence of function values $f(x_0), f(x_1), \ldots$ is lower-bounded by A6.5 and non-increasing, hence it converges to some f_∞. Combining the results above with $f(x_m) - f_\infty \geq f(x_m) - f(x_{\ell+1})$, we find for all m that

$$f(x_m) - f_\infty \geq \frac{\rho' c_2}{4 \max(L_{gn}, c_0)} \|\mathrm{grad} f(x_m)\|_{x_m}^2. \tag{6.48}$$

Take the limit $m \to \infty$ to conclude, using $f(x_m) - f_\infty \to 0$. $\square$

6.4.6 Practical aspects

We list some practical considerations in a nutshell:

1. A typical value for ρ' is $\frac{1}{10}$.
2. Possible default settings for $\bar{\Delta}$ are $\sqrt{\dim \mathcal{M}}$ or the diameter of the manifold if it is bounded; and $\Delta_0 = \frac{1}{8}\bar{\Delta}$.
3. H_k is often taken to be $\mathrm{Hess} f(x_k)$ when available, regardless of whether or not the retraction is second order. This does not affect local convergence rates since close to critical points the distinction between first- and second-order retraction is irrelevant for us.

4. Practical stopping criteria for RTR typically involve an upper-bound on the total number of iterations and a threshold on the gradient norm such as: terminate if $\|\mathrm{grad} f(x_k)\|_{x_k} \leq \varepsilon_g$. Typically, $\varepsilon_g = 10^{-8}\|\mathrm{grad} f(x_0)\|_{x_0}$ is a good value. It is rare that one would explicitly check the eigenvalues of $\mathrm{Hess} f(x_k)$ before termination.
5. Computing ρ_k (6.22) can be delicate close to convergence, as it involves the computation of $f(x_k) - f(x_k^+)$: a difference of two potentially large numbers that could be dangerously close to one another. Specifically, say we compute $f(x_k)$ and we store it in memory in the variable f_1. Even if $f(x_k)$ is computed with maximal accuracy, it must eventually be rounded to one of the real numbers that are exactly representable in, say, double precision, that is, on 64 bits following the IEEE 754 standard. This standard guarantees a relative accuracy of $\varepsilon_\mathrm{M} \approx 10^{-16}$, so that $f_1 = f(x_k)(1 + \varepsilon_1)$ with $|\varepsilon_1| \leq \varepsilon_\mathrm{M}$. This is a relative accuracy guarantee since

   $$\frac{|f_1 - f(x_k)|}{|f(x_k)|} \leq \varepsilon_\mathrm{M}.$$

   (In practice, computing $f(x_k)$ would involve further errors leading to a larger right-hand side.) Likewise, $f_2 = f(x_k^+)(1 + \varepsilon_2)$ with $|\varepsilon_2| \leq \varepsilon_\mathrm{M}$.

   Assuming the difference between f_1 and f_2 is exactly representable in memory,[1] in computing the numerator for ρ_k we truly compute

   $$f_1 - f_2 = f(x_k) - f(x_k^+) + \varepsilon_1 f(x_k) - \varepsilon_2 f(x_k^+).$$

   The best we can claim in general about the relative error is:

   $$\frac{|(f_1 - f_2) - (f(x_k) - f(x_k^+))|}{|f(x_k) - f(x_k^+)|} \leq \varepsilon_\mathrm{M} \frac{|f(x_k)| + |f(x_k^+)|}{|f(x_k) - f(x_k^+)|}.$$

   The right-hand side can be catastrophically large. Indeed, if $f(x_k)$ and $f(x_k^+)$ are large in absolute value yet their difference is very small (which may happen near convergence), the relative error on the computation of the numerator of ρ_k may make it useless. For example, with $f(x_k) = 10^4$ and $f(x_k) - f(x_k^+) = 10^{-12}$, the relative error bound is close to 1, meaning *none* of the digits in the computed numerator can be trusted. In turn, this can lead to wrong decisions in RTR regarding step rejections and trust-region radius updates.

   No such issues plague the denominator, provided it is appropriately computed. Indeed,

   $$m_k(0) - m_k(s_k) = -\langle s_k, \mathrm{grad} f(x_k)\rangle_{x_k} - \frac{1}{2}\langle s_k, H_k(s_k)\rangle_{x_k}. \tag{6.49}$$

   Using the right-hand side for computation, if the step s_k is small and the gradient is small, then we combine two small real numbers, which is not as dangerous as computation of the left-hand side.

[1] By the *Sterbenz lemma*, this is true if f_1, f_2 are within a factor 2 of each other.

A standard fix [CGT00, §17.4.2] to these numerical issues is to regularize the computation of ρ_k as

$$\rho_k = \frac{f(x_k) - f(x_k^+) + \delta_k}{-\langle s_k, \mathrm{grad} f(x_k)\rangle_{x_k} - \frac{1}{2}\langle s_k, H_k(s_k)\rangle_{x_k} + \delta_k}, \tag{6.50}$$

with

$$\delta_k = \max(1, |f(x_k)|)\varepsilon_{\mathrm{M}}\rho_{\mathrm{reg}}. \tag{6.51}$$

The parameter ρ_{reg} can be set to 10^3 for example. When both the true numerator and denominator of ρ_k become very small near convergence, the regularization nudges (6.50) toward 1, which leads to step acceptance as expected. This is a heuristic to (try to) address an inescapable limitation of inexact arithmetic, though a detailed analysis by Sun and Nocedal provides insight into what one may reasonably guarantee with it [SN22].

6. Care should be put in implementations to minimize the number of calls to the map H_k. For example, in the subproblem solver described in Section 6.5 below, exactly one call to H_k is needed per iteration, and furthermore the vector $H_k(s_k)$ is a by-product of that algorithm when H_k is linear (Exercise 6.28), so that computing the denominator of ρ_k does not require further calls to H_k.

6.5 The trust-region subproblem: truncated CG

The trust-region subproblem (6.21) consists in approximately solving a problem of the form

$$\min_{s \in \mathrm{T}_x\mathcal{M}} m(s) \text{ subject to } \|s\|_x \leq \Delta \quad \text{where} \quad m(s) = \frac{1}{2}\langle s, Hs\rangle_x - \langle b, s\rangle_x, \tag{6.52}$$

with a map $H\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$, a tangent vector $b \in \mathrm{T}_x\mathcal{M}$ and a radius $\Delta > 0$. At iteration k of RTR, these objects are $H = H_k$, $b = -\mathrm{grad} f(x_k)$ and $\Delta = \Delta_k$.

We consider the important particular case where H is a linear, self-adjoint map (for example, $H_k = \mathrm{Hess} f(x_k)$.) Then, $m\colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}$ is a quadratic function. Aside from the constraint $\|s\|_x \leq \Delta$, if H is furthermore positive definite, then we know from Section 6.3 that conjugate gradients (CG, Algorithm 6.2) can be used to compute a global minimizer of m: simply compare functions in (6.4) and (6.52).

The general idea of the *truncated CG* method (tCG), Algorithm 6.4, is to run CG on $m(s)$ (6.52) while

1. Keeping an eye out for signs that H may not be positive definite;
2. Checking whether we left the trust region; and
3. Looking for opportunities to terminate early even if neither of those events happen.

Recall that CG generates directions p_i. If the scalars $\langle p_i, Hp_i \rangle_x$ are positive for $i = 0, \ldots, n-2$, then $p_0, \ldots, p_{n-2}$ are linearly independent hence they form a basis for a subspace of $\mathrm{T}_x\mathcal{M}$. Moreover, H is positive definite on that subspace. Thus, up to that point, all the properties of CG hold. If, however, upon considering p_{n-1} we determine that $\langle p_{n-1}, Hp_{n-1} \rangle_x$ is non-positive, then this is proof that H is not positive definite. In such situation, tCG computes the next step v_n by moving away from v_{n-1} along p_{n-1} so as to minimize the model m, that is, tCG sets $v_n = v_{n-1} + tp_{n-1}$ with t such that $m(v_n)$ is minimized, under the constraint $\|v_n\|_x \leq \Delta$. There are two candidates for the value of t, namely, the two roots of the quadratic

$$\|v_{n-1} + tp_{n-1}\|_x^2 - \Delta^2 = \|p_{n-1}\|_x^2 t^2 + 2t \langle v_{n-1}, p_{n-1} \rangle_x + \|v_{n-1}\|_x^2 - \Delta^2. \quad (6.53)$$

The product of these roots is negative since $\|v_{n-1}\|_x < \Delta$ (otherwise we would have already terminated), hence one root is positive and the other is negative. It can be shown that selecting the positive root leads to the smallest value in the model [ABG07, §3].

Now assuming $\langle p_{n-1}, Hp_{n-1} \rangle_x$ is positive, we consider the tentative new step $v_{n-1}^+ = v_{n-1} + \alpha_n p_{n-1}$. If this step lies outside the trust region, it seems at first that we face a dilemma. Indeed, a priori, it might happen that later iterates re-enter the trust region, in which case it would be unwise to stop. Fortunately, this cannot happen. Specifically, it can be shown that steps grow in norm, so that if one iterate leaves the trust region, then no future iterate re-enters it [CGT00, Thm. 7.5.1], [NW06, Thm. 7.3]. Thus, it is reasonable to act now: tCG proceeds by reducing how much we move along p_{n-1}, setting $v_n = v_{n-1} + tp_{n-1}$ instead with $t \geq 0$ being the largest value that fulfills the trust-region constraint. This happens to correspond exactly to the positive root of the quadratic in eq. (6.53). In the unlikely event that v_{n-1}^+ lies exactly on the boundary of the trust region, it makes sense to stop by the same argument: this is why we test for $\|v_{n-1}^+\|_x \geq \Delta$ with a non-strict inequality.

Finally, if neither non-positive $\langle p_i, Hp_i \rangle_x$ are encountered nor do the steps leave the trust region, we rely on a stopping criterion to terminate tCG early. The principle is that we should only work hard to solve the subproblem when RTR is already close to convergence. Specifically, with $r_0 = b = -\mathrm{grad}f(x_k)$, the chosen stopping criterion with parameters θ and κ allows tCG to terminate if

$$\|r_n\|_{x_k} \leq \|\mathrm{grad}f(x_k)\|_{x_k} \cdot \min(\|\mathrm{grad}f(x_k)\|_{x_k}^{\theta}, \kappa). \quad (6.54)$$

It is only when the gradient of f is small that tCG puts in the extra effort to reach residuals as small as $\|\mathrm{grad}f(x_k)\|_{x_k}^{1+\theta}$. This is key to obtain superlinear convergence, of order $\min(1+\theta, 2)$ (in particular, quadratic convergence for $\theta = 1$), see Theorem 6.30 below. Intuitively, superlinear convergence occurs because when x_k is close to a critical point with positive definite Hessian, and with $H_k = \mathrm{Hess}f(x_k)$, steps produced by tCG are increasingly similar to Newton steps.

The comments at the end of Section 6.3 regarding how to run CG in practice

Algorithm 6.4 tCG: truncated conjugate gradients on a tangent space

Parameters: $\kappa \geq 0, \theta \in (0, 1]$, e.g., $\kappa = \frac{1}{10}$, $\theta = 1$
Input: self-adjoint H on $\mathrm{T}_x\mathcal{M}$, $b \in \mathrm{T}_x\mathcal{M}$ and radius $\Delta > 0$
Output: approximate minimizer of $m(s) = \frac{1}{2}\langle s, Hs\rangle_x - \langle b, s\rangle_x$ subject to $\|s\|_x \leq \Delta$
Set $v_0 = 0, r_0 = b, p_0 = r_0$
If $r_0 = 0$
  output $s = v_0$
For $n = 1, 2, \ldots$
  Compute Hp_{n-1} (this is the only call to H)
  Compute $\langle p_{n-1}, Hp_{n-1}\rangle_x$
  $\alpha_n = \frac{\|r_{n-1}\|_x^2}{\langle p_{n-1}, Hp_{n-1}\rangle_x}$
  $v_{n-1}^+ = v_{n-1} + \alpha_n p_{n-1}$
  If $\langle p_{n-1}, Hp_{n-1}\rangle_x \leq 0$ **or** $\|v_{n-1}^+\|_x \geq \Delta$
    Set $v_n = v_{n-1} + tp_{n-1}$ with $t \geq 0$ such that $\|v_n\|_x = \Delta$
    (t is the positive root of the quadratic in (6.53).)
    output $s = v_n$
  $v_n = v_{n-1}^+$
  $r_n = r_{n-1} - \alpha_n Hp_{n-1}$
  If $\|r_n\|_x \leq \|r_0\|_x \min(\|r_0\|_x^\theta, \kappa)$
    output $s = v_n$
  $\beta_n = \frac{\|r_n\|_x^2}{\|r_{n-1}\|_x^2}$
  $p_n = r_n + \beta_n p_{n-1}$

apply to tCG as well. Specifically, it is common to set a hard limit on the maximum number of iterations, and it is beneficial to ensure tangent vectors remain tangent numerically.

Just like regular CG, tCG can be *preconditioned* [CGT00, §5.1.6]: this can improve performance dramatically. In a precise sense, preconditioning tCG is equivalent to changing the Riemannian metric [MS16].

Finally, it is good to know that the trust-region subproblem, despite being non-convex, can be solved to global optimality efficiently. See [Vav91] and [CGT00, §7] for pointers to a vast literature.

Exercise 6.26. *Show that v_1 as computed by Algorithm 6.4 is the Cauchy point as constructed in Lemma 6.15. Since iterates monotonically improve $m(v_n)$* (6.52) *this implies that tCG guarantees A6.3 (p135) with $c_2 = \frac{1}{2}$.*

Exercise 6.27. *Consider using tCG within RTR, so that $b = -\mathrm{grad} f(x_k)$ and $H = H_k$ at iteration k of RTR. If $b = 0$, tCG terminates immediately with $s = 0$ (this leads RTR to set $s_k = 0$, so that $\rho_k = \frac{0}{0}$ (not-a-number); standard extended arithmetic conventions then lead RTR to set $x_{k+1} = x_k$ and $\Delta_{k+1} = \Delta_k$). Check* *that this may violate A6.4 (p135) if H_k has negative eigenvalues (in particular, tCG does* not *compute a global minimum of the trust-region subproblem in this case). Explain why it is necessary for tCG to terminate immediately if $b = 0$, that is, explain why even if we skip the initial "if" statement the rest of the algorithm would not be able to exploit the negative eigenvalues of H_k. See [CGT00, §7.5.4] for a fix based on Lanczos iterations.*

Exercise 6.28. *Algorithm 6.4 terminates with a vector s as output. Show that the same algorithm can also output Hs as a by-product without requiring additional calls to H. Explicitly, if the second "output" statement triggers, then $Hs = b - r_{n-1} + tHp_{n-1}$; and if the third "output" statement triggers, then $Hs = b - r_n$. This is useful to compute the denominator of ρ_k* (6.22) *in the trust-region method via* (6.49).

6.6 Local convergence of RTR with tCG*

Under suitable assumptions, once iterates of RTR are close enough to a critical point where the Hessian is positive definite, RTR converges superlinearly to that point provided subproblems are solved with sufficient accuracy (for example, using tCG). The two theorems below make this precise: they are (in some ways, restricted) variations of claims found in [ABG07] and [AMS08, §7]. The proofs are omitted.

The first result is a variation of [AMS08, Thm. 7.4.10]: it is a type of capture theorem for RTR with tCG. It involves a special assumption on the retraction that prevents undue distance distortions. It holds in particular if the retraction is the exponential map (with $c_5 = 1$), and it also holds if $\mathcal{M}$ is compact (see Lemma 6.32 below).

A6.9. *There exist positive constants c_4, c_5 such that, for all $(x, v) \in \mathrm{T}\mathcal{M}$, if $\|v\|_x \le c_4$ then $\mathrm{dist}(x, \mathrm{R}_x(v)) \le c_5 \|v\|_x$.*

Below, we require that $\mathrm{Hess} f$ is continuous.

Theorem 6.29. *Let $S = \{(x_0, s_0), (x_1, s_1), \ldots\}$ be the pairs of iterates and tentative steps generated by RTR with tCG as subproblem solver, with models $H_k = \mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0)$ or $H_k = \mathrm{Hess} f(x_k)$. Assume $\|H_k\| \le c_0$ with some constant c_0 for all k, so that A6.1 and A6.3 hold, and also assume f is lower-bounded as in A6.5. Further assume A6.6 and A6.8 hold on S. (In particular, the assumptions of Proposition 6.25 hold, so that $\|\mathrm{grad} f(x_k)\|_{x_k} \to 0$.) Let the retraction satisfy A6.9.*

Let $x_\star \in \mathcal{M}$ satisfy $\mathrm{grad} f(x_\star) = 0$ and $\mathrm{Hess} f(x_\star) \succ 0$—in particular, it is a local minimizer. There exists a neighborhood $\mathcal{U}$ of $x_\star$ such that, if x_k is in $\mathcal{U}$ for some k, then all subsequent iterates are in $\mathcal{U}$ and they converge to $x_\star$.

The second result is a restriction of [AMS08, Thm. 7.4.11]. It establishes superlinear local convergence (recall Definitions 4.14 and 4.15). We require that $\mathrm{Hess}f$ is continuously differentiable.

Theorem 6.30. *Let $S = \{(x_0, s_0), (x_1, s_1), \ldots\}$ be the pairs of iterates and tentative steps generated by RTR with tCG as subproblem solver, either with models $H_k = \mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0)$ or with models $H_k = \mathrm{Hess}f(x_k)$. If the latter, assume there exists a constant c_6 such that $\|c''(0)\|_x \leq c_6$ for all curves of the type $c(t) = \mathrm{R}_x(ts)$ with $s \in \mathrm{T}_x\mathcal{M}$ of unit norm and $x = x_k$ for some k—this holds with $c_6 = 0$ if the retraction is second order.*

Let $x_\star \in \mathcal{M}$ satisfy $\mathrm{grad}f(x_\star) = 0$ and $\mathrm{Hess}f(x_\star) \succ 0$. If the sequence $x_0, x_1, x_2, \ldots$ converges to $x_\star$ (as Theorem 6.29 might provide), then there exist a constant $c_7 > 0$ and an index K such that, for all $k \geq K$, we have

$$\mathrm{dist}(x_{k+1}, x_\star) \leq c_7 \,\mathrm{dist}(x_k, x_\star)^{\min(\theta+1,2)},$$

where $\theta > 0$ is a parameter in the stopping criterion of tCG. In particular, with $\theta = 1$ convergence is at least quadratic.

6.7 Simplified assumptions for RTR with tCG*

The main theorems of Sections 6.4 and 6.6 involve a number of assumptions that need to be checked in order to claim convergence guarantees for RTR. In this section, we restrict the discussion to RTR with tCG as subproblem solver and include simple assumptions that simplify the process of verifying that all other assumptions hold. The resulting statements are more restrictive than above, but they can often be applied directly in applications. This is especially simple if $\mathcal{M}$ is compact, as is the case for the Stiefel and the Grassman manifolds for example.

Throughout, we require that $\mathrm{Hess}f$ is continuously differentiable.

Proposition 6.31. *Let $S = \{(x_0, s_0), (x_1, s_1), \ldots\}$ be the pairs of iterates and tentative steps generated by RTR with models $H_k = \mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0)$ and tCG as subproblem solver. (If the retraction is second order, the models coincide with $\mathrm{Hess}f(x_k)$.) Assume the iterates $x_0, x_1, \ldots$ are contained in a compact subset of $\mathcal{M}$. (This holds in particular if $\mathcal{M}$ is compact, or if any of the sublevel sets $\{x \in \mathcal{M} : f(x) \leq f(x_k)\}$ is compact.) Then, A6.1 holds with some $c_0 \geq 0$, A6.2 holds with $c_1 = 0$, A6.3 holds with $c_2 = \frac{1}{2}$ (Exercise 6.26), A6.4 may not hold (Exercise 6.27), A6.5 holds with $f_{\mathrm{low}} = \inf_k f(x_k) > -\infty$, A6.6 and A6.7 hold on S with some constants L_g and L_H (Lemma 10.57), and A6.8 holds on S with some constant L_{gn}.*

If the subproblem solver is replaced by one which solves the trust-region subproblem to optimality, then all of the above remain true except A6.4 is also satisfied with $c_3 = \frac{1}{2}$ (Corollary 6.17).

Proof. All claims are clear except for the gradient-norm Lipschitz-type assumption A6.8 which we now verify explicitly. In so doing, we use concepts from

Sections 10.1 (length of curves), 10.3 (parallel transport) and 10.4 (Lipschitz continuity).

Since the iterates x_k are contained in a compact set $\mathcal{K}$ we know that S is included in $\mathcal{T} = \{(x, s) : x \in \mathcal{K} \text{ and } \|s\|_x \leq \bar{\Delta}\}$ which is compact in $\mathrm{T}\mathcal{M}$ (Exercise 10.31). For each $(x, s) \in \mathrm{T}\mathcal{M}$, the map $\mathrm{DR}_x(s)$ is linear from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_{\mathrm{R}_x(s)}\mathcal{M}$. Its operator norm is continuous as a function of (x, s) since R is smooth, that is, the function $(x, s) \mapsto \|\mathrm{DR}_x(s)\|$ is continuous on $\mathcal{T}$. Since $\mathcal{T}$ is compact, we deduce that there exists a constant r such that $\|\mathrm{DR}_x(s)\| \leq r$ for all $(x, s) \in \mathcal{T}$. Consequently, with $(x, s) \in \mathcal{T}$ arbitrary and $c(t) = \mathrm{R}_x(ts)$, we find that the length of the curve c on the interval $[0, 1]$ satisfies:

$$\begin{aligned} L(c) &= \int_0^1 \|c'(t)\|_{c(t)}\mathrm{d}t \\ &= \int_0^1 \|\mathrm{DR}_x(ts)[s]\|_{c(t)}\mathrm{d}t \leq \int_0^1 r\|s\|_x\mathrm{d}t = r\|s\|_x. \end{aligned} \tag{6.55}$$

The set $\mathrm{R}(\mathcal{T}) = \{\mathrm{R}_x(s) : (x, s) \in \mathcal{T}\}$ is compact in $\mathcal{M}$ since it is the image of a compact set through a continuous map. Thus, $\mathrm{Hess}f$ is continuous hence bounded (in operator norm) by some constant q on $\mathrm{R}(\mathcal{T})$. Writing $\mathrm{PT}^c_{1\leftarrow 0}$ for parallel transport along a curve c from $t = 0$ to $t = 1$ (this is an isometry from $\mathrm{T}_{c(0)}\mathcal{M}$ to $\mathrm{T}_{c(1)}\mathcal{M}$), it follows with $c(t) = \mathrm{R}_x(ts)$ and using Proposition 10.46 that, for all $(x, s) \in \mathcal{T}$,

$$\begin{aligned} \|\mathrm{grad}f(\mathrm{R}_x(s))\|_{\mathrm{R}_x(s)} &= \|\mathrm{grad}f(\mathrm{R}_x(s)) - \mathrm{PT}^c_{1\leftarrow 0}\mathrm{grad}f(x) + \mathrm{PT}^c_{1\leftarrow 0}\mathrm{grad}f(x)\|_{\mathrm{R}_x(s)} \\ &\leq \|\mathrm{grad}f(\mathrm{R}_x(s)) - \mathrm{PT}^c_{1\leftarrow 0}\mathrm{grad}f(x)\|_{\mathrm{R}_x(s)} + \|\mathrm{grad}f(x)\|_x \\ &\leq qL(c) + \|\mathrm{grad}f(x)\|_x \\ &\leq qr\|s\|_x + \|\mathrm{grad}f(x)\|_x. \end{aligned}$$

Thus, $\|\mathrm{grad}f(\mathrm{R}_x(s))\|_{\mathrm{R}_x(s)} - \|\mathrm{grad}f(x)\|_x \leq qr\|s\|_x$. A similar argument shows that $\|\mathrm{grad}f(x)\|_x - \|\mathrm{grad}f(\mathrm{R}_x(s))\|_{\mathrm{R}_x(s)} \leq qr\|s\|_x$, so that A6.8 holds on $\mathcal{T}$ with $L_{gn} = qr$. □

Lemma 6.32. *Any retraction* R *on a compact manifold* $\mathcal{M}$ *satisfies A6.9.*

Proof. For all $c_4 > 0$ the set $\mathcal{T} = \{(x, v) \in \mathrm{T}\mathcal{M} : \|v\|_x \leq c_4\}$ is compact (Exercise 10.31) hence there exists $c_5 > 0$ such that $\|\mathrm{DR}_x(v)\| \leq c_5$ for all $(x, v) \in \mathcal{T}$ (by continuity of the operator norm and smoothness of the retraction). It then follows from (6.55) and from the definitions of distance and length (Section 10.1) that

$$\mathrm{dist}(x, \mathrm{R}_x(v)) \leq L(c) \leq c_5\|v\|_x, \tag{6.56}$$

where $c(t) = \mathrm{R}_x(tv)$ is a curve from $c(0) = x$ to $c(1) = \mathrm{R}_x(v)$. □

Corollary 6.33. *Let $S = \{(x_0, s_0), (x_1, s_1), \ldots\}$ be the pairs of iterates and tentative steps generated by RTR with models $H_k = \mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0)$ and tCG as subproblem solver.*

If the sublevel set $\{x \in \mathcal{M} : f(x) \leq f(x_0)\}$ is compact (which holds if $\mathcal{M}$ is *compact), then the sequence of iterates $x_0, x_1, x_2, \ldots$ has at least one accumulation point and all of its accumulation points are critical points.*

Further assume the retraction satisfies A6.9 (this holds if $\mathcal{M}$ is compact by Lemma 6.32, or if $\mathcal{M}$ is complete and the retraction is the exponential map). If one of the accumulation points has a positive definite Hessian, then the sequence converges to that point with a superlinear local convergence rate (quadratic if $\theta = 1$ in tCG).

Proof. RTR is a descent method ($f(x_{k+1}) \leq f(x_k)$ for all k) hence the sequence $x_0, x_1, \ldots$ is contained in a compact set: this ensures that it has at least one accumulation point. All of these accumulation points are critical points owing to Proposition 6.25, whose assumptions are satisfied owing to Proposition 6.31. If A6.9 holds too, then Theorem 6.29 applies, guaranteeing that if any of the accumulation points has a positive definite Hessian then that critical point is attractive: eventually, the sequence enters any neighborhood of that point and converges to it as a result. The rate of convergence follows from Theorem 6.30. □

6.8 Numerically checking a Hessian*

In Section 4.8, we considered a numerical method to check whether code to compute the Riemannian gradient is correct. Similarly, we now describe a method to check code for the Riemannian Hessian. In the Matlab toolbox Manopt, this method is implemented as `checkhessian`.

The two first points to check are:

1. That $\mathrm{Hess} f(x)$ indeed maps $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_x\mathcal{M}$ linearly, and
2. That it is indeed a self-adjoint map.

This can be done numerically by generating a random $x \in \mathcal{M}$ and two random tangent vectors $u, v \in \mathrm{T}_x\mathcal{M}$, computing both $\mathrm{Hess} f(x)[u]$ and $\mathrm{Hess} f(x)[v]$, verifying that these are tangent, checking that

$$\mathrm{Hess} f(x)[au + bv] = a\mathrm{Hess} f(x)[u] + b\mathrm{Hess} f(x)[v]$$

for some random scalars a, b, and finally confirming that

$$\langle u, \mathrm{Hess} f(x)[v] \rangle_x = \langle \mathrm{Hess} f(x)[u], v \rangle_x ,$$

all up to machine precision.

This being secured, consider the Taylor expansion (5.28): if R is a second-order retraction, or if x is a critical point, then

$$f(\mathrm{R}_x(tv)) = f(x) + t \langle \mathrm{grad} f(x), v \rangle_x + \frac{t^2}{2} \langle \mathrm{Hess} f(x)[v], v \rangle_x + O(t^3). \tag{6.57}$$

This says that, under the stated conditions,

$$E(t) \triangleq \left| f(\mathrm{R}_x(tv)) - f(x) - t \left\langle \mathrm{grad} f(x), v \right\rangle_x - \frac{t^2}{2} \left\langle \mathrm{Hess} f(x)[v], v \right\rangle_x \right| = O(t^3).$$

Taking the logarithm on both sides, we find that $\log(E(t))$ must grow approximately linearly in $\log(t)$, with a slope of three (or more) when t is small:

$$\log(E(t)) \approx 3\log(t) + \text{constant}.$$

This suggests a procedure to check the Hessian numerically:

1. Check that the gradient is correct (Section 4.8);
2. Run the preliminary checks (tangency, linearity and symmetry);
3. If using a second-order retraction, generate a random point $x \in \mathcal{M}$; otherwise, find an (approximate) critical point $x \in \mathcal{M}$, for example using Riemannian gradient descent;
4. Generate a random tangent vector $v \in \mathrm{T}_x\mathcal{M}$ with $\|v\|_x = 1$;
5. Compute $f(x), \left\langle \mathrm{grad} f(x), v \right\rangle_x$ and $\left\langle \mathrm{Hess} f(x)[v], v \right\rangle_x$;
6. Compute $E(t)$ for several values of t logarithmically spaced on the interval $[10^{-8}, 10^0]$;
7. Plot $E(t)$ as a function of t, in a log–log plot;
8. Check that the plot exhibits a slope of three (or more) over several orders of magnitude.

Again, we do not expect to see a slope of three over the whole range, but we do expect to see this over a range of values of t covering at least one or two orders of magnitude. Of course, the test is less conclusive if it has to be run at a critical point. Even if computing second-order retractions turns out to be expensive for the manifold at hand, its use here as part of a diagnostics tool is worthwhile: we are free to use any other retraction for the optimization algorithm.

6.9 Notes and references

First- and second-order optimality conditions are further studied in [YZS14, BH19], notably to include the case of constrained optimization on manifolds. Newton's method on manifolds is analyzed in most treatments of optimization on manifolds; see for example [AMS08, §6] and the many references therein, including [ADM$^+$02, Man02]. In particular, the convergence result Theorem 6.7 and Exercise 6.9 correspond to [AMS08, Thm. 6.3.2]. The reference material for the discussion of conjugate gradients in Section 6.3 is [TB97, Lect. 38].

Assumption A6.9 in Section 6.6 parallels an assumption made for the same reasons in [AMS08, eq. (7.25)].

Trust-region methods in Euclidean space are discussed in great detail by Conn et al. [CGT00]; see also [NW06] for a shorter treatment. Absil et al. [ABG07] introduced the Riemannian version of the trust-region method. Their analysis also appears in [AMS08, §7]. The global convergence analysis which shows RTR computes approximate first- and second-order critical points in a bounded number of iterations is mostly the same as in [BAC18]. Certain parts appear almost verbatim in that reference (in particular, the proofs of Lemmas 6.22 and 6.23). It is itself based on a similar analysis of the Euclidean version proposed by Cartis et al. [CGT12], who also show examples for which the worst-case is attained. The global convergence results in terms of limit inferior and limit of gradient norm (Corollary 6.24 and Proposition 6.25) appear with somewhat different assumptions as [AMS08, Thm. 7.4.2, Thm. 7.4.4]: the proofs are adapted accordingly.

The RTR method presented here generates sequences whose accumulation points are first-order critical (under some assumptions). It can also find approximate second-order critical points up to any tolerance, but the theory does not guarantee accumulation at exact second-order critical points. A somewhat more theoretical variant of RTR presented in [LKB21] does accumulate at second-order critical points. It mirrors a Euclidean construction by Curtis et al. [CLR18].

For local convergence results, the capture theorem (Theorem 6.29) and the superlinear local convergence result (Theorem 6.30) appear with proofs as [AMS08, Thm. 7.4.10, Thm. 7.4.11]. The statements here are somewhat different but the same proofs apply. In particular, for Theorem 6.30, the reference statement [AMS08, Thm. 7.4.11] makes the two following assumptions (among others). First, there exists $c_6 > 0$ such that, for all k,

$$\|H_k - \mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0)\| \le c_6 \|\mathrm{grad} f(x_k)\|_{x_k}. \tag{6.58}$$

This is clear if $H_k = \mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0)$. If $H_k = \mathrm{Hess} f(x_k)$ the above follows from the assumptions in Theorem 6.30 and from the following formula (see for example Exercise 10.73):

$$\langle v, \mathrm{Hess}(f \circ \mathrm{R}_x)(0)[v] \rangle_x = \langle v, \mathrm{Hess} f(x)[v] \rangle_x + \langle \mathrm{grad} f(x), c''(0) \rangle_x \tag{6.59}$$

where $(x, v) \in \mathrm{T}\mathcal{M}$ is arbitrary and $c(t) = \mathrm{R}_x(tv)$. Second, there exist positive c_8, c_9, c_{10} such that, for all $(x, v) \in \mathrm{T}\mathcal{M}$ with $\mathrm{dist}(x, x_\star) \le c_8$ and $\|v\|_x \le c_9$ it holds

$$\|\mathrm{Hess}(f \circ \mathrm{R}_x)(v) - \mathrm{Hess}(f \circ \mathrm{R}_x)(0)\| \le c_{10} \|v\|_x. \tag{6.60}$$

This always holds if $\mathrm{Hess} f$ is continuous, by Lemma 10.57.

To some extent, the trust-region method is a fix of Newton's method to make it globally convergent. At its core, it is based on putting a hard limit on how far one trusts a certain quadratic model for the (pullback of the) cost function. Alternatively, one may resort to a soft limit by adding a cubic regularization term to a quadratic model. In the same way that the trust-region radius is updated adaptively, the weight of the regularization term can also be updated adaptively, leading to the *adaptive regularization with cubics* (ARC) method. In the Euclidean case, it dates back to seminal work by Griewank [Gri81] and Nesterov and Polyak [NP06]. Cartis et al. give a thorough treatment including complexity bounds [CGT11b, CGT11a]. Qi proposed a first extension of ARC to

Riemannian manifolds [Qi11]. Iteration complexity analyses akin to the one we give here for RTR appear in [ZZ18, ABBC20]. As a theoretical strength, ARC is an optimal method for cost functions with Lipschitz continuous gradient and Hessian.

7 Embedded submanifolds: examples

In this chapter, we describe several embedded submanifolds of linear spaces that occur in applications. For each one, we rely on Chapters 3 and 5 to derive the geometric tools that are relevant to optimize over them. See Table 7.1 for a list of the manifolds discussed in this chapter (and a few more), together with pointers to Matlab implementations in the toolbox Manopt [BMAS14]. PyManopt [TKW16] and Manopt.jl [Ber21] provide similar implementations in Python and Julia. All three toolboxes are available from manopt.org.

Remember from Section 3.2 that products of embedded submanifolds are embedded submanifolds. This extends to general manifolds. Throughout the book, we show how to build the geometric toolbox of a product using the geometric toolboxes of its parts. See Table 7.2 for pointers. Manopt builds these toolboxes automatically for products $\mathcal{M}_1 \times \cdots \times \mathcal{M}_k$ and powers $\mathcal{M}^k = \mathcal{M} \times \cdots \times \mathcal{M}$ with the tools `productmanifold` and `powermanifold`.

This chapter is meant to be consulted periodically for illustration while reading earlier chapters.

7.1 Euclidean spaces as manifolds

Optimization on manifolds generalizes unconstrained optimization: the tools and algorithms we develop here apply just as well to optimization on linear spaces. For good measure, we spell out the relevant geometric tools.

Let $\mathcal{E}$ be a real linear space, such as $\mathbb{R}^n, \mathbb{R}^{m\times n}, \mathbb{C}^n, \mathbb{C}^{m\times n}$, etc.: see Section 3.1. We think of $\mathcal{E}$ as a (linear) manifold. Its dimension as a manifold is the same as its dimension as a linear space. All tangent spaces are the same: for $x \in \mathcal{E}$,

$$\mathrm{T}_x\mathcal{E} = \mathcal{E}. \tag{7.1}$$

An obvious (and reasonable) choice of retraction is

$$\mathrm{R}_x(v) = x + v, \tag{7.2}$$

though Definition 3.47 allows for more exotic choices as well.

Equipped with an inner product, $\mathcal{E}$ is a Euclidean space, and also a (linear) Riemannian manifold. The orthogonal projector from $\mathcal{E}$ to a tangent space is of

$\mathcal{M}$	Set	Manopt tools	Section
$\mathcal{E}$	$\mathbb{R}^n, \mathbb{R}^{m\times n}, \ldots$	`euclideanfactory`	7.1
	$\mathbb{C}^n, \mathbb{C}^{m\times n}, \ldots$	`euclideancomplexfactory`	
	$\mathrm{Sym}(n)$	`symmetricfactory`	
	$\mathrm{Skew}(n)$	`skewsymmetricfactory`	
	Subspace	`euclideansubspacefactory`	
S^{d-1}	Sphere in $\mathbb{R}^{m\times n}$	`spherefactory`	7.2
	Sphere in $\mathbb{C}^{m\times n}$	`spherecomplexfactory`	
$\mathrm{OB}(d, n)$	Oblique manifold	`obliquefactory`	
	Complex oblique	`obliquecomplexfactory`	
$\mathbb{C}_1^n$	n complex phases	`complexcirclefactory`	
$\mathrm{St}(n, p)$	Stiefel	`stiefelfactory`	7.3
	Complex Stiefel	`stiefelcomplexfactory`	
$\mathrm{O}(n)$	Orthogonal group	(see $\mathrm{St}(n, n)$ or $\mathrm{SO}(n)$)	7.4
$\mathrm{SO}(n)$	Rotation group	`rotationsfactory`	
$\mathrm{U}(n)$	Unitary group	`unitaryfactory`	
$\mathbb{R}_r^{m\times n}$	Fixed rank	`fixedrankembeddedfactory`	7.5
H^n	Hyperbolic space	`hyperbolicfactory`	7.6
	$\{x \in \mathcal{E} : h(x) = 0\}$		7.7
$\mathrm{Gr}(n, p)$	Set of subspaces	`grassmannfactory`	9.16
	in $\mathbb{R}^n$ or $\mathbb{C}^n$	`grassmanncomplexfactory`	
$\mathrm{Sym}(n)^+$	Positive definite	`sympositivedefinitefactory`	11.7
$\mathcal{M}_1 \times \mathcal{M}_2$	Product manifold	`productmanifold`	
$\mathcal{M}^k$	Power manifold	`powermanifold`	

Table 7.1 List of manifolds described in this chapter (and a few more), with pointers to implementations in Manopt (Matlab). The toolbox offers more, as documented on the website manopt.org. The latter also points to PyManopt and Manopt.jl with implementations in Python and Julia. Section 7.8 points to additional manifolds of interest. Details regarding product manifolds are given throughout the book: see Table 7.2.

course the identity map:

$$\mathrm{Proj}_x(u) = u. \tag{7.3}$$

Smoothness of a function $f : \mathcal{E} \to \mathbb{R}$ is defined in the usual sense; its classical gradient and its Riemannian gradient coincide.

More generally, we may consider a linear manifold $\mathcal{M}$ embedded in a Euclidean space $\mathcal{E}$, that is: $\mathcal{M}$ is a *linear subspace* of $\mathcal{E}$. For example, we may consider $\mathrm{Sym}(n)$—the space of real symmetric matrices of size n—to be a submanifold of $\mathbb{R}^{n\times n}$. It still holds that $\mathrm{T}_x\mathcal{M} = \mathcal{M}$ for all $x \in \mathcal{M}$, and $\mathrm{R}_x(v) = x + v$ is still a good choice for a retraction. Numerically, points and tangent vectors of $\mathcal{M}$ are typically stored as elements of $\mathcal{E}$. In this more general setup, Proj_x denotes the orthogonal projection from $\mathcal{E}$ to $\mathrm{T}_x\mathcal{M}$, that is: orthogonal projection from $\mathcal{E}$ to $\mathcal{M}$. In particular, it does not depend on x: we write $\mathrm{Proj}_{\mathcal{M}}$. If we make $\mathcal{M}$ into a

Product of manifolds is a manifold (embedded)	Proposition 3.20
Product of manifolds is a manifold (general)	Exercise 8.31
Differential of $F\colon \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{N}$	Exercise 3.40
Tangent bundle of $\mathcal{M}_1 \times \mathcal{M}_2$	Equation (3.31)
Retraction for $\mathcal{M}_1 \times \mathcal{M}_2$	Exercise 3.50
Product of Riemannian metrics is Riemannian	Example 3.57
Gradient of $f\colon \mathcal{M}_1 \times \mathcal{M}_2 \to \mathbb{R}$	Exercise 3.67
Product connection ∇ on $\mathcal{M}_1 \times \mathcal{M}_2$	Exercise 5.4
Product of Riemannian connections is Riemannian	Exercise 5.13
Hessian of $f\colon \mathcal{M}_1 \times \mathcal{M}_2 \to \mathbb{R}$	Example 5.19
Covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$ induced by product ∇	Exercise 5.34
Geodesics on Riemannian product manifold	Exercise 5.39
Exponential map on Riemannian product manifold	Exercise 10.32
Riemannian distance on Riemannian product	Exercise 10.14
Parallel transport on product manifold	Exercise 10.39

Table 7.2 The product $\mathcal{M}_1 \times \mathcal{M}_2$ of two manifolds is a manifold. Moreover, if we know how to work on $\mathcal{M}_1$ and $\mathcal{M}_2$ separately, then it is easy to work on their product as well. This table points to the relevant facts to do that in various places of this book.

Riemannian submanifold of $\mathcal{E}$, that is, if the inner product on $\mathcal{M}$ is the same as the inner product on $\mathcal{E}$ (appropriately restricted), then Proposition 3.61 states the following: given a smooth $f\colon \mathcal{M} \to \mathbb{R}$ with smooth extension $\bar{f}\colon U \to \mathbb{R}$ defined on a neighborhood U of $\mathcal{M}$ in $\mathcal{E}$,

$$\mathrm{grad} f(x) = \mathrm{Proj}_{\mathcal{M}}(\mathrm{grad} \bar{f}(x)). \tag{7.4}$$

For example, with the usual inner product on $\mathcal{E} = \mathbb{R}^{n\times n}$ (3.14), with $\mathcal{M} = \mathrm{Sym}(n)$ as a Riemannian submanifold, $\mathrm{Proj}_{\mathcal{M}}(Z) = \frac{Z+Z^\top}{2}$ so that the gradient of a function on $\mathrm{Sym}(n)$ is simply the symmetric part of its classical gradient on all of $\mathbb{R}^{n\times n}$.

Of course, we could endow $\mathcal{E}$ with a non-Euclidean Riemannian metric, that is, with a Riemannian metric which varies from point to point: see Exercise 7.1.

Second-order tools

Covariant derivatives (∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$) on a Euclidean space $\mathcal{E}$ coincide with the usual vector field derivatives. The Riemannian Hessian of a function $f\colon \mathcal{E} \to \mathbb{R}$ coincides with its Euclidean Hessian. The retraction $\mathrm{R}_x(v) = x + v$ is a second-order retraction (Definition 5.42). In fact, it is the exponential map (Section 10.2).

Further consider the case where $\mathcal{M}$ is a linear subspace and a Riemannian submanifold of $\mathcal{E}$. Then, continuing with the same notation as above, ∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$ are still the usual vector field derivatives, and the Hessian of f is related to that of $\bar{f}$ through

$$\mathrm{Hess} f(x)[v] = \mathrm{Proj}_{\mathcal{M}}\big(\mathrm{Hess} \bar{f}(x)[v]\big) \tag{7.5}$$

for all $x, v \in \mathcal{M}$. We can also write this symmetrically as:

$$\mathrm{Hess} f(x) = \mathrm{Proj}_{\mathcal{M}} \circ \mathrm{Hess} \bar{f}(x) \circ \mathrm{Proj}_{\mathcal{M}}. \tag{7.6}$$

The retraction $\mathrm{R}_x(v) = x + v$ retains the aforementioned properties on $\mathcal{M}$.

Exercise 7.1. *We could endow a linear space with a non-Euclidean Riemannian metric, that is, with a Riemannian metric which varies from point to point. To be explicit, let $\mathcal{M}$ denote the manifold $\mathbb{R}^n$ with the Riemannian metric*

$$\langle u, v \rangle_x = u^\top G(x) v,$$

where $G(x) \in \mathrm{Sym}(n)$ is a positive definite matrix which varies smoothly with x. The retraction $\mathrm{R}_x(v) = x + v$ is still acceptable since retractions are defined independently of the Riemannian structure.

Given a smooth function $\bar{f} \colon \mathbb{R}^n \to \mathbb{R}$, we can formally define $f \colon \mathcal{M} \to \mathbb{R}$ through $f(x) = \bar{f}(x)$ for all x. This way, $\mathrm{grad} f$ denotes the Riemannian gradient of f on $\mathcal{M}$ and $\mathrm{grad} \bar{f}$ denotes the Euclidean gradient of $\bar{f}$ on $\mathbb{R}^n$, where the latter is equipped with the canonical inner product $\langle u, v \rangle = u^\top v$. Give a formula for $\mathrm{grad} f(x)$ in terms of $\mathrm{grad} \bar{f}(x)$.

Consider the special case where the Hessian of $\bar{f}$ is everywhere positive definite ($\bar{f}$ is strictly convex) and we let $G(x) = \mathrm{Hess} \bar{f}(x)$. Compare the classical Newton method on $\bar{f}$ and Riemannian gradient descent on f.

7.2 The unit sphere in a Euclidean space

Let $\mathcal{E}$ be a Euclidean space endowed with an inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$. For example, this could be $\mathbb{R}^d$ with the metric $\langle u, v \rangle = u^\top v$, or it could be $\mathbb{R}^{n \times p}$ with the metric $\langle U, V \rangle = \mathrm{Tr}(U^\top V)$. With $d = \dim \mathcal{E}$, we define the unit sphere in $\mathcal{E}$ as

$$\mathrm{S}^{d-1} = \{x \in \mathcal{E} : \|x\| = 1\}. \tag{7.7}$$

A defining function is $h(x) = \langle x, x \rangle - 1$. Its differential is $\mathrm{D}h(x)[v] = 2 \langle x, v \rangle$, so that

$$\mathrm{T}_x \mathrm{S}^{d-1} = \{v \in \mathcal{E} : \langle x, v \rangle = 0\}, \tag{7.8}$$

and $\dim \mathrm{S}^{d-1} = \dim \mathcal{E} - 1 = d - 1$. One possible retraction is

$$\mathrm{R}_x(v) = \frac{x + v}{\|x + v\|} = \frac{x + v}{\sqrt{1 + \|v\|^2}}. \tag{7.9}$$

The orthogonal projector to the tangent space at x is

$$\mathrm{Proj}_x \colon \mathcal{E} \to \mathrm{T}_x \mathrm{S}^{d-1} \colon u \mapsto \mathrm{Proj}_x(u) = u - \langle x, u \rangle x. \tag{7.10}$$

Equip S^{d-1} with the induced Riemannian metric to turn it into a Riemannian submanifold. Then, for a smooth function $f \colon \mathrm{S}^{d-1} \to \mathbb{R}$ with smooth extension

$\bar{f}: U \to \mathbb{R}$ in a neighborhood U of S^{d-1} in $\mathcal{E}$, the gradient of f is given by Proposition 3.61 as

$$\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad}\bar{f}(x)) = \mathrm{grad}\bar{f}(x) - \left\langle x, \mathrm{grad}\bar{f}(x)\right\rangle x. \tag{7.11}$$

In particular, x is a critical point of f if and only if $\mathrm{grad}\bar{f}(x)$ is parallel to x.

In Manopt, formulas such as (7.11) which convert the Euclidean gradient of a smooth extension into a Riemannian gradient are available for each manifold as `egrad2rgrad`.

A product of k spheres is called an *oblique manifold*. For example, the product of k spheres in $\mathbb{R}^d$ is denoted by $\mathrm{OB}(d,k) = (\mathrm{S}^{d-1})^k$. Its elements are typically represented using matrices in $\mathbb{R}^{d\times k}$ (or $\mathbb{R}^{k\times d}$) whose columns (or rows) have unit norm. The same can be done for complex matrices. An often useful particular case is the *complex circle*, which consists of all complex numbers of unit modulus (called *phases*): this is nothing but an alternative way of representing S^1.

Second-order tools

With S^{d-1} as a Riemannian submanifold of the Euclidean space $\mathcal{E}$, covariant derivatives (∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$) on S^{d-1} coincide with the usual vector field derivatives (of smooth extensions) in $\mathcal{E}$, followed by orthogonal projection to tangent spaces (Theorem 5.9, Proposition 5.31).

We can use this to obtain a formula for the Riemannian Hessian of $f: \mathrm{S}^{d-1} \to \mathbb{R}$, with smooth extension $\bar{f}: U \to \mathbb{R}$ defined on a neighborhood U of S^{d-1} in $\mathcal{E}$. Following Example 5.17, we let

$$\bar{G}(x) = \mathrm{grad}\bar{f}(x) - \left\langle x, \mathrm{grad}\bar{f}(x)\right\rangle x$$

denote a smooth extension of the vector field $\mathrm{grad} f$ to a neighborhood of S^{d-1} in $\mathcal{E}$. Then,

$$\begin{aligned}
\mathrm{Hess} f(x)[v] &= \nabla_v \mathrm{grad} f \\
&= \mathrm{Proj}_x\big(\mathrm{D}\bar{G}(x)[v]\big) \\
&= \mathrm{Proj}_x\Big(\mathrm{Hess}\bar{f}(x)[v] - \big[\left\langle v, \mathrm{grad}\bar{f}(x)\right\rangle + \left\langle x, \mathrm{Hess}\bar{f}(x)[v]\right\rangle\big]\, x \\
&\qquad\qquad - \left\langle x, \mathrm{grad}\bar{f}(x)\right\rangle v\Big) \\
&= \mathrm{Proj}_x(\mathrm{Hess}\bar{f}(x)[v]) - \left\langle x, \mathrm{grad}\bar{f}(x)\right\rangle v.
\end{aligned} \tag{7.12}$$

In Manopt, formulas such as (7.12) which convert the Euclidean gradient and Hessian of a smooth extension into a Riemannian Hessian are available for each manifold as `ehess2rhess`.

The retraction (7.9) is a second-order retraction (see Definition 5.42, Example 5.43 and Proposition 5.55). Geodesics on S^{d-1} are given in Example 5.37.

7.3 The Stiefel manifold: orthonormal matrices

For $p \leq n$, let $\mathbb{R}^{n\times p}$ be endowed with the standard inner product $\langle U, V\rangle = \mathrm{Tr}(U^\top V)$. The (compact)[1] *Stiefel manifold* is the set of matrices in $\mathbb{R}^{n\times p}$ whose columns are *orthonormal* in $\mathbb{R}^n$ with respect to the inner product $\langle u, v\rangle = u^\top v$. This can be written conveniently as:[2]

$$\mathrm{St}(n,p) = \left\{X \in \mathbb{R}^{n\times p} : X^\top X = I_p\right\}, \tag{7.13}$$

where I_p is the identity matrix of size p. In particular, $\mathrm{St}(n,1)$ is the unit sphere in $\mathbb{R}^n$. We call matrices in $\mathrm{St}(n,p)$ *orthonormal matrices* and we reserve the word *orthogonal matrix* for square orthonormal matrices.

Consider the following function:

$$h\colon \mathbb{R}^{n\times p} \to \mathrm{Sym}(p)\colon X \mapsto h(X) = X^\top X - I_p, \tag{7.14}$$

where $\mathrm{Sym}(p)$ is the linear space of symmetric matrices of size p. The latter has dimension $k = \frac{p(p+1)}{2}$, so that we may identify it with $\mathbb{R}^k$ if desired. We can verify that h is a defining function for $\mathrm{St}(n,p)$. Indeed, h is smooth and $h^{-1}(0) = \mathrm{St}(n,p)$: it remains to check that the differential of h has rank k for all $X \in \mathrm{St}(n,p)$. To this end, consider $\mathrm{D}h(X)\colon \mathbb{R}^{n\times p} \to \mathrm{Sym}(p)$:

$$\begin{aligned}\mathrm{D}h(X)[V] &= \lim_{t\to 0} \frac{h(X+tV) - h(X)}{t}\\ &= \lim_{t\to 0} \frac{(X+tV)^\top(X+tV) - X^\top X}{t}\\ &= X^\top V + V^\top X.\end{aligned} \tag{7.15}$$

To show $\mathrm{D}h(X)$ has rank k, we must show its image (or range) is a linear subspace of dimension k. Since the codomain $\mathrm{Sym}(p)$ has dimension k, we must show that the image of $\mathrm{D}h(X)$ is all of $\mathrm{Sym}(p)$, that is, $\mathrm{D}h(X)$ is *surjective*. To do so, consider $V = \frac{1}{2}XA$ with $A \in \mathrm{Sym}(p)$ arbitrary. Then,

$$\mathrm{D}h(X)[V] = \frac{1}{2}X^\top XA + \frac{1}{2}A^\top X^\top X = A.$$

In other words: for any matrix $A \in \mathrm{Sym}(p)$, there exists a matrix $V \in \mathbb{R}^{n\times p}$ such that $\mathrm{D}h(X)[V] = A$. This confirms the image of $\mathrm{D}h(X)$ is all of $\mathrm{Sym}(p)$, so that it has rank k. Thus, h is a defining function for $\mathrm{St}(n,p)$, making it an embedded submanifold of $\mathbb{R}^{n\times p}$ of dimension

$$\dim \mathrm{St}(n,p) = \dim \mathbb{R}^{n\times p} - \dim \mathrm{Sym}(p) = np - \frac{p(p+1)}{2}. \tag{7.16}$$

The tangent spaces are subspaces of $\mathbb{R}^{n\times p}$:

$$\mathrm{T}_X\mathrm{St}(n,p) = \ker \mathrm{D}h(X) = \left\{V \in \mathbb{R}^{n\times p} : X^\top V + V^\top X = 0\right\} \tag{7.17}$$

[1] The *non-compact* Stiefel manifold refers to the open subset of matrices of rank p in $\mathbb{R}^{n\times p}$. We always mean compact.

[2] Many authors use the notation $\mathrm{St}(p,n)$ for the same set—we prefer the notation $\mathrm{St}(n,p)$ as it is reminiscent of the size of the matrices.

It is sometimes convenient to parameterize tangent vectors in explicit form. First, complete[3] the orthonormal basis formed by the columns of X with a matrix $X_\perp \in \mathbb{R}^{n\times(n-p)}$ such that $[\,X\ X_\perp\,] \in \mathbb{R}^{n\times n}$ is orthogonal:

$$X^\top X = I_p, \qquad X_\perp^\top X_\perp = I_{n-p}, \qquad \text{and} \qquad X^\top X_\perp = 0. \tag{7.18}$$

Since $[\,X\ X_\perp\,]$ is, in particular, invertible, any matrix $V \in \mathbb{R}^{n\times p}$ can be written as

$$V = \begin{bmatrix} X & X_\perp \end{bmatrix} \begin{bmatrix} \Omega \\ B \end{bmatrix} = X\Omega + X_\perp B, \tag{7.19}$$

for a unique choice of $\Omega \in \mathbb{R}^{p\times p}$ and $B \in \mathbb{R}^{(n-p)\times p}$. Using this decomposition, V is a tangent vector at X if and only if

$$0 = \mathrm{D}h(X)[V] = X^\top(X\Omega + X_\perp B) + (X\Omega + X_\perp B)^\top X = \Omega + \Omega^\top.$$

In other words, Ω must be skew-symmetric, while B is free. Thus,

$$\mathrm{T}_X\mathrm{St}(n,p) = \left\{ X\Omega + X_\perp B : \Omega \in \mathrm{Skew}(p), B \in \mathbb{R}^{(n-p)\times p} \right\}, \tag{7.20}$$

where we used the decomposition (7.19) with respect to an arbitrary choice of $X_\perp \in \mathbb{R}^{n\times(n-p)}$ satisfying (7.18), and

$$\mathrm{Skew}(p) = \{\Omega \in \mathbb{R}^{p\times p} : \Omega^\top = -\Omega\} \tag{7.21}$$

is the set of skew-symmetric matrices of size p.

One popular retraction for $\mathrm{St}(n,p)$ is the *Q-factor retraction*:[4]

$$\mathrm{R}_X(V) = Q, \tag{7.22}$$

where $QR = X + V$ is a (thin) QR decomposition: $Q \in \mathrm{St}(n,p)$ and $R \in \mathbb{R}^{p\times p}$ upper triangular with nonnegative diagonal entries. This is well defined since, for a tangent vector $V \in \mathrm{T}_X\mathrm{St}(n,p)$,

$$(X+V)^\top(X+V) = I_p + V^\top V \tag{7.23}$$

is positive definite, showing $X+V$ has full rank p: under that condition, the QR decomposition is indeed unique. This retraction can be computed in $\sim np^2$ basic arithmetic operations $(+,-,\times,/,\sqrt{\cdot})$ using the modified Gram–Schmidt algorithm or a Householder triangularization. The defining properties of a retraction are satisfied: Surely, $\mathrm{R}_X(0) = X$; Furthermore, inspecting the Gram–Schmidt algorithm reveals that it maps full-rank matrices in $\mathbb{R}^{n\times p}$ to their Q-factor through a sequence of smooth operations, so that R is smooth (by composition); Finally, an expression for $\mathrm{DR}_X(V)$ is derived in [AMS08, Ex. 8.1.5], from which it is straightforward to verify that $\mathrm{DR}_X(0)$ is the identity map.

[3] The matrix $X_\perp$ is *never* built explicitly: it is merely a useful mathematical tool.

[4] Some software packages offer a built-in `qr` routine which may not enforce nonnegativity of diagonal entries of R—this is the case of Matlab for example. It is important to flip the signs of the columns of Q accordingly. In Manopt, call `qr_unique`.

Another popular retraction for $\mathrm{St}(n, p)$ is the *polar retraction*:

$$\begin{aligned}\mathrm{R}_X(V) &= (X+V)\left((X+V)^\top(X+V)\right)^{-1/2} \\ &= (X+V)(I_p + V^\top V)^{-1/2}, \end{aligned} \tag{7.24}$$

where $M^{-1/2}$ denotes the inverse matrix square root of M. This can be computed through eigenvalue decomposition of the matrix $I_p + V^\top V$, or (better) through SVD of $X+V$. Indeed, if $X+V = U\Sigma W^\top$ is a thin singular value decomposition, the polar factor of $X + V$ is $UW^\top$ and that is equivalent to (7.24). Clearly, $\mathrm{R}_X(0) = X$ and R is smooth. It is straightforward to check that $\mathrm{DR}_X(0)$ is the identity map. In fact, the polar retraction is the metric projection retraction (see Section 5.12 and [Sch66]), globally well defined since $X + V$ has full rank for all $(X, V) \in \mathrm{T}\mathcal{M}$ as argued above [AM12, Prop. 7].

Yet another often used retraction for the Stiefel manifold is the *Cayley transform* [WY13, JD15].

The orthogonal projector to a tangent space of $\mathrm{St}(n, p)$ must be such that $U - \mathrm{Proj}_X(U)$ is orthogonal to $\mathrm{T}_X\mathrm{St}(n, p)$, that is, the difference must be in the orthogonal complement of the tangent space in $\mathbb{R}^{n\times p}$. The latter is called the *normal space* to $\mathrm{St}(n, p)$ at X:

$$\begin{aligned}\mathrm{N}_X\mathrm{St}(n,p) &= (\mathrm{T}_X\mathrm{St}(n,p))^\perp \\ &= \left\{U \in \mathbb{R}^{n\times p} : \langle U, V\rangle = 0 \text{ for all } V \in \mathrm{T}_X\mathrm{St}(n,p)\right\} \\ &= \left\{U \in \mathbb{R}^{n\times p} : \langle U, X\Omega + X_\perp B\rangle = 0 \right. \\ &\qquad\qquad \left. \text{for all } \Omega \in \mathrm{Skew}(p), B \in \mathbb{R}^{(n-p)\times p}\right\}.\end{aligned}$$

Expand normal vectors as $U = XA + X_\perp C$ with some $A \in \mathbb{R}^{p\times p}$ and $C \in \mathbb{R}^{(n-p)\times p}$; then:

$$\begin{aligned}\mathrm{N}_X\mathrm{St}(n,p) &= \left\{U \in \mathbb{R}^{n\times p} : \langle XA + X_\perp C, X\Omega + X_\perp B\rangle = 0 \right. \\ &\qquad\qquad \left. \text{for all } \Omega \in \mathrm{Skew}(p), B \in \mathbb{R}^{(n-p)\times p}\right\} \\ &= \left\{U \in \mathbb{R}^{n\times p} : \langle A, \Omega\rangle = 0 \text{ and } \langle C, B\rangle = 0 \right. \\ &\qquad\qquad \left. \text{for all } \Omega \in \mathrm{Skew}(p), B \in \mathbb{R}^{(n-p)\times p}\right\} \\ &= \left\{XA : A \in \mathrm{Sym}(p)\right\}, \end{aligned} \tag{7.25}$$

where we used that the orthogonal complement of $\mathrm{Skew}(p)$ in $\mathbb{R}^{p\times p}$ is $\mathrm{Sym}(p)$. Thus, orthogonal projection of $U \in \mathbb{R}^{n\times p}$ satisfies

$$U - \mathrm{Proj}_X(U) = XA$$

for some symmetric matrix A. Furthermore, the projected vector must lie in $\mathrm{T}_X\mathrm{St}(n, p)$, hence

$$\mathrm{Proj}_X(U)^\top X + X^\top \mathrm{Proj}_X(U) = 0.$$

Plugging the former into the latter yields

$$(U - XA)^\top X + X^\top(U - XA) = 0,$$

that is, $U^\top X + X^\top U = 2A$. Hence,

$$\mathrm{Proj}_X(U) = U - X\frac{X^\top U + U^\top X}{2} \tag{7.26}$$
$$= (I - XX^\top)U + X\frac{X^\top U - U^\top X}{2}. \tag{7.27}$$

One convenient way to turn $\mathrm{St}(n, p)$ into a Riemannian manifold is to make it a Riemannian submanifold of $\mathbb{R}^{n\times p}$, in which case the projector yields a convenient formula for the gradient of smooth functions f in terms of a smooth extension $\bar{f}$, by Proposition 3.61:

$$\mathrm{grad} f(X) = \mathrm{Proj}_X(\mathrm{grad}\bar{f}(X)) = \mathrm{grad}\bar{f}(X) - X\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X)), \tag{7.28}$$

where $\mathrm{sym}(M) = \frac{M+M^\top}{2}$ extracts the symmetric part of a matrix.

Other Riemannian metrics are sometimes used: see for example the so-called *canonical metric* in [EAS98].

Second-order tools

With $\mathrm{St}(n, p)$ as a Riemannian submanifold of $\mathbb{R}^{n\times p}$, covariant derivatives (∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$) on $\mathrm{St}(n, p)$ coincide with the usual vector field derivatives (of smooth extensions) in $\mathbb{R}^{n\times p}$, followed by orthogonal projection to tangent spaces (Theorem 5.9, Proposition 5.31).

We use this to obtain a formula for the Riemannian Hessian of $f\colon \mathrm{St}(n, p) \to \mathbb{R}$, with smooth extension $\bar{f}$ defined on a neighborhood of $\mathrm{St}(n, p)$ in $\mathbb{R}^{n\times p}$. Let

$$\bar{G}(X) = \mathrm{grad}\bar{f}(X) - X\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X))$$

denote a smooth extension of the vector field $\mathrm{grad} f$ to a neighborhood of $\mathrm{St}(n, p)$ in $\mathbb{R}^{n\times p}$. Then,

$$\begin{aligned}
\mathrm{Hess} f(X)[V] &= \nabla_V \mathrm{grad} f \\
&= \mathrm{Proj}_X\big(\mathrm{D}\bar{G}(X)[V]\big) \\
&= \mathrm{Proj}_X\Big(\mathrm{Hess}\bar{f}(X)[V] - V\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X)) - XS\Big) \\
&= \mathrm{Proj}_X\Big(\mathrm{Hess}\bar{f}(X)[V] - V\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X))\Big),
\end{aligned} \tag{7.29}$$

where $S = \mathrm{sym}(V^\top \mathrm{grad}\bar{f}(X) + X^\top \mathrm{Hess}\bar{f}(X)[V])$, and XS vanishes through Proj_X. The polar retraction (7.24) is a second-order retraction (Definition 5.42) because it is the metric projection retraction (Proposition 5.55), but the Q-factor retraction (7.22) is not. Geodesics on $\mathrm{St}(n, p)$ are given in [AMS08, eq. (5.26)].

Exercise 7.2. *Show that the polar retraction* R *on* $\mathcal{M} = \mathrm{St}(n, p)$ (7.24) *is such that* $E(X, V) = (X, \mathrm{R}_X(V))$ *from* $\mathrm{T}\mathcal{M}$ *to* $E(\mathrm{T}\mathcal{M})$ *has a smooth inverse.*

7.4 The orthogonal group and rotation matrices

As a special case of the Stiefel manifold, matrices in $\mathrm{St}(n, n)$ form the *orthogonal group*, that is, the set of orthogonal matrices in $\mathbb{R}^{n\times n}$:

$$\mathrm{O}(n) = \{X \in \mathbb{R}^{n\times n} : X^\top X = XX^\top = I_n\}. \tag{7.30}$$

It is a group equipped with matrix multiplication as its group operation. Being a special case of the Stiefel manifold, $\mathrm{O}(n)$ is also an embedded submanifold, this time of $\mathbb{R}^{n\times n}$. As a set which is both a manifold and a group, it is known as a *Lie group* (more about this in Section 9.2). It has dimension

$$\dim \mathrm{O}(n) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}, \tag{7.31}$$

and tangent spaces given by

$$\mathrm{T}_X\mathrm{O}(n) = \left\{X\Omega \in \mathbb{R}^{n\times n} : \Omega \in \mathrm{Skew}(n)\right\} = X\mathrm{Skew}(n). \tag{7.32}$$

Notice how $\mathrm{T}_{I_n}\mathrm{O}(n) = \mathrm{Skew}(n)$, so that $\mathrm{T}_X\mathrm{O}(n) = X\mathrm{T}_{I_n}\mathrm{O}(n)$: tangent spaces are essentially "translated" versions of the tangent space at the identity matrix, which is also the identity element of $\mathrm{O}(n)$ as a group. In Lie group parlance, we call $\mathrm{T}_{I_n}\mathrm{O}(n)$ the *Lie algebra* of $\mathrm{O}(n)$.

Numerically, it is convenient to represent tangent vectors at X simply by their skew-symmetric factor Ω, keeping in mind that we mean to represent the tangent vector $X\Omega$. More generally, it is important to mind the distinction between how we represent points and vectors in the ambient space, and how we represent points and tangent vectors on the manifold.

Both the Q-factor and the polar retractions of $\mathrm{St}(n, p)$ are valid retractions for $\mathrm{O}(n)$.

The orthogonal projector is given by

$$\mathrm{Proj}_X(U) = X\frac{X^\top U - U^\top X}{2} = X\,\mathrm{skew}(X^\top U), \tag{7.33}$$

where $\mathrm{skew}(M) = \frac{M - M^\top}{2}$ extracts the skew-symmetric part of a matrix. Turning $\mathrm{O}(n)$ into a Riemannian submanifold of $\mathbb{R}^{n\times n}$ with the standard Euclidean metric, this once more gives a direct formula for the gradient of a smooth function on $\mathrm{O}(n)$, through Proposition 3.61:

$$\mathrm{grad} f(X) = X\,\mathrm{skew}(X^\top \mathrm{grad}\bar{f}(X)). \tag{7.34}$$

Of course, this is equivalent to the corresponding formula (7.28) for Stiefel.

An important feature of $\mathrm{O}(n)$, relevant for optimization, is that it is *disconnected*. Specifically, it has two components, corresponding to orthogonal matrices of determinant $+1$ and -1:

$$1 = \det(I_n) = \det(XX^\top) = \det(X)^2.$$

Indeed: since the determinant is a continuous function from $\mathbb{R}^{n\times n}$ to $\mathbb{R}$, by the intermediate value theorem, any continuous curve connecting a matrix with determinant $+1$ to a matrix with determinant -1 must pass through a matrix with determinant zero, hence, must leave $\mathrm{O}(n)$.

Our optimization algorithms move along continuous curves (retraction curves). As a result, when we initialize such an algorithm in a certain connected component, it cannot "jump" to another connected component. Therefore, it is important to initialize in the appropriate component. Geometrically, orthogonal matrices of size n correspond to rotations of $\mathbb{R}^n$, possibly composed with a reflection for those matrices that have determinant -1. In situations where only rotations are relevant, it makes sense to consider the *special orthogonal group*, also known as the *group of rotations*:

$$\mathrm{SO}(n) = \{X \in \mathrm{O}(n) : \det(X) = +1\}. \tag{7.35}$$

This is still an embedded submanifold of $\mathbb{R}^{n\times n}$ of course. To verify it, consider the defining function $h(X) = X^\top X - I_n$ defined on $\{X \in \mathbb{R}^{n\times n} : \det(X) > 0\}$, which is an open subset of $\mathbb{R}^{n\times n}$.

As a connected component of $\mathrm{O}(n)$, all the tools we developed so far apply just as well to $\mathrm{SO}(n)$. This includes eq. (7.34) for gradients as well as

$$\dim \mathrm{SO}(n) = \frac{n(n-1)}{2}, \tag{7.36}$$

$$\mathrm{T}_X\mathrm{SO}(n) = X\mathrm{Skew}(n), \tag{7.37}$$

$$\mathrm{Proj}_X(U) = X\,\mathrm{skew}(X^\top U). \tag{7.38}$$

It is clear that retractions on $\mathrm{O}(n)$ yield retractions on $\mathrm{SO}(n)$ since, being smooth, they cannot leave a connected component.

Second-order tools

With $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ as Riemannian submanifolds of the Euclidean space $\mathbb{R}^{n\times n}$, covariant derivatives (∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$) coincide with the usual vector field derivatives (of smooth extensions) in $\mathbb{R}^{n\times n}$, followed by orthogonal projection to tangent spaces (Theorem 5.9, Proposition 5.31).

We use this to obtain a formula for the Riemannian Hessian of a real function f on $\mathrm{O}(n)$ or $\mathrm{SO}(n)$, with smooth extension $\bar{f}$ in $\mathbb{R}^{n\times n}$. Of course, exactly the same developments as for the Stiefel manifold hold, so that by (7.29) we get:

$$\mathrm{Hess}f(X)[V] = \mathrm{Proj}_X\left(\mathrm{Hess}\bar{f}(X)[V] - V\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X))\right). \tag{7.39}$$

Writing $V = X\Omega$ for some $\Omega \in \mathrm{Skew}(n)$, this also reads

$$\mathrm{Hess}f(X)[X\Omega] = X\,\mathrm{skew}\left(X^\top \mathrm{Hess}\bar{f}(X)[V] - \Omega\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X))\right),$$

making the skew-symmetric representation of $\mathrm{Hess}f(X)[X\Omega]$ clearer.

For both $\mathrm{O}(n)$ and $\mathrm{SO}(n)$, the polar retraction (7.24) is the metric projection retraction (because it was so for the Stiefel manifold) hence it is a second-order retraction (see Section 5.12), but the Q-factor retraction (7.22) is not. It is an exercise to show that

$$c(t) = X \exp(t\Omega) \tag{7.40}$$

is a geodesic on $\mathrm{O}(n)$ (or $\mathrm{SO}(n)$) such that $c(0) = X$ and $c'(0) = X\Omega$. (This happens because the Riemannian metric is bi-invariant, so that the Lie exponential map and the Riemannian exponential map coincide, and it is known that the Lie exponential map is given by the matrix exponential exp.)

Exercise 7.3. *Show that $c(t)$ as defined by (7.40) is indeed a curve on $\mathrm{O}(n)$, and verify that $\frac{\mathrm{d}}{\mathrm{d}t}c(t) = c(t)\Omega$. Deduce that $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}}{\mathrm{d}t}c(t)\right) = c(t)\Omega^2$ and, eventually, that $c''(t) = \frac{\mathrm{D}}{\mathrm{d}t}c'(t) = 0$, which confirms c is a geodesic. Hint: use (4.33) to express the differential of the matrix exponential, and use the fact that $\exp(A + B) = \exp(A)\exp(B)$ if A and B commute.*

Exercise 7.4. *Work out a geometric toolbox for the* unitary group

$$\mathrm{U}(n) = \{X \in \mathbb{C}^{n\times n} : X^*X = I_n\} \tag{7.41}$$

as a Riemannian submanifold of $\mathbb{C}^{n\times n}$ with the usual inner product (3.17).

7.5 Fixed-rank matrices

The set of real matrices of size $m \times n$ and rank r,

$$\mathbb{R}_r^{m\times n} = \{X \in \mathbb{R}^{m\times n} : \mathrm{rank}(X) = r\}, \tag{7.42}$$

is an embedded submanifold of $\mathbb{R}^{m\times n}$, as we now show. Importantly, this is only true for *fixed* rank r: the set of matrices in $\mathbb{R}^{m\times n}$ with rank *up to* r is *not* an embedded submanifold of $\mathbb{R}^{m\times n}$. It is, however, an *algebraic variety* and a *stratified space*—we do not consider optimization on such spaces. Moreover, in contrast to the examples discussed earlier in this chapter, $\mathbb{R}_r^{m\times n}$ is neither open nor closed in $\mathbb{R}^{m\times n}$.

For an arbitrary $X \in \mathbb{R}_r^{m\times n}$, we now build a local defining function. We cannot use $h(X) = \mathrm{rank}(X) - r$ as a defining function because it is not continuous, let alone smooth. Instead, we proceed as follows. Since X has rank r, it contains an invertible submatrix of size $r \times r$, that is, it is possible to extract r columns and r rows of X such that the resulting matrix in $\mathbb{R}^{r\times r}$ is invertible. For notational convenience, assume for now that this is the case for the first r rows and columns, so that X can be written in block form as

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$$

with $X_{11} \in \mathbb{R}^{r\times r}$ invertible, and $X_{12} \in \mathbb{R}^{r\times(n-r)}$, $X_{21} \in \mathbb{R}^{(m-r)\times r}$ and $X_{22} \in$

$\mathbb{R}^{(m-r)\times(n-r)}$. Since X has rank r, its $n-r$ last columns must be linear combinations of its r first columns, that is, there exists $W \in \mathbb{R}^{r\times(n-r)}$ such that

$$\begin{bmatrix} X_{12} \\ X_{22} \end{bmatrix} = \begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} W.$$

Consequently, $W = X_{11}^{-1}X_{12}$ and $X_{22} = X_{21}W = X_{21}X_{11}^{-1}X_{12}$. Under our assumption that X_{11} is invertible, this relationship between the blocks of X is necessary and sufficient for X to have rank r.

This suggests a candidate local defining function. Let $\mathcal{U}$ be the subset of $\mathbb{R}^{m\times n}$ consisting of all matrices whose upper-left submatrix of size $r \times r$ is invertible: X is in $\mathcal{U}$, and $\mathcal{U}$ is open in $\mathbb{R}^{m\times n}$ since its complement—the set of matrices whose upper-left submatrix has determinant equal to zero—is closed. Consider

$$h\colon \mathcal{U} \to \mathbb{R}^{(m-r)\times(n-r)} : Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \mapsto h(Y) = Y_{22} - Y_{21}Y_{11}^{-1}Y_{12},$$

with the same block-matrix structure as before. By the above, $h^{-1}(0) = \mathbb{R}_r^{m\times n} \cap \mathcal{U}$. Furthermore, h is smooth in $\mathcal{U}$. Finally, its differential at Y is ($V \in \mathbb{R}^{m\times n}$ has the same block structure as Y):

$$\mathrm{D}h(Y)[V] = V_{22} - V_{21}Y_{11}^{-1}Y_{12} + Y_{21}Y_{11}^{-1}V_{11}Y_{11}^{-1}Y_{12} - Y_{21}Y_{11}^{-1}V_{12},$$

where we used the following identity for the differential of the matrix inverse (recall Example 4.24):

$$\mathrm{D}\big(M \mapsto M^{-1}\big)(M)[H] = -M^{-1}HM^{-1}. \tag{7.43}$$

The codomain of $\mathrm{D}h(Y)$ is $\mathbb{R}^{(m-r)\times(n-r)}$. Any matrix in that codomain can be attained with some input V (simply consider setting V_{11}, V_{12}, V_{21} to zero, so that $\mathrm{D}h(Y)[V] = V_{22}$). Thus, the differential of h is surjective everywhere in $\mathcal{U}$: it is a local defining function for $\mathbb{R}_r^{m\times n}$ around X. If the upper-left submatrix of size $r \times r$ of X is not invertible, we can construct another local defining function using the same procedure: one for each choice of submatrix.

Together, these local defining functions cover the whole set, showing that $\mathbb{R}_r^{m\times n}$ is an embedded submanifold of $\mathbb{R}^{m\times n}$ with dimension

$$\begin{aligned} \dim \mathbb{R}_r^{m\times n} &= \dim \mathbb{R}^{m\times n} - \dim \mathbb{R}^{(m-r)\times(n-r)} \\ &= mn - (m-r)(n-r) \\ &= r(m+n-r). \end{aligned} \tag{7.44}$$

Notice that, for a given rank r, the dimension of $\mathbb{R}_r^{m\times n}$ grows *linearly* with $m+n$, as opposed to the dimension of the embedding space $\mathbb{R}^{m\times n}$ which grows much faster, as mn. To exploit this key feature in numerical algorithms, we must *represent* X appropriately in memory: this should make it possible to store matrices with an amount of memory that grows linearly in $m+n$ even though their size is $m \times n$. One convenient choice is as a thin singular value decomposition:

$$X = U\Sigma V^\top, \qquad U \in \mathrm{St}(m,r), \qquad \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}, \tag{7.45}$$
$$V \in \mathrm{St}(n,r),$$

where $\sigma_1 \geq \cdots \geq \sigma_r > 0$ are the singular values of X. To identify X uniquely, it is only necessary to store U, Σ, V in memory. We stress that this is only about representation: the use of orthonormal matrices is only for convenience, and has no bearing on the geometry of $\mathbb{R}_r^{m\times n}$.

The tangent space to $\mathbb{R}_r^{m\times n}$ at X is given by the kernel of $\mathrm{D}h(X)$, with an appropriately chosen h. However, this characterization is impractical because it requires one to identify an invertible submatrix of X in order to determine which local defining function to use. Besides, it is more convenient to aim for a representation of the tangent space $\mathrm{T}_X\mathbb{R}_r^{m\times n}$ that is compatible with the practical representation of X (7.45).

Since we know that each tangent space has dimension as in (7.44), it is sufficient to exhibit a linear subspace of that dimension which is included in the tangent space. Going back to the definition of tangent space (3.23), we do so by explicitly constructing smooth curves on $\mathbb{R}_r^{m\times n}$.

Given $X = U\Sigma V^\top$ as above, let $U(t)$ be a smooth curve on $\mathrm{St}(m,r)$ such that $U(0) = U$, let $V(t)$ be a smooth curve on $\mathrm{St}(n,r)$ such that $V(0) = V$, and let $\Sigma(t)$ be a smooth curve in the set of invertible matrices of size $r \times r$ (this is an open submanifold of $\mathbb{R}^{r\times r}$) such that $\Sigma(0) = \Sigma$. Then,

$$c(t) = U(t)\Sigma(t)V(t)^\top$$

is a smooth curve on $\mathbb{R}_r^{m\times n}$ such that $c(0) = X$. Hence, its velocity at zero is a tangent vector at X:

$$c'(0) = U'(0)\Sigma V^\top + U\Sigma'(0)V^\top + U\Sigma V'(0)^\top \in \mathrm{T}_X\mathbb{R}_r^{m\times n}.$$

Since $U(t)$ is a smooth curve on $\mathrm{St}(m,r)$ through U, its velocity $U'(0)$ is in the tangent space to $\mathrm{St}(m,r)$ at U. The other way around, for any vector in $\mathrm{T}_U\mathrm{St}(m,r)$, there is a smooth curve $U(t)$ with that velocity at $t = 0$. From (7.20), this means that for any $\Omega \in \mathrm{Skew}(r)$ and $B \in \mathbb{R}^{(m-r)\times r}$ we can arrange to have

$$U'(0) = U\Omega + U_\perp B,$$

where $U_\perp$ is such that $\begin{bmatrix} U & U_\perp \end{bmatrix}$ is orthogonal. Likewise, for any $\Omega' \in \mathrm{Skew}(r)$ and $C \in \mathbb{R}^{(n-r)\times r}$ we can arrange to have

$$V'(0) = V\Omega' + V_\perp C,$$

with $V_\perp$ such that $\begin{bmatrix} V & V_\perp \end{bmatrix}$ is orthogonal. Finally, since $\Sigma(t)$ is a smooth curve in an open submanifold of $\mathbb{R}^{r\times r}$, we can arrange for $\Sigma'(0)$ to be any matrix $A \in \mathbb{R}^{r\times r}$.

Overall, this shows that all of the following velocities are in the tangent space of $\mathbb{R}_r^{m\times n}$ at X:

$$\begin{aligned} c'(0) &= (U\Omega + U_\perp B)\Sigma V^\top + UAV^\top + U\Sigma(V\Omega' + V_\perp C)^\top \\ &= U(\underbrace{\Omega\Sigma + A - \Sigma\Omega'}_{M})V^\top + \underbrace{U_\perp B\Sigma}_{U_p} V^\top + U(\underbrace{V_\perp C\Sigma^\top}_{V_p})^\top. \end{aligned} \tag{7.46}$$

Since Σ is invertible, we find that any matrix of the form

$$UMV^\top + U_p V^\top + UV_p^\top$$

with $M \in \mathbb{R}^{r\times r}$ arbitrary and $U_p \in \mathbb{R}^{m\times r}, V_p \in \mathbb{R}^{n\times r}$ such that $U^\top U_p = V^\top V_p = 0$ is tangent at X. The conditions on U_p and V_p amount to $2r^2$ linear constraints, hence we have found a linear subspace of $\mathrm{T}_X\mathbb{R}_r^{m\times n}$ of dimension

$$r^2 + mr + nr - 2r^2 = r(m+n-r).$$

This coincides with the dimension of $\mathrm{T}_X\mathbb{R}_r^{m\times n}$ by (7.44). Thus, we have found the whole tangent space:

$$\begin{aligned} \mathrm{T}_X\mathbb{R}_r^{m\times n} = \big\{UMV^\top + U_pV^\top + UV_p^\top :& \\ & M \in \mathbb{R}^{r\times r}, U_p \in \mathbb{R}^{m\times r}, V_p \in \mathbb{R}^{n\times r}, \text{ and} \\ & U^\top U_p = 0, V^\top V_p = 0\big\}. \end{aligned} \tag{7.47}$$

Notice how, if X is already identified by the triplet (U, Σ, V), then to represent a tangent vector at X we only need small matrices M, U_p, V_p. These require essentially the same amount of memory as for storing X. Sometimes, it is convenient (for analysis, not computation) to write tangent vectors as follows:

$$\mathrm{T}_X\mathbb{R}_r^{m\times n} = \left\{ \begin{bmatrix} U & U_\perp \end{bmatrix} \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} \begin{bmatrix} V & V_\perp \end{bmatrix}^\top : A, B, C \text{ are arbitrary} \right\}. \tag{7.48}$$

This reveals the dimension of the tangent space even more explicitly.

To build a retraction for $\mathbb{R}_r^{m\times n}$, one possibility is to use metric projection (Section 5.12): make the step in the ambient space, then project back to the manifold. To project from $\mathbb{R}^{m\times n}$ to $\mathbb{R}_r^{m\times n}$, we first need to endow $\mathbb{R}^{m\times n}$ with a Euclidean metric: we choose the standard inner product, $\langle U, V\rangle = \mathrm{Tr}(U^\top V)$, with its induced norm $\|U\| = \sqrt{\langle U, U\rangle}$ (the Frobenius norm). Then, we construct the retraction as:

$$\mathrm{R}_X(H) = \operatorname*{arg\,min}_{Y \in \mathbb{R}_r^{m\times n}} \|X + H - Y\|^2. \tag{7.49}$$

Following the well-known Eckart–Young–Mirsky theorem, the solution to this optimization problem (when it exists) is given by the singular value decomposition of $X + H$ truncated at rank r. With X and H represented as above, this can be computed efficiently. Indeed, consider

$$\begin{aligned} X + H &= U(\Sigma + M)V^\top + U_pV^\top + UV_p^\top \\ &= \begin{bmatrix} U & U_p \end{bmatrix} \begin{bmatrix} \Sigma + M & I_r \\ I_r & 0 \end{bmatrix} \begin{bmatrix} V & V_p \end{bmatrix}^\top. \end{aligned}$$

This notably reveals that $X + H$ has rank at most $2r$. Compute[5] thin QR factorizations of the left and right matrices:

$$Q_U R_U = \begin{bmatrix} U & U_p \end{bmatrix}, \qquad\qquad Q_V R_V = \begin{bmatrix} V & V_p \end{bmatrix},$$

with $Q_U \in \mathrm{St}(m, 2r)$, $Q_V \in \mathrm{St}(n, 2r)$ and $R_U, R_V \in \mathbb{R}^{2r\times 2r}$ upper triangular (assuming $2r \leq m, n$; otherwise, the procedure is easily adapted). This costs $\sim 8(m + n)r^2$ arithmetic operations. Then,

$$X + H = Q_U \underbrace{R_U \begin{bmatrix} \Sigma + M & I_r \\ I_r & 0 \end{bmatrix} R_V^\top}_{\approx \tilde{U}\tilde{\Sigma}\tilde{V}^\top} Q_V^\top.$$

Compute a singular value decomposition $\tilde{U}\tilde{\Sigma}\tilde{V}^\top$ of the middle part as indicated, *truncated at rank* r: $\tilde{U}, \tilde{V} \in \mathrm{St}(2r, r)$, and $\tilde{\Sigma} \in \mathbb{R}^{r\times r}$ diagonal with decreasing, nonnegative diagonal entries. This costs essentially some multiple of $\sim r^3$ arithmetic operations, and reveals the truncated singular value decomposition of $X + H$:

$$\mathrm{R}_X(H) = (Q_U\tilde{U})\tilde{\Sigma}(Q_V\tilde{V})^\top. \tag{7.50}$$

Computing the products $Q_U\tilde{U}$ and $Q_V\tilde{V}$ costs $\sim 4(m + n)r^2$ arithmetic operations. The triplet $(Q_U\tilde{U}, \tilde{\Sigma}, Q_V\tilde{V})$ represents the retracted point on $\mathbb{R}_r^{m\times n}$.

Notice that if we wish to compute $\mathrm{R}_X(tH)$ for several different values of t (as would happen in a line-search procedure), then we can save the QR computations and replace the matrix $\left[\begin{smallmatrix} \Sigma+M & I_r \\ I_r & 0 \end{smallmatrix}\right]$ with $\left[\begin{smallmatrix} \Sigma+tM & tI_r \\ tI_r & 0 \end{smallmatrix}\right]$. After a first retraction, subsequent retractions along the same direction could be up to three times faster.

It is clear that $\mathrm{R}_X(0) = X$. That this retraction is indeed well defined and smooth (locally) and that $\mathrm{DR}_X(0)$ is the identity map follow from general properties of metric projection retractions (Section 5.12). See also [AO15] for details on this and several other retractions on $\mathbb{R}_r^{m\times n}$.

⋆ There is an important caveat with the retraction detailed above. Specifically, since $\mathbb{R}_r^{m\times n}$ is not (the shell of) a convex set in $\mathbb{R}^{m\times n}$, projection to $\mathbb{R}_r^{m\times n}$ is not globally well defined. This fact is apparent in the step where we compute the rank-r truncated singular value decomposition of a matrix of size $2r \times 2r$: depending on the vector being retracted, that operation may not have a solution (if the matrix has rank less than r), or the solution may not be unique (if its

[5] Here, the signs on the diagonals of R_U, R_V are irrelevant. In principle, some work can be saved using that U has orthonormal columns and that the columns of U_p are orthogonal to those of U, but this is numerically delicate when U_p is ill conditioned; likewise for V, V_p.

rth and $(r+1)$st singular values are positive and equal). Overall, this means we must be careful when we use this retraction.

With $\mathbb{R}^{m\times n}$ still endowed with the standard inner product, we now turn to the orthogonal projectors of $\mathbb{R}_r^{m\times n}$. From (7.48), it is clear that the normal space at $X = U\Sigma V^\top$ is given by:

$$\mathrm{N}_X\mathbb{R}_r^{m\times n} = \left\{ U_\perp W V_\perp^\top : W \in \mathbb{R}^{(m-r)\times(n-r)} \right\}. \tag{7.51}$$

Then, the orthogonal projection of $Z \in \mathbb{R}^{m\times n}$ to $\mathrm{T}_X\mathbb{R}_r^{m\times n}$ satisfies both

$$Z - \mathrm{Proj}_X(Z) = U_\perp W V_\perp^\top$$

for some W and, following (7.47),

$$\mathrm{Proj}_X(Z) = UMV^\top + U_pV^\top + UV_p^\top \tag{7.52}$$

for some M, U_p, V_p with $U^\top U_p = V^\top V_p = 0$. Combined, these state

$$Z = UMV^\top + U_pV^\top + UV_p^\top + U_\perp W V_\perp^\top.$$

Define $P_U = UU^\top, P_V = VV^\top, P_U^\perp = I_m - P_U$ and $P_V^\perp = I_n - P_V$. Then, we find in turn:

$$P_U Z P_V = UMV^\top, \qquad P_U^\perp Z P_V = U_pV^\top, \qquad \text{and} \qquad P_U Z P_V^\perp = UV_p^\top.$$

Hence,

$$\begin{aligned} \mathrm{Proj}_X(Z) &= P_U Z P_V + P_U^\perp Z P_V + P_U Z P_V^\perp \qquad (7.53)\\ &= U(U^\top Z V)V^\top + (I_m - UU^\top)ZVV^\top + UU^\top Z(I_n - VV^\top). \end{aligned}$$

In the notation of (7.52), this is a tangent vector at X represented by

$$M = U^\top Z V, \qquad U_p = ZV - UM, \qquad \text{and} \qquad V_p = Z^\top U - VM^\top. \tag{7.54}$$

If Z is structured so that $U^\top Z$ and ZV can be computed efficiently, its projection can also be computed efficiently: this is crucial in practice.

Turning $\mathbb{R}_r^{m\times n}$ into a Riemannian submanifold of $\mathbb{R}^{m\times n}$ with the standard Euclidean metric, the gradient of a smooth $f\colon \mathbb{R}_r^{m\times n} \to \mathbb{R}$ with smooth extension $\bar{f}$ to a neighborhood of $\mathbb{R}_r^{m\times n}$ in $\mathbb{R}^{m\times n}$ is given by Proposition 3.61 as

$$\mathrm{grad} f(X) = \mathrm{Proj}_X(\mathrm{grad} \bar{f}(X)),$$

to be computed using (7.52) and (7.54). In applications, $\mathrm{grad}\bar{f}(X)$ is often a sparse matrix, or a low-rank matrix available in factored form, or a sum of such structured matrices. In those cases, the projection can (and should) be computed efficiently.

Let $X = U\Sigma V^\top \in \mathbb{R}_r^{m\times n}$ be a matrix represented by the triplet (U, Σ, V), and let $\dot{X}, \dot{X}'$ be two tangent vectors at X represented as in (7.47) by triplets

(M, U_p, V_p) and (M', U_p', V_p') (respectively). With the stated Riemannian structure on $\mathbb{R}_r^{m\times n}$, we can compute the inner product $\dot{X}$ and $\dot{X}'$ as follows:

$$\left\langle \dot{X}, \dot{X}' \right\rangle_X = \langle UMV^\top + U_pV^\top + UV_p^\top, UM'V^\top + U_p'V^\top + U(V_p')^\top \rangle$$
$$= \langle M, M' \rangle + \langle U_p, U_p' \rangle + \langle V_p, V_p' \rangle, \tag{7.55}$$

where $\langle \cdot, \cdot \rangle$ refers to the usual Frobenius inner products over the appropriate matrix spaces. Notice how the cancellations that occurred above make it possible to compute inner products of tangent vectors using only the triplets that represent them, for a moderate computational cost.

Second-order tools

With $\mathbb{R}_r^{m\times n}$ as a Riemannian submanifold of the Euclidean space $\mathbb{R}^{m\times n}$, covariant derivatives (∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$) coincide with the usual vector field derivatives (of smooth extensions), followed by orthogonal projection to tangent spaces (Theorem 5.9, Proposition 5.31).

We use this to obtain a formula for the Riemannian Hessian of $f\colon \mathbb{R}_r^{m\times n} \to \mathbb{R}$ with smooth extension $\bar{f}$. Let O be the subset of $\mathbb{R}^{m\times n}$ containing all matrices whose rth and $(r+1)$st singular values are distinct: this is a neighborhood of $\mathbb{R}_r^{m\times n}$. Given a matrix X in O, let P_U be the orthogonal projector from $\mathbb{R}^m$ to the subspace spanned by the r dominant left singular vectors of X: this is smooth in X. In particular, if $X = U\Sigma V^\top$ has rank r (with factors as in (7.53)), then $P_U = UU^\top$. Likewise, let P_V be the orthogonal projector from $\mathbb{R}^n$ to the subspace spanned by the r dominant right singular vectors of X, also smooth in X, so that for $X = U\Sigma V^\top \in \mathbb{R}_r^{m\times n}$ we have $P_V = VV^\top$. The projectors to the orthogonal complements are $P_U^\perp = I_m - P_U$ and $P_V^\perp = I_n - P_V$. Then, we define a smooth extension of $\mathrm{grad} f(X)$ to O in $\mathbb{R}^{m\times n}$ with the shorthand $Z = \mathrm{grad}\bar{f}(X)$ as

$$\begin{aligned}\bar{G}(X) &= P_U Z P_V + P_U^\perp Z P_V + P_U Z P_V^\perp \\ &= P_U Z P_V + Z P_V - P_U Z P_V + P_U Z - P_U Z P_V \\ &= Z P_V + P_U Z - P_U Z P_V.\end{aligned}$$

In order to differentiate $\bar{G}(X)$, we must determine the differentials of P_U and P_V as a function of X. To this end, consider any tangent vector $H = UMV^\top + U_pV^\top + UV_p^\top$ at $X \in \mathbb{R}_r^{m\times n}$. We aim to design a smooth curve c on $\mathbb{R}_r^{m\times n}$ such that $c(0) = X$ and $c'(0) = H$. Then, we can use

$$\mathrm{D}\bar{G}(X)[H] = (\bar{G} \circ c)'(0)$$

to reach our conclusion.

Taking inspiration from (7.46), pick a smooth curve $U(t)$ on $\mathrm{St}(m, r)$ such that $U(0) = U$ and $U'(0) = U_p\Sigma^{-1}$. Similarly, pick a smooth curve $V(t)$ on $\mathrm{St}(n, r)$ such that $V(0) = V$ and $V'(0) = V_p\Sigma^{-1}$, and set $\Sigma(t) = \Sigma + tM$. By design, this ensures that $c(t) = U(t)\Sigma(t)V(t)^\top$ satisfies $c(0) = X$ and $c'(0) = H$. Define $\dot{P}_U$—the derivative of P_U at X along H—through:

$$\begin{aligned}
P_{U(t)} &= U(t)U(t)^\top, \text{ and} \\
\dot{P}_U &\triangleq \frac{\mathrm{d}}{\mathrm{d}t} P_{U(t)}\bigg|_{t=0} = U(0)U'(0)^\top + U'(0)U(0)^\top = U\Sigma^{-1}U_p^\top + U_p\Sigma^{-1}U^\top.
\end{aligned}$$

Likewise, define $\dot{P}_V$ through

$$P_{V(t)} = V(t)V(t)^\top, \qquad \text{and} \qquad \dot{P}_V \triangleq \frac{\mathrm{d}}{\mathrm{d}t} P_{V(t)}\bigg|_{t=0} = V\Sigma^{-1}V_p^\top + V_p\Sigma^{-1}V^\top.$$

With $\dot{Z} = \mathrm{Hess}\bar{f}(X)[H]$ for short, this allows us to write

$$\begin{aligned}
\mathrm{D}\bar{G}(X)[H] &= \dot{Z}P_V + Z\dot{P}_V + \dot{P}_U Z + P_U\dot{Z} - \dot{P}_U Z P_V - P_U \dot{Z} P_V - P_U Z \dot{P}_V \\
&= (P_U + P_U^\perp)\dot{Z}P_V + P_U\dot{Z}(P_V + P_V^\perp) - P_U\dot{Z}P_V + P_U^\perp Z\dot{P}_V + \dot{P}_U Z P_V^\perp \\
&= P_U\dot{Z}P_V + P_U^\perp(\dot{Z}P_V + Z\dot{P}_V) + (P_U\dot{Z} + \dot{P}_U Z)P_V^\perp.
\end{aligned}$$

We can now use the fact that $\mathbb{R}_r^{m\times n}$ is a Riemannian submanifold of $\mathbb{R}^{m\times n}$ together with (7.52) to claim

$$\mathrm{Hess}f(X)[H] = \mathrm{Proj}_X\big(\mathrm{D}\bar{G}(X)[H]\big) = U\tilde{M}V^\top + \tilde{U}_pV^\top + U\tilde{V}_p^\top, \tag{7.56}$$

for matrices $\tilde{M}, \tilde{U}_p, \tilde{V}_p$ given as in (7.54). Explicitly,

$$\begin{aligned}
\tilde{M} &= U^\top \mathrm{D}\bar{G}(X)[H]V = U^\top \dot{Z} V, \\
\tilde{U}_p &= \mathrm{D}\bar{G}(X)[H]V - U\tilde{M} = P_U^\perp\left(\dot{Z}V + ZV_p\Sigma^{-1}\right), \\
\tilde{V}_p &= (\mathrm{D}\bar{G}(X)[H])^\top U - V\tilde{M}^\top = P_V^\perp\left(\dot{Z}^\top U + Z^\top U_p\Sigma^{-1}\right),
\end{aligned} \tag{7.57}$$

where $Z = \mathrm{grad}\bar{f}(X)$, $\dot{Z} = \mathrm{Hess}\bar{f}(X)[H]$, $X = U\Sigma V^\top$ and H is represented by the triplet (M, U_p, V_p).

Once more, Z and $\dot{Z}$ are matrices in $\mathbb{R}^{m\times n}$ whose structure (if any) should be exploited to compute the products ZV_p, $Z^\top U_p$, $\dot{Z}V$ and $\dot{Z}^\top U$ efficiently.

We may reorganize the above as:

$$\begin{aligned}
\mathrm{Hess}f(X)[H] &= \mathrm{Proj}_X(\mathrm{Hess}\bar{f}(X)[H]) \\
&\quad + \left[P_U^\perp \mathrm{grad}\bar{f}(X)V_p\Sigma^{-1}\right]V^\top + U\left[P_V^\perp(\mathrm{grad}\bar{f}(X))^\top U_p\Sigma^{-1}\right]^\top.
\end{aligned} \tag{7.58}$$

This highlights the Riemannian Hessian as the projection of the Euclidean Hessian with additional corrections to $\tilde{U}_p$ and $\tilde{V}_p$ (between brackets): compare with Corollary 5.47. Notice the Σ^{-1} factors: these indicate that Riemannian Hessians are likely to behave poorly close to the "brink", that is, if some of the top r singular values of X are near zero.

In closing, we note that the retraction (7.50) is second order because it is the metric projection retraction (Proposition 5.55).

Exercise 7.5. *Taking inspiration from the discussion of retraction* (7.50)*, propose an algorithm to compute an SVD representation of a tangent vector. More precisely: given a point* $X \in \mathbb{R}_r^{m\times n}$ *represented by a triplet* (U, Σ, V) *as in* (7.45) *and a tangent vector* $\dot{X} \in \mathrm{T}_X\mathbb{R}_r^{m\times n}$ *represented by a triplet* (M, U_p, V_p) *as in* (7.47)*, explain how you compute a triplet* $(\tilde{U}, \tilde{\Sigma}, \tilde{V})$ *which is a representation of* $\dot{X} = \tilde{U}\tilde{\Sigma}\tilde{V}^\top$*, where* $\tilde{U}, \tilde{V}$ *have* $2r$ *orthonormal columns and* $\tilde{\Sigma}$ *has nonnegative (but not necessarily positive) diagonal entries, with overall complexity linear in* $m+n$*. (In Manopt, such functions are called* `tangent2ambient`*.)*

Exercise 7.6. *In this section, we have developed representations of points and tangent vectors on the manifold* $\mathbb{R}_r^{m\times n}$ *which allow for efficient computation. To develop theory however, it is sometimes more convenient to work with the points and tangent vectors directly, rather than in terms of particular matrix decompositions. It is indeed possible to find such expressions.*

Let P_X *and* $P_X^\perp$ *denote orthogonal projectors to the image (the range) of* X *and to its orthogonal complement, respectively. Thus, if* $X = U\Sigma V^\top$ *is an SVD of* $X \in \mathbb{R}_r^{m\times n}$*, then*

$$P_X = UU^\top, \qquad P_X^\perp = I_m - UU^\top, \qquad P_{X^\top} = VV^\top, \qquad P_{X^\top}^\perp = I_n - VV^\top.$$

Verify the following:

$$\begin{aligned} \mathrm{T}_X\mathbb{R}_r^{m\times n} &= \{\dot{X} \in \mathbb{R}^{m\times n} : P_X^\perp \dot{X} P_{X^\top}^\perp = 0\}, \\ \mathrm{Proj}_X(Z) &= Z - P_X^\perp Z P_{X^\top}^\perp \\ &= P_X Z + Z P_{X^\top} - P_X Z P_{X^\top}. \end{aligned}$$

With $\bar{f}$ *a smooth extension of* $f \colon \mathbb{R}_r^{m\times n} \to \mathbb{R}$*, further verify that*

$$\begin{aligned} \mathrm{grad} f(X) &= \mathrm{Proj}_X(\mathrm{grad}\bar{f}(X)) \text{ and} \\ \mathrm{Hess} f(X)[\dot{X}] &= \mathrm{Proj}_X\left(\mathrm{Hess}\bar{f}(X)[\dot{X}] + N\dot{X}^\top(X^\dagger)^\top + (X^\dagger)^\top\dot{X}^\top N\right), \end{aligned}$$

where $X^\dagger$ *is the Moore–Penrose pseudo-inverse of* X*, and* N *is shorthand for the normal part of the gradient of* $\bar{f}$ *at* X*:*

$$N = P_X^\perp \mathrm{grad}\bar{f}(X) P_{X^\top}^\perp.$$

Of course, X *is second-order critical for* f *if and only if* $\mathrm{grad} f(X) = 0$ *and* $\mathrm{Hess} f(X) \succeq 0$*. The latter is equivalent to the condition that*

$$\left\langle \dot{X}, \mathrm{Hess}\bar{f}(X)[\dot{X}] + N\dot{X}^\top(X^\dagger)^\top + (X^\dagger)^\top\dot{X}^\top N \right\rangle \geq 0$$

for all $\dot{X} \in \mathrm{T}_X\mathbb{R}_r^{m\times n}$*.*

7.6 The hyperboloid model

Consider the bilinear map $\langle\cdot,\cdot\rangle_M$ on $\mathbb{R}^{n+1}$ defined by

$$\langle u, v\rangle_M = -u_0v_0 + u_1v_1 + \cdots + u_nv_n = u^\top Jv \tag{7.59}$$

with $J = \mathrm{diag}(-1, 1, \ldots, 1)$. This is not a Euclidean inner product because J has one negative eigenvalue, but it is a *pseudo-inner product* because all eigenvalues of J are nonzero. It is called the *Minkowski pseudo-inner product* on $\mathbb{R}^{n+1}$.

Consider the following subset of $\mathbb{R}^{n+1}$ (sometimes denoted by H^n):

$$\begin{aligned}\mathcal{M} &= \left\{x \in \mathbb{R}^{n+1} : \langle x, x\rangle_M = -1 \text{ and } x_0 > 0\right\} \\ &= \left\{x \in \mathbb{R}^{n+1} : x_0^2 = 1 + x_1^2 + \cdots + x_n^2 \text{ and } x_0 > 0\right\}. \end{aligned} \tag{7.60}$$

The equation $\langle x, x\rangle_M = -1$ defines two connected components determined by the sign of x_0. The condition $x_0 > 0$ selects one of them. The defining function $h(x) = \langle x, x\rangle_M + 1$ has differential

$$\mathrm{D}h(x)[u] = 2\langle x, u\rangle_M = (2Jx)^\top u.$$

Notice that $x_0 \neq 0$ for all $x \in \mathcal{M}$; hence, $2Jx \neq 0$ for all $x \in \mathcal{M}$. We deduce that $\mathcal{M}$ is an embedded submanifold of $\mathbb{R}^{n+1}$ of dimension n with tangent spaces

$$\mathrm{T}_x\mathcal{M} = \left\{u \in \mathbb{R}^{n+1} : \langle x, u\rangle_M = 0\right\}. \tag{7.61}$$

For $n = 2$, the manifold $\mathcal{M}$ is one sheet of a hyperboloid of two sheets in $\mathbb{R}^3$.

While $\langle \cdot, \cdot\rangle_M$ is only a pseudo-inner product on $\mathbb{R}^{n+1}$, it is an inner product when restricted to the tangent spaces of $\mathcal{M}$. Indeed, for all $(x, u) \in \mathrm{T}\mathcal{M}$,

$$\begin{aligned}\langle u, u\rangle_M &= u_1^2 + \cdots + u_n^2 - u_0^2 \\ &= u_1^2 + \cdots + u_n^2 - \frac{1}{x_0^2}(x_1u_1 + \cdots + x_nu_n)^2 \\ &\geq u_1^2 + \cdots + u_n^2 - \frac{1}{x_0^2}(x_1^2 + \cdots + x_n^2)(u_1^2 + \cdots + u_n^2) \\ &= (u_1^2 + \cdots + u_n^2)\left(1 - \frac{x_0^2 - 1}{x_0^2}\right) \\ &= \frac{1}{x_0^2}(u_1^2 + \cdots + u_n^2) \\ &\geq 0.\end{aligned}$$

Above, we used in turn: $\langle x, u\rangle_M = 0$ to eliminate u_0, then Cauchy–Schwarz, then $\langle x, x\rangle_M = -1$ to claim $x_1^2 + \cdots + x_n^2 = x_0^2 - 1$. As a result, $\|u\|_M = \sqrt{\langle u, u\rangle_M}$ is a well-defined norm on any tangent space. This is despite the fact that $\langle u, u\rangle_M$ can be negative if u does not belong to any tangent space of $\mathcal{M}$.

It is easy to check that the restriction of $\langle \cdot, \cdot\rangle_M$ to each tangent space $\mathrm{T}_x\mathcal{M}$ defines a Riemannian metric on $\mathcal{M}$, turning it into a Riemannian manifold. With this Riemannian structure, we call $\mathcal{M}$ a *hyperbolic space* in the *hyperboloid model*. The main geometric trait of $\mathcal{M}$ with $n \geq 2$ is that its *sectional curvatures* are constant, equal to -1. Manifolds with that property are called hyperbolic spaces. There are several other models that share this trait, namely the *Beltrami–Klein model*, the *Poincaré ball model* and the *Poincaré half-space model*. For more about curvature and these models, see [Lee18, p62].

The tangent space $\mathrm{T}_x\mathcal{M}$ is an n-dimensional subspace of $\mathbb{R}^{n+1}$. Its orthogonal complement with respect to $\langle\cdot,\cdot\rangle_M$ is the one-dimensional normal space

$$\mathrm{N}_x\mathcal{M} = \left\{v \in \mathbb{R}^{n+1} : \langle u, v\rangle_M = 0 \text{ for all } u \in \mathrm{T}_x\mathcal{M}\right\} = \mathrm{span}(x). \tag{7.62}$$

Thus, orthogonal projection from $\mathbb{R}^{n+1}$ to $\mathrm{T}_x\mathcal{M}$ with respect to $\langle\cdot,\cdot\rangle_M$ takes the form $\mathrm{Proj}_x(z) = z + \alpha x$ with $\alpha \in \mathbb{R}$ chosen so that $z + \alpha x$ is in $\mathrm{T}_x\mathcal{M}$, that is, so that $0 = \langle x, z + \alpha x\rangle_M = \langle x, z\rangle_M - \alpha$. In other words:

$$\mathrm{Proj}_x(z) = z + \langle x, z\rangle_M \cdot x. \tag{7.63}$$

With this tool in hand, we can construct a useful formula to compute gradients of functions on $\mathcal{M}$.

Proposition 7.7. *Let $\bar{f}\colon \mathbb{R}^{n+1} \to \mathbb{R}$ be a smooth function on the Euclidean space $\mathbb{R}^{n+1}$ with the usual inner product $\langle u, v\rangle = u^\top v$. Let $f = \bar{f}|_\mathcal{M}$ be the restriction of f to $\mathcal{M}$ with the Riemannian structure as described above. The gradient of f is related to that of $\bar{f}$ as follows:*

$$\mathrm{grad} f(x) = \mathrm{Proj}_x\!\left(J \mathrm{grad}\bar{f}(x)\right), \tag{7.64}$$

where $J = \mathrm{diag}(-1, 1, \ldots, 1)$ and Proj_x is defined by (7.63).

Proof. By definition, $\mathrm{grad} f(x)$ is the unique vector in $\mathrm{T}_x\mathcal{M}$ such that $\mathrm{D}f(x)[u] = \langle \mathrm{grad} f(x), u\rangle_M$ for all $u \in \mathrm{T}_x\mathcal{M}$. Since $\bar{f}$ is a smooth extension of f, we can compute

$$\begin{aligned} \mathrm{D}f(x)[u] &= \mathrm{D}\bar{f}(x)[u] \\ &= \left\langle \mathrm{grad}\bar{f}(x), u \right\rangle \\ &= \left\langle J\mathrm{grad}\bar{f}(x), u \right\rangle_M \\ &= \left\langle J\mathrm{grad}\bar{f}(x), \mathrm{Proj}_x(u) \right\rangle_M \\ &= \left\langle \mathrm{Proj}_x\!\left(J\mathrm{grad}\bar{f}(x)\right), u \right\rangle_M. \end{aligned}$$

Above, the second line is by definition of $\mathrm{grad}\bar{f}(x)$; the third by definition of $\langle\cdot,\cdot\rangle_M$; the fourth because u is tangent at x; and the fifth because Proj_x is self-adjoint with respect to $\langle\cdot,\cdot\rangle_M$, as are all orthogonal projectors. The claim follows by uniqueness. □

As a remark, note that $J\mathrm{grad}\bar{f}(x)$ which appears in (7.64) is the gradient of $\bar{f}$ in the Minkowski space $\mathbb{R}^{n+1}$ with pseudo-inner product $\langle\cdot,\cdot\rangle_M$. See O'Neill [O'N83] for a general treatment of submanifolds of spaces equipped with pseudo-inner products.

Second-order tools

For all smooth vector fields V on $\mathcal{M}$ and all $(x, u) \in \mathrm{T}\mathcal{M}$, define the operator ∇ as

$$\nabla_u V = \mathrm{Proj}_x(\mathrm{D}\bar{V}(x)[u]), \tag{7.65}$$

where $\bar{V}$ is any smooth extension of V to a neighborhood of $\mathcal{M}$ in $\mathbb{R}^{n+1}$ and $\mathrm{D}\bar{V}(x)[u]$ is the usual directional derivative. It is an exercise to check that ∇ is the Riemannian connection for $\mathcal{M}$. It is instructive to compare this with Theorem 5.9 where we make the same claim under the assumption that the embedding space is Euclidean. Here, the embedding space is not Euclidean, but the result stands. Again, see O'Neill [O'N83] for a general treatment.

The covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$ (induced by ∇) for a smooth vector field Z along a smooth curve $c\colon I \to \mathcal{M}$ is given by

$$\frac{\mathrm{D}}{\mathrm{d}t}Z(t) = \mathrm{Proj}_{c(t)}\left(\frac{\mathrm{d}}{\mathrm{d}t}Z(t)\right) \tag{7.66}$$

where $\frac{\mathrm{d}}{\mathrm{d}t}Z(t)$ is the usual derivative of Z understood as a map from I to $\mathbb{R}^{n+1}$—this makes use of the fact that $Z(t) \in \mathrm{T}_{c(t)}\mathcal{M} \subset \mathbb{R}^{n+1}$. Compare this with Proposition 5.31.

It is an exercise to check that, for arbitrary $(x, u) \in \mathrm{T}\mathcal{M}$,

$$\begin{aligned} c(t) = \mathrm{Exp}_x(tu) &\triangleq \cosh(\|tu\|_M)x + \frac{\sinh(\|tu\|_M)}{\|tu\|_M}tu \\ &= \cosh(t\|u\|_M)x + \frac{\sinh(t\|u\|_M)}{\|u\|_M}u \end{aligned} \tag{7.67}$$

defines the unique geodesic on $\mathcal{M}$ such that $c(0) = x$ and $c'(0) = u$. Notice that this is defined for all t: Exp is a second-order retraction defined on the whole tangent bundle (see also Section 10.2). Compare with the geodesics on the sphere, Example 5.37.

We proceed to construct a formula for the Hessian of a function on $\mathcal{M}$ based on the gradient and Hessian of a smooth extension.

Proposition 7.8. *(Continued from Proposition 7.7.) The Hessian of f is related to that of $\bar{f}$ as follows:*

$$\mathrm{Hess}f(x)[u] = \mathrm{Proj}_x\big(J\mathrm{Hess}\bar{f}(x)[u]\big) + \big\langle x, J\mathrm{grad}\bar{f}(x)\big\rangle_M \cdot u, \tag{7.68}$$

where $J = \mathrm{diag}(-1, 1, \ldots, 1)$ and Proj_x is defined by (7.63).

Proof. Consider the following smooth vector field in $\mathbb{R}^{n+1}$:

$$\bar{G}(x) = J\mathrm{grad}\bar{f}(x) + \big\langle J\mathrm{grad}\bar{f}(x), x\big\rangle_M \cdot x.$$

This is a smooth extension of $\mathrm{grad}f$ from $\mathcal{M}$ to $\mathbb{R}^{n+1}$. Thus, for all $(x, u) \in \mathrm{T}\mathcal{M}$ we have

$$\begin{aligned} \mathrm{Hess}f(x)[u] &= \nabla_u \mathrm{grad}f \\ &= \mathrm{Proj}_x\big(\mathrm{D}\bar{G}(x)[u]\big) \\ &= \mathrm{Proj}_x\big(J\mathrm{Hess}\bar{f}(x)[u] + qx + \big\langle J\mathrm{grad}\bar{f}(x), x\big\rangle_M \cdot u\big) \\ &= \mathrm{Proj}_x\big(J\mathrm{Hess}\bar{f}(x)[u]\big) + \big\langle J\mathrm{grad}\bar{f}(x), x\big\rangle_M \cdot u, \end{aligned}$$

where q is the derivative of $\big\langle J\mathrm{grad}\bar{f}(x), x\big\rangle_M$ at x along u—and we do not need to compute it since qx is in the normal space, hence it vanishes through the projector. $\square$

Exercise 7.9. *Check that $\langle \cdot, \cdot \rangle_M$ indeed defines a Riemannian metric on $\mathcal{M}$. Verify that ∇ (7.65) is the Riemannian connection for $\mathcal{M}$, that $\frac{\mathrm{D}}{\mathrm{d}t}$ (7.66) is the covariant derivative induced by ∇ and that $c(t)$ (7.67) is a geodesic on $\mathcal{M}$ satisfying $c(0) = x$ and $c'(0) = u$ (that this is the unique such geodesic is a consequence of general results, see Section 10.2).*

7.7 Manifolds defined by $h(x) = 0$

Let $h\colon \mathcal{E} \to \mathbb{R}^k$ be a smooth function on a Euclidean space of dimension strictly larger than k with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$. If $\mathrm{D}h(x)$ has full rank k for all $x \in \mathcal{M}$, the set

$$\mathcal{M} = \{x \in \mathcal{E} : h(x) = 0\} \tag{7.69}$$

is an embedded submanifold of $\mathcal{E}$ of dimension $\dim \mathcal{E} - k$: we assume so here. In contrast with Definition 3.10, we require the whole manifold to be defined with a single defining function h. Notwithstanding, everything below still holds if $\mathcal{M}$ is only locally defined by h: we focus on the case of a global h for notational simplicity and because it covers several of the examples we have encountered.

With the notation $h(x) = (h_1(x), \ldots, h_k(x))^\top$ to highlight the k constraint functions $h_i\colon \mathcal{E} \to \mathbb{R}$, we can spell out the linear map

$$\mathrm{D}h(x)[v] = \big(\langle \mathrm{grad} h_1(x), v \rangle, \ldots, \langle \mathrm{grad} h_k(x), v \rangle \big)^\top \tag{7.70}$$

and its adjoint

$$\mathrm{D}h(x)^*[\alpha] = \sum_{i=1}^{k} \alpha_i \mathrm{grad} h_i(x). \tag{7.71}$$

The tangent spaces are given by

$$\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x) = \{v \in \mathcal{E} : \langle \mathrm{grad} h_i(x), v \rangle = 0 \text{ for all } i\}. \tag{7.72}$$

The fact that $\mathrm{D}h(x)$ has full rank k means that the gradients of the constraints at x are linearly independent. In other words, they form a basis for the normal space at x:

$$\mathrm{N}_x\mathcal{M} = (\ker \mathrm{D}h(x))^\perp = \mathrm{span}\big(\mathrm{grad} h_1(x), \ldots, \mathrm{grad} h_k(x)\big). \tag{7.73}$$

Let $\mathrm{Proj}_x\colon \mathcal{E} \to \mathrm{T}_x\mathcal{M}$ denote orthogonal projection from $\mathcal{E}$ to $\mathrm{T}_x\mathcal{M}$. Then, for any vector v in $\mathcal{E}$ there exists a unique choice of coefficients $\alpha \in \mathbb{R}^k$ such that

$$v = \mathrm{Proj}_x(v) + \mathrm{D}h(x)^*[\alpha]. \tag{7.74}$$

This decomposes v into its tangent and normal parts at x. Explicitly, α is the unique solution to the following least-squares problem:

$$\alpha = \arg\min_{\alpha \in \mathbb{R}^k} \|v - \mathrm{D}h(x)^*[\alpha]\|^2 = \left(\mathrm{D}h(x)^*\right)^\dagger [v],$$

where the dagger denotes Moore–Penrose pseudo-inversion, so that

$$\mathrm{Proj}_x(v) = v - \mathrm{D}h(x)^* \left[\left(\mathrm{D}h(x)^*\right)^\dagger [v]\right]. \tag{7.75}$$

This formula is expected since $\mathrm{im}\, \mathrm{D}h(x)^* = (\ker \mathrm{D}h(x))^\perp = \mathrm{N}_x\mathcal{M}$.

One possible retraction for $\mathcal{M}$ is metric projection as studied in Section 5.12. It relies on the Euclidean metric to define:

$$\mathrm{R}_x(v) = \arg\min_{y \in \mathcal{E}} \|x + v - y\| \text{ subject to } h(y) = 0. \tag{7.76}$$

This is well defined for small enough v, but $\mathrm{R}_x(v)$ may not be uniquely defined for all v. It may be difficult to compute in general.

Let $\mathcal{M}$ be a Riemannian submanifold of $\mathcal{E}$. Then, R (7.76) is a second-order retraction. Given a smooth function $\bar{f} \colon \mathcal{E} \to \mathbb{R}$ and its restriction $f = \bar{f}|_{\mathcal{M}}$, the Riemannian gradient follows from (7.75) as

$$\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad} \bar{f}(x)) = \mathrm{grad} \bar{f}(x) - \sum_{i=1}^{k} \lambda_i(x) \mathrm{grad} h_i(x) \tag{7.77}$$

with $\lambda(x) = \left(\mathrm{D}h(x)^*\right)^\dagger [\mathrm{grad} \bar{f}(x)]$.

Second-order tools

With $\mathcal{M}$ as a Riemannian submanifold of the Euclidean space $\mathcal{E}$, covariant derivatives (∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$) coincide with the usual vector field derivatives (of smooth extensions), followed by orthogonal projection to tangent spaces (Theorem 5.9, Proposition 5.31). We use this to determine the Riemannian Hessian of $f = \bar{f}|_{\mathcal{M}}$.

Notice that

$$\lambda(x) = (\mathrm{D}h(x)^*)^\dagger [\mathrm{grad} \bar{f}(x)] \tag{7.78}$$

is a smooth function on the open subset of $\mathcal{E}$ consisting of all points x where $\mathrm{D}h(x)$ has full rank k. Thus, we can differentiate $\mathrm{grad} f(x)$ (7.77) as follows:

$$\begin{aligned} \mathrm{D}\mathrm{grad} f(x)[v] &= \mathrm{Hess} \bar{f}(x)[v] \\ &\quad - \sum_{i=1}^{k} \mathrm{D}\lambda_i(x)[v] \cdot \mathrm{grad} h_i(x) - \sum_{i=1}^{k} \lambda_i(x) \mathrm{Hess} h_i(x)[v]. \end{aligned}$$

Then, since $\mathrm{Hess} f(x)[v]$ is nothing but the orthogonal projection of $\mathrm{D}\mathrm{grad} f(x)[v]$ to $\mathrm{T}_x\mathcal{M}$ and since each $\mathrm{grad} h_i(x)$ is orthogonal to $\mathrm{T}_x\mathcal{M}$, it follows that

$$\mathrm{Hess} f(x)[v] = \mathrm{Proj}_x\left(\mathrm{Hess} \bar{f}(x)[v] - \sum_{i=1}^{k} \lambda_i(x) \mathrm{Hess} h_i(x)[v]\right). \tag{7.79}$$

This can be summarized with pleasantly symmetric identities:

$$\mathrm{grad} f(x) = \mathrm{grad} \bar{f}(x) - \sum_{i=1}^{k} \lambda_i(x) \mathrm{grad} h_i(x), \tag{7.80}$$

$$\mathrm{Hess} f(x) = \mathrm{Proj}_x \circ \left(\mathrm{Hess} \bar{f}(x) - \sum_{i=1}^{k} \lambda_i(x) \mathrm{Hess} h_i(x) \right) \circ \mathrm{Proj}_x, \tag{7.81}$$

with $\lambda(x)$ as defined in (7.78), and with the understanding that the linear map on the right-hand side of (7.81) is restricted to $\mathrm{T}_x\mathcal{M}$. Notice that $\lambda(x)$ depends on $\mathrm{grad}\bar{f}(x)$ only through its normal component: compare with the Hessian formulas in Section 5.11. The work above easily yields an expression for the Weingarten map (5.38) of $\mathcal{M}$.

Exercise 7.10. *Consider the equality constrained optimization problem*

$$\min_{x \in \mathcal{E}} \bar{f}(x) \qquad \textit{subject to} \qquad h(x) = 0, \tag{7.82}$$

where $\bar{f} \colon \mathcal{E} \to \mathbb{R}$ *and* $h \colon \mathcal{E} \to \mathbb{R}^k$ *are smooth on a Euclidean space* $\mathcal{E}$ *with* $\dim \mathcal{E} > k$*. The* Lagrangian function $L \colon \mathcal{E} \times \mathbb{R}^k \to \mathbb{R}$ *for this problem is:*

$$L(x, \lambda) = \bar{f}(x) - \langle \lambda, h(x) \rangle .$$

A classical result is that if $x \in \mathcal{E}$ *is such that* $\mathrm{D}h(x)$ *has rank* k *and* x *is a local minimizer for* (7.82) *then* x *satisfies* KKT conditions *of order one and two; explicitly: there exists a unique* $\lambda \in \mathbb{R}^k$ *such that*

1. $\mathrm{grad} L(x, \lambda) = 0$*, and*
2. $\langle \mathrm{Hess}_x L(x, \lambda)[v], v \rangle \geq 0$ *for all* $v \in \ker \mathrm{D}h(x)$*,*

where the Hessian of L *is taken with respect to* x *only. These are the classical first- and second-order necessary optimality condition for* (7.82)*.*

The full-rank requirement on $\mathrm{D}h(x)$ *is known as the* linear independence constraint qualification *(*LICQ*), because it amounts to the requirement that the gradients of the constraints at* x *be linearly independent.*

We know $\mathcal{M} = \{x \in \mathcal{E} : h(x) = 0\}$ *is an embedded submanifold of* $\mathcal{E}$ *if* $\mathrm{D}h(x)$ *has rank* k *for* all $x \in \mathcal{M}$*. Assuming this holds, show that* $x \in \mathcal{E}$ *satisfies the first-order KKT conditions if and only if* x *is in* $\mathcal{M}$ *and* $\mathrm{grad} f(x) = 0$*, where* $f = \bar{f}|_{\mathcal{M}}$ *is restricted to* $\mathcal{M}$ *equipped with the Riemannian submanifold structure. Additionally, show that* x *satisfies both first- and second-order KKT conditions if and only if* $x \in \mathcal{M}$*,* $\mathrm{grad} f(x) = 0$ *and* $\mathrm{Hess} f(x) \succeq 0$*.*

This confirms that the classical necessary optimality conditions are equivalent to the conditions we established in Sections 4.2 and 6.1 when LICQ holds globally. (Of course, this reasoning can also be applied locally around any point x*.) This gives KKT conditions a natural geometric interpretation. These considerations form part of the basis of Luenberger's seminal paper [Lue72] which started the field of optimization on manifolds.*

7.8 Notes and references

Much of the material in this chapter is standard, though some of it rarely appears in as much detail.

For the Stiefel manifold in particular, we follow mostly [AMS08].

The construction of tools for optimization on $\mathbb{R}_r^{m\times n}$ as a Riemannian submanifold of $\mathbb{R}^{m\times n}$ follows work by Vandereycken [Van13]. Similarly, one can derive tools for optimization over fixed-rank tensors in tensor train (TT) and Tucker format [UV13, KSV14, HS18, UV20]. Fine properties of curves generated by the metric projection retraction to the real algebraic variety of matrices of rank upper-bounded by r appear in [Lev20, Thm. 3.1, Cor. 3.3]. One popular technique to optimize over matrices of rank up to r (rather than equal to r) is to set $X = AB^\top$ and to optimize over the factors $A \in \mathbb{R}^{m\times r}$, $B \in \mathbb{R}^{n\times r}$ (this is an over-parameterization since the factorization is not unique). There exist other such smooth over-parameterizations of the variety of bounded rank matrices, see for example [LKB21].

Applications of optimization on hyperbolic space in machine learning include hierarchical embeddings [NK17, JMM19, KMU$^+$20].

Here are a few other manifolds of interest for applications:

- The Stiefel manifold with the canonical metric [EAS98];
- The Grassmann manifold $\mathrm{Gr}(n, p)$ of subspaces of dimension p in $\mathbb{R}^n$. It can be viewed as a quotient manifold of $\mathrm{St}(n, p)$, or as an embedded submanifold of $\mathbb{R}^{n\times n}$ where each subspace is identified with an orthogonal projector of rank p: see the discussion around eq. (9.90) in Section 9.16;
- Matrices with positive entries (see Section 11.6);
- Positive definite matrices (see Section 11.7);
- Positive semidefinite matrices with a fixed rank [VAV09, JBAS10, MA20]—see also Example 9.57;
- Multinomial manifolds; the simplex; stochastic matrices [DH19];
- The rigid motion group (special Euclidean group) $\mathrm{SE}(n)$: this is a manifold as the product of the manifolds $\mathbb{R}^n$ and $\mathrm{SO}(n)$, providing a parameterization of all possible rigid motions in $\mathbb{R}^n$ as a combination of a translation and a rotation (to add reflections, use $\mathrm{O}(n)$ instead of $\mathrm{SO}(n)$);
- The essential manifold for camera pose descriptions (epipolar constraint between projected points in two perspective views) [TD14];
- Shape space as the manifold of shapes in $\mathbb{R}^2, \mathbb{R}^3, \ldots$ (see [FCPJ04] or [MS20] and many references therein).

At times, the search space of an optimization problem in a linear space $\mathcal{E}$ is defined through two sets of equality constraints, $h(x) = 0$ and $g(x) = 0$, in such a way that $\{x : h(x) = 0\}$ defines an embedded submanifold of $\mathcal{E}$ but the intersection $\{x : h(x) = 0 \text{ and } g(x) = 0\}$ does not. Then, it may be beneficial to optimize over the manifold $\{x : h(x) = 0\}$ and to move the constraint $g(x) = 0$ to the cost function as a penalty. This can be done in several ways, for example using a quadratic penalty—$f(x)+\lambda\|g(x)\|^2$—or using a type of augmented Lagrangian method [LB20]. One can also attempt to handle inequality constraints in this fashion.

8 General manifolds

In this chapter, we consider the general definition of a (smooth) manifold. Following Brickell and Clark [BC70], we initially give a (too) broad definition, devoid of topological considerations. To avoid confusion, we refer to these objects as *manifolds**, with a star. Promptly after that, in order to exclude topological curiosities that are of little interest to optimization, we restrict the definition and call the remaining objects *manifolds*. This final definition is standard.

Of course, embedded submanifolds of linear spaces—as we have considered so far—are manifolds: we shall verify this. Interestingly, the general perspective enables us to consider new manifolds. In particular, we touch upon the Grassmann manifold which consists of all linear subspaces of a given dimension in some linear space. Chapter 9 discusses such manifolds in more depth.

We then revisit our geometric toolbox to generalize smooth maps, tangent spaces, vector fields, retractions, Riemannian metrics, gradients, connections, Hessians, etc. By design, Chapters 4 and 6 regarding optimization algorithms apply verbatim to the general setting.

8.1 A permissive definition

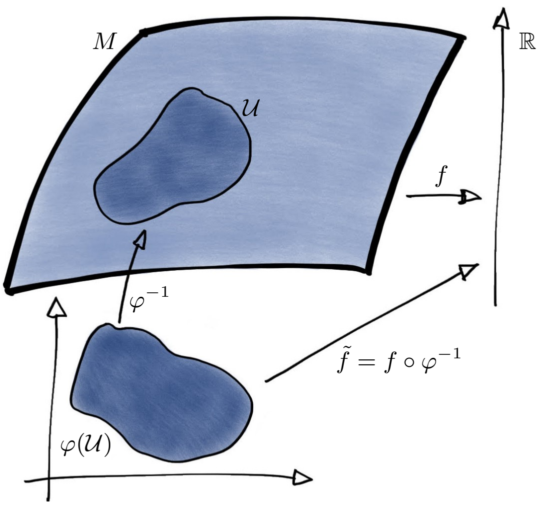
Given a set M (without any particular structure so far), the first step toward defining a smooth manifold structure on M is to model M after $\mathbb{R}^d$. To do so, we introduce the concept of *chart*. A chart establishes a one-to-one correspondence between a subset of M and an *open* subset of $\mathbb{R}^d$. This allows us to leverage the powerful tools we have at our disposal on $\mathbb{R}^d$ to work on M.

As the terms *chart* and (later) *atlas* suggest, it helps to think of M as the Earth (a sphere), of charts as two-dimensional, flat maps of parts of the Earth, and of atlases as collections of maps that cover the Earth.

Definition 8.1. *A d-dimensional* chart *on a set M is a pair $(\mathcal{U}, \varphi)$ consisting of a subset $\mathcal{U}$ of M (called the* domain*) and a map $\varphi \colon \mathcal{U} \to \mathbb{R}^d$ such that:*

1. *$\varphi(\mathcal{U})$ is open in $\mathbb{R}^d$, and*
2. *φ is invertible between $\mathcal{U}$ and $\varphi(\mathcal{U})$.*

The numbers $(\varphi(x)_1, \ldots, \varphi(x)_d)$ are the coordinates *of the point $x \in \mathcal{U}$ in the chart φ. The map $\varphi^{-1} \colon \varphi(\mathcal{U}) \to \mathcal{U}$ is a* local parameterization *of M.*



$$\begin{tikzcd} M \arrow[r, "f"] & \mathbb{R} \\ \varphi(\mathcal{U}) \subseteq \mathbb{R}^d \arrow[u, "\varphi^{-1}"] \arrow[ur, "\tilde{f}"'] & \end{tikzcd}$$

Figure 8.1 Illustration and matching commutative diagram for eq. (8.1) expressing a real function f on a set M through a chart $(\mathcal{U}, \varphi)$.

When the domain is clear, we often call φ itself a chart. Given a point x in M, we say φ is a chart *around* x if x is in the domain of φ.

For a function from (an open subset of) $\mathbb{R}^d$ to $\mathbb{R}$, we readily have a notion of smoothness: it is *smooth* at x if it is infinitely differentiable at x, in the usual sense. One of the goals of differential geometry is to generalize this notion to functions $f: M \to \mathbb{R}$ on a more general class of sets M. Let $(\mathcal{U}, \varphi)$ be a d-dimensional chart around $x \in M$. Then,

$$\tilde{f} = f \circ \varphi^{-1} : \varphi(\mathcal{U}) \to \mathbb{R} \tag{8.1}$$

is called a *coordinate representative* of f in this chart. Since $\varphi(\mathcal{U})$ is open in $\mathbb{R}^d$, it makes sense to talk of differentiability of $\tilde{f}$. In particular, we may want to define that, with respect to this chart, f is smooth at x if $\tilde{f}$ is smooth at $\varphi(x)$.

Two d-dimensional charts $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ on M around x are *compatible* if they yield the same conclusions regarding smoothness of functions at x. Restricted to the appropriate domains, the coordinate representatives

$$\tilde{f} = f \circ \varphi^{-1} \qquad \text{and} \qquad \hat{f} = f \circ \psi^{-1}$$

are related by

$$\tilde{f} = \hat{f} \circ (\psi \circ \varphi^{-1}) \qquad \text{and} \qquad \hat{f} = \tilde{f} \circ (\varphi \circ \psi^{-1}).$$

Thus, the differentiability properties of $\tilde{f}$ and $\hat{f}$ are the same if the domains involved are open in $\mathbb{R}^d$ and if $\psi \circ \varphi^{-1}$ and its inverse are smooth. This is made precise in the following definition. There, we could allow non-overlapping charts to have different dimensions, but this serves little purpose in optimization.

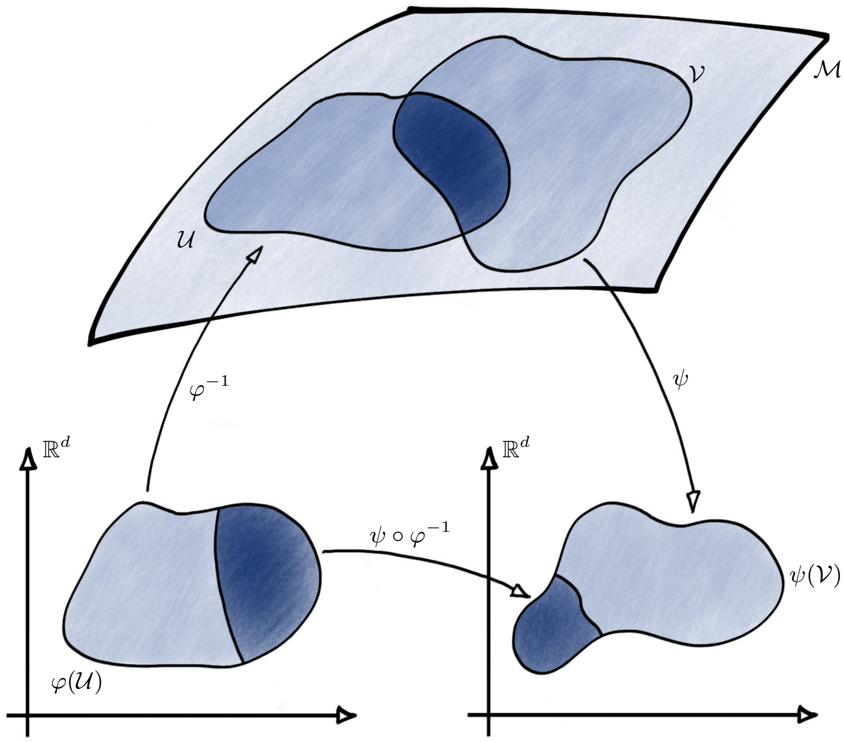


Figure 8.2 Overlapping charts $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ on a manifold of dimension d. The darker area on the manifold corresponds to the intersection $\mathcal{U} \cap \mathcal{V}$ of the chart domains. In the coordinate spaces (bottom), the darker areas correspond to the open images $\varphi(\mathcal{U} \cap \mathcal{V})$ and $\psi(\mathcal{U} \cap \mathcal{V})$: the coordinate change map $\psi \circ \varphi^{-1}$ is a diffeomorphism between these two.

Definition 8.2. *Two charts $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ of M are* compatible *if they have the same dimension d and either $\mathcal{U} \cap \mathcal{V} = \emptyset$, or $\mathcal{U} \cap \mathcal{V} \neq \emptyset$ and:*

1. *$\varphi(\mathcal{U} \cap \mathcal{V})$ is open in $\mathbb{R}^d$;*
2. *$\psi(\mathcal{U} \cap \mathcal{V})$ is open in $\mathbb{R}^d$; and*
3. *$\psi \circ \varphi^{-1} \colon \varphi(\mathcal{U} \cap \mathcal{V}) \to \psi(\mathcal{U} \cap \mathcal{V})$ is a smooth invertible function whose inverse is also smooth (i.e., it is a diffeomorphism, see Definition 3.11).*

A collection of charts is compatible if each pair of charts in that collection is compatible. Compatible charts that cover the whole set M form an *atlas*.

Definition 8.3. *An* atlas *$\mathcal{A}$ on a set M is a compatible collection of charts on M whose domains cover M. In particular, for every $x \in M$, there is a chart $(\mathcal{U}, \varphi) \in \mathcal{A}$ such that $x \in \mathcal{U}$.*

Given an atlas $\mathcal{A}$, it is an exercise to show that the collection $\mathcal{A}^+$ of all charts of M which are compatible with $\mathcal{A}$ is itself an atlas of M, called a *maximal atlas.*

$$\begin{tikzcd}
\mathcal{U} \subseteq \mathcal{M} \arrow[r, "F"] & \mathcal{V} \subseteq \mathcal{M}' \arrow[d, "\psi"] \\
\varphi(\mathcal{U}) \subseteq \mathbb{R}^d \arrow[u, "\varphi^{-1}"] \arrow[r, "\tilde{F}"] & \psi(\mathcal{V}) \subseteq \mathbb{R}^{d'}
\end{tikzcd}$$

Figure 8.3 Commutative diagram for Definition 8.5 expressing a map through charts.

Thus, any atlas uniquely defines a maximal atlas: we use the latter to define manifolds* (the star is a reminder that topological concerns are delayed to a later section.) We say that the maximal atlas defines a *smooth structure* on M.

Definition 8.4. *A* manifold* *is a pair* $\mathcal{M} = (M, \mathcal{A}^+)$*, consisting of a set* M *and a maximal atlas* $\mathcal{A}^+$ *on* M*. The* dimension of $\mathcal{M}$ *is the dimension of any of its charts. When the atlas is clear from context, we often conflate notation for* $\mathcal{M}$ *and* M*.*

We can now define smoothness of maps between manifolds*. Below, smoothness of $\tilde{F}$ is understood in the usual sense for maps between open subsets of linear spaces (see Section 3.1).

Definition 8.5. *A map* $F \colon \mathcal{M} \to \mathcal{M}'$ *is* smooth at $x \in \mathcal{M}$ *if*

$$\tilde{F} = \psi \circ F \circ \varphi^{-1} \colon \varphi(\mathcal{U}) \to \psi(\mathcal{V})$$

is smooth at $\varphi(x)$*, where* $(\mathcal{U}, \varphi)$ *is a chart of* $\mathcal{M}$ *around* x *and* $(\mathcal{V}, \psi)$ *is a chart of* $\mathcal{M}'$ *around* $F(x)$*. The map* F *is* smooth *if it is smooth at every point* x *in* $\mathcal{M}$*. We call* $\tilde{F}$ *a* coordinate representative *of* F*.*

Remark 8.6. *By extension, we say a map* $F \colon \mathcal{M} \to \mathcal{M}'$ *is* k times (continuously) differentiable *if its coordinate representatives are so. Smoothness corresponds to* $k = \infty$*. Later, we endow* $\mathcal{M}$ *with a Riemannian metric so that* $f \colon \mathcal{M} \to \mathbb{R}$ *is (continuously) differentiable if and only if it has a (continuous) Riemannian gradient, and* f *is twice (continuously) differentiable if and only if it has a (continuous) Riemannian Hessian.*

It is an exercise to verify that Definition 8.5 is independent of the choice of charts, and that composition preserves smoothness.

Example 8.7. *Let* $\mathcal{E}$ *be a linear space of dimension* d*. We can equip* $\mathcal{E}$ *with a smooth structure as follows: choose a basis for* $\mathcal{E}$*; set* $\mathcal{U} = \mathcal{E}$ *and let* $\varphi(x) \in \mathbb{R}^d$ *denote the coordinates of* x *in the chosen basis; the maximal atlas generated by* $(\mathcal{U}, \varphi)$ *yields the usual smooth structure on* $\mathcal{E}$*. For example, if* $\mathcal{E} = \mathbb{R}^d$*, we can choose* $\varphi(x) = x$*. By default, we always use this smooth structure on* $\mathbb{R}^d$*.*

Example 8.8. *Let* M *be an open subset of a linear space* $\mathcal{E}$ *of dimension* d*.*

With the same chart as in the previous example, only restricted to $\mathcal{U} = M$, it is clear that $\varphi(M)$ is open in $\mathbb{R}^d$, so that $(\mathcal{U}, \varphi)$ is a chart for M, and it covers all of M hence it defines an atlas on M. We conclude that any open subset of a linear space is a manifold with a natural atlas. By default, we always use this smooth structure on open subsets of linear spaces.*

Example 8.9. *The local parameterization $\varphi^{-1}\colon \varphi(\mathcal{U}) \to \mathcal{M}$ associated to a chart $(\mathcal{U}, \varphi)$ is a smooth map. Likewise, with the previous example in mind, the chart $\varphi\colon \mathcal{U} \to \mathbb{R}^d$ is a smooth map. Indeed, in both cases, we can arrange for their coordinate representative to be the identity map.*

Example 8.10. *Consider the unit circle, $\mathrm{S}^1 = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$. One possible atlas is made of four charts, each defined on a half circle—dubbed North, East, South and West—as follows:*

$$\begin{aligned}
\mathcal{U}_N &= \{x \in \mathrm{S}^1 : x_2 > 0\}, & \varphi_N(x) &= x_1,\\
\mathcal{U}_E &= \{x \in \mathrm{S}^1 : x_1 > 0\}, & \varphi_E(x) &= x_2,\\
\mathcal{U}_S &= \{x \in \mathrm{S}^1 : x_2 < 0\}, & \varphi_S(x) &= x_1,\\
\mathcal{U}_W &= \{x \in \mathrm{S}^1 : x_1 < 0\}, & \varphi_W(x) &= x_2.
\end{aligned}$$

It is clear that these are one-dimensional charts. For example, checking the North chart we find that $\varphi_N\colon \mathcal{U}_N \to \varphi_N(\mathcal{U}_N)$ is invertible and $\varphi_N(\mathcal{U}_N) = (-1, 1)$ is open in $\mathbb{R}$, as required. Furthermore, these charts are compatible. For example, checking for the North and East charts, we find that:

1. *$\mathcal{U}_N \cap \mathcal{U}_E = \{x \in \mathrm{S}^1 : x_1 > 0$ and $x_2 > 0\}$;*
2. *$\varphi_N(\mathcal{U}_N \cap \mathcal{U}_E) = (0, 1)$ is open;*
3. *$\varphi_E(\mathcal{U}_N \cap \mathcal{U}_E) = (0, 1)$ is open; and*
4. *$\varphi_E^{-1}(z) = (\sqrt{1 - z^2}, z)$, so that $(\varphi_N \circ \varphi_E^{-1})(z) = \sqrt{1 - z^2}$, which is smooth and smoothly invertible on $(0, 1)$.*

The charts also cover the whole set S^1, so that together they form an atlas $\mathcal{A}$ for S^1. As a result, $(\mathrm{S}^1, \mathcal{A}^+)$ is a manifold.*

Earlier, using Definition 3.10, we called S^1 an embedded submanifold of $\mathbb{R}^2$. In Section 8.3, we argue more generally that embedded submanifolds of linear spaces (as per that early definition) are manifolds.*

Example 8.11. *We now discuss a new example: the $(n-1)$-dimensional* real projective space, *RP^{n-1}. This is the set of lines through the origin (that is, one-dimensional linear subspaces) of $\mathbb{R}^n$. To any nonzero point $x \in \mathbb{R}^n$, we associate a linear subspace as follows:*

$$\pi\colon \mathbb{R}^n \backslash \{0\} \to \mathrm{RP}^{n-1}\colon x \mapsto \pi(x) = \{\alpha x : \alpha \in \mathbb{R}\}.$$

The classical atlas for RP^{n-1} is built from the following charts. For a given i in $\{1, \ldots, n\}$, consider the following subset of RP^{n-1}:

$$\mathcal{U}_i = \{\pi(x) : x \in \mathbb{R}^n \text{ and } x_i \neq 0\}.$$

This is the set of lines through the origin that are not parallel to the plane P_i defined by $x_i = 1$. In other words, this is the set of lines through the origin that intersect that plane. This allows us to define the map φ_i on the domain $\mathcal{U}_i$ into $\mathbb{R}^{n-1}$, as the coordinates of the intersection of the line $\pi(x)$ with the plane P_i:

$$\varphi_i(\pi(x)) = \left(\frac{x_1}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_n}{x_i}\right).$$

The map φ_i is indeed well defined because the right-hand side depends only on $\pi(x)$ and not on x itself—this is key. The range $\varphi_i(\mathcal{U}_i)$ is all of $\mathbb{R}^{n-1}$ (since there exists a line through the origin and any point of P_i), hence it is open. Furthermore, φ_i is invertible:

$$\varphi_i^{-1}(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n) = \pi(z_1, \ldots, z_{i-1}, 1, z_{i+1}, \ldots, z_n).$$

Thus, $\{(\mathcal{U}_i, \varphi_i)\}_{i=1,\ldots,n}$ are charts for RP^{n-1}. They cover RP^{n-1} since no line can be parallel to all planes $P_1, \ldots, P_n$. Thus, it remains to verify that the charts are compatible. For all pairs $i \neq j$, consider the following:

1. *$\mathcal{U}_i \cap \mathcal{U}_j = \{\pi(x) : x \in \mathbb{R}^n, x_i \neq 0$ and $x_j \neq 0\}$;*
2. *$\varphi_i(\mathcal{U}_i \cap \mathcal{U}_j)$ and $\varphi_j(\mathcal{U}_i \cap \mathcal{U}_j)$ are both subsets of $\mathbb{R}^{n-1}$ defined by one coordinate being nonzero: they are indeed open;*
3. *Without loss of generality, consider $i < j$. Then,*

$$\begin{aligned}&(\varphi_j \circ \varphi_i^{-1})(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n) \\ &\qquad = \left(\frac{z_1}{z_j}, \ldots, \frac{z_{i-1}}{z_j}, \frac{1}{z_j}, \frac{z_{i+1}}{z_j}, \ldots, \frac{z_{j-1}}{z_j}, \frac{z_{j+1}}{z_j}, \ldots, \frac{z_n}{z_j}\right)\end{aligned}$$

is indeed smooth on the appropriate domain, and similarly for $\varphi_i \circ \varphi_j^{-1}$.

As a result, the charts form an atlas for RP^{n-1}, turning it into a manifold.*

In Chapter 9, we discuss a generalization of this idea: the Grassmann manifold, which consists of all linear subspaces of a given dimension.

It is important to note that, in general, a set M may admit two (or more) distinct atlases $\mathcal{A}$ and $\mathcal{A}'$ that are not compatible (their union is not an atlas), so that their corresponding maximal atlases are distinct. These two atlases then lead to different smooth structures on M, which shows that it is not sufficient to specify the set M: an atlas must also be specified—see Exercise 8.14.

Exercise 8.12. *Given an atlas $\mathcal{A}$ for a set M, show that the collection $\mathcal{A}^+$ of all charts of M which are compatible with $\mathcal{A}$ is a well-defined atlas of M.*

Exercise 8.13. *Show Definition 8.5 is independent of the choice of charts. Furthermore, show that if $F\colon \mathcal{M} \to \mathcal{M}'$ and $G\colon \mathcal{M}' \to \mathcal{M}''$ are smooth, then their composition $G \circ F$ is smooth. More broadly, establish the smoothness rules from Exercises 3.37, 3.38, 3.39 and 3.40 for general manifolds. We study the claims about differentials later in Exercise 8.40.*

Exercise 8.14. *For the set $M = \mathbb{R}$, consider the two following charts, both defined on all of M: $\varphi(x) = x$ and $\psi(x) = \sqrt[3]{x}$. Verify that these are indeed charts, and that they are* not *compatible. Let $\mathcal{A}^+$ be the maximal atlas generated by φ and let $\mathcal{M} = (M, \mathcal{A}^+)$ denote $\mathbb{R}$ with the resulting smooth structure (this is the usual structure on $\mathbb{R}$). Likewise, let $\mathcal{B}^+$ be the maximal atlas generated by ψ and write $\mathcal{M}' = (M, \mathcal{B}^+)$. Give an example of a function $f: \mathbb{R} \to \mathbb{R}$ which is* not *smooth as a function from $\mathcal{M}$ to $\mathbb{R}$ yet which* is *smooth as a function from $\mathcal{M}'$ to $\mathbb{R}$.*

8.2 The atlas topology, and a final definition

In the above section, we have equipped a set M with a smooth structure. This affords us the notion of smooth functions between properly endowed sets. As we now show, this structure further induces a topology on M, that is, a notion of open sets, called the *atlas topology*. In turn, having a topology on M is useful in optimization to define concepts such as local optima and convergence.

We start with a few reminders. After discussing two desirable properties of topologies, we restrict the definition of manifold to those whose atlas topology enjoy those properties.

The usual notion of open sets in $\mathbb{R}^d$ can be abstracted to arbitrary sets as topologies. Essentially, in defining a topology, we declare certain subsets to be open, while making sure that certain basic properties hold, as specified below.

Definition 8.15. *A* topology *on a set M is a collection $\mathcal{T}$ of subsets of M with the following properties. A subset of M is called* open *if and only if it is in $\mathcal{T}$, and:*

1. *M and $\emptyset$ are open;*
2. *The union of any collection of open sets is open; and*
3. *The intersection of any finite collection of open sets is open.*

A subset C of M is called closed *if it is the complement of an open set in M, that is, $M \backslash C$ is open. In particular, M and $\emptyset$ are both open and closed. Some subsets of M may be neither open nor closed.*

A *topological space* is a pair $(M, \mathcal{T})$ consisting of a set with a topology. Given two topological spaces $(M, \mathcal{T})$, $(M', \mathcal{T}')$ and a map $F: M \to M'$, we define that F is *continuous* if for every open set O' in M' the pre-image

$$F^{-1}(O') = \{x \in M : F(x) \in O'\}$$

is open in M.

In defining a topology on a manifold* $\mathcal{M} = (M, \mathcal{A}^+)$, it is natural to require that the chart functions be continuous in that topology. In particular, since for any chart $(\mathcal{U}, \varphi)$ of $\mathcal{M}$ we have that $\varphi(\mathcal{U})$ is open in $\mathbb{R}^d$ (assuming $\dim \mathcal{M} = d$), we should require that $\varphi^{-1}(\varphi(\mathcal{U})) = \mathcal{U}$ be open, that is: chart domains should be deemed open. It is easy to check with the following definitions that this collection of sets forms a *basis* for a topology consisting in the collection of all unions of chart domains [BC70, Prop. 2.4.2].

Definition 8.16. *A collection $\mathcal{B}$ of subsets of a set M is a* basis *for a topology on M if*

1. *For each $x \in M$, there is a set $B \in \mathcal{B}$ such that $x \in B$; and*
2. *If $x \in B_1 \cap B_2$ for $B_1, B_2 \in \mathcal{B}$, there exists $B_3 \in \mathcal{B}$ such that $x \in B_3$ and $B_3 \subseteq B_1 \cap B_2$.*

The topology $\mathcal{T}$ defined by $\mathcal{B}$ is the collection of all unions of elements of $\mathcal{B}$.

In the following definition, it is important to consider the *maximal* atlas as otherwise we may miss open some sets.

Definition 8.17. *Given a maximal atlas $\mathcal{A}^+$ on a set M, the* atlas topology *on M states that a subset of M is open if and only if it is the union of a collection of chart domains.*

A subset S of a topological space $\mathcal{T}$ inherits a topology called the *subspace topology*: it consists in the collection of all open sets of $\mathcal{T}$ intersected with S. By default, when we consider a subset of a topological space, we tacitly equip it with the subspace topology. With this in mind, we get the following convenient fact, true by design [BC70, Prop. 2.4.3].

Proposition 8.18. *In the atlas topology, any chart $\varphi\colon \mathcal{U} \to \varphi(\mathcal{U})$ is continuous and its inverse is also continuous (i.e., it is a* homeomorphism*.)*

A welcome consequence of the latter proposition is that, with the atlas topologies on manifolds* $\mathcal{M}$ and $\mathcal{M}'$, any function $F\colon \mathcal{M} \to \mathcal{M}'$ which is smooth in the sense of Definition 8.5 is also continuous in the topological sense [BC70, Prop. 2.4.4].

One of the reasons we need to discuss topologies in some detail is that, in general, atlas topologies may lack certain desirable properties: we must require them explicitly. The first such property is called *Hausdorff* (or T_2).

Definition 8.19. *A topology on a set M is* Hausdorff *if all pairs of distinct points have disjoint neighborhoods, that is: for all x, x' distinct in M there exist open sets O and O' such that $x \in O$, $x' \in O'$ and $O \cap O' = \emptyset$.*

Recall that a sequence $x_0, x_1, \ldots$ on a topological space is said to *converge* to x if, for every neighborhood $\mathcal{U}$ of x, there exists an index k such that $x_k, x_{k+1}, \ldots$ are all in $\mathcal{U}$: we then say that the sequence is convergent and that x is its limit. Crucially for optimization, in a Hausdorff topology, any convergent sequence of points has a unique limit [Lee12, p600]. This may not be the case otherwise (consider for example the trivial topology, in which the only open sets are the empty set and the set itself.)

The second desirable property is called *second-countable*.

Definition 8.20. *A topology is* second-countable *if there is a countable basis for its topology.*

At last, we can give a proper definition of manifolds.

Definition 8.21. *A* manifold *is a pair $\mathcal{M} = (M, \mathcal{A}^+)$ consisting of a set M and a maximal atlas $\mathcal{A}^+$ on M such that the atlas topology is Hausdorff and second-countable.*

A manifold* is indeed not always a manifold: the atlas topology is not always Hausdorff (see Examples 3.2.1–3 in [BC70]), and it may also not be second-countable (see Example 3.3.2 in the same reference). The following proposition gives a convenient way of ensuring a (not necessarily maximal) atlas induces a suitable topology [Lee12, Lem. 1.35].

Proposition 8.22. *Let $\mathcal{A}$ be an atlas for the set M. Assume both:*

1. *For all $x, y \in M$ distinct, either both x and y are in the domain of some chart, or there exist two disjoint chart domains $\mathcal{U}$ and $\mathcal{V}$ such that $x \in \mathcal{U}$ and $y \in \mathcal{V}$; and*
2. *Countably many of the chart domains suffice to cover M.*

Then, the atlas topology of $\mathcal{A}^+$ is Hausdorff (by property 1) and second-countable (by property 2), so that $\mathcal{M} = (M, \mathcal{A}^+)$ is a manifold.

The following proposition provides yet another way of assessing the atlas topology [BC70, Prop. 3.1.1]. We use it in Section 8.3. The "only if" direction is a direct consequence of Proposition 8.18.

Proposition 8.23. *Let the set M be equipped with both a maximal atlas $\mathcal{A}^+$ and a topology $\mathcal{T}$. The atlas topology on M coincides with $\mathcal{T}$ if and only if the charts of one atlas of M in $\mathcal{A}^+$ are homeomorphisms with respect to $\mathcal{T}$.*

Open subsets of manifolds are manifolds in a natural way by restriction of the chart domains, called *open submanifolds*. Unless otherwise specified, when working with an open subset of a manifold (often, a chart domain), we implicitly mean to use the open submanifold geometry. See also Section 8.14 for further facts about open submanifolds.

Definition 8.24. *Let $\mathcal{M}$ be a manifold and let $\mathcal{V}$ be open in $\mathcal{M}$ in the atlas topology. For any chart $(\mathcal{U}, \varphi)$ of $\mathcal{M}$ such that $\mathcal{U} \cap \mathcal{V} \neq \emptyset$, build the chart $(\mathcal{U} \cap \mathcal{V}, \varphi)$ on $\mathcal{V}$. The collection of these charts forms an atlas for $\mathcal{V}$, turning it into a manifold in its own right. Equipped with this atlas, we call $\mathcal{V}$ an* open submanifold *of $\mathcal{M}$.*

Example 8.25. *In all examples from Section 8.1, we have constructed atlases with a finite number of charts. Hence, by Proposition 8.22, their atlas topologies are second-countable. Furthermore, for linear spaces and open subsets of linear spaces, we have used only one chart, so that the same proposition guarantees the* *resulting topologies are Hausdorff. We conclude that linear spaces and their open subsets are manifolds.*

Part of the motivation for the topological restrictions introduced in this section is that a manifold* carries *partitions of unity* if (and essentially only if) the topology is as prescribed—see [BC70, §3.4]. Partitions of unity are useful in particular to show existence of Riemannian metrics (see Section 8.9). In short: every manifold can be turned into a Riemannian manifold [Lee12, Prop. 13.3].

We close this section with the definition of compact manifolds, for which we first recall a few topological notions. (See also Theorem 10.8.)

Definition 8.26. *Let $\mathcal{M} = (M, \mathcal{T})$ be a topological space (for example, a manifold with its atlas topology). An* open cover *of a subset S of $\mathcal{M}$ is a collection of open sets of $\mathcal{M}$ whose union contains S. We say S is* compact *if, for each open cover of S, one can select a finite number of open sets from that open cover whose union still contains S (called a* finite subcover*). The space $\mathcal{M}$ itself is compact if $S = M$ is compact.*

Definition 8.27. *A* compact manifold *is a manifold which is compact as a topological space with its atlas topology.*

Example 8.28. *An embedded submanifold $\mathcal{M}$ of a Euclidean space $\mathcal{E}$ is a compact manifold if and only if $\mathcal{M}$ is a compact subset of $\mathcal{E}$, that is, $\mathcal{M}$ is closed and bounded as a subset of $\mathcal{E}$. This is because $\mathcal{M}$ inherits its topology from $\mathcal{E}$, as we shall discuss in Section 8.3. In particular, the unit sphere, the Stiefel manifold, the orthogonal group and the special orthogonal group as discussed in Chapter 7 all are compact manifolds.*

Exercise 8.29. *To show that the circle S^1 and the real projective space RP^{n-1} are manifolds, it remains to verify that their atlases (as constructed in Section 8.1) induce Hausdorff topologies. Do this using Proposition 8.22. You may need to add a few charts to the atlases.*

Exercise 8.30. *Check that Definition 8.24 is legitimate, that is, show that the proposed charts are indeed charts, that they form an atlas, and that the atlas topology is Hausdorff and second-countable.*

Exercise 8.31. *Let $\mathcal{M}$ and $\mathcal{N}$ be two manifolds. For any pair of charts $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ of $\mathcal{M}$ and $\mathcal{N}$ respectively, consider the map ϕ defined on $\mathcal{U} \times \mathcal{V}$ by $\phi(x, y) = (\varphi(x), \psi(y))$. Show that these maps define a smooth structure on the product space $\mathcal{M} \times \mathcal{N}$, called the* product manifold structure*. Deduce that $\dim(\mathcal{M} \times \mathcal{N}) = \dim \mathcal{M} + \dim \mathcal{N}$, and that open subsets of the product manifold are unions of products of open subsets of $\mathcal{M}$ and $\mathcal{N}$.*

8.3 Embedded submanifolds are manifolds

All the way back in Chapter 3, we defined embedded submanifolds of linear spaces with Definition 3.10. In this section, we show that all sets we have thus far called embedded submanifolds are indeed manifolds. To do so, we equip them with an atlas, and we confirm that the corresponding atlas topology coincides with the topology we have been using so far. In Section 8.14, we shall also see that our early notion of smooth maps between embedded submanifolds of linear spaces agrees with the more general notion of smooth maps between manifolds.

Proposition 8.32. *A subset $\mathcal{M}$ of a linear space $\mathcal{E}$ which is an embedded submanifold as per Definition 3.10 admits an atlas which makes it a manifold in the sense of Definition 8.21. The corresponding atlas topology coincides with the subspace topology as given in Definition 3.21.*

Proof. Let $d = \dim \mathcal{E}$ and $n = \dim \mathcal{M} = d - k$. The claim has two parts.

Part 1.
We construct an atlas for $\mathcal{M}$ to make it a manifold*. Let $x \in \mathcal{M}$ be arbitrary. By Theorem 3.12, there exists a neighborhood U of x in $\mathcal{E}$, an open set W in $\mathbb{R}^d$ and a diffeomorphism $F \colon U \to W$ such that $F(\mathcal{M} \cap U) = E \cap W$ where $E = \{y \in \mathbb{R}^d : y_{n+1} = \cdots = y_d = 0\}$ is a linear subspace of $\mathbb{R}^d$. We use F to propose a tentative chart $(\mathcal{U}, \varphi)$ for $\mathcal{M}$ around x: let

$$\mathcal{U} = \mathcal{M} \cap U \qquad \text{and} \qquad \varphi \colon \mathcal{U} \to \varphi(\mathcal{U}) \colon y \mapsto \varphi(y) = \mathrm{trim}(F(y)), \tag{8.2}$$

where $\mathrm{trim} \colon \mathbb{R}^d \to \mathbb{R}^n$ discards the last k components of a vector. This map is invertible since the k entries removed by trim are identically zero on $\mathcal{U}$, so that

$$\varphi^{-1}(z) = F^{-1}(\mathrm{zpad}(z)), \tag{8.3}$$

where $\mathrm{zpad} \colon \mathbb{R}^n \to \mathbb{R}^d$ pads a vector with k zeros at the end. The composition $\mathrm{trim} \circ \mathrm{zpad}$ is identity on $\mathbb{R}^n$ while $\mathrm{zpad} \circ \mathrm{trim}$ is identity on E. Notice that W is open in $\mathbb{R}^d$ and

$$\varphi(\mathcal{U}) = \mathrm{trim}(F(\mathcal{M} \cap U)) = \mathrm{trim}(E \cap W).$$

One can then verify that $\varphi(\mathcal{U})$ is open in $\mathbb{R}^n$ using standard properties of the topologies on $\mathbb{R}^n$ and $\mathbb{R}^d$. Thus, $(\mathcal{U}, \varphi)$ is an n-dimensional chart for $\mathcal{M}$ around x. Such a chart can be constructed around every point $x \in \mathcal{M}$, so that we cover the whole set. The last step is to verify that the charts are compatible. To this end, consider two charts as above, $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$, with overlapping domains and associated diffeomorphisms $F \colon U \to F(U) \subseteq \mathbb{R}^d$ and $G \colon V \to G(V) \subseteq \mathbb{R}^d$. Then, the change of coordinates map is

$$\psi \circ \varphi^{-1} = \mathrm{trim} \circ G \circ F^{-1} \circ \mathrm{zpad},$$

from $\varphi(\mathcal{U} \cap \mathcal{V})$ to $\psi(\mathcal{U} \cap \mathcal{V})$. These domains are open because $U \cap V$ is open, hence so are $F(U \cap V)$ and $G(U \cap V)$ and we have

$$\begin{aligned}\varphi(\mathcal{U} \cap \mathcal{V}) &= \operatorname{trim}(E \cap F(U \cap V)), \text{ and}\\ \psi(\mathcal{U} \cap \mathcal{V}) &= \operatorname{trim}(E \cap G(U \cap V)).\end{aligned}$$

To check the first identity, verify that $F(\mathcal{U} \cap \mathcal{V}) = E \cap F(U \cap V)$ as follows, then compose with trim:

- $F(\mathcal{U} \cap \mathcal{V}) \subseteq E \cap F(U \cap V)$ as $F(\mathcal{U} \cap \mathcal{V}) = F(\mathcal{M} \cap U \cap V) \subseteq F(U \cap V)$ and $F(\mathcal{U} \cap \mathcal{V}) \subseteq F(\mathcal{U}) = E \cap F(U) \subseteq E$, and
- $E \cap F(U \cap V) \subseteq F(\mathcal{U} \cap \mathcal{V})$ since for all $y \in E \cap F(U \cap V)$ there exists $x \in U \cap V$ such that $F(x) = y$, hence x is in U and $F(x)$ is in E, which implies that x is in $\mathcal{U} = \mathcal{M} \cap U$; since x is also in V we deduce that x is in $\mathcal{U} \cap \mathcal{V}$, hence y is in $F(\mathcal{U} \cap \mathcal{V})$.

Overall, we find that the change of coordinates map $\psi \circ \varphi^{-1}$ is smooth (by composition) and its inverse $\varphi \circ \psi^{-1} = \operatorname{trim} \circ F \circ G^{-1} \circ \operatorname{zpad}$ is also smooth, so that the charts are compatible. This finishes the construction of our atlas, turning $\mathcal{M}$ into a manifold*.

Part 2.
That the atlas and subspace topologies coincide follows from Proposition 8.23. Indeed, we only need to show that the charts constructed above are homeomorphisms with respect to the subspace topology on $\mathcal{M}$. By definition, $\mathcal{U} = \mathcal{M} \cap U$ is open in that topology. Furthermore, $\varphi(\mathcal{U})$ is open in $\mathbb{R}^n$ as we argued above. Since the map $\varphi \colon \mathcal{U} \to \varphi(\mathcal{U})$ is invertible, it remains to argue that it and its inverse are continuous in the subspace topology. That φ is continuous is clear since it is the restriction of the continuous map $\operatorname{trim} \circ F$ from U to $\mathcal{U}$. That φ^{-1} is continuous is also clear since it is equal to the continuous map $F^{-1} \circ \operatorname{zpad}$, only with the codomain restricted to $\mathcal{U}$.

The topology on $\mathcal{E}$ is Hausdorff and second-countable, and it is easy to see that the subspace topology inherits these properties. Thus, we conclude that $\mathcal{M}$ equipped with the above atlas is a manifold. □

Additionally, the constructed atlas yields the *unique* smooth structure on $\mathcal{M}$ for which the atlas topology coincides with the subspace topology—see Section 8.14. This is why, even though in general it does not make sense to say that a set is or is not a manifold, it does make sense to say that a subset of a linear space is or is not an embedded submanifold of that linear space.

8.4 Tangent vectors and tangent spaces

In defining tangent vectors to a manifold in Section 3.2, we relied heavily on the linear embedding space. In the general setting however, we do not have this luxury. We must turn to a more general, intrinsic definition. Here, we present one general definition of tangent vectors on manifolds as equivalence classes of curves. Another (equivalent) definition is through the notion of *derivation* (at a point): we do not discuss it.

Let x be a point on a d-dimensional manifold $\mathcal{M}$. Consider the set C_x of smooth curves on $\mathcal{M}$ passing through x at $t = 0$:

$$C_x = \{c \mid c\colon I \to \mathcal{M} \text{ is smooth and } c(0) = x\}.$$

Smoothness of c on an open interval $I \subseteq \mathbb{R}$ around 0 is to be understood through Definition 8.5.

We define an equivalence relation on C_x denoted by $\sim$. Let $(\mathcal{U}, \varphi)$ be a chart of $\mathcal{M}$ around x and consider $c_1, c_2 \in C_x$. Then, $c_1 \sim c_2$ if and only if $\varphi \circ c_1$ and $\varphi \circ c_2$ have the same derivative at $t = 0$, that is:

$$c_1 \sim c_2 \qquad \iff \qquad (\varphi \circ c_1)'(0) = (\varphi \circ c_2)'(0). \tag{8.4}$$

These derivatives are well defined as $\varphi \circ c_i$ is a smooth function (by composition) from some open interval around 0 to an open subset of $\mathbb{R}^d$. It is an exercise to prove that this equivalence relation is independent of the choice of chart.

The equivalence relation partitions C_x into equivalence classes: we call them *tangent vectors*. The rationale is that all the curves in a same equivalence class (and only those) pass through x with the same "velocity", as judged by their velocities through $\varphi(x)$ in coordinates.

Definition 8.33. *The equivalence class of a curve $c \in C_x$ is the set of curves that are equivalent to c as per* (8.4)*:*

$$[c] = \{\hat{c} \in C_x : c \sim \hat{c}\}.$$

Each equivalence class is called a tangent vector *to $\mathcal{M}$ at x. The* tangent space *to $\mathcal{M}$ at x, denoted by $\mathrm{T}_x\mathcal{M}$, is the quotient set*

$$\mathrm{T}_x\mathcal{M} = C_x/\sim\; = \{[c] : c \in C_x\},$$

that is, the set of all equivalence classes.

Given a chart $(\mathcal{U}, \varphi)$ around x, the map

$$\theta_x^\varphi \colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}^d \colon [c] \mapsto \theta_x^\varphi([c]) = (\varphi \circ c)'(0) \tag{8.5}$$

is well defined by construction: the expression $(\varphi \circ c)'(0)$ does not depend on the choice of representative c in $[c]$. It is an exercise to show that θ_x^φ is bijective. This bijection naturally induces a linear space structure over $\mathrm{T}_x\mathcal{M}$, by copying the linear structure of $\mathbb{R}^d$:

$$a \cdot [c_1] + b \cdot [c_2] \triangleq (\theta_x^\varphi)^{-1}\big(a \cdot \theta_x^\varphi([c_1]) + b \cdot \theta_x^\varphi([c_2])\big). \tag{8.6}$$

This structure, again, is independent of the choice of chart. Thus, the tangent space is a linear space in its own right.

Theorem 8.34. *Tangent spaces are linear spaces of dimension* $\dim \mathcal{M}$ *with the linear structure given through* (8.6).

When $\mathcal{M}$ is an embedded submanifold of a linear space, the two definitions of tangent spaces we have seen are compatible in the sense that they yield the same vector space structure, so that we always use the simpler one. In particular, the tangent spaces of (an open subset of) a linear space $\mathcal{E}$ (for example, $\mathbb{R}^d$) are identified with $\mathcal{E}$ itself.

Theorem 8.35. *For* $\mathcal{M}$ *embedded in a linear space* $\mathcal{E}$*, there exists a linear space isomorphism (that is, an invertible linear map) showing that Definitions 3.14 and 8.33 are compatible.*

Proof. Pick $x \in \mathcal{M}$. Let $(\mathcal{U}, \varphi)$ be a chart around x as built in (8.2) from a diffeomorphism F so that $\varphi = \operatorname{trim} \circ F|_{\mathcal{U}}$ and $\varphi^{-1} = F^{-1} \circ \operatorname{zpad}|_{\varphi(\mathcal{U})}$. Pick an arbitrary smooth curve c on $\mathcal{M}$ satisfying $c(0) = x$. This is also a curve in $\mathcal{E}$: let $v = c'(0) \in \mathcal{E}$. Passing to coordinates, define $\tilde{c}(t) = \varphi(c(t))$. Write $c = \varphi^{-1} \circ \tilde{c} = F^{-1} \circ \operatorname{zpad} \circ \tilde{c}$ to see that $F \circ c = \operatorname{zpad} \circ \tilde{c}$. Thus,

$$
\begin{aligned}
F(x) &= F(c(0)) = \operatorname{zpad}(\tilde{c}(0)), \text{ and} \\
\mathrm{D}F(x)[v] &= (F \circ c)'(0) = \operatorname{zpad}(\tilde{c}'(0)).
\end{aligned}
$$

Moreover, $\theta_x^\varphi([c]) = (\varphi \circ c)'(0) = \tilde{c}'(0)$. Therefore, with $v = c'(0)$,

$$
v = \mathrm{D}F(x)^{-1}[\operatorname{zpad}(\theta_x^\varphi([c]))]. \tag{8.7}
$$

This is a linear map converting the tangent vector $[c]$ in the sense of Definition 8.33 to the tangent vector v in the sense of Definition 3.14. This map is one-to-one, with inverse given by:

$$
[c] = (\theta_x^\varphi)^{-1}(\operatorname{trim}(\mathrm{D}F(x)[v])). \tag{8.8}
$$

Thus, the two definitions of tangent spaces are compatible. □

Exercise 8.36. *Show that the equivalence relation* (8.4) *is independent of the choice of chart* $(\mathcal{U}, \varphi)$ *around* x*. Show that* θ_x^φ (8.5) *is bijective. Show that the linear structure on* $\mathrm{T}_x\mathcal{M}$ *defined by* (8.6) *is independent of the choice of chart, so that it makes sense to talk of linear combinations of tangent vectors without specifying a chart.*

8.5 Differentials of smooth maps

By design, the notion of tangent vector induces a notion of directional derivatives. Let $F \colon \mathcal{M} \to \mathcal{M}'$ be a smooth map. For any tangent vector $v \in \mathrm{T}_x\mathcal{M}$, pick a representative curve c (formally, $c \in v$) and consider the map $t \mapsto F(c(t))$: this is a smooth curve on $\mathcal{M}'$ passing through $F(x)$ at $t = 0$. The equivalence class of that curve is a tangent vector to $\mathcal{M}'$ at $F(x)$. The equivalence relation (8.4) is specifically crafted so that this map between tangent spaces does not depend on the choice of c in v. This yields a notion of differential for maps between manifolds. In equation (8.9) below, brackets on the right-hand side select an equivalence class of curves, whereas brackets on the left-hand side merely distinguish between x (the point at which we differentiate) and v (the direction along which we differentiate) as per usual.

Definition 8.37. *Given manifolds $\mathcal{M}$ and $\mathcal{M}'$, the differential of a smooth map $F\colon \mathcal{M} \to \mathcal{M}'$ at x is a linear map $\mathrm{D}F(x)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_{F(x)}\mathcal{M}'$ defined by:*

$$\mathrm{D}F(x)[v] = [t \mapsto F(c(t))], \tag{8.9}$$

where c is a smooth curve on $\mathcal{M}$ passing through x at $t = 0$ such that $v = [c]$.

When the codomain of F is (an embedded submanifold of) a linear space, Theorem 8.35 provides an identification of the abstract tangent spaces of that codomain with the concrete tangent spaces from Chapter 3. In this way, we can confirm that Definitions 8.37 and 3.34 are compatible.

Proposition 8.38. *For a smooth map $F\colon \mathcal{M} \to \mathcal{N}$ where $\mathcal{N}$ is an embedded submanifold of a linear space $\mathcal{E}$, we identify the tangent spaces of $\mathcal{N}$ to subspaces of $\mathcal{E}$ as provided by Theorem 8.35. Then, with $v = [c]$ a tangent vector at $x \in \mathcal{M}$, we can write*

$$\mathrm{D}F(x)[v] = (F \circ c)'(0), \tag{8.10}$$

where $F \circ c$ is seen as a map into $\mathcal{E}$.

In particular, let $\mathfrak{F}(\mathcal{M})$ denote the set of smooth scalar fields on $\mathcal{M}$, that is, the set of smooth functions $f\colon \mathcal{M} \to \mathbb{R}$. Then, identifying the tangent spaces of $\mathbb{R}$ with $\mathbb{R}$ itself, we write

$$\mathrm{D}f(x)[v] = (f \circ c)'(0) \tag{8.11}$$

for the differential $\mathrm{D}f(x)\colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}$, where $v = [c]$.

Proof. This is essentially a tautology. Let us write the proof for a map $G\colon \mathcal{M} \to \mathcal{N}$ so we can use F to denote the diffeomorphism appearing in the conversion formula (8.7) for a chart φ of $\mathcal{N}$ around $G(x)$. On the one hand, since $G \circ c$ is a curve on $\mathcal{N}$ passing through $G(x)$, formula (8.7) provides

$$(G \circ c)'(0) = \mathrm{D}F(G(x))^{-1}[\mathrm{zpad}(\theta^{\varphi}_{G(x)}([G \circ c]))]. \tag{8.12}$$

On the other hand, Definition 8.37 states that $\mathrm{D}G(x)[v] = [G \circ c]$, and the concrete representation of $[G \circ c]$ is obtained through (8.7) as:

$$\mathrm{D}F(G(x))^{-1}[\mathrm{zpad}(\theta^{\varphi}_{G(x)}([G \circ c]))].$$

Thus, the concrete representation of $\mathrm{D}G(x)[v]$ is $(G \circ c)'(0)$. $\square$

Exercise 8.39. *Verify that equation (8.9) is well defined, that is, the right-hand side does not depend on the choice of c representing v. Additionally, show that $\mathrm{D}F(x)$ is indeed a linear map with respect to the linear structure (8.6) on tangent spaces.*

Exercise 8.40. *(Continued from Exercise 8.13.) For smooth maps $F_1, F_2 \colon \mathcal{M} \to \mathcal{E}$ (with $\mathcal{E}$ a linear space) and real numbers a_1, a_2, show that $F \colon x \mapsto a_1 F_1(x) + a_2 F_2(x)$ is smooth and we have linearity:*

$$\mathrm{D}F(x) = a_1 \mathrm{D}F_1(x) + a_2 \mathrm{D}F_2(x).$$

For smooth maps $f \in \mathfrak{F}(\mathcal{M})$ and $G \colon \mathcal{M} \to \mathcal{E}$, show that the product map $fG \colon x \mapsto f(x)G(x)$ is smooth from $\mathcal{M}$ to $\mathcal{E}$ and we have a product rule:

$$\mathrm{D}(fG)(x)[v] = G(x)\mathrm{D}f(x)[v] + f(x)\mathrm{D}G(x)[v].$$

Let $F \colon \mathcal{M} \to \mathcal{M}'$ and $G \colon \mathcal{M}' \to \mathcal{M}''$ be smooth. Establish the chain rule for the differential of their composition:

$$\mathrm{D}(G \circ F)(x)[v] = \mathrm{D}G(F(x))[\mathrm{D}F(x)[v]].$$

Generalize the claim of Exercise 3.40 too.

8.6 Tangent bundles and vector fields

Identically to Definition 3.42, we define the *tangent bundle* as the disjoint union of all tangent spaces, now provided by Definition 8.33.

Definition 8.41. *The* tangent bundle *of a manifold $\mathcal{M}$ is the set:*

$$\mathrm{T}\mathcal{M} = \{(x, v) : x \in \mathcal{M} \text{ and } v \in \mathrm{T}_x\mathcal{M}\}.$$

We often conflate notation for (x, v) and v when the context is clear.

Definition 8.42. *The projection $\pi \colon \mathrm{T}\mathcal{M} \to \mathcal{M}$ extracts the base of a vector, that is, $\pi(x, v) = x$. At times, we may write $\pi(v) = x$.*

Just like tangent bundles of embedded submanifolds are themselves embedded submanifolds (Theorem 3.43), tangent bundles of manifolds are manifolds in a natural way. (Smoothness of π is understood through Definition 8.5.)

Theorem 8.43. *For any manifold $\mathcal{M}$ of dimension d, the tangent bundle $\mathrm{T}\mathcal{M}$ is itself a manifold of dimension $2d$, in such a way that the projection $\pi \colon \mathrm{T}\mathcal{M} \to \mathcal{M}$ is smooth.*

Proof. From any chart $(\mathcal{U}, \varphi)$ of $\mathcal{M}$, we construct a chart $(\tilde{\mathcal{U}}, \tilde{\varphi})$ of $\mathrm{T}\mathcal{M}$ as follows. Define the domain $\tilde{\mathcal{U}} = \pi^{-1}(\mathcal{U})$ to be the set of all tangent vectors to any point in $\mathcal{U}$. Then, define $\tilde{\varphi} \colon \tilde{\mathcal{U}} \to \tilde{\varphi}(\tilde{\mathcal{U}}) \subseteq \mathbb{R}^{2d}$ as

$$\tilde{\varphi}(x, v) = (\varphi(x), \theta_x^{\varphi}(v)), \tag{8.13}$$

where θ_x^{φ} is defined by (8.5). See [Lee12, Prop. 3.18] for details. $\square$

The smooth structure on tangent bundles makes differentials of smooth maps be smooth maps themselves.

Proposition 8.44. *Consider a smooth map $F: \mathcal{M} \to \mathcal{M}'$ and its differential $\mathrm{D}F: \mathrm{T}\mathcal{M} \to \mathrm{T}\mathcal{M}'$ defined by $\mathrm{D}F(x, v) = \mathrm{D}F(x)[v]$. With the natural smooth structures on $\mathrm{T}\mathcal{M}$ and $\mathrm{T}\mathcal{M}'$, the map $\mathrm{D}F$ is smooth.*

Proof. Write $\mathrm{D}F$ in coordinates using charts from Theorem 8.43, then use Proposition 8.51 below. Details in [Lee12, Prop. 3.21]. □

The manifold structure on $\mathrm{T}\mathcal{M}$ makes it possible to define smooth vector fields on manifolds as smooth maps from $\mathcal{M}$ to $\mathrm{T}\mathcal{M}$.

Definition 8.45. *A* vector field *V is a map from $\mathcal{M}$ to $\mathrm{T}\mathcal{M}$ such that $\pi \circ V$ is the identity map. The vector at x is written $V(x)$ and lies in $\mathrm{T}_x\mathcal{M}$. If V is also a smooth map, then it is a* smooth vector field*. The set of smooth vector fields on $\mathcal{M}$ is denoted by $\mathfrak{X}(\mathcal{M})$.*

In Section 8.8, we use the following characterization of smooth vector fields to construct *coordinate vector fields.*

Proposition 8.46. *A vector field V on $\mathcal{M}$ is smooth if and only if, for every chart $(\mathcal{U}, \varphi)$ of $\mathcal{M}$, the map $x \mapsto \theta_x^\varphi(V(x))$ is smooth on $\mathcal{U}$.*

Proof. Using Definition 8.5 about smooth maps and the charts of $\mathrm{T}\mathcal{M}$ defined by (8.13), we conclude that V is smooth if and only if, for every chart $(\mathcal{U}, \varphi)$ of $\mathcal{M}$,

$$\tilde{V} = \tilde{\varphi} \circ V \circ \varphi^{-1} : \varphi(\mathcal{U}) \subseteq \mathbb{R}^d \to \mathbb{R}^{2d}$$

is smooth, where $\tilde{\varphi}(x, v) = (\varphi(x), \theta_x^\varphi(v))$. For $z = \varphi(x)$, we have

$$\tilde{V}(z) = (z, \theta_x^\varphi(V(x)))$$

so that $\tilde{V}$ is smooth if and only if $x \mapsto \theta_x^\varphi(V(x))$ is smooth on $\mathcal{U}$. □

Let V be a vector field on $\mathcal{M}$. As we did in Definition 5.5, we define the action of V on a smooth function $f \in \mathfrak{F}(\mathcal{U})$ with $\mathcal{U}$ open in $\mathcal{M}$ as the function $Vf : \mathcal{U} \to \mathbb{R}$ determined by

$$(Vf)(x) = \mathrm{D}f(x)[V(x)]. \tag{8.14}$$

Based on the latter, we mention a characterization of smooth vector fields which is sometimes useful. The proof in the direction we need is an exercise in Section 8.8. See [Lee12, Prop. 8.14] for the other one.

Proposition 8.47. *A vector field V on a manifold $\mathcal{M}$ is smooth if and only if Vf is smooth for all $f \in \mathfrak{F}(\mathcal{M})$.*

Exercise 8.48. *Show that for $V, W \in \mathfrak{X}(\mathcal{M})$ and $f, g \in \mathfrak{F}(\mathcal{M})$ the vector field $fV + gW$ is smooth.*

8.7 Retractions and velocity of a curve

Now equipped with broader notions of smooth maps, tangent vectors and tangent bundles for a manifold $\mathcal{M}$, we can generalize the notion of retraction from Definition 3.47.

Definition 8.49. *A* retraction *on $\mathcal{M}$ is a smooth map*

$$\mathrm{R}\colon \mathrm{T}\mathcal{M} \to \mathcal{M}\colon (x, v) \mapsto \mathrm{R}_x(v)$$

such that for each $(x, v) \in \mathrm{T}\mathcal{M}$ the curve $c(t) = \mathrm{R}_x(tv)$ satisfies $v = [c]$, where $[c]$ is the equivalence class of the curve c as per Definition 8.33.

The latter definition is somewhat abstract. We can give it a more familiar look by defining the notion of velocity of a curve on a general manifold.

Definition 8.50. *Let $c\colon I \to \mathcal{M}$ be a smooth curve. The* velocity of c at t*, denoted by $c'(t)$, is the tangent vector in $\mathrm{T}_{c(t)}\mathcal{M}$ given by*

$$c'(t) = [\tau \mapsto c(t + \tau)],$$

where the brackets on the right-hand side take the equivalence class of the shifted curve, as per Definition 8.33.

Observe that $c'(0) = [c]$. Thus, a smooth map $\mathrm{R}\colon \mathrm{T}\mathcal{M} \to \mathcal{M}$ is a retraction exactly if each curve $c(t) = \mathrm{R}_x(tv)$ satisfies $c'(0) = [c] = v$ and (as implicitly required by the latter) $c(0) = x$. This characterization matches Definition 3.47.

Moreover, it is equivalent still to define retractions as smooth maps $\mathrm{R}\colon \mathrm{T}\mathcal{M} \to \mathcal{M}\colon (x, v) \mapsto \mathrm{R}_x(v)$ such that, for all $(x, v) \in \mathrm{T}\mathcal{M}$, we have

1. $\mathrm{R}_x(0) = x$, and
2. $\mathrm{DR}_x(0)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ is the identity map: $\mathrm{DR}_x(0)[v] = v$.

To be clear, here, 0 denotes the zero tangent vector at x, that is, the equivalence class of smooth curves on $\mathcal{M}$ that pass through x at $t = 0$ with zero velocity, as judged through any chart around x. Also, the differential $\mathrm{DR}_x(0)$ makes sense as $\mathrm{R}_x\colon \mathrm{T}_x\mathcal{M} \to \mathcal{M}$ is a smooth map, and we identify the tangent spaces of $\mathrm{T}_x\mathcal{M}$ (a linear space) with itself (using Theorem 8.35), so that $\mathrm{T}_0(\mathrm{T}_x\mathcal{M})$—the domain of $\mathrm{DR}_x(0)$—is identified with $\mathrm{T}_x\mathcal{M}$.

8.8 Coordinate vector fields as local frames

Let $(\mathcal{U}, \varphi)$ be a chart on a d-dimensional manifold $\mathcal{M}$. Here and in many places, we use that $\mathcal{U}$ itself is a manifold; specifically, an open submanifold of $\mathcal{M}$: see Definition 8.24. Consider the following vector fields on $\mathcal{U}$, called *coordinate vector fields*:

$$W_i(x) = \left[t \mapsto \varphi^{-1}(\varphi(x) + te_i)\right], \qquad i = 1, \ldots, d, \tag{8.15}$$

where $e_1, \ldots, e_d$ are the canonical basis vectors for $\mathbb{R}^d$ (that is, the columns of the identity matrix of size d). The defining property of these vector fields is that, when pushed through θ_x^φ (8.5), they correspond to the constant coordinate vector fields of $\mathbb{R}^d$:

$$\theta_x^\varphi(W_i(x)) = \frac{\mathrm{d}}{\mathrm{d}t} \varphi(\varphi^{-1}(\varphi(x) + te_i))\bigg|_{t=0} = e_i. \tag{8.16}$$

As a corollary, we obtain a generalization of Proposition 3.69: local frames exist around any point on a manifold (see Definition 3.68).

Proposition 8.51. *Coordinate vector fields* (8.15) *are smooth on $\mathcal{U}$, that is, $W_1, \ldots, W_d$ belong to $\mathfrak{X}(\mathcal{U})$. Furthermore, they form a local frame, that is, for all $x \in \mathcal{U}$, the tangent vectors $W_1(x), \ldots, W_d(x)$ are linearly independent.*

Proof. Smoothness follows from (8.16) and Proposition 8.46. Now consider the linear structure on $\mathrm{T}_x\mathcal{M}$ defined by (8.6): $W_1(x), \ldots, W_d(x)$ are linearly independent if and only if they are so after being pushed through θ_x^φ, which is clearly the case owing to (8.16). □

To interpret the corollary below, use the fact that a vector field is smooth on $\mathcal{M}$ if and only if it is smooth when restricted to each chart domain $\mathcal{U}$.

Corollary 8.52. *Given a vector field V on $\mathcal{M}$ and a chart $(\mathcal{U}, \varphi)$, there exist unique functions $g_1, \ldots, g_d \colon \mathcal{U} \to \mathbb{R}$ such that $V|_\mathcal{U} = g_1 W_1 + \cdots + g_d W_d$. These functions are smooth if and only if $V|_\mathcal{U}$ is smooth.*

Proof. That functions $g_i \colon \mathcal{U} \to \mathbb{R}$ such that $V|_\mathcal{U} = \sum_i g_i W_i$ exist and are unique follows from linear independence of $W_1(x), \ldots, W_d(x)$. The smoothness equivalence follows from Proposition 8.46 and

$$\theta_x^\varphi(V(x)) = \sum_{i=1}^d g_i(x) \theta_x^\varphi(W_i(x)) = (g_1(x), \ldots, g_d(x))^\top, \tag{8.17}$$

where we used $\theta_x^\varphi(W_i(x)) = e_i$ by (8.16). □

Exercise 8.53. *Show that for all $V \in \mathfrak{X}(\mathcal{M})$ and $f \in \mathfrak{F}(\mathcal{M})$ the function Vf is smooth on $\mathcal{M}$ (this is one direction of Proposition 8.47).*

8.9 Riemannian metrics and gradients

Since tangent spaces are linear spaces, we can define inner products on them. The following definitions already appeared in the context of embedded submanifolds in Sections 3.7 and 3.8: they extend verbatim to the general case.

Definition 8.54. *An inner product on $\mathrm{T}_x\mathcal{M}$ is a bilinear, symmetric, positive definite function $\langle \cdot, \cdot \rangle_x : \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M} \to \mathbb{R}$. It induces a norm for tangent vectors: $\|u\|_x = \sqrt{\langle u, u \rangle_x}$. A* metric *on $\mathcal{M}$ is a choice of inner product $\langle \cdot, \cdot \rangle_x$ for each $x \in \mathcal{M}$.*

Definition 8.55. *A metric* $\langle \cdot, \cdot \rangle_x$ *on* $\mathcal{M}$ *is a* Riemannian metric *if it varies smoothly with* x*, in the sense that for all smooth vector fields* V, W *on* $\mathcal{M}$ *the function* $x \mapsto \langle V(x), W(x) \rangle_x$ *is smooth from* $\mathcal{M}$ *to* $\mathbb{R}$*.*

Definition 8.56. *A manifold with a Riemannian metric is a* Riemannian manifold.

Definition 8.57. *Let* $f \colon \mathcal{M} \to \mathbb{R}$ *be a smooth function on a Riemannian manifold* $\mathcal{M}$*. The* Riemannian gradient *of* f *is the vector field* $\mathrm{grad} f$ *on* $\mathcal{M}$ *uniquely defined by the following identities:*

$$\forall (x, v) \in \mathrm{T}\mathcal{M}, \qquad \mathrm{D}f(x)[v] = \langle v, \mathrm{grad} f(x) \rangle_x, \tag{8.18}$$

where $\mathrm{D}f(x)$ *is as in Proposition 8.38 and* $\langle \cdot, \cdot \rangle_x$ *is the Riemannian metric.*

The gradient of a smooth function is a smooth vector field: the proof of Proposition 3.70 extends as is, using local frames provided by Proposition 8.51 for example.

Proposition 8.58. *For* $f \in \mathfrak{F}(\mathcal{M})$*, the gradient* $\mathrm{grad} f$ *is smooth.*

Proposition 3.59 also holds true in the general case, with the same proof. We restate the claim here. See also Exercise 10.73.

Proposition 8.59. *Let* $f \colon \mathcal{M} \to \mathbb{R}$ *be a smooth function on a Riemannian manifold* $\mathcal{M}$ *equipped with a retraction* R*. Then, for all* $x \in \mathcal{M}$*,*

$$\mathrm{grad} f(x) = \mathrm{grad}(f \circ \mathrm{R}_x)(0), \tag{8.19}$$

where $f \circ \mathrm{R}_x \colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}$ *is defined on a Euclidean space (*$\mathrm{T}_x\mathcal{M}$ *with the inner product* $\langle \cdot, \cdot \rangle_x$*), hence its gradient is a "classical" gradient.*

Likewise, Example 3.57 and Exercise 3.67 regarding Riemannian product manifolds generalize verbatim for product manifolds as defined in Exercise 8.31.

8.10 Lie brackets as vector fields

Recall Definition 5.5 where we introduced the notion of Lie bracket of smooth vector fields $U, V \in \mathfrak{X}(\mathcal{M})$: for all $f \in \mathfrak{F}(\mathcal{U})$ with $\mathcal{U}$ open in $\mathcal{M}$, the Lie bracket $[U, V]$ acts on f and produces a smooth function on $\mathcal{U}$ defined by:

$$[U, V]f = U(Vf) - V(Uf). \tag{8.20}$$

We now extend Proposition 5.10 to show that $[U, V]$ acts on $\mathfrak{F}(\mathcal{M})$ in the exact same way that a specific smooth vector field does, which allows us to think of $[U, V]$ itself as being that smooth vector field. To this end, we first show a special property of coordinate vector fields.

Proposition 8.60. *Lie brackets of coordinate vector fields* (8.15) *vanish identically, that is,*

$$[W_i, W_j]f = 0$$

for all $1 \leq i, j \leq d$ *and all* $f \in \mathfrak{F}(\mathcal{U})$.

Proof. Writing f in coordinates as $\tilde{f} = f \circ \varphi^{-1}$ (smooth from $\varphi(\mathcal{U})$ open in $\mathbb{R}^d$ to $\mathbb{R}$ by Definition 8.5), we find using Proposition 8.38:

$$\begin{aligned}(W_i f)(x) &= \mathrm{D}f(x)[W_i(x)] \\ &= \frac{\mathrm{d}}{\mathrm{d}t} f\left(\varphi^{-1}\left(\varphi(x) + te_i\right)\right)\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \tilde{f}(\varphi(x) + te_i)\Big|_{t=0} \\ &= \mathrm{D}\tilde{f}(\varphi(x))[e_i] \\ &= \langle \mathrm{grad}\tilde{f}(\varphi(x)), e_i \rangle, \end{aligned} \tag{8.21}$$

where we use the canonical inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^d$ to define the Euclidean gradient of $\tilde{f}$. Using this result twice, we obtain

$$\begin{aligned}(W_j(W_i f))(x) &= \mathrm{D}\left((W_i f) \circ \varphi^{-1}\right)(\varphi(x))[e_j] \\ &= \mathrm{D}\left(\langle \mathrm{grad}\tilde{f}, e_i \rangle\right)(\varphi(x))[e_j] \\ &= \left\langle \mathrm{Hess}\tilde{f}(\varphi(x))[e_j], e_i \right\rangle. \end{aligned}$$

Since the Euclidean Hessian $\mathrm{Hess}\tilde{f}$ is self-adjoint, we find that

$$(W_i(W_j f))(x) = (W_j(W_i f))(x),$$

hence $([W_i, W_j]f)(x) = 0$ for all $x \in \mathcal{U}$ and for all i, j. □

Proposition 8.61. *Let* U, V *be two smooth vector fields on a manifold* $\mathcal{M}$. *There exists a unique smooth vector field* W *on* $\mathcal{M}$ *such that* $[U, V]f = Wf$ *for all* $f \in \mathfrak{F}(\mathcal{M})$. *We identify* $[U, V]$ *with that smooth vector field.*

Proof. We first show the claim on a chart domain. Let $(\mathcal{U}, \varphi)$ be a chart of $\mathcal{M}$, and let $W_1, \ldots, W_d$ be the corresponding coordinate vector fields (8.15). By Corollary 8.52, any two vector fields $U, V \in \mathfrak{X}(\mathcal{M})$ can be expressed on $\mathcal{U}$ as

$$U|_{\mathcal{U}} = \sum_{i=1}^{d} g_i W_i, \qquad V|_{\mathcal{U}} = \sum_{j=1}^{d} h_j W_j,$$

for a unique set of smooth functions $g_i, h_j \in \mathfrak{F}(\mathcal{U})$. For all $f \in \mathfrak{F}(\mathcal{U})$,

$$Vf = \sum_{j=1}^{d} h_j W_j f.$$

Using linearity and Leibniz' rule (Exercise 5.11),

$$U(Vf) = \sum_{i,j} g_i W_i(h_j W_j f) = \sum_{i,j} g_i (W_i h_j)(W_j f) + g_i h_j W_i(W_j f).$$

With similar considerations for $V(Uf)$, namely,

$$V(Uf) = \sum_{i,j} h_j W_j(g_i W_i f) = \sum_{i,j} h_j (W_j g_i)(W_i f) + h_j g_i W_j(W_i f),$$

we find

$$\begin{aligned}[U,V]f &= U(Vf) - V(Uf) \\ &= \sum_{i,j} g_i (W_i h_j)(W_j f) - h_j (W_j g_i)(W_i f) + \sum_{i,j} g_i h_j [W_i, W_j] f.\end{aligned}$$

Since $[W_i, W_j]f = 0$ by Proposition 8.60, it follows that, on the domain $\mathcal{U}$, there is a unique smooth vector field, specifically,

$$\sum_{i,j} g_i (W_i h_j) W_j - h_j (W_j g_i) W_i, \tag{8.22}$$

which acts on $\mathfrak{F}(\mathcal{U})$ in the exact same way as does $[U, V]$. This construction can be repeated on a set of charts whose domains cover $\mathcal{M}$. By uniqueness, the constructions on overlapping chart domains are compatible. Hence, this defines a smooth vector field on all of $\mathcal{M}$. We identify it with $[U, V]$. □

8.11 Riemannian connections and Hessians

The notion of connection applies in the general case. For convenience we repeat Definition 5.20 here. (Definition 5.1 also extends as is.)

Definition 8.62. *An (affine)* connection *on $\mathcal{M}$ is an operator*

$$\nabla \colon \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}) \colon (U, V) \mapsto \nabla_U V$$

which has three properties for all $U, V, W \in \mathfrak{X}(\mathcal{M})$, $f, g \in \mathfrak{F}(\mathcal{M})$ and $a, b \in \mathbb{R}$:

1. *$\mathfrak{F}(\mathcal{M})$-linearity in U: $\nabla_{fU+gW} V = f \nabla_U V + g \nabla_W V$;*
2. *$\mathbb{R}$-linearity in V: $\nabla_U(aV + bW) = a \nabla_U V + b \nabla_U W$; and*
3. *Leibniz rule: $\nabla_U(fV) = (Uf)V + f \nabla_U V$.*

The field $\nabla_U V$ is the covariant derivative *of V along U with respect to ∇.*

Likewise, Theorem 5.6 regarding the existence and uniqueness of a Riemannian connection extends without difficulty. We use Proposition 8.61 (stating Lie brackets are vector fields) to state the symmetry condition in a more standard way.

Theorem 8.63. *On a Riemannian manifold $\mathcal{M}$, there exists a unique connection ∇ which satisfies two additional properties for all $U, V, W \in \mathfrak{X}(\mathcal{M})$:*

4. Symmetry: $[U, V] = \nabla_U V - \nabla_V U$*; and*
5. Compatibility with the metric: $U\langle V, W\rangle = \langle \nabla_U V, W\rangle + \langle V, \nabla_U W\rangle$.

This connection is called the Levi-Civita *or* Riemannian connection.

As we showed in Proposition 5.21 in the embedded case, connections are pointwise operators in U. The proof from the embedded case extends to the general case with two changes: first, we now use the more general proof of existence of local frames provided by Proposition 8.51; second, we must reaffirm the technical Lemma 5.27 which allows us to make sense of ∇ when applied to locally defined smooth vector fields (such as coordinate vector fields for example).

Proposition 8.64. *For any connection* ∇ *and smooth vector fields* U, V *on a manifold* $\mathcal{M}$*, the vector field* $\nabla_U V$ *at* x *depends on* U *only through* $U(x)$*. Thus, we can write* $\nabla_u V$ *to mean* $(\nabla_U V)(x)$ *for any* $U \in \mathfrak{X}(\mathcal{M})$ *such that* $U(x) = u$*, without ambiguity.*

These observations allow us to extend Definition 5.14 for Riemannian Hessians to general manifolds.

Definition 8.65. *Let* $\mathcal{M}$ *be a Riemannian manifold with its Riemannian connection* ∇*. The* Riemannian Hessian *of* $f \in \mathfrak{F}(\mathcal{M})$ *at* $x \in \mathcal{M}$ *is the linear map* $\mathrm{Hess} f(x) \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ *defined as follows:*

$$\mathrm{Hess} f(x)[u] = \nabla_u \mathrm{grad} f.$$

Equivalently, $\mathrm{Hess} f$ *maps* $\mathfrak{X}(\mathcal{M})$ *to* $\mathfrak{X}(\mathcal{M})$ *as* $\mathrm{Hess} f[U] = \nabla_U \mathrm{grad} f$.

The proof that the Riemannian Hessian is self-adjoint, given for embedded submanifolds in Proposition 5.15, extends verbatim.

Proposition 8.66. *The Riemannian Hessian is self-adjoint with respect to the Riemannian metric. That is, for all* $x \in \mathcal{M}$ *and* $u, v \in \mathrm{T}_x\mathcal{M}$,

$$\langle \mathrm{Hess} f(x)[u], v\rangle_x = \langle u, \mathrm{Hess} f(x)[v]\rangle_x .$$

Likewise, considerations for connections on product manifolds from Exercises 5.4 and 5.13 also extend to the general case.

8.12 Covariant derivatives and geodesics

Recall Definition 5.28: given a smooth curve $c \colon I \to \mathcal{M}$ on a manifold $\mathcal{M}$, the map $Z \colon I \to \mathrm{T}\mathcal{M}$ is a *smooth vector field on* c if $Z(t)$ is in $\mathrm{T}_{c(t)}\mathcal{M}$ for all $t \in I$ and Z is smooth as a map from I (open in $\mathbb{R}$) to $\mathrm{T}\mathcal{M}$. The set of smooth vector fields on c is denoted by $\mathfrak{X}(c)$.

Theorem 5.29, both a definition of covariant derivatives and a statement of their existence and uniqueness, extends to general manifolds as is. So does its proof, provided we use local frames on general manifolds (Proposition 8.51) and we reaffirm the notation (5.14) justified in the embedded case.

Theorem 8.67. *Let $c\colon I \to \mathcal{M}$ be a smooth curve on a manifold equipped with a connection ∇. There exists a unique operator $\frac{\mathrm{D}}{\mathrm{d}t}\colon \mathfrak{X}(c) \to \mathfrak{X}(c)$ which satisfies the following properties for all $Y, Z \in \mathfrak{X}(c)$, $U \in \mathfrak{X}(\mathcal{M})$, $g \in \mathfrak{F}(I)$, and $a, b \in \mathbb{R}$:*

1. *$\mathbb{R}$-linearity: $\frac{\mathrm{D}}{\mathrm{d}t}(aY + bZ) = a\frac{\mathrm{D}}{\mathrm{d}t}Y + b\frac{\mathrm{D}}{\mathrm{d}t}Z$;*
2. *Leibniz rule: $\frac{\mathrm{D}}{\mathrm{d}t}(gZ) = g'Z + g\frac{\mathrm{D}}{\mathrm{d}t}Z$;*
3. *Chain rule: $\left(\frac{\mathrm{D}}{\mathrm{d}t}(U \circ c)\right)(t) = \nabla_{c'(t)}U$ for all $t \in I$.*

We call $\frac{\mathrm{D}}{\mathrm{d}t}$ the induced covariant derivative*. If moreover $\mathcal{M}$ is a Riemannian manifold and ∇ is compatible with its metric $\langle \cdot, \cdot \rangle$ (e.g., if ∇ is the Riemannian connection), then the induced covariant derivative also satisfies:*

4. *Product rule: $\frac{\mathrm{d}}{\mathrm{d}t}\langle Y, Z\rangle = \left\langle \frac{\mathrm{D}}{\mathrm{d}t}Y, Z\right\rangle + \left\langle Y, \frac{\mathrm{D}}{\mathrm{d}t}Z\right\rangle$,*

where $\langle Y, Z\rangle \in \mathfrak{F}(I)$ is defined by $\langle Y, Z\rangle(t) = \langle Y(t), Z(t)\rangle_{c(t)}$.

Recall the notion of velocity c' of a smooth curve c stated in Definition 8.50. Clearly, c' is a smooth vector field along c, that is, $c' \in \mathfrak{X}(c)$. Then, using the induced covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$, we may define acceleration along a curve similarly to Definition 5.36, and geodesics as in Definition 5.38.

Definition 8.68. *Let $c\colon I \to \mathcal{M}$ be a smooth curve. The* acceleration *of c is the smooth vector field $c'' \in \mathfrak{X}(c)$ defined by:*

$$c'' = \frac{\mathrm{D}}{\mathrm{d}t}c'.$$

A geodesic *is a smooth curve $c\colon I \to \mathcal{M}$ such that $c''(t) = 0$ for all $t \in I$.*

Exercises 5.34 and 5.39 regarding covariant derivatives on product manifolds extend as is, as does Exercise 5.35 for reparameterizations.

8.13 Taylor expansions and second-order retractions

Using the general tools constructed thus far, the reasoning that led to second-order Taylor expansions for embedded submanifolds and which culminated in eq. (5.26) extends to a general Riemannian manifold $\mathcal{M}$. Hence, we can state in general that, for $f \in \mathfrak{F}(\mathcal{M})$ and any smooth curve c on $\mathcal{M}$ such that $c(0) = x$ and $c'(0) = v$,

$$f(c(t)) = f(x) + t\,\langle \mathrm{grad} f(x), v\rangle_x + \frac{t^2}{2}\,\langle \mathrm{Hess} f(x)[v], v\rangle_x + \frac{t^2}{2}\,\langle \mathrm{grad} f(x), c''(0)\rangle_x + O(t^3). \quad (8.23)$$

Definition 5.42 extends as is to the general case.

Definition 8.69. *A* second-order retraction R *on a Riemannian manifold $\mathcal{M}$ is a retraction such that, for all $x \in \mathcal{M}$ and all $v \in \mathrm{T}_x\mathcal{M}$, the curve $c(t) = \mathrm{R}_x(tv)$ has zero acceleration at $t = 0$, that is, $c''(0) = 0$.*

In turn, this allows us to extend Propositions 5.44 and 5.45 to the general case with the same proofs, verbatim.

Proposition 8.70. *Consider a Riemannian manifold $\mathcal{M}$ equipped with any retraction* R, *and a smooth function $f\colon \mathcal{M} \to \mathbb{R}$. If x is a critical point of f (that is, if* $\mathrm{grad} f(x) = 0$*), then*

$$f(\mathrm{R}_x(s)) = f(x) + \frac{1}{2} \langle \mathrm{Hess} f(x)[s], s \rangle_x + O(\|s\|_x^3). \tag{8.24}$$

If R *is a second-order retraction, then for any point $x \in \mathcal{M}$ we have*

$$f(\mathrm{R}_x(s)) = f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{1}{2} \langle \mathrm{Hess} f(x)[s], s \rangle_x + O(\|s\|_x^3). \tag{8.25}$$

Proposition 8.71. *If the retraction is second order or if* $\mathrm{grad} f(x) = 0$, *then*

$$\mathrm{Hess} f(x) = \mathrm{Hess}(f \circ \mathrm{R}_x)(0),$$

where the right-hand side is the Hessian of $f \circ \mathrm{R}_x \colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}$ at $0 \in \mathrm{T}_x\mathcal{M}$.

8.14 Submanifolds embedded in manifolds

In Chapter 3, we defined our first class of smooth sets, which we called embedded submanifolds of linear spaces. In Section 8.3, we showed that embedded submanifolds of linear spaces are manifolds. Now, we define the concept of embedded submanifold of a manifold: this includes embedded submanifolds of linear spaces as a special case. This will serve us well in Chapter 9.

Given a subset $\mathcal{M}$ of a manifold $\overline{\mathcal{M}}$, there may exist many smooth structures for $\mathcal{M}$. These may or may not interact nicely with the smooth structure of $\overline{\mathcal{M}}$. Let us make this precise.

Consider the *inclusion map* $i\colon \mathcal{M} \to \overline{\mathcal{M}}$: it maps points of $\mathcal{M}$ to themselves in $\overline{\mathcal{M}}$, that is, $i(x) = x$. Depending on the smooth structure we choose for $\mathcal{M}$, this map may or may not be smooth. If it is, then we can differentiate it and $\mathrm{D}i(x)$ is a linear map from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_x\overline{\mathcal{M}}$. If that map is injective (for all x), we call $\mathcal{M}$ a *submanifold* of $\overline{\mathcal{M}}$.

Below, notice how, in order to define whether or not $\mathcal{M}$ is a submanifold of $\overline{\mathcal{M}}$, we first need $\mathcal{M}$ to be a manifold in its own right.

Definition 8.72. *Consider two manifolds, $\mathcal{M}$ and $\overline{\mathcal{M}}$, such that $\mathcal{M}$ (as a set) is included in $\overline{\mathcal{M}}$. If the inclusion map $i\colon \mathcal{M} \to \overline{\mathcal{M}}$ is smooth and* $\mathrm{D}i(x)$ *has rank equal to* $\dim \mathcal{M}$ *for all $x \in \mathcal{M}$, we say $\mathcal{M}$ is an* (immersed) submanifold *of $\overline{\mathcal{M}}$.*

Under the rank condition, $\dim \mathcal{M} \leq \dim \overline{\mathcal{M}}$ and the kernel of $\mathrm{D}i(x)$ is trivial. This is just as well, because otherwise there exists a smooth curve $c\colon I \to \mathcal{M}$ passing through $c(0) = x$ with nonzero velocity $c'(0)$, yet the 'same' curve $\bar{c} = i \circ c\colon I \to \overline{\mathcal{M}}$ on $\overline{\mathcal{M}}$ (smooth by composition) passes through x with zero velocity $\bar{c}'(0) = \mathrm{D}i(x)[c'(0)]$.

Among the submanifold structures of $\mathcal{M}$ (if any), there may exist at most one such that the atlas topology on $\mathcal{M}$ coincides with the subspace topology induced by $\overline{\mathcal{M}}$ [Lee12, Thm. 5.31]. When $\mathcal{M}$ admits such a smooth structure, we call $\mathcal{M}$ (with that structure) an *embedded submanifold of* $\overline{\mathcal{M}}$. (The 'figure-eight' example shows this is not always the case [Lee12, Fig. 4.3].)

Definition 8.73. *If $\mathcal{M}$ is a submanifold of $\overline{\mathcal{M}}$ and its atlas topology coincides with the subspace topology of $\mathcal{M} \subseteq \overline{\mathcal{M}}$ (that is, every open set of $\mathcal{M}$ is the intersection of some open set of $\overline{\mathcal{M}}$ with $\mathcal{M}$), then $\mathcal{M}$ is called an* embedded submanifold *of $\overline{\mathcal{M}}$, while $\overline{\mathcal{M}}$ is called the* ambient *or* embedding space.

Theorem 8.74. *A subset $\mathcal{M}$ of a manifold $\overline{\mathcal{M}}$ admits at most one smooth structure that makes $\mathcal{M}$ an embedded submanifold of $\overline{\mathcal{M}}$.*

Hence, it makes sense to say that a subset of a manifold is or is not an embedded submanifold, where in the affirmative we implicitly mean to endow $\mathcal{M}$ with that (unique) smooth structure.

The next result gives a complete characterization of embedded submanifolds. It reduces to Definition 3.10 when $\overline{\mathcal{M}}$ is a linear space $\mathcal{E}$.

Theorem 8.75. *Let $\overline{\mathcal{M}}$ be a manifold. A subset $\mathcal{M}$ of $\overline{\mathcal{M}}$ is an* embedded submanifold of $\overline{\mathcal{M}}$ *if and only if either of the following holds:*

1. *$\mathcal{M}$ is an open subset of $\overline{\mathcal{M}}$. Then,* $\dim \mathcal{M} = \dim \overline{\mathcal{M}}$ *and we also call this an* open submanifold *as in Definition 8.24; or*
2. *For a fixed integer $k \geq 1$ and for each $x \in \mathcal{M}$, there exists a neighborhood $\overline{\mathcal{U}}$ of x in $\overline{\mathcal{M}}$ and a smooth function $h \colon \overline{\mathcal{U}} \to \mathbb{R}^k$ such that*

$$h^{-1}(0) = \mathcal{M} \cap \overline{\mathcal{U}} \qquad \text{and} \qquad \operatorname{rank} \mathrm{D}h(x) = k.$$

Then, $\dim \mathcal{M} = \dim \overline{\mathcal{M}} - k$ *and h is called a* local defining function.

The tangent spaces of $\mathcal{M}$ are linear subspaces of those of $\overline{\mathcal{M}}$:

$$\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x) \subseteq \mathrm{T}_x\overline{\mathcal{M}}, \tag{8.26}$$

where h is any local defining function for $\mathcal{M}$ around x. Formally, the identification is done through $\mathrm{D}i(x) \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\overline{\mathcal{M}}$.

In Theorem 8.75, there is nothing special about $\mathbb{R}^k$: we could just as well consider local defining maps into an arbitrary manifold of dimension k, as this is locally equivalent to $\mathbb{R}^k$ through a chart. In particular, it often happens that an embedded submanifold can be defined with a single defining map, motivating the next corollary.

Corollary 8.76. *Let $h \colon \overline{\mathcal{M}} \to \mathcal{N}$ be a smooth map and consider its non-empty level set $\mathcal{M} = h^{-1}(\alpha)$. If $\mathrm{D}h(x)$ has rank equal to* $\dim \mathcal{N}$ *for all $x \in \mathcal{M}$, then $\mathcal{M}$ is closed in $\overline{\mathcal{M}}$, it is an embedded submanifold of $\overline{\mathcal{M}}$ with dimension* $\dim \mathcal{M} = \dim \overline{\mathcal{M}} - \dim \mathcal{N}$*, and* $\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x)$*.*

Above, the set $h^{-1}(\alpha)$ is closed since it is the pre-image of the singleton $\{\alpha\}$ through the continuous map h, and a singleton is closed in atlas topology since it maps to a singleton through a chart. An embedded submanifold which is closed in the embedding space is called *properly embedded* [Lee12, Prop. 5.5].

If the differential of h is not surjective at all points of $\mathcal{M}$, a version of Corollary 8.76 still holds provided the rank of the differential is constant *in a neighborhood* of $\mathcal{M}$. Crucially, it is *not* sufficient for this condition to hold just on $\mathcal{M}$: see Section 3.10.

Proposition 8.77. *Let $\mathcal{M} = h^{-1}(\alpha)$ be a non-empty level set of the smooth map $h\colon \overline{\mathcal{M}} \to \mathcal{N}$. If $\operatorname{rank} \mathrm{D}h(x) = r$ for all x in a neighborhood of $\mathcal{M}$ in $\overline{\mathcal{M}}$, then $\mathcal{M}$ is closed, it is an embedded submanifold of $\overline{\mathcal{M}}$ with dimension $\dim \mathcal{M} = \dim \overline{\mathcal{M}} - r$, and $\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x)$.*

In Definition 3.30, we defined smooth maps to and from embedded submanifolds of linear spaces as those maps which admit a smooth extension to and from the embedding spaces. Now that we understand embedded submanifolds as manifolds, we must verify that our early definition of smooth map agrees with the general notion in Definition 8.5. That is indeed true: Propositions 8.79 and 8.80 below assert as much for the general case of embedded submanifolds of manifolds. To prove them, we introduce a powerful technical result first.

Lemma 8.78. *If $\mathcal{M}$ is an embedded submanifold of $\overline{\mathcal{M}}$, there exists a neighborhood $\overline{\mathcal{U}}$ of $\mathcal{M}$ in $\overline{\mathcal{M}}$ and a smooth map $r\colon \overline{\mathcal{U}} \to \mathcal{M}$ such that $r(x) = x$ for all $x \in \mathcal{M}$.*

Proof sketch. Endow $\overline{\mathcal{M}}$ with a Riemannian metric: this is always doable [Lee18, Prop. 2.4]. Since $\mathcal{M}$ is embedded in $\overline{\mathcal{M}}$, it has a *tubular neighborhood* $\overline{\mathcal{U}}$ [Lee18, Thm. 5.25]. It is straightforward to construct r from the properties of tubular neighborhoods. Note: r is a *(topological) retraction*: this is different from (but related to) our retractions. $\square$

Proposition 8.79. *Let $\mathcal{M}$ be an embedded submanifold of $\overline{\mathcal{M}}$ and let $\mathcal{N}$ be a manifold.*

1. *If $\bar{F}\colon \overline{\mathcal{M}} \to \mathcal{N}$ is smooth (at $x \in \mathcal{M}$), then $F = \bar{F}|_{\mathcal{M}}$ is smooth (at x).*
2. *There exists a neighborhood $\overline{\mathcal{U}}$ of $\mathcal{M}$ in $\overline{\mathcal{M}}$ such that any map $F\colon \mathcal{M} \to \mathcal{N}$ can be extended to a map $\bar{F}\colon \overline{\mathcal{U}} \to \mathcal{N}$ with the property that $\bar{F}$ is smooth if F is smooth, and $\bar{F}$ is smooth at $x \in \mathcal{M}$ if F is smooth at x.*

Proof. The first part holds because the inclusion map $i\colon \mathcal{M} \to \overline{\mathcal{M}}$ is smooth for submanifolds hence $\bar{F}|_{\mathcal{M}} = \bar{F} \circ i$ inherits the smoothness of $\bar{F}$ by composition. For the second part, summon the map $r\colon \overline{\mathcal{U}} \to \mathcal{M}$ provided by Lemma 8.78. Define $\bar{F} = F \circ r$; note that $\bar{F}|_{\mathcal{M}} = F$. $\square$

As a side note, a map defined on any subset of a manifold is said to be smooth if it can be smoothly extended to a neighborhood of its domain. This is compatible with the notion of smooth maps on embedded submanifolds.

Proposition 8.80. *Let $\mathcal{M}$ be an embedded submanifold of $\overline{\mathcal{M}}$ and let $\mathcal{N}$ be a manifold. A map $F\colon \mathcal{N} \to \mathcal{M}$ is smooth (at x) if and only if $\bar{F}\colon \mathcal{N} \to \overline{\mathcal{M}}$, defined by $\bar{F}(y) = F(y)$, is smooth (at x).*

Proof. Smoothness of F implies smoothness of $\bar{F}$ since $\bar{F} = i \circ F$, where $i\colon \mathcal{M} \to \overline{\mathcal{M}}$ is the inclusion map. The other way around, summon the map $r\colon \overline{\mathcal{U}} \to \mathcal{M}$ provided by Lemma 8.78. Through charts, it is easy to confirm that we may restrict the codomain of $\bar{F}$ to $\overline{\mathcal{U}}$ without affecting its smoothness. Since $F = r \circ \bar{F}$, it follows that smoothness of $\bar{F}$ implies that of F. See also [Lee12, Cor. 5.30]. □

As we discovered in Chapters 3 and 5, geometric tools for Riemannian submanifolds of Euclidean spaces are related to their counterparts in that Euclidean space in a straightforward way. This is true more generally for Riemannian submanifolds of manifolds, and the proofs we have considered extend to the general case with little friction. We now summarize these results.

Assume $\overline{\mathcal{M}}$ is a Riemannian manifold and $\mathcal{M}$ is embedded in $\overline{\mathcal{M}}$. We know from eq. (8.26) that $\mathrm{T}_x\mathcal{M}$ is a linear subspace of $\mathrm{T}_x\overline{\mathcal{M}}$. Equip the submanifold $\mathcal{M}$ with a Riemannian metric by restricting the metric $\langle \cdot, \cdot \rangle_x$ of $\overline{\mathcal{M}}$ to the tangent spaces of $\mathcal{M}$. This makes $\mathcal{M}$ a *Riemannian submanifold* of $\overline{\mathcal{M}}$. Assume these structures for the remainder of the section.

Let Proj_x denote the linear map which projects vectors from $\mathrm{T}_x\overline{\mathcal{M}}$ to $\mathrm{T}_x\mathcal{M}$ orthogonally with respect to $\langle \cdot, \cdot \rangle_x$. This object features abundantly in the formulas below.

Consider a smooth function $f\colon \mathcal{M} \to \mathbb{R}$ and any smooth extension $\bar{f}\colon \overline{\mathcal{U}} \to \mathbb{R}$ defined on a neighborhood $\overline{\mathcal{U}}$ of $\mathcal{M}$ in $\overline{\mathcal{M}}$. Then, for all $x \in \mathcal{M}$,

$$\mathrm{grad} f(x) = \mathrm{Proj}_x\big(\mathrm{grad} \bar{f}(x)\big)\,. \tag{8.27}$$

For any two smooth vector fields $U, V \in \mathfrak{X}(\mathcal{M})$ and corresponding smooth extensions $\bar{U}, \bar{V} \in \mathfrak{X}(\overline{\mathcal{U}})$, the Riemannian connection ∇ on $\mathcal{M}$ is related to the Riemannian connection $\bar{\nabla}$ on $\overline{\mathcal{M}}$ through the identity (valid along $\mathcal{M}$):

$$\nabla_U V = \mathrm{Proj}(\bar{\nabla}_{\bar{U}} \bar{V}). \tag{8.28}$$

On a technical note: $\bar{U}, \bar{V}$ are not necessarily defined on all of $\overline{\mathcal{M}}$. We interpret $\bar{\nabla}_{\bar{U}} \bar{V}$ in the usual way, using the fact that $(\bar{\nabla}_{\bar{U}} \bar{V})(x)$ depends on $\bar{U}, \bar{V}$ only locally around x. See also [Lee18, Thm. 8.2]. In pointwise notation, we have for all $u \in \mathrm{T}_x\mathcal{M}$:

$$\nabla_u V = \mathrm{Proj}_x(\bar{\nabla}_u \bar{V}). \tag{8.29}$$

As a result, the Hessian of the function f above is related to the gradient and Hessian of $\bar{f}$ through these relations: let $G(x) = \mathrm{grad} f(x)$ be the gradient vector field of f on $\mathcal{M}$, and let $\bar{G}$ be a smooth extension of G to a neighborhood of $\mathcal{M}$ in $\overline{\mathcal{M}}$. Then, for all $u \in \mathrm{T}_x\mathcal{M} \subseteq \mathrm{T}_x\overline{\mathcal{M}}$,

$$\mathrm{Hess} f(x)[u] = \nabla_u \mathrm{grad} f = \mathrm{Proj}_x\big(\bar{\nabla}_u \bar{G}\big)\,. \tag{8.30}$$

A similarly simple expression is valid for covariant derivatives of vector fields along curves, in analogy to (5.18):

$$\frac{\mathrm{D}}{\mathrm{d}t}Z(t) = \mathrm{Proj}_{c(t)}\left(\frac{\bar{\mathrm{D}}}{\mathrm{d}t}Z(t)\right), \tag{8.31}$$

where c is a smooth curve on $\mathcal{M}$ (hence also on $\overline{\mathcal{M}}$), Z is a smooth vector field on c (which can be understood both in $\mathcal{M}$ and in $\overline{\mathcal{M}}$), $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}$ is the covariant derivative for vector fields on c in $\overline{\mathcal{M}}$, and $\frac{\mathrm{D}}{\mathrm{d}t}$ is the covariant derivative for vector fields on c in $\mathcal{M}$. From this expression we also recover a convenient formula for the acceleration $c'' = \frac{\mathrm{D}}{\mathrm{d}t}c'$ of a curve c on $\mathcal{M}$ in terms of its acceleration $\ddot{c} = \frac{\bar{\mathrm{D}}}{\mathrm{d}t}\dot{c}$ in the embedding space $\overline{\mathcal{M}}$, akin to (5.23):

$$c''(t) = \mathrm{Proj}_{c(t)}(\ddot{c}(t)). \tag{8.32}$$

Moreover, the objects and results presented in Section 5.11 extend to the general case of Riemannian submanifolds of Riemannian manifolds. In particular, the second fundamental form II and the Weingarten map $\mathcal{W}$ are defined in the same way and lead to the same formulas for the Hessian and for the decomposition of $\bar{\nabla}$ and $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}$ in tangent and normal parts.

8.15 Notes and references

Main references for this chapter are the books by Lee [Lee12, Lee18], Brickell and Clark [BC70], O'Neill [O'N83], and Absil et al. [AMS08].

Brickell and Clark define manifolds to be what we call manifolds*. As a result, topological assumptions are always stated explicitly, which is instructive to track their importance in various aspects of the theory. O'Neill defines a manifold to be a Hausdorff topological space equipped with a maximal atlas, without requiring second-countability (though see pp21–22 of that reference). Lee defines *topological* manifolds first—imposing both Hausdorff and second-countability—and defines smooth manifolds as an additional layer of structure on those spaces, requiring the atlas topology to match the existing topology. We use the same definition as Absil et al.: this is compatible with Lee's definitions.

All of these references also lay out basics of topology. The relevance of the topological conditions imposed in Section 8.2 for optimization is spelled out in [AMS08, §3.1.2].

A *closed manifold* is a compact manifold (Definition 8.27) without boundary [Lee12, p27]. As we do not discuss manifolds with boundary, compact and closed manifolds coincide in our treatment, but the latter terminology may be confusing as the manifold itself is always closed with respect to its own topology (by Definition 8.15).

We defined tangent vectors as equivalence classes of curves, which is one of the standard approaches. Another standard definition of tangent vectors, favored notably by Lee and O'Neill, is through the notion of derivation. These definitions are equivalent. A (brief) discussion of the link between these two definitions appears in [Lee12, p72].

Embedded submanifolds are called *regular submanifolds* by Brickell and Clark, and simply *submanifolds* by O'Neill. Furthermore, we mean Riemannian submanifolds to be embedded (as does O'Neill), whereas Lee allows them to be merely immersed, pointing out when it is necessary for them to be embedded [Lee18, p15].

Theorem 8.75 for embedded submanifolds follows [Lee12, Prop. 4.1 and 5.16]. The ensuing characterization of tangent spaces stated in (8.26) matches [Lee12, Prop. 5.38]. Corollary 8.76 for embedded submanifolds defined by a single defining map appears as [Lee12, Cor. 5.14]. The related Proposition 8.77 is a consequence of the constant-rank level set theorem [Lee12, Thm. 5.12].

9 Quotient manifolds

The Grassmannian $\mathrm{Gr}(n,p)$ is the set of linear subspaces of dimension p in $\mathbb{R}^n$. Perhaps the best-known example of an optimization problem over $\mathrm{Gr}(n,p)$ is principal component analysis (PCA). Given k points $y_1, \ldots, y_k \in \mathbb{R}^n$, the goal is to find a linear subspace $L \in \mathrm{Gr}(n,p)$ which fits the data as well as possible, in the following sense:

$$\min_{L \in \mathrm{Gr}(n,p)} \sum_{i=1}^{k} \mathrm{dist}(L, y_i)^2, \tag{9.1}$$

where $\mathrm{dist}(L,y)$ is the Euclidean distance between y and the point in L closest to y. This particular formulation of the problem admits an explicit solution involving the SVD of the data matrix $M = [y_1, \ldots, y_k]$. This is not the case for other cost functions, which may be more accommodating of outliers in the data, or more amenable to the inclusion of priors. For these, we may need more general optimization algorithms to address (9.1). Thus we ask: how can one solve optimization problems over $\mathrm{Gr}(n,p)$?

Any iterative algorithm to minimize a function $f\colon \mathrm{Gr}(n,p) \to \mathbb{R}$ generates a sequence of subspaces $L_0, L_1, L_2, \ldots$ The first point of order is to choose how these subspaces are to be represented in memory. A reasonable idea is to represent $L \in \mathrm{Gr}(n,p)$ with a matrix $X \in \mathbb{R}^{n\times p}$ whose columns form a basis for L. For each L, many matrices X fit this requirement. For numerical reasons, it is often beneficial to use orthonormal bases. Thus, we decide to represent L with a matrix X in $\mathrm{St}(n,p)$, that is, $L = \mathrm{span}(X)$ and $X^\top X = I_p$.

Even working with orthonormal bases to represent subspaces, there are still many possible choices. To be definite, we define an equivalence relation $\sim$ over $\mathrm{St}(n,p)$: two matrices $X, Y \in \mathrm{St}(n,p)$ are deemed equivalent if their columns span the same subspace:

$$X \sim Y \iff \mathrm{span}(X) = \mathrm{span}(Y) \iff X = YQ \text{ for some } Q \in \mathrm{O}(p),$$

where $\mathrm{O}(p)$ is the orthogonal group: the set of orthogonal matrices of size $p \times p$. Formally, this allows us to *identify* subspaces with *equivalence classes*: if $L = \mathrm{span}(X)$, we identify L with

$$[X] = \{Y \in \mathrm{St}(n,p) : Y \sim X\} = \{XQ : Q \in \mathrm{O}(p)\}.$$

This identification establishes a one-to-one correspondence between $\mathrm{Gr}(n,p)$ and the set of equivalence classes, called the *quotient set*:

$$\mathrm{St}(n,p)/\!\sim\; = \{[X] : X \in \mathrm{St}(n,p)\}. \tag{9.2}$$

It is also common to denote this quotient set by $\mathrm{St}(n,p)/\mathrm{O}(p)$, to highlight the special role of the orthogonal group in the equivalence relation: we discuss this more below.

Given $X \in \mathrm{St}(n,p)$ such that $L = \mathrm{span}(X)$, the distance function in (9.1) admits an explicit expression: $XX^\top$ is the matrix which represents orthogonal projection from $\mathbb{R}^n$ to L, so that, in the Euclidean norm $\|\cdot\|$,

$$\mathrm{dist}(L,y)^2 = \|y - XX^\top y\|^2 = \|y\|^2 - \|X^\top y\|^2.$$

Hence, with $A = MM^\top$, $\|\cdot\|$ denoting the Frobenius norm for matrices and

$$\bar{f}(X) = \sum_{i=1}^{k} \|X^\top y_i\|^2 = \|X^\top M\|^2 = \mathrm{Tr}(X^\top A X), \tag{9.3}$$

we may rewrite (9.1) equivalently as

$$\max_{[X]\in \mathrm{St}(n,p)/\sim} f([X]), \qquad \text{with} \qquad f([X]) = \bar{f}(X). \tag{9.4}$$

Crucially, $f\colon \mathrm{St}(n,p)/\!\sim\; \to \mathbb{R}$ is well defined on the quotient set since $\bar{f}(X) = \bar{f}(Y)$ whenever $X \sim Y$: we say $\bar{f}$ is *invariant under* $\sim$.

On the one hand, problem (9.4) is closely related to

$$\max_{X\in \mathrm{St}(n,p)} \bar{f}(X), \tag{9.5}$$

which we know how to handle using our optimization tools for embedded submanifolds, generating a sequence of matrices $X_0, X_1, \ldots$ in $\mathrm{St}(n,p)$.

On the other hand, a practical implementation of a (yet to be determined) optimization algorithm on $\mathrm{St}(n,p)/\!\sim$, which generates a sequence of equivalence classes $[X_0], [X_1], \ldots$, would also actually generate matrices $X_0, X_1, \ldots$ in Stiefel to represent these equivalence classes. One wonders then: in practical terms, what distinguishes an algorithm on $\mathrm{St}(n,p)/\!\sim$ from one on $\mathrm{St}(n,p)$?

The key consideration is *preservation of invariance*. To illustrate this notion, let us consider how gradient descent proceeds to minimize $\bar{f}$ on $\mathrm{St}(n,p)$ as a Riemannian submanifold of $\mathbb{R}^{n\times p}$ with the usual Euclidean metric. Using the projector to the tangent spaces of Stiefel, $\mathrm{Proj}_X^{\mathrm{St}}$ (7.27), the gradient is given by

$$\begin{aligned}\frac{1}{2}\mathrm{grad}\bar{f}(X) &= \mathrm{Proj}_X^{\mathrm{St}}(AX) \qquad\qquad (9.6)\\ &= (I_n - XX^\top)AX + X\frac{X^\top AX - X^\top AX}{2} = (I_n - XX^\top)AX.\end{aligned}$$

(Notice how the second term vanishes: we will see that this is not by accident.) Assuming constant step-size α for simplicity, Riemannian gradient descent iterates

$$X_{k+1} = G(X_k) \triangleq \mathrm{R}_{X_k}(-\alpha \mathrm{grad}\bar{f}(X_k)).$$

When is it legitimate to think of this sequence of iterates as corresponding to a sequence on the quotient set? *Exactly when the equivalence class of* X_{k+1} *depends only on the equivalence class of* X_k*, and* not *on* X_k *itself.* Indeed, only then can we claim that the algorithm iterates from $[X_k]$ to $[X_{k+1}]$.

To assess the latter, we must determine how $[X_{k+1}]$ changes if X_k is replaced by another representative of the same equivalence class, that is, if X_k is replaced by X_kQ for some orthogonal Q. A first observation is that

$$\forall X \in \mathrm{St}(n,p), Q \in \mathrm{O}(p), \qquad \mathrm{grad}\bar{f}(XQ) = \mathrm{grad}\bar{f}(X) \cdot Q.$$

Hence, if the retraction has the property[1] that

$$\forall (X,V) \in \mathrm{TSt}(n,p), Q \in \mathrm{O}(p), \qquad [\mathrm{R}_{XQ}(VQ)] = [\mathrm{R}_X(V)], \tag{9.7}$$

then it follows that, for all $Q \in \mathrm{O}(p)$,

$$[G(XQ)] = [\mathrm{R}_{XQ}(-\alpha \mathrm{grad}\bar{f}(XQ))] = [\mathrm{R}_X(-\alpha \mathrm{grad}\bar{f}(X))] = [G(X)].$$

Thus, under that condition, $[X_{k+1}]$ is indeed a function of $[X_k]$:

$$X \sim Y \implies G(X) \sim G(Y). \tag{9.8}$$

We already know retractions which satisfy property (9.7). For example, the polar retraction (7.24) can be written as

$$\mathrm{R}_X^{\mathrm{pol}}(V) = (X+V)(I_p + V^\top V)^{-1/2},$$

so that

$$\mathrm{R}_{XQ}^{\mathrm{pol}}(VQ) = (X+V)Q \cdot \left(Q^\top [I_p + V^\top V] Q\right)^{-1/2} = \mathrm{R}_X^{\mathrm{pol}}(V) \cdot Q. \tag{9.9}$$

Also, the QR retraction (7.22) is such that $\mathrm{R}_X^{\mathrm{QR}}(V)$ is a matrix whose columns form an orthonormal basis for $\mathrm{span}(X+V)$. As a result, $\mathrm{R}_{XQ}^{\mathrm{QR}}(VQ)$ is a matrix whose columns form a basis for $\mathrm{span}((X+V)Q)$, which of course is the same subspace (it does not, however, satisfy the stronger property that $\mathrm{R}_{XQ}(VQ) = \mathrm{R}_X(V) \cdot Q$ as the polar one did).

These considerations allow us to conclude that Riemannian gradient descent for $\bar{f}$ on $\mathrm{St}(n,p)$ with either of these retractions induces a well-defined sequence on the quotient set $\mathrm{St}(n,p)/\sim$, defined by the map

$$[X_{k+1}] = F([X_k]) \triangleq [G(X_k)].$$

At this point, a few questions come naturally:

1. Is the sequence defined by $[X_{k+1}] = F([X_k])$ itself a "gradient descent" sequence of sorts for the optimization problem (9.4) on the quotient set?
2. Can we devise more sophisticated algorithms such as the trust-regions method to operate on the quotient set?
3. Are other quotient sets similarly amenable to optimization?

[1] This property makes sense because if V is tangent to $\mathrm{St}(n,p)$ at X then VQ is tangent to $\mathrm{St}(n,p)$ at XQ.

We answer all questions in the affirmative. The crux of this chapter is to argue that quotient sets such as $\mathrm{St}(n,p)/\sim$ are themselves Riemannian manifolds in a natural way, called *Riemannian quotient manifolds*. This identification of $\mathrm{Gr}(n,p)$ with $\mathrm{St}(n,p)/\sim$ gives meaning to the claim that $\mathrm{Gr}(n,p)$ is a quotient manifold; it is called the *Grassmann manifold*. All the tools and algorithms we have developed for optimization on general manifolds apply in particular to quotient manifolds. The iterative method described above turns out to be a bona fide Riemannian gradient descent method in that geometry, and with more work we can similarly describe second-order optimization algorithms.

Parts of this chapter focus on a particular class of Riemannian quotient manifolds obtained through *group actions* on manifolds, as is the case for $\mathrm{Gr}(n,p)$ constructed here. Particular attention is given to the practical representation of points and tangent vectors for quotient manifolds, and to the computation of objects such as gradients and Hessians.

What do we stand to gain from the quotient approach? First, it should be clear that nothing is lost: Riemannian quotient manifolds are Riemannian manifolds, hence all algorithms and accompanying theory apply. Second, optimization on the quotient achieves a natural goal: if the cost function of an optimization problem is insensitive to certain transformations, then it is reasonable to require an algorithm for that problem to be similarly unfazed.

Sometimes, this property leads to computational advantages. Even when it does not, the quotient perspective can yield better theoretical understanding. Specifically, consider the local convergence rates we discussed for gradient descent (Theorem 4.20), Newton's method (Theorem 6.7) and trust regions (Theorem 6.30): for all of these, the fast convergence guarantees hold provided the algorithm converges to a critical point where the Hessian of the cost function is positive definite. It is easy to come up with counter-examples showing that the condition is necessary in general. For example, with $f(x) = x^4$ on the real line, gradient descent with (appropriate) constant step-size converges sublinearly, and Newton's method converges only linearly to zero.

As we show in Lemma 9.41, if the cost function on the *total space* (the set before we pass to the quotient) is invariant under the quotient, then its Hessian cannot possibly be positive definite at critical points. This is because the cost function is constant along the equivalence classes: directions tangent to these equivalence classes are necessarily in the kernel of the Hessian. Thus, the standard fast convergence results do not ever apply on the total space.

Yet, it often happens that we do see fast convergence on the total space empirically. This is the case notably for problem (9.5) above, on the Stiefel manifold. Why is that?

As transpires from the discussion above and as we detail further in this chapter, the reason is that, under certain circumstances, optimization algorithms on the total space can be interpreted as matching algorithms on the quotient manifold. Moreover, the spurious directions tangent to equivalence classes are quotiented out in their own way, so that they do not appear in the kernel of the Hessian on the quotient manifold: that Hessian can be positive definite. In that scenario, the quotient approach does not confer a computational advantage over the total space approach (the two are algorithmically equivalent or close), but it does provide the stronger theoretical perspective, aptly explaining why we do get fast local convergence. As for second-order methods, the quotient perspective can deliver genuinely new and crisply motivated algorithms.

Throughout the chapter, we use the Grassmann manifold as a running example. For convenience, we collect our findings about it in Section 9.16. In that section, we also show that the same geometry for $\mathrm{Gr}(n,p)$ can be realized as a Riemannian submanifold of a Euclidean space, so that we could have also discussed this manifold in Chapter 7. The hope is that by then it will be clear that the quotient perspective carries many conceptual (and aesthetic) advantages, and that it provides a firm grasp of symmetries beyond the Grassmann manifold.

9.1 A definition and a few facts

Let $\sim$ be an equivalence relation on a manifold $\overline{\mathcal{M}}$ with equivalence classes

$$[x] = \{y \in \overline{\mathcal{M}} : x \sim y\},$$

and let

$$\mathcal{M} = \overline{\mathcal{M}}/\sim \; = \{[x] : x \in \overline{\mathcal{M}}\} \tag{9.10}$$

be the resulting *quotient set*. The *canonical projection* or *natural projection* links the *total space* $\overline{\mathcal{M}}$ to its quotient $\mathcal{M}$:

$$\pi \colon \overline{\mathcal{M}} \to \mathcal{M} \colon x \mapsto \pi(x) = [x]. \tag{9.11}$$

The quotient set $\mathcal{M}$ inherits a topology from $\overline{\mathcal{M}}$ called the *quotient topology*, turning $\mathcal{M}$ into a *quotient space*. This topology is defined as follows:

$$\mathcal{U} \subseteq \mathcal{M} \text{ is open} \iff \pi^{-1}(\mathcal{U}) \text{ is open in } \overline{\mathcal{M}}.$$

This notably ensures that π, then called the *quotient map*, is continuous.

Say we equip the quotient space $\mathcal{M}$ with a smooth structure as in Chapter 8 (assuming this is possible). Then, it makes sense to ask whether π is smooth and, accordingly, whether its differential at some point has full rank. These considerations enter into the definition of *quotient manifold*.

Definition 9.1. *The quotient set $\mathcal{M} = \overline{\mathcal{M}}/\sim$ equipped with a smooth structure is a* quotient manifold *of $\overline{\mathcal{M}}$ if the projection π* (9.11) *is smooth and its differential $\mathrm{D}\pi(x) \colon \mathrm{T}_x\overline{\mathcal{M}} \to \mathrm{T}_{[x]}\mathcal{M}$ has rank $\dim \mathcal{M}$ for all $x \in \overline{\mathcal{M}}$.*

As an exercise, one can show that the projective space RP^{n-1} with smooth structure as in Example 8.11 is a quotient manifold of $\mathbb{R}^n\backslash\{0\}$ with the equivalence relation that deems two points to be equivalent if they belong to the same line through the origin. However, this way of identifying quotient manifolds is impractical, as it requires first to know that the quotient space is a manifold with a certain atlas, then to check explicitly that π has the required properties using that particular smooth structure. In this chapter, we discuss more convenient tools.

By construction, π is continuous with respect to the quotient space topology. With a quotient manifold structure on $\mathcal{M}$, π is smooth, hence a fortiori continuous with respect to the atlas topology. In fact, the atlas topology coincides with the quotient topology in that case. We have the following remarkable result [Lee12, Thm. 4.31].

Theorem 9.2. *A quotient space $\mathcal{M} = \overline{\mathcal{M}}/\sim$ admits at most one smooth structure that makes it a quotient manifold of $\overline{\mathcal{M}}$. When this is the case, the atlas topology of $\mathcal{M}$ is the quotient topology.*

This statement should be compared to Theorem 8.74 for embedded submanifolds: a subset of a manifold admits at most one smooth structure that makes it an embedded submanifold. Thus, just as it made sense to say that a subset of a manifold is or is not an embedded submanifold, so it makes sense to say that a quotient space of a manifold is or is not a quotient manifold.

A direct consequence of Definition 9.1 and Corollary 8.76 is that equivalence classes are embedded submanifolds of the total space. As we discuss this, it appears that we must sometimes distinguish between $[x]$ as a point of $\mathcal{M}$ and $[x]$ as a subset of $\overline{\mathcal{M}}$. When in need, we adopt this convention: $[x] = \pi(x)$ is a point of $\mathcal{M}$, whereas $[x] = \pi^{-1}(\pi(x))$ is a subset of $\overline{\mathcal{M}}$.

Proposition 9.3. *Let $\mathcal{M} = \overline{\mathcal{M}}/\sim$ be a quotient manifold. For any $x \in \overline{\mathcal{M}}$, the equivalence class $\mathcal{F} = \pi^{-1}(\pi(x))$, also called a* fiber*, is closed in $\overline{\mathcal{M}}$ and it is an embedded submanifold of $\overline{\mathcal{M}}$. Its tangent spaces are given by*

$$\mathrm{T}_y\mathcal{F} = \ker \mathrm{D}\pi(y) \subseteq \mathrm{T}_y\overline{\mathcal{M}}. \tag{9.12}$$

In particular, $\dim \mathcal{F} = \dim \overline{\mathcal{M}} - \dim \mathcal{M}$.

Proof. Apply Corollary 8.76 with $\pi \colon \overline{\mathcal{M}} \to \mathcal{M}$ as the defining map, and $\mathcal{F}$ as the level set $\{y \in \overline{\mathcal{M}} : \pi(y) = [x]\}$. □

Thus, when an equivalence relation yields a quotient manifold, that equivalence relation partitions the total space into closed, embedded submanifolds called fibers. In particular, notice that all fibers have the same dimension. This sometimes allows one to determine quickly that a given quotient space cannot possibly be a quotient manifold—see Exercise 9.8. In the following example, we illustrate Proposition 9.3 through the spaces that featured in the introduction.

Example 9.4. *Consider the set $\mathcal{M} = \mathrm{St}(n, p)/\sim$ as in (9.2). We have not yet argued that this is a quotient manifold: we do so in the next section. For now, let us assume that $\mathcal{M}$ indeed is a quotient manifold. Then, given a point*

$X \in \mathrm{St}(n,p)$, Proposition 9.3 tells us that the fiber

$$\mathcal{F} = \{Y \in \mathrm{St}(n,p) : X \sim Y\} = \{XQ : Q \in \mathrm{O}(p)\}$$

is an embedded submanifold of $\mathrm{St}(n,p)$. (We could also show this directly.)

The tangent space to $\mathcal{F}$ at X is a subspace of $\mathrm{T}_X\mathrm{St}(n,p)$, corresponding to the kernel of the differential of π at X (9.12). As $\mathrm{D}\pi(x)$ is an abstract object, it is often more convenient to approach $\mathrm{T}_X\mathcal{F}$ as follows: all tangent vectors in $\mathrm{T}_X\mathcal{F}$ are of the form $\bar{c}'(0)$ for some smooth curve $\bar{c} \colon I \to \mathcal{F}$ with $\bar{c}(0) = X$. Moreover, any such curve is necessarily of the form $\bar{c}(t) = XQ(t)$ with $Q \colon I \to \mathrm{O}(p)$ a smooth curve on the manifold $\mathrm{O}(p)$ with $Q(0) = I_p$. Thus, all tangent vectors in $\mathrm{T}_X\mathcal{F}$ are of the form $XQ'(0)$. Now we recall that the tangent space to $\mathrm{O}(p)$ at $Q(0) = I_p$ is the set of skew-symmetric matrices of size p (7.32) to conclude that

$$\mathrm{T}_X\mathcal{F} = \{X\Omega : \Omega + \Omega^\top = 0\} \subset \mathrm{T}_X\mathrm{St}(n,p).$$

We can connect this to π: by design, $c(t) \triangleq \pi(\bar{c}(t)) = [X]$ is a constant curve on $\mathrm{St}(n,p)/\sim$. Since we are assuming $\mathcal{M}$ is a quotient manifold, π is smooth too. This allows us to use the chain rule, writing $Q'(0) = \Omega$:

$$0 = c'(0) = \mathrm{D}\pi(\bar{c}(0))[\bar{c}'(0)] = \mathrm{D}\pi(X)[X\Omega].$$

This confirms that any matrix of the form $X\Omega$ is in the kernel of $\mathrm{D}\pi(X)$.

Theorem 9.2 tells us that a quotient space may be a quotient manifold in at most one way. When it is, we sometimes want to have access to charts of the resulting smooth structure on the quotient manifold. The next result provides such charts. It constitutes one part of the rank theorem in differential geometry [Lee12, Thm. 4.12].

Proposition 9.5. *Let $\mathcal{M} = \overline{\mathcal{M}}/\sim$ be a quotient manifold with canonical projection π and $\dim \overline{\mathcal{M}} = n + k$, $\dim \mathcal{M} = n$. For all $x \in \overline{\mathcal{M}}$, there exists a chart $(\overline{\mathcal{U}}, \bar{\varphi})$ of $\overline{\mathcal{M}}$ around x and a chart $(\mathcal{U}, \varphi)$ of $\mathcal{M}$ around $\pi(x) = [x]$ such that $\pi(\overline{\mathcal{U}}) \subseteq \mathcal{U}$ and the coordinate representation of π,*[2]

$$\tilde{\pi} = \varphi \circ \pi \circ \bar{\varphi}^{-1} \colon \bar{\varphi}(\overline{\mathcal{U}}) \subseteq \mathbb{R}^{n+k} \to \varphi(\mathcal{U}) \subseteq \mathbb{R}^n, \tag{9.13}$$

is simply the function $\tilde{\pi}(z_1, \ldots, z_{n+k}) = (z_1, \ldots, z_n)$.

It is an exercise to check that π is an *open map*, that is: it maps open sets of $\overline{\mathcal{M}}$ to open sets of $\mathcal{M}$ [Lee12, Prop. 4.28]. We may thus replace $\mathcal{U}$ with $\pi(\overline{\mathcal{U}})$ in Proposition 9.5 when convenient. By Definition 9.1, we also know that π is surjective, and that its differentials $\mathrm{D}\pi(x)$ are surjective as well. (The latter makes it a *submersion*.)

Given our focus on optimization, when facing quotient manifolds, we are naturally led to consider pairs of problems: one which consists in minimizing $f \colon \mathcal{M} \to \mathbb{R}$, and a companion problem which consists in minimizing $\bar{f} = f \circ \pi \colon \overline{\mathcal{M}} \to \mathbb{R}$.

[2] Note that $\overline{\mathcal{U}}$ is not guaranteed to contain whole fibers, that is, $\pi^{-1}(\pi(\overline{\mathcal{U}}))$ may not be included in $\overline{\mathcal{U}}$.

The properties of π provide strong links between the salient points of both problems. The following proposition states those links beyond the context of quotients. Notice that it does not require any Riemannian structures. The proof is deferred to Section 9.17.

Proposition 9.6. *Let $\overline{\mathcal{M}}$ and $\mathcal{M}$ be two manifolds with a map $\pi \colon \overline{\mathcal{M}} \to \mathcal{M}$. Consider a function $f \colon \mathcal{M} \to \mathbb{R}$ and its lift $\bar{f} = f \circ \pi \colon \overline{\mathcal{M}} \to \mathbb{R}$. The two optimization problems $\min_{y \in \mathcal{M}} f(y)$ and $\min_{x \in \overline{\mathcal{M}}} \bar{f}(x)$ are related as follows:*

1. *If π is surjective, then x is a global minimizer of $\bar{f}$ if and only if $\pi(x)$ is a global minimizer of f.*
2. *If π is continuous and open, then x is a local minimizer of $\bar{f}$ if and only if $\pi(x)$ is a local minimizer of f.*
3. *If π is smooth and all differentials of π are surjective, then:*
   - *(a) x is a first-order critical point of $\bar{f}$ if and only if $\pi(x)$ is a first-order critical point of f, and*
   - *(b) x is a second-order critical point of $\bar{f}$ if and only if $\pi(x)$ is a second-order critical point of f.*

If $\mathcal{M}$ is a quotient manifold of $\overline{\mathcal{M}}$ with projection π, all of the above hold.

Exercise 9.7. *Show that the projective space RP^{n-1} with the smooth structure of Example 8.11 is a quotient manifold of $\mathbb{R}^n \backslash \{0\}$ with the equivalence relation $x \sim y \iff x = \alpha y$ for some $\alpha \in \mathbb{R}$.*

Exercise 9.8. *Consider the following equivalence relation over $\overline{\mathcal{M}} = \mathbb{R}^{n \times p}$, with $1 \leq p < n$: $X \sim Y$ if and only if $Y = XQ$ for some $Q \in \mathrm{O}(p)$. Argue that $\overline{\mathcal{M}}/\sim$ is* not *a quotient manifold. (Contrast with the introduction of this chapter, where $\overline{\mathcal{M}} = \mathrm{St}(n, p)$.)*

Exercise 9.9. *Let $\mathcal{M}$ be a quotient manifold of $\overline{\mathcal{M}}$ with canonical projection π. Show that π is an open map, that is, if $\overline{\mathcal{U}}$ is open in $\overline{\mathcal{M}}$, then $\pi(\overline{\mathcal{U}})$ is open in $\mathcal{M}$.*

9.2 Quotient manifolds through group actions

There exists an explicit characterization of which equivalence relations on a manifold $\overline{\mathcal{M}}$ yield quotient manifolds (see Section 9.17). Using this characterization, however, is not straightforward. Fortunately, there exists a special class of suitable equivalence relations defined through *group actions* on manifolds that are both simple to identify and important in practice. This covers our approach to the Grassmann manifold in the introduction of this chapter (where the group is $\mathrm{O}(p)$) and other examples we discuss below (see Exercise 9.20).

We start with a few definitions regarding groups, Lie groups and group actions. A set G equipped with an operation $\cdot \colon G \times G \to G$ is a *group* if:

1. The operation is associative: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in G$;
2. There exists a unique element $e \in G$ (called the identity) such that $e \cdot g = g \cdot e = g$ for all $g \in G$; and
3. For each $g \in G$ there is an element $g^{-1} \in G$ (called the inverse of g) such that $g \cdot g^{-1} = g^{-1} \cdot g = e$.

If the set G is further equipped with a smooth structure—making it a manifold $\mathcal{G}$—and the group operation plays nicely with the smooth structure, then we call $\mathcal{G}$ a *Lie group*.

Definition 9.10. *Let $\mathcal{G}$ be both a manifold and a group with operation $\cdot$. If the product map*

$$\mathrm{prod}\colon \mathcal{G} \times \mathcal{G} \to \mathcal{G}\colon (g, h) \mapsto \mathrm{prod}(g, h) = g \cdot h$$

and the inverse map

$$\mathrm{inv}\colon \mathcal{G} \to \mathcal{G}\colon g \mapsto \mathrm{inv}(g) = g^{-1}$$

are smooth, then $\mathcal{G}$ is a Lie group. *Smoothness of* prod *is understood with respect to the product manifold structure on $\mathcal{G} \times \mathcal{G}$ (see Exercise 8.31).*

Example 9.11. *Some examples of Lie groups include* $\mathrm{O}(n)$ *(the orthogonal group),* $\mathrm{SO}(n)$ *(the rotation group) and* $\mathrm{GL}(n)$ *(the general linear group, which is the set of invertible matrices of size $n \times n$), with group operation given by the matrix product, and smooth structure as embedded submanifolds of $\mathbb{R}^{n \times n}$. Their identity is the identity matrix I_n. Another example is the group of translations, $\mathbb{R}^n$, whose group operation is vector addition. Its identity is the zero vector. Yet another common example is the* special Euclidean group, $\mathrm{SE}(n)$, *whose elements are of the form $(R, t) \in \mathrm{SO}(n) \times \mathbb{R}^n$, with group operation $(R, t) \cdot (R', t') = (RR', Rt' + t)$. The identity element is $(I_n, 0)$. Equivalently, we may represent (R, t) as the matrix $\left[\begin{smallmatrix} R & t \\ 0 & 1 \end{smallmatrix}\right]$, in which case the group operation is the matrix product.*

Elements of a group can sometimes be used to transform points of a manifold. For example, $X \in \mathrm{St}(n, p)$ can be transformed into another element of $\mathrm{St}(n, p)$ by right-multiplication with an orthogonal matrix $Q \in \mathrm{O}(p)$. Under some conditions, these transformations are called *group actions*.

Definition 9.12. *Given a Lie group $\mathcal{G}$ and a manifold $\overline{\mathcal{M}}$, a* left group action *is a map $\theta\colon \mathcal{G} \times \overline{\mathcal{M}} \to \overline{\mathcal{M}}$ such that:*

1. *For all $x \in \overline{\mathcal{M}}$, $\theta(e, x) = x$ (identity), and*
2. *For all $g, h \in \mathcal{G}$ and $x \in \overline{\mathcal{M}}$, $\theta(g \cdot h, x) = \theta(g, \theta(h, x))$ (compatibility).*

As a consequence, for all $g \in \mathcal{G}$, the map $x \mapsto \theta(g, x)$ is invertible on $\overline{\mathcal{M}}$, with inverse $x \mapsto \theta(g^{-1}, x)$. The group action is smooth *if θ is smooth as a map on the product manifold $\mathcal{G} \times \overline{\mathcal{M}}$ to the manifold $\overline{\mathcal{M}}$. We then say the group $\mathcal{G}$ acts smoothly on $\overline{\mathcal{M}}$.*

Similarly, a right group action *is a map* $\theta \colon \overline{\mathcal{M}} \times \mathcal{G} \to \overline{\mathcal{M}}$ *such that* $\theta(x, e) = x$ *and* $\theta(x, g \cdot h) = \theta(\theta(x, g), h)$*, for all* $g, h \in \mathcal{G}$ *and* $x \in \overline{\mathcal{M}}$*, and this action is smooth if* θ *is smooth as a map between manifolds.*

A group action induces an equivalence relation, as follows.

Definition 9.13. *The* orbit *of* $x \in \overline{\mathcal{M}}$ *through the left action* θ *of* $\mathcal{G}$ *is the set* $\mathcal{G}x \triangleq \{\theta(g, x) : g \in \mathcal{G}\}$*. This induces an equivalence relation* $\sim$ *on* $\overline{\mathcal{M}}$*:*

$$x \sim y \iff y = \theta(g, x) \text{ for some } g \in \mathcal{G},$$

that is, two points of $\overline{\mathcal{M}}$ *are equivalent if they belong to the same orbit. As such, orbits and equivalence classes coincide. We denote the quotient space* $\overline{\mathcal{M}}/\sim$ *as* $\overline{\mathcal{M}}/\mathcal{G}$ *(also called the* orbit space*), where the specific group action is indicated by context. The definition is similar for right actions.*

Some authors write $\mathcal{G}\backslash\overline{\mathcal{M}}$ or $\overline{\mathcal{M}}/\mathcal{G}$ to distinguish between left and right action quotients. We always write $\overline{\mathcal{M}}/\mathcal{G}$.

Example 9.14. *The map* $\theta(X, Q) = XQ$ *defined on* $\mathrm{St}(n, p) \times \mathrm{O}(p)$ *is a smooth, right group action. Its orbits are the equivalence classes we have considered thus far, namely,* $[X] = \{XQ : Q \in \mathrm{O}(p)\}$*. Thus,* $\mathrm{St}(n, p)/\mathrm{O}(p)$ *is one-to-one with the Grassmann manifold* $\mathrm{Gr}(n, p)$*.*

We have already discussed that not all equivalence relations on manifolds lead to quotient manifolds. Unfortunately, neither do all smooth group actions: further properties are required. Specifically, it is sufficient for the actions also to be *free* and *proper*.

Definition 9.15. *A group action* θ *is* free *if, for all* x*, acting on* x *with any group element which is not the identity results in a point different from* x*. For instance, a left action is free if, for all* $x \in \overline{\mathcal{M}}$*,* $\theta(g, x) = x \implies g = e$*.*

If the action is not free, then different orbits could have different dimensions, which is impossible for a quotient manifold.

For the following definition, recall Definition 8.26 for compact sets.

Definition 9.16. *A left group action* θ *is* proper *if*

$$\vartheta \colon \mathcal{G} \times \overline{\mathcal{M}} \to \overline{\mathcal{M}} \times \overline{\mathcal{M}} \colon (g, x) \mapsto \vartheta(g, x) = (\theta(g, x), x)$$

is a proper map*, that is, all compact subsets of* $\overline{\mathcal{M}} \times \overline{\mathcal{M}}$ *map to compact subsets of* $\mathcal{G} \times \overline{\mathcal{M}}$ *through* ϑ^{-1}*. The definition is similar for right actions.*

The reason we require the action to be proper is topological: if the action is smooth and proper, then the quotient topology is Hausdorff [Lee12, Prop. 21.4]. On the other hand, there exist smooth, free actions that are not proper and for which the quotient topology ends up not being Hausdorff [Lee12, Ex. 21.3, Pb. 21-5].

Checking whether an action is free is often straightforward. Checking for properness, on the other hand, can be more delicate. Fortunately, if the group $\mathcal{G}$ is compact (which is the case for $\mathrm{SO}(n)$ and $\mathrm{O}(n)$), then every smooth action is proper [Lee12, Cor. 21.6].

Proposition 9.17. *Every smooth action by a compact Lie group is proper.*

If the group is not compact, see Section 9.17 for some pointers, specifically Proposition 9.60.

We can now state the main theorem of this section. This is our tool of choice to identify quotient manifolds [Lee12, Thm. 21.10].

Theorem 9.18. *If the Lie group $\mathcal{G}$ acts smoothly, freely and properly on the smooth manifold $\overline{\mathcal{M}}$, then the quotient space $\overline{\mathcal{M}}/\mathcal{G}$ is a quotient manifold of dimension $\dim \overline{\mathcal{M}} - \dim \mathcal{G}$; orbits (that is, fibers) have dimension $\dim \mathcal{G}$.*

Example 9.19. *Continuing our running example, we now check that the Grassmann manifold, seen as the quotient space $\mathrm{St}(n,p)/\mathrm{O}(p)$, is indeed a quotient manifold. We already checked that the action $\theta(X,Q) = XQ$ is smooth. By Proposition 9.17, it is proper since $\mathrm{O}(p)$ is compact. It is also free since $XQ = X$ implies $Q = I_p$ (by left-multiplying with $X^\top$). Thus, Theorem 9.18 implies $\mathrm{Gr}(n,p)$, identified with $\mathrm{St}(n,p)/\mathrm{O}(p)$, is a quotient manifold. More explicitly: the theorem tells us there exists a unique smooth structure which turns $\mathrm{Gr}(n,p)$ into a manifold such that*

$$\pi \colon \mathrm{St}(n,p) \to \mathrm{Gr}(n,p) \colon X \mapsto \pi(X) \triangleq [X] = \{XQ : Q \in \mathrm{O}(p)\}$$

has the properties laid out in Definition 9.1. Additionally, we know that

$$\dim \mathrm{Gr}(n,p) = \dim \mathrm{St}(n,p) - \dim \mathrm{O}(p) = p(n-p).$$

By Proposition 9.3, the fibers are closed, embedded submanifolds of $\mathrm{St}(n,p)$ with dimension $\frac{p(p-1)}{2}$: this is compatible with our work in Example 9.4 where we showed the tangent space to a fiber at X is $\{X\Omega : \Omega \in \mathrm{Skew}(p)\}$.

In contrast, one can check that the group action underlying Exercise 9.8 is smooth and proper, but it is not free: not all orbits have the same dimension, hence the quotient space is not a quotient manifold.

Exercise 9.20. *In each item below, a Lie group $\mathcal{G}$ (recall Example 9.11) acts on a manifold $\overline{\mathcal{M}}$ through some action θ (the first one is a right action, the other two are left actions). Check that these are indeed group actions and that the quotient spaces $\overline{\mathcal{M}}/\mathcal{G}$ are quotient manifolds. (Recall that $\mathbb{R}_k^{m\times n}$ is the set of matrices of size $m \times n$ and rank k.)*

1. $\overline{\mathcal{M}} = \mathrm{SO}(n)^k$, $\mathcal{G} = \mathrm{SO}(n)$, $\theta((R_1, \ldots, R_k), Q) = (R_1 Q, \ldots, R_k Q)$.
2. $\overline{\mathcal{M}} = \mathbb{R}_r^{m\times r} \times \mathbb{R}_r^{n\times r}$, $\mathcal{G} = \mathrm{GL}(r)$, $\theta(J, (L,R)) = (LJ^{-1}, RJ^\top)$.
3. $\overline{\mathcal{M}} = \mathbb{R}_d^{d\times n}$, $\mathcal{G} = \mathrm{SE}(d)$, $\theta((R,t), X) = RX + t\mathbf{1}^\top$.

Describe the equivalence classes and the significance of the quotient manifolds.

$$\begin{tikzcd} \overline{\mathcal{M}} \arrow[d, "\pi"'] \arrow[dr, "\bar{F} = F \circ \pi"] & \\ \mathcal{M} \arrow[r, "F"'] & \mathcal{N} \end{tikzcd} \qquad \begin{tikzcd} \mathcal{N} \arrow[r, "\bar{F}"] \arrow[dr, "F"'] & \overline{\mathcal{M}} \arrow[d, "\pi"] \\ & \mathcal{M} \end{tikzcd}$$

Figure 9.1 Commutative diagrams for Theorem 9.21 (left) and Proposition 9.23 (right) about lifting a map F on or to a quotient manifold.

9.3 Smooth maps to and from quotient manifolds

Smooth maps on a quotient manifold $\overline{\mathcal{M}}/\sim$ can be understood entirely through smooth maps on the corresponding total space $\overline{\mathcal{M}}$.

Theorem 9.21. *Given a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with canonical projection π and any manifold $\mathcal{N}$, a map $F : \mathcal{M} \to \mathcal{N}$ is smooth if and only if $\bar{F} = F \circ \pi : \overline{\mathcal{M}} \to \mathcal{N}$ is smooth.*

One direction is clear: if F is smooth on the quotient manifold, then $\bar{F} = F \circ \pi$ is smooth on the total space by composition: we call $\bar{F}$ the *lift* of F. Consider the other direction: if $\bar{F}$ on $\overline{\mathcal{M}}$ is invariant under $\sim$, then it is of the form $\bar{F} = F \circ \pi$ for some map F on the quotient and we say $\bar{F}$ *descends* to the quotient. To argue that F is smooth if $\bar{F}$ is smooth we introduce the notion of *local section*: a map S which smoothly selects a representative of each equivalence class in some neighborhood. One can establish their existence using the special charts afforded by Proposition 9.5—see [Lee12, Thm. 4.26].

Proposition 9.22. *For any $x \in \overline{\mathcal{M}}$ there exists a neighborhood $\mathcal{U}$ of $[x]$ on the quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ and a smooth map $S : \mathcal{U} \to \overline{\mathcal{M}}$ (called a* local section*) such that $\pi \circ S$ is the identity map on $\mathcal{U}$ and $S([x]) = x$.*

Proof of Theorem 9.21. If F is smooth, then $\bar{F}$ is smooth by composition. The other way around, if $\bar{F}$ is smooth, let us show that F is smooth at an arbitrary $[x]$. Use Proposition 9.22 to pick a local section S defined on a neighborhood $\mathcal{U}$ of $[x]$. Since $\bar{F} = F \circ \pi$, we find that $F|_{\mathcal{U}} = \bar{F} \circ S$: this is smooth by composition. Thus, F is smooth in some neighborhood around any point $[x]$, that is, F is smooth. $\square$

We note another result, this one about maps *into* quotient manifolds.

Proposition 9.23. *Let $\bar{F} : \mathcal{N} \to \overline{\mathcal{M}}$ be a map from one manifold into another, and let $\mathcal{M} = \overline{\mathcal{M}}/\sim$ be a quotient manifold of $\overline{\mathcal{M}}$ with projection π. If $\bar{F}$ is smooth, then $F = \pi \circ \bar{F} : \mathcal{N} \to \mathcal{M}$ is smooth. The other way around, if $F : \mathcal{N} \to \mathcal{M}$ is smooth, then for all $[x] \in \mathcal{M}$ there exists a neighborhood $\mathcal{U}$ of $[x]$ such that $F|_{F^{-1}(\mathcal{U})} = \pi \circ \bar{F}$ for some smooth map $\bar{F} : F^{-1}(\mathcal{U}) \to \overline{\mathcal{M}}$.*

Proof. The first part is through composition of smooth maps. For the second part, consider $\bar{F} = S \circ F$ with a local section S defined on $\mathcal{U}$, as provided by Proposition 9.22: the domain $F^{-1}(\mathcal{U})$ is open in $\mathcal{N}$ since F is continuous, $\bar{F}$ is smooth by composition, and it is indeed the case that $\pi \circ \bar{F}$ is equal to F on $F^{-1}(\mathcal{U})$ since $\pi \circ S$ is the identity on $\mathcal{U}$. □

The first part states that a smooth map into the total space yields a smooth map into the quotient manifold after composition with π. In particular, if $\bar{c}\colon I \to \overline{\mathcal{M}}$ is a smooth curve on the total space, then $c = \pi \circ \bar{c}$ is a smooth curve on the quotient manifold.

The second part offers a partial converse. For example, if $c\colon I \to \mathcal{M}$ is a smooth curve on the quotient, then for any $t_0 \in I$ there exists an interval $J \subseteq I$ around t_0 and a smooth curve $\bar{c}\colon J \to \overline{\mathcal{M}}$ such that $c(t) = \pi(\bar{c}(t))$.

9.4 Tangent, vertical and horizontal spaces

Tangent vectors to a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ are rather abstract objects: a point $[x] \in \mathcal{M}$ is an equivalence class for $\sim$, and a tangent vector $\xi \in \mathrm{T}_{[x]}\mathcal{M}$ is an equivalence class of smooth curves on $\mathcal{M}$ passing through $[x]$ as defined in Section 8.4. Fortunately, we can put the total space $\overline{\mathcal{M}}$ to good use to select concrete representations of tangent vectors.

In all that follows, it is helpful to think of the case where $\overline{\mathcal{M}}$ is itself an embedded submanifold of a linear space $\mathcal{E}$ (in our running example, $\overline{\mathcal{M}} = \mathrm{St}(n, p)$ is embedded in $\mathbb{R}^{n\times p}$). Then, tangent vectors to $\overline{\mathcal{M}}$ can be represented easily as matrices, as we did in early chapters.

Accordingly, our goal is to establish one-to-one correspondences between certain tangent vectors of $\overline{\mathcal{M}}$ and tangent vectors of $\mathcal{M}$. Owing to Definition 9.1, a tool of choice for this task is $\mathrm{D}\pi(x)$: it maps $\mathrm{T}_x\overline{\mathcal{M}}$ onto $\mathrm{T}_{[x]}\mathcal{M}$. This map, however, is not one-to-one. To resolve this issue, we proceed to restrict its domain

Consider a point $x \in \overline{\mathcal{M}}$ and its fiber $\mathcal{F}$. We know from Proposition 9.3 that $\mathrm{T}_x\mathcal{F}$ is a subspace of $\mathrm{T}_x\overline{\mathcal{M}}$ and that it coincides with the kernel of $\mathrm{D}\pi(x)$. We call it the *vertical space* V_x. In some sense, vertical directions are the "uninteresting" directions of $\mathrm{T}_x\overline{\mathcal{M}}$ from the standpoint of the quotient manifold. If $\overline{\mathcal{M}}$ is a Riemannian manifold, we have access to an inner product $\langle \cdot, \cdot \rangle_x$ on $\mathrm{T}_x\overline{\mathcal{M}}$. This naturally suggests that we also consider the orthogonal complement of V_x.

Definition 9.24. *For a quotient manifold* $\mathcal{M} = \overline{\mathcal{M}}/\sim$*, the* vertical space *at* $x \in \overline{\mathcal{M}}$ *is the subspace*

$$\mathrm{V}_x = \mathrm{T}_x\mathcal{F} = \ker \mathrm{D}\pi(x)$$

where $\mathcal{F} = \{y \in \overline{\mathcal{M}} : y \sim x\}$ *is the fiber of* x*. If* $\overline{\mathcal{M}}$ *is Riemannian, we call the orthogonal complement of* V_x *the* horizontal space *at* x*:*

$$\mathrm{H}_x = (\mathrm{V}_x)^\perp = \left\{u \in \mathrm{T}_x\overline{\mathcal{M}} : \langle u, v \rangle_x = 0 \textit{ for all } v \in \mathrm{V}_x\right\}.$$

Then, $\mathrm{T}_x\overline{\mathcal{M}} = \mathrm{V}_x \oplus \mathrm{H}_x$ *is a direct sum of linear spaces.*

Since $\ker \mathrm{D}\pi(x) = \mathrm{V}_x$, the restricted linear map

$$\mathrm{D}\pi(x)|_{\mathrm{H}_x} \colon \mathrm{H}_x \to \mathrm{T}_{[x]}\mathcal{M} \tag{9.14}$$

is bijective. Via this map, we may use (concrete) horizontal vectors at x to represent (abstract) tangent vectors at $[x]$ unambiguously. The former are called *horizontal lifts* of the latter.

Definition 9.25. *Consider a point* $x \in \overline{\mathcal{M}}$ *and a tangent vector* $\xi \in \mathrm{T}_{[x]}\mathcal{M}$*. The* horizontal lift *of* ξ *at* x *is the (unique) horizontal vector* $u \in \mathrm{H}_x$ *such that* $\mathrm{D}\pi(x)[u] = \xi$*. We write*

$$u = \left(\mathrm{D}\pi(x)|_{\mathrm{H}_x}\right)^{-1}[\xi] = \mathrm{lift}_x(\xi). \tag{9.15}$$

The following compositions are often useful:

$$\mathrm{D}\pi(x) \circ \mathrm{lift}_x = \mathrm{Id} \qquad \textit{and} \qquad \mathrm{lift}_x \circ \mathrm{D}\pi(x) = \mathrm{Proj}_x^{\mathrm{H}}, \tag{9.16}$$

where $\mathrm{Proj}_x^{\mathrm{H}}$ *is the orthogonal projector from* $\mathrm{T}_x\overline{\mathcal{M}}$ *to* H_x*.*

Conveniently, this definition also allows us to understand smooth curves that represent ξ on the quotient manifold. Indeed, since the horizontal lift u of ξ at x is a tangent vector to $\overline{\mathcal{M}}$ at x, there exists a smooth curve $\bar{c} \colon I \to \overline{\mathcal{M}}$ such that $\bar{c}(0) = x$ and $\bar{c}'(0) = u$. Push this curve to the quotient manifold as follows:

$$c = \pi \circ \bar{c} \colon I \to \mathcal{M}.$$

This is a curve on $\mathcal{M}$, smooth by composition. Moreover, $c(0) = [x]$ and, by the chain rule,

$$c'(0) = \mathrm{D}\pi(\bar{c}(0))[\bar{c}'(0)] = \mathrm{D}\pi(x)[u] = \xi. \tag{9.17}$$

In other words: $c = \pi \circ \bar{c}$ is a smooth curve on the quotient space which passes through $[x]$ with velocity ξ.

Of course, the horizontal lift depends on the point at which the vector is lifted, but there is no ambiguity as to which abstract tangent vector it represents. Specifically, for a tangent vector $\xi \in \mathrm{T}_{[x]}\mathcal{M}$, if $x \sim y$, we may consider horizontal lifts $u_x \in \mathrm{H}_x$ and $u_y \in \mathrm{H}_y$. While u_x and u_y are generally different objects, they represent the same tangent vector of $\mathcal{M}$:

$$\mathrm{D}\pi(x)[u_x] = \xi = \mathrm{D}\pi(y)[u_y]. \tag{9.18}$$

The following example illustrates the concept of vertical and horizontal spaces for $\mathrm{St}(n,p)/\mathrm{O}(p)$ and shows how horizontal lifts of a same vector at two different points of a fiber are related.

Example 9.26. *In Example 9.19 we determined the tangent spaces of fibers of* $\mathcal{M} = \mathrm{St}(n,p)/\mathrm{O}(p)$*. That reveals the vertical spaces:*

$$\mathrm{V}_X = \{X\Omega : \Omega \in \mathrm{Skew}(p)\}.$$

With the usual Riemannian metric on $\mathrm{St}(n,p)$*, namely,* $\langle U, V\rangle_X = \mathrm{Tr}(U^\top V)$*, we can determine the horizontal spaces:*

$$\mathrm{H}_X = \{U \in \mathrm{T}_X\mathrm{St}(n,p) : \langle U, X\Omega\rangle_X = 0 \text{ for all } \Omega \in \mathrm{Skew}(p)\}.$$

In this definition, we conclude that $X^\top U$ *must be symmetric. Yet, from* (7.17) *we also know that* $U \in \mathbb{R}^{n\times p}$ *is in* $\mathrm{T}_X\mathrm{St}(n,p)$ *exactly when* $X^\top U$ *is skew-symmetric. Hence, we deduce that*

$$\mathrm{H}_X = \{U \in \mathbb{R}^{n\times p} : X^\top U = 0\}.$$

For a given $\xi \in \mathrm{T}_{[X]}\mathcal{M}$*, say* U_X *is its horizontal lift at* X*. Consider another point in* $[X]$*, namely,* $Y = XQ$ *for some* $Q \in \mathrm{O}(p)$*. What is the horizontal lift of* ξ *at* Y*? To determine this, as a first step, we select a smooth curve* $\bar{c}$ *on* $\mathrm{St}(n,p)$ *such that* $\bar{c}(0) = X$ *and* $\bar{c}'(0) = U_X$*. From eq.* (9.17) *we know that*

$$\xi = (\pi \circ \bar{c})'(0).$$

Now, consider another smooth curve on $\mathrm{St}(n,p)$*:* $\tilde{c}(t) = \bar{c}(t)Q$*. Clearly,* $\tilde{c}(0) = XQ = Y$ *and* $\tilde{c}'(0) = \bar{c}'(0)Q = U_XQ$ *. Since by construction* $\pi\circ\bar{c}$ *and* $\pi\circ\tilde{c}$ *are the* same *curve on* $\mathcal{M}$*, we may conclude that*

$$\xi = (\pi \circ \bar{c})'(0) = (\pi \circ \tilde{c})'(0) = \mathrm{D}\pi(\tilde{c}(0))[\tilde{c}'(0)] = \mathrm{D}\pi(Y)[U_XQ].$$

Crucially, U_XQ *is a horizontal vector at* Y *since* $Y^\top U_XQ = Q^\top X^\top U_XQ = 0$ *owing to* $X^\top U_X = 0$*. Uniqueness of horizontal lifts then tells us that* $U_Y = U_XQ$ *is the horizontal lift of* ξ *at* Y*, that is,*

$$\mathrm{lift}_{XQ}(\xi) = \mathrm{lift}_X(\xi)\cdot Q. \tag{9.19}$$

That formula proves useful later on.

9.5 Vector fields

A vector field on a quotient manifold $\mathcal{M}$ is defined in the usual way as a suitable map from $\mathcal{M}$ to the tangent bundle $\mathrm{T}\mathcal{M}$. In light of the above discussion regarding horizontal lifts of vectors, it is natural to relate vector fields on $\mathcal{M}$ to *horizontal vector fields* on $\overline{\mathcal{M}}$, that is, vector fields on the total space whose tangent vectors are horizontal.

Specifically, if V is a vector field on $\mathcal{M} = \overline{\mathcal{M}}/\sim$, then

$$\bar{V}(x) = \mathrm{lift}_x(V([x])) \tag{9.20}$$

uniquely defines a vector field $\bar{V}$ on $\overline{\mathcal{M}}$ called the *horizontal lift* of V. We also write more compactly

$$\bar{V} = \mathrm{lift}(V). \tag{9.21}$$

Conveniently, a vector field is smooth exactly if its horizontal lift is smooth.

$$\begin{tikzcd} \overline{\mathcal{M}} \arrow[r, "\bar{V}"] \arrow[d, "\pi"'] & \mathrm{T}\overline{\mathcal{M}} \arrow[d, "\mathrm{D}\pi"] \\ \mathcal{M} \arrow[r, "V"] & \mathrm{T}\mathcal{M} \end{tikzcd}$$

Figure 9.2 Commutative diagram for Theorem 9.27 about lifted vector fields.

Theorem 9.27. *A vector field V on a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with canonical projection π is related to its horizontal lift $\bar{V}$ by:*

$$V \circ \pi = \mathrm{D}\pi \circ \bar{V}. \tag{9.22}$$

Moreover, V is smooth on $\mathcal{M}$ if and only if $\bar{V}$ is smooth on $\overline{\mathcal{M}}$.

The "if" direction of this proposition is fairly direct. To establish the "only if" part, we need one additional technical result first. From Proposition 9.5, recall that for all $x' \in \overline{\mathcal{M}}$ there exists a chart $(\overline{\mathcal{U}}, \bar{\varphi})$ of $\overline{\mathcal{M}}$ around x' and a chart $(\mathcal{U}, \varphi)$ of $\mathcal{M}$ around $\pi(x') = [x']$ such that $\tilde{\pi} = \varphi \circ \pi \circ \bar{\varphi}^{-1}$ is simply $\tilde{\pi}(z_1, \ldots, z_{n+k}) = (z_1, \ldots, z_n)$, with $\dim \overline{\mathcal{M}} = n + k$ and $\dim \mathcal{M} = n$. Recall also the definition of coordinate vector fields given by equation (8.15).

Proposition 9.28. *The coordinate vector fields $\bar{W}_1, \ldots, \bar{W}_{n+k}$ for the chart $\bar{\varphi}$ have the property that (with $e_1, \ldots, e_n$ the canonical basis vectors of $\mathbb{R}^n$):*

$$\mathrm{D}\pi(x)\big[\bar{W}_i(x)\big] = \begin{cases} (\mathrm{D}\varphi(\pi(x)))^{-1}[e_i] & \text{if } i \in \{1, \ldots n\}, \\ 0 & \text{if } i \in \{n+1, \ldots, n+k\}. \end{cases}$$

In particular, $\bar{W}_{n+1}, \ldots, \bar{W}_{n+k}$ are vertical.

Proof. Each coordinate vector field $\bar{W}_i$ is defined for $z \in \bar{\varphi}(\overline{\mathcal{U}})$ by (8.15):

$$\bar{W}_i(\bar{\varphi}^{-1}(z)) = \mathrm{D}\bar{\varphi}^{-1}(z)[\bar{e}_i], \tag{9.23}$$

where $\bar{e}_i$ is the ith canonical basis vector of $\mathbb{R}^{n+k}$. Differentiate $\tilde{\pi}$ at z along the direction $\dot{z} \in \mathbb{R}^{n+k}$: using the simple expression of $\tilde{\pi}$ on one side, and the chain rule for $\tilde{\pi} = \varphi \circ \pi \circ \bar{\varphi}^{-1}$ on the other side, we get

$$\begin{aligned} (\dot{z}_1, \ldots, \dot{z}_n) &= \mathrm{D}\tilde{\pi}(z)[\dot{z}] \\ &= \mathrm{D}\varphi(\pi(\bar{\varphi}^{-1}(z)))\big[\mathrm{D}\pi(\bar{\varphi}^{-1}(z))\big[\mathrm{D}\bar{\varphi}^{-1}(z)[\dot{z}]\big]\big]. \end{aligned}$$

Introducing the notation $x = \bar{\varphi}^{-1}(z)$, the expression simplifies to:

$$(\dot{z}_1, \ldots, \dot{z}_n) = \mathrm{D}\varphi(\pi(x))\big[\mathrm{D}\pi(x)\big[\mathrm{D}\bar{\varphi}^{-1}(z)[\dot{z}]\big]\big]. \tag{9.24}$$

In particular, for $\dot{z} = \bar{e}_i$ we recognize the coordinate vector fields as in (9.23) so that

$$\mathrm{D}\varphi(\pi(x))\big[\mathrm{D}\pi(x)\big[\bar{W}_i(x)\big]\big] = \begin{cases} e_i & \text{if } i \in \{1, \ldots, n\}, \\ 0 & \text{if } i \in \{n+1, \ldots, n+k\}. \end{cases}$$

To conclude, note that $\mathrm{D}\varphi(\pi(x))$ is invertible since φ is a chart. $\square$

We can now give a proof for Theorem 9.27 characterizing smoothness of vector fields on quotient manifolds.

Proof of Theorem 9.27. Equation (9.22) follows from the definition of $\bar{V}$ (9.21) and from the properties of horizontal lifts (9.16). Using Theorem 9.21 then equation (9.22), we find the following equivalences:

$$V \text{ is smooth} \iff V \circ \pi \text{ is smooth} \iff \mathrm{D}\pi \circ \bar{V} \text{ is smooth.}$$

Since $\mathrm{D}\pi$ is smooth by Proposition 8.44, if $\bar{V}$ is smooth, then V is smooth by composition.

The other way around, if V is smooth, then $\mathrm{D}\pi \circ \bar{V}$ is smooth. We want to deduce that $\bar{V}$ is smooth. To this end, for any $x' \in \overline{\mathcal{M}}$, summon the coordinate vector fields $\bar{W}_1, \ldots, \bar{W}_{n+k}$ afforded by Proposition 9.28 and defined on some neighborhood $\overline{\mathcal{U}}$ of x'. By Corollary 8.52, there exist unique functions $g_i \colon \overline{\mathcal{U}} \to \mathbb{R}$ such that, on the domain $\overline{\mathcal{U}}$,

$$\bar{V}(x) = \sum_{i=1}^{n+k} g_i(x) \bar{W}_i(x), \tag{9.25}$$

and $\bar{V}$ is smooth on $\overline{\mathcal{U}}$ if (and only if) these functions are smooth.

We first show $g_1, \ldots, g_n$ are smooth. Since $\mathrm{D}\pi \circ \bar{V}$ is smooth,

$$x \mapsto \sum_{i=1}^{n+k} g_i(x) \mathrm{D}\pi(x)[\bar{W}_i(x)]$$

is smooth. Using properties of $\mathrm{D}\pi \circ \bar{W}_i$ specified by Proposition 9.28, we further find that

$$x \mapsto \sum_{i=1}^{n} g_i(x) \left(\mathrm{D}\varphi(\pi(x))\right)^{-1} [e_i] = \left(\mathrm{D}\varphi(\pi(x))\right)^{-1} [(g_1(x), \ldots, g_n(x))]$$

is smooth, where e_i is the ith canonical basis vector of $\mathbb{R}^n$. Since $\mathrm{D}\varphi(\pi(x))$ is smooth as a function of x (because it is part of a chart for $\mathrm{T}\mathcal{M}$, see Theorem 8.43), it follows that $g_1, \ldots, g_n$ are smooth.

It only remains to show $g_{n+1}, \ldots, g_{n+k}$ are also smooth. To this end, we establish linear equations relating the coordinate functions. With $j \in \{1, \ldots, k\}$, we get k equations by taking an inner product of (9.25) against $\bar{W}_{n+j}$. Since $\bar{V}$

is horizontal and each $\bar{W}_{n+j}$ is vertical, we find:

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \langle \bar{W}_1, \bar{W}_{n+1} \rangle & \cdots & \langle \bar{W}_n, \bar{W}_{n+1} \rangle \\ \vdots & & \vdots \\ \langle \bar{W}_1, \bar{W}_{n+k} \rangle & \cdots & \langle \bar{W}_n, \bar{W}_{n+k} \rangle \end{bmatrix} \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix} + \begin{bmatrix} \langle \bar{W}_{n+1}, \bar{W}_{n+1} \rangle & \cdots & \langle \bar{W}_{n+k}, \bar{W}_{n+1} \rangle \\ \vdots & & \vdots \\ \langle \bar{W}_{n+1}, \bar{W}_{n+k} \rangle & \cdots & \langle \bar{W}_{n+k}, \bar{W}_{n+k} \rangle \end{bmatrix} \begin{bmatrix} g_{n+1} \\ \vdots \\ g_{n+k} \end{bmatrix}.$$

Since (a) the functions $\langle \bar{W}_i, \bar{W}_j \rangle$ are smooth for all i and j (by definition of Riemannian metrics and smoothness of coordinate vector fields), (b) the coordinate functions $g_1, \ldots, g_n$ are smooth, and (c) the $k \times k$ coefficient matrix is positive definite (by linear independence of the coordinate vector fields) and thus smoothly invertible, we conclude that $g_{n+1}, \ldots, g_{n+k}$ are indeed smooth. This confirms that $\bar{V}$ is smooth at an arbitrary x', hence $\bar{V}$ is smooth. □

In light of the above result, actions (recall (8.14)) of smooth vector fields on smooth functions on the quotient manifold are easily understood in the total space.

Proposition 9.29. *For a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with canonical projection π, consider a smooth vector field $V \in \mathfrak{X}(\mathcal{M})$ and a smooth function $f \in \mathfrak{F}(\mathcal{M})$ together with their lifts $\bar{V} \in \mathfrak{X}(\overline{\mathcal{M}})$ and $\bar{f} \in \mathfrak{F}(\overline{\mathcal{M}})$. Then, the lift of Vf is $\bar{V}\bar{f}$, that is:*

$$(Vf) \circ \pi = \bar{V}\bar{f}. \tag{9.26}$$

In words: we may lift then act, or act then lift.

Proof. By definition of the action of a smooth vector field on a smooth function (8.14), for all $[x] \in \mathcal{M}$,

$$(Vf)([x]) = \mathrm{D}f([x])[V([x])].$$

On the other hand, by the chain rule on $\bar{f} = f \circ \pi$ and (9.22),

$$(\bar{V}\bar{f})(x) = \mathrm{D}\bar{f}(x)[\bar{V}(x)] = \mathrm{D}f([x])[\mathrm{D}\pi(x)[\bar{V}(x)]] = \mathrm{D}f([x])[V([x])].$$

Hence, $(\bar{V}\bar{f})(x) = (Vf)(\pi(x))$ for all $x \in \overline{\mathcal{M}}$. □

Another useful consequence of Theorem 9.27 is that we can construct local frames (Definition 3.68) for $\overline{\mathcal{M}}$ that separate into horizontal and vertical parts. We use this and the next proposition to argue smoothness of certain retractions in Theorem 9.33 below. Recall the definition of orthonormal local frame given in Exercise 3.72.

Proposition 9.30. *Let $\mathcal{M} = \overline{\mathcal{M}}/\sim$ be a quotient manifold with canonical projection π and $\dim \mathcal{M} = n$, $\dim \overline{\mathcal{M}} = n + k$. For every $x \in \overline{\mathcal{M}}$, there exists an* *orthonormal local frame* $\hat{W}_1, \ldots, \hat{W}_{n+k}$ *of* $\overline{\mathcal{M}}$ *smoothly defined on a neighborhood* $\overline{\mathcal{U}}$ *of* x *in* $\overline{\mathcal{M}}$ *such that*

1. $\hat{W}_1, \ldots, \hat{W}_n$ *are horizontal vector fields, and*
2. $\hat{W}_{n+1}, \ldots, \hat{W}_{n+k}$ *are vertical vector fields.*

Also, $W_i = \mathrm{D}\pi[\hat{W}_i]$ *for* $i = 1, \ldots, n$ *form a local frame for* $\mathcal{M}$ *on* $\pi(\overline{\mathcal{U}})$.

Proof. By Proposition 8.51, the coordinate vector fields $\bar{W}_1, \ldots, \bar{W}_{n+k}$ provided by Proposition 9.28 already form a local frame. Moreover, $\bar{W}_{n+1}, \ldots, \bar{W}_{n+k}$ are already vertical. However, $\bar{W}_1, \ldots, \bar{W}_n$ may not be horizontal. To build $\hat{W}_1, \ldots, \hat{W}_{n+k}$, apply Gram–Schmidt orthogonalization to $\bar{W}_1, \ldots, \bar{W}_{n+k}$ in reverse order: it is an exercise to verify that this achieves the desired result and preserves smoothness. $\square$

See Section 9.7 for a comment regarding orthonormality of the local frame $W_1, \ldots, W_n$ constructed in Proposition 9.30.

We can use this last proposition to show that the lift map is smooth. To make sense of this statement, we resort to local sections (Proposition 9.22).

Proposition 9.31. *For every* $[x']$ *on a quotient manifold* $\mathcal{M} = \overline{\mathcal{M}}/\sim$, *there exists a local section* $S\colon \mathcal{U} \to \overline{\mathcal{M}}$ *on a neighborhood* $\mathcal{U}$ *of* $[x']$ *such that*

$$\ell\colon \mathrm{T}\mathcal{U} \to \mathrm{T}\overline{\mathcal{M}}\colon ([x], \xi) \mapsto \ell([x], \xi) = \mathrm{lift}_{S([x])}(\xi)$$

is smooth.

Proof. Using Proposition 9.30, select a local frame $\hat{W}_1, \ldots, \hat{W}_{n+k}$ on a neighborhood $\overline{\mathcal{U}}$ of x' in $\overline{\mathcal{M}}$. This also yields a corresponding local frame $W_1, \ldots, W_n$ on $\mathcal{U} = \pi(\overline{\mathcal{U}})$ (a neighborhood of $[x']$) defined by $W_i = \mathrm{D}\pi[\hat{W}_i]$. By construction, $\hat{W}_1, \ldots, \hat{W}_n$ are horizontal and $\hat{W}_{n+1}, \ldots, \hat{W}_{n+k}$ are vertical. Select a local section $S\colon \mathcal{U} \to \overline{\mathcal{M}}$ such that $S([x']) = x'$ (if this requires reducing the domain $\mathcal{U}$, do so and reduce $\overline{\mathcal{U}}$ as well to preserve the relation $\mathcal{U} = \pi(\overline{\mathcal{U}})$). Notice that these domains are still neighborhoods of $[x']$ and x' respectively, and the local frames are well defined on them. Further select a chart $(\mathcal{U}, \varphi)$ of $\mathcal{M}$ around $[x']$ and a chart $(\overline{\mathcal{U}}, \bar{\varphi})$ of $\overline{\mathcal{M}}$ around x' such that $S(\mathcal{U}) \subseteq \overline{\mathcal{U}}$; again, reduce domains if necessary.

Use the local frame $W_1, \ldots, W_n$ to build a chart of $\mathrm{T}\mathcal{U}$ as follows:

$$([x], \xi) \mapsto (\varphi([x]), a_1, \ldots, a_n),$$

where $a_1, \ldots, a_n$ are defined through $\xi = a_1 W_1([x]) + \cdots + a_n W_n([x])$. That this is a chart follows from the fact that basic charts of the tangent bundle are defined using coordinate vector fields (see Theorem 8.43), and changing coordinates between these and any local frame is a diffeomorphism. Likewise, build a chart of $\mathrm{T}\overline{\mathcal{M}}$ on the domain $\mathrm{T}\overline{\mathcal{U}}$ as

$$(x, u) \mapsto (\bar{\varphi}(x), a_1, \ldots, a_{n+k}),$$

where a_is are uniquely defined by $u = a_1 \hat{W}_1(x) + \cdots + a_{n+k} \hat{W}_{n+k}(x)$.

We can write ℓ through these charts as $\tilde{\ell} \colon \mathbb{R}^{2n} \to \mathbb{R}^{2(n+k)}$: since

$$\begin{aligned}\mathrm{lift}_{S([x])}(\xi) &= \mathrm{lift}_{S([x])}\left(a_1 W_1([x]) + \cdots + a_n W_n([x])\right) \\ &= a_1 \hat{W}_1(S([x])) + \cdots + a_n \hat{W}_n(S([x])),\end{aligned}$$

the coordinates of the vector part of $\ell([x], \xi)$ are obtained simply by appending k zeros, so that

$$\tilde{\ell}(z_1, \ldots, z_n, a_1, \ldots, a_n) = \left(\bar{\varphi}(S(\varphi^{-1}(z))), a_1, \ldots, a_n, 0, \ldots, 0\right).$$

This is a smooth function, hence ℓ is smooth. $\square$

Exercise 9.32. *Work out the details for the proof of Proposition 9.30.*

9.6 Retractions

Given a retraction $\overline{\mathrm{R}}$ on the total space $\overline{\mathcal{M}}$, we may try to define a retraction R on the quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ as follows:

$$\mathrm{R}_{[x]}(\xi) = \left[\overline{\mathrm{R}}_x(\mathrm{lift}_x(\xi))\right]. \tag{9.27}$$

If this is well defined, that is, if the right-hand side does not depend on the choice of lifting point $x \in [x]$, this is indeed a retraction.

Theorem 9.33. *If the retraction $\overline{\mathrm{R}}$ on the total space $\overline{\mathcal{M}}$ satisfies*

$$x \sim y \implies \overline{\mathrm{R}}_x(\mathrm{lift}_x(\xi)) \sim \overline{\mathrm{R}}_y(\mathrm{lift}_y(\xi)) \tag{9.28}$$

for all $x, y \in \overline{\mathcal{M}}$ and $\xi \in \mathrm{T}_{[x]}\mathcal{M}$, then (9.27) defines a retraction R on $\mathcal{M}$.

Proof. Since $\mathrm{lift}_x(0) = 0$, it holds that $\mathrm{R}_{[x]}(0) = [x]$. Assuming for now that R is indeed smooth, by the chain rule, for all $\xi \in \mathrm{T}_{[x]}\mathcal{M}$ we have

$$\mathrm{DR}_{[x]}(0)[\xi] = \mathrm{D}\pi(x)\left[\mathrm{D}\overline{\mathrm{R}}_x(0)[\mathrm{Dlift}_x(0)[\xi]]\right] = \xi,$$

where we used $\mathrm{Dlift}_x(0) = \mathrm{lift}_x$ since it is a linear map, $\mathrm{D}\overline{\mathrm{R}}_x(0)$ is identity since $\overline{\mathrm{R}}$ is a retraction, and $\mathrm{D}\pi(x) \circ \mathrm{lift}_x$ is identity. This confirms that $\mathrm{DR}_{[x]}(0)$ is the identity map on $\mathrm{T}_{[x]}\mathcal{M}$.

To verify smoothness, invoke Proposition 9.31 to select a local section $S \colon \mathcal{U} \to \overline{\mathcal{M}}$; then, $\mathrm{R}_{[x]}(\xi) = \pi\left(\overline{\mathrm{R}}_{S([x])}\left(\mathrm{lift}_{S([x])}(\xi)\right)\right)$ is smooth on $\mathrm{T}\mathcal{U}$ since π, S and $\overline{\mathrm{R}}$ are smooth, and so is the map $([x], \xi) \mapsto \mathrm{lift}_{S([x])}(\xi)$. To conclude, repeat this argument on a collection of domains $\mathcal{U}$ to cover all of $\mathrm{T}\mathcal{M}$. $\square$

Example 9.34. *As we can guess from the introduction of this chapter, both the QR retraction and the polar retraction on $\mathrm{St}(n, p)$ satisfy the condition in Theorem 9.33. Indeed, from (9.19) we know that, for all $Q \in \mathrm{O}(p)$,*

$$XQ + \mathrm{lift}_{XQ}(\xi) = (X + \mathrm{lift}_X(\xi))\, Q.$$

As a result, these are valid retractions on the quotient manifold $\mathrm{St}(n,p)/\mathrm{O}(p)$*:*

$$\mathrm{R}^{\mathrm{QR}}_{[X]}(\xi) = [\mathrm{qfactor}(X + \mathrm{lift}_X(\xi))]\,, \quad \textit{and} \tag{9.29}$$

$$\mathrm{R}^{\mathrm{pol}}_{[X]}(\xi) = [\mathrm{pfactor}(X + \mathrm{lift}_X(\xi))]\,, \tag{9.30}$$

where qfactor *extracts the Q-factor of a QR decomposition of a matrix in* $\mathbb{R}^{n\times p}$*, and* pfactor *extracts its polar factor: see* (7.22) *and* (7.24)*.*

9.7 Riemannian quotient manifolds

A quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\!\sim$ is a manifold in its own right. As such, we may endow it with a Riemannian metric of our choosing. As is the case for Riemannian submanifolds—which inherit their metric from the embedding space—endowing the quotient manifold with a metric inherited from the total space leads to nice formulas for objects such as gradients, connections, Hessians and covariant derivatives.

What does it take for the Riemannian metric of $\overline{\mathcal{M}}$ to induce a Riemannian metric on the quotient manifold $\mathcal{M}$? A natural idea is to try to work with horizontal lifts. Specifically, consider $\xi, \zeta \in \mathrm{T}_{[x]}\mathcal{M}$. It is tempting to (tentatively) define an inner product $\langle\cdot,\cdot\rangle_{[x]}$ on $\mathrm{T}_{[x]}\mathcal{M}$ by

$$\langle\xi,\zeta\rangle_{[x]} = \langle\mathrm{lift}_x(\xi), \mathrm{lift}_x(\zeta)\rangle_x\,, \tag{9.31}$$

where $\langle\cdot,\cdot\rangle_x$ is the Riemannian metric on $\mathrm{T}_x\overline{\mathcal{M}}$. For this to make sense, the definition of $\langle\xi,\zeta\rangle_{[x]}$ must not depend on our choice of x: the point at which tangent vectors are lifted. That is, for all $\xi,\zeta \in \mathrm{T}_{[x]}\mathcal{M}$, we must have

$$x \sim y \implies \langle\mathrm{lift}_x(\xi), \mathrm{lift}_x(\zeta)\rangle_x = \langle\mathrm{lift}_y(\xi), \mathrm{lift}_y(\zeta)\rangle_y\,. \tag{9.32}$$

If this condition holds for all $[x]$, we may ask whether (9.31) defines a *Riemannian* metric on $\mathcal{M}$. The answer is yes. Indeed, recall from Definition 8.55 that a metric is Riemannian if for every pair of smooth vector fields $U, V \in \mathfrak{X}(\mathcal{M})$ the function

$$f([x]) = \langle U([x]), V([x])\rangle_{[x]}$$

is smooth on $\mathcal{M}$. To see that this is the case, consider the horizontal lifts $\bar{U}, \bar{V}$ of U, V, and the function $\bar{f} = \langle\bar{U},\bar{V}\rangle$ on $\overline{\mathcal{M}}$:

$$\bar{f}(x) = \left\langle\bar{U}(x), \bar{V}(x)\right\rangle_x = \langle\mathrm{lift}_x(U([x])), \mathrm{lift}_x(V([x]))\rangle_x \overset{(9.31)}{=} \langle U([x]), V([x])\rangle_{[x]} = f([x]).$$

The function $\bar{f}$ is smooth since $\bar{U}$ and $\bar{V}$ are smooth by Theorem 9.27. Furthermore, $\bar{f} = f\circ\pi$, which shows f is smooth by Theorem 9.21. Thus, if (9.31) is well defined, then it is a Riemannian metric, and lifting commutes with taking inner products, that is,

$$\forall U, V \in \mathfrak{X}(\mathcal{M}), \qquad \langle U, V\rangle\circ\pi = \langle\mathrm{lift}(U), \mathrm{lift}(V)\rangle\,. \tag{9.33}$$

The above discussion supports the following result, which doubles as a definition of *Riemannian quotient manifold* (coined in [AMS08, p53]).

Theorem 9.35. *If the Riemannian metric on* $\overline{\mathcal{M}}$ *satisfies* (9.32)*, then* (9.31) *defines a Riemannian metric on the quotient manifold* $\mathcal{M} = \overline{\mathcal{M}}/\sim$*. With this metric,* $\mathcal{M}$ *is called a* Riemannian quotient manifold *of* $\overline{\mathcal{M}}$*.*

For a Riemannian quotient manifold, $\mathrm{D}\pi(x)|_{\mathrm{H}_x}$ and its inverse lift_x are isometries for all $x \in \overline{\mathcal{M}}$; the canonical projection π is then called a *Riemannian submersion* [O'N83, Def. 7.44]. One particular consequence is that the vector fields W_i in Proposition 9.30 are orthonormal.

Example 9.36. *With the usual trace inner product on* $\mathrm{St}(n,p)$ *to make it a Riemannian submanifold of* $\mathbb{R}^{n\times p}$*, we consider the following tentative metric for* $\mathrm{St}(n,p)/\mathrm{O}(p)$*:*

$$\langle \xi, \zeta \rangle_{[X]} = \langle U, V \rangle_X = \mathrm{Tr}(U^\top V),$$

where $U = \mathrm{lift}_X(\xi)$ *and* $V = \mathrm{lift}_X(\zeta)$*. Using the relationship* (9.19) *between lifts at different points of a fiber, we find that, for all* $Q \in \mathrm{O}(p)$*,*

$$\begin{aligned}\langle \mathrm{lift}_{XQ}(\xi), \mathrm{lift}_{XQ}(\zeta) \rangle_{XQ} &= \langle UQ, VQ \rangle_{XQ} = \mathrm{Tr}((UQ)^\top (VQ)) \\ &= \mathrm{Tr}(U^\top V) = \langle U, V \rangle_X = \langle \mathrm{lift}_X(\xi), \mathrm{lift}_X(\zeta) \rangle_X .\end{aligned}$$

This confirms that the tentative metric is invariant under the choice of lifting point: condition (9.32) *is fulfilled, thus Theorem 9.35 tells us this is a Riemannian metric on the quotient manifold, turning it into a Riemannian quotient manifold.*

For the special case of a quotient manifold defined through a Lie group action, condition (9.32) holds if the group action plays nicely with the metric.

Definition 9.37. *A smooth left group action* θ *of a Lie group* $\mathcal{G}$ *on a Riemannian manifold* $\overline{\mathcal{M}}$ *is* isometric *if for all* $g \in \mathcal{G}$ *the map* $F\colon \overline{\mathcal{M}} \to \overline{\mathcal{M}}$ *defined by* $F(x) = \theta(g,x)$ *is isometric, in the sense that*

$$\forall x \in \overline{\mathcal{M}}, \forall u, v \in \mathrm{T}_x\overline{\mathcal{M}}, \qquad \langle \mathrm{D}F(x)[u], \mathrm{D}F(x)[v] \rangle_{F(x)} = \langle u, v \rangle_x .$$

The definition is similar for right actions.

Theorem 9.38. *If the Lie group* $\mathcal{G}$ *acts smoothly, freely, properly and isometrically on the Riemannian manifold* $\overline{\mathcal{M}}$*, then the quotient space* $\overline{\mathcal{M}}/\mathcal{G}$ *is a Riemannian quotient manifold with the Riemannian metric defined in* (9.31)*.*

Proof. In light of Theorem 9.18, it remains to verify condition (9.32). Fix an arbitrary $g \in \mathcal{G}$. Consider the map $F\colon \overline{\mathcal{M}} \to \overline{\mathcal{M}}$ defined by $F(x) = \theta(g,x)$ (assuming without loss of generality that we are dealing with a left action). Fix an arbitrary point $x \in \overline{\mathcal{M}}$. Let $c\colon I \to \overline{\mathcal{M}}$ be an arbitrary smooth curve on $\overline{\mathcal{M}}$

satisfying $c(0) = x$, and let $u \triangleq c'(0)$. It holds that $\pi \circ c = \pi \circ F \circ c$. In particular, their derivatives at $t = 0$ are equal; by the chain rule, this translates into:

$$\mathrm{D}\pi(x)[u] = \mathrm{D}\pi(F(x))[\mathrm{D}F(x)[u]].$$

This holds for all $u \in \mathrm{T}_x\overline{\mathcal{M}}$. Apply $\mathrm{lift}_{F(x)}$ to both sides on the left to deduce that

$$\mathrm{lift}_{F(x)} \circ \mathrm{D}\pi(x) = \mathrm{Proj}^{\mathrm{H}}_{F(x)} \circ \mathrm{D}F(x). \tag{9.34}$$

From this expression, it is clear that $\mathrm{D}F(x)$ transforms vertical vectors at x into vertical vectors at $F(x)$. Moreover, since $\mathrm{D}F(x)$ is a linear isometry and since V_x and $\mathrm{V}_{F(x)}$ have the same dimension, it follows that $\mathrm{D}F(x)$ maps any orthonormal basis of V_x to an orthonormal basis of $\mathrm{V}_{F(x)}$. Consequently, $\mathrm{D}F(x)$ also transforms horizontal vectors at x into horizontal vectors at $F(x)$. Therefore,

$$\mathrm{lift}_{F(x)} \circ \mathrm{D}\pi(x)|_{\mathrm{H}_x} = \mathrm{D}F(x)|_{\mathrm{H}_x}, \tag{9.35}$$

where $\mathrm{D}F(x)|_{\mathrm{H}_x} \colon \mathrm{H}_x \to \mathrm{H}_{F(x)}$ is a linear isometry. It is now clear that condition (9.32) holds: given $y \sim x$, pick g such that $y = F(x)$ and use $\xi = \mathrm{D}\pi(x)[u], \zeta = \mathrm{D}\pi(x)[v]$ for some horizontal vectors $u, v \in \mathrm{H}_x$. $\square$

With this last result, we can revisit Example 9.36. For any $Q \in \mathrm{O}(p)$, consider the map $F \colon \mathrm{St}(n,p) \to \mathrm{St}(n,p)$ defined by $F(X) = XQ$—this captures the group action. Of course, $\mathrm{D}F(X)[U] = UQ$. Thus,

$$\forall X \in \mathrm{St}(n,p), \forall U, V \in \mathrm{T}_X\mathrm{St}(n,p),$$
$$\langle \mathrm{D}F(X)[U], \mathrm{D}F(X)[V] \rangle_{F(X)} = \langle UQ, VQ \rangle_{XQ} = \langle U, V \rangle_X .$$

This confirms that F is an isometry: apply Theorem 9.38 to conclude.

9.8 Gradients

The gradient of a smooth function on a quotient manifold equipped with a Riemannian metric is defined in the same way as for any manifold, see Definition 8.57. Being a smooth vector field on the quotient manifold, the gradient is an abstract object. Prompted by the discussions of the past few sections, we aim to represent the gradient via a horizontal lift. In the important special case of a Riemannian quotient manifold as defined through Theorem 9.35, this can be done rather easily, as we now show.

Consider $f \colon \mathcal{M} \to \mathbb{R}$ and its lift $\bar{f} = f \circ \pi$. On the one hand, the gradient of f with respect to the metric on $\mathcal{M}$ satisfies:

$$\forall ([x], \xi) \in \mathrm{T}\mathcal{M}, \qquad \mathrm{D}f([x])[\xi] = \langle \mathrm{grad} f([x]), \xi \rangle_{[x]} .$$

On the other hand, the gradient of the lifted function $\bar{f}$ with respect to the metric on the total space $\overline{\mathcal{M}}$ obeys:

$$\forall (x, u) \in \mathrm{T}\overline{\mathcal{M}}, \qquad \mathrm{D}\bar{f}(x)[u] = \left\langle \mathrm{grad}\bar{f}(x), u \right\rangle_x .$$

Fix a point $x \in \overline{\mathcal{M}}$. Starting with the latter, using the chain rule on $\bar{f} = f \circ \pi$, and concluding with the former we find:

$$\forall u \in \mathrm{T}_x\overline{\mathcal{M}}, \qquad \left\langle \mathrm{grad}\bar{f}(x), u \right\rangle_x = \mathrm{D}\bar{f}(x)[u] \\ = \mathrm{D}f(\pi(x))[\mathrm{D}\pi(x)[u]] \\ = \langle \mathrm{grad}f([x]), \mathrm{D}\pi(x)[u] \rangle_{[x]}.$$

This holds for all tangent vectors. Thus, for horizontal vectors in particular, using the definition of metric on a Riemannian quotient manifold given by (9.31) and the relations (9.16), we get

$$\forall u \in \mathrm{H}_x, \qquad \left\langle \mathrm{grad}\bar{f}(x), u \right\rangle_x = \langle \mathrm{grad}f([x]), \mathrm{D}\pi(x)[u] \rangle_{[x]} \\ = \langle \mathrm{lift}_x(\mathrm{grad}f([x])), u \rangle_x.$$

This tells us that the horizontal part of $\mathrm{grad}\bar{f}(x)$ is equal to the lift of $\mathrm{grad}f([x])$ at x. What about the vertical part? That one is necessarily zero, owing to the fact that $\bar{f}$ is constant along fibers. Indeed,

$$\forall v \in \mathrm{V}_x, \qquad \left\langle \mathrm{grad}\bar{f}(x), v \right\rangle_x = \mathrm{D}\bar{f}(x)[v] = \mathrm{D}f([x])[\mathrm{D}\pi(x)[v]] = 0, \tag{9.36}$$

since $\mathrm{D}\pi(x)[v] = 0$ for $v \in \mathrm{V}_x$. This leads to a simple conclusion.

Proposition 9.39. *The Riemannian gradient of f on a Riemannian quotient manifold is related to the Riemannian gradient of the lifted function $\bar{f} = f \circ \pi$ on the total space via*

$$\mathrm{lift}_x(\mathrm{grad}f([x])) = \mathrm{grad}\bar{f}(x), \tag{9.37}$$

for all $x \in \overline{\mathcal{M}}$.

In words: to compute the horizontal lift of the gradient of a smooth function f on a Riemannian quotient manifold, we only need to compute the gradient of the lifted function, $\bar{f} = f \circ \pi$. In other words: taking gradients commutes with lifting. Compare with (8.27) for Riemannian submanifolds.

Example 9.40. *In the introduction of this chapter, we considered the cost function $\bar{f}$ (9.3) defined on $\mathrm{St}(n, p)$:*

$$\bar{f}(X) = \mathrm{Tr}(X^\top AX).$$

This function has the invariance $\bar{f}(XQ) = \bar{f}(X)$ for all $Q \in \mathrm{O}(p)$. Thus, there is a well-defined smooth function f on the Riemannian quotient manifold $\mathrm{St}(n, p)/\mathrm{O}(p)$ related to $\bar{f}$ by $\bar{f} = f \circ \pi$. Remembering the expression (9.6) for the gradient of $\bar{f}$ with respect to the usual Riemannian metric on $\mathrm{St}(n, p)$, and applying Proposition 9.39 to relate it to the gradient of f on the Riemannian quotient manifold, we find:

$$\mathrm{lift}_X(\mathrm{grad}f([X])) = \mathrm{grad}\bar{f}(X) = 2(I_n - XX^\top)AX. \tag{9.38}$$

Notice how the gradient of $\bar{f}$ is necessarily horizontal: comparing with the explicit *description of the horizontal and vertical spaces given in Example 9.26, we see why one of the terms in* (9.6) *had to cancel.*

9.9 A word about Riemannian gradient descent

Consider a smooth cost function f on a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ endowed with a Riemannian metric and a retraction R. Given an initial guess $[x_0] \in \mathcal{M}$, RGD on f iterates

$$[x_{k+1}] = \mathrm{R}_{[x_k]}(-\alpha_k \mathrm{grad} f([x_k])) \tag{9.39}$$

with step-sizes α_k determined in some way. How can we run this abstract algorithm numerically, in practice?

The first step is to decide how to store the iterates $[x_0], [x_1], [x_2], \ldots$ in memory. An obvious choice is to store $x_0, x_1, x_2, \ldots$ themselves. These are points of $\overline{\mathcal{M}}$: if the latter is an embedded submanifold of a Euclidean space for example, this should be straightforward.

With access to x_k as a representative of $[x_k]$, we turn to computing $\mathrm{grad} f([x_k])$. In the spirit of Section 9.4, we consider its horizontal lift at x_k. This is a tangent vector to $\overline{\mathcal{M}}$ at x_k: it should be straightforward to store in memory as well. If $\mathcal{M}$ is a Riemannian quotient manifold of $\overline{\mathcal{M}}$, then Proposition 9.39 conveniently tells us that

$$\mathrm{lift}_{x_k}(\mathrm{grad} f([x_k])) = \mathrm{grad} \bar{f}(x_k), \tag{9.40}$$

where $\bar{f} = f \circ \pi$ and $\pi \colon \overline{\mathcal{M}} \to \mathcal{M}$ is the canonical projection. Thus, by computing $\mathrm{grad} \bar{f}(x_k)$, we get a hold of $\mathrm{grad} f([x_k])$.

With these ingredients in memory, it remains to discuss how we can compute x_{k+1}. Let us assume that R is related to a retraction $\overline{\mathrm{R}}$ on $\overline{\mathcal{M}}$ through (9.27). Then, proceeding from (9.39) we deduce that

$$\begin{aligned} [x_{k+1}] &= \mathrm{R}_{[x_k]}(-\alpha_k \mathrm{grad} f([x_k])) \\ &= \left[\overline{\mathrm{R}}_{x_k}(-\alpha_k \, \mathrm{lift}_{x_k}(\mathrm{grad} f([x_k])))\right] \\ &= \left[\overline{\mathrm{R}}_{x_k}(-\alpha_k \mathrm{grad} \bar{f}(x_k))\right]. \end{aligned} \tag{9.41}$$

Thus, if $\mathcal{M}$ is a Riemannian quotient manifold of $\overline{\mathcal{M}}$ and if R and $\overline{\mathrm{R}}$ are related via (9.27), then numerically iterating

$$x_{k+1} = \overline{\mathrm{R}}_{x_k}(-\alpha_k \mathrm{grad} \bar{f}(x_k)) \tag{9.42}$$

on $\overline{\mathcal{M}}$ is *equivalent* to running the abstract iteration (9.39) on $\mathcal{M}$. Interestingly, iteration (9.42) is nothing but RGD on $\bar{f}$. In this sense, and under the stated assumptions, RGD on $\mathcal{M}$ and on $\overline{\mathcal{M}}$ are identical.

As a technical point, note that for (9.42) to be a proper instantiation of (9.39), we must make sure that the chosen step-size α_k depends on x_k only through the equivalence class $[x_k]$. Under the same assumptions as above, this is indeed the case so long as α_k is determined based on the line-search function (and possibly other invariant quantities)—this covers typical line-search methods. Explicitly, the line-search functions for $\bar{f}$ at x and for f at $[x]$ are the same:

$$\forall t, \qquad \bar{f}\big(\overline{\mathrm{R}}_x(-t\mathrm{grad}\bar{f}(x))\big) = f\big(\mathrm{R}_{[x]}(-t\mathrm{grad}f([x]))\big).$$

Though running RGD on $\mathcal{M}$ or on $\overline{\mathcal{M}}$ may be the same, we still reap a theoretical benefit from the quotient perspective. We discussed the local convergence behavior of RGD in Section 4.6, noting that we may expect linear convergence to a local minimizer if the Hessian of the cost function at that point is positive definite. Crucially, the cost function $\bar{f}$ on the total space $\overline{\mathcal{M}}$ *cannot* admit such critical points because of its invariance under $\sim$.

Lemma 9.41. *Let $\overline{\mathcal{M}}$ be a Riemannian manifold and let $f\colon \mathcal{M} \to \mathbb{R}$ be smooth on a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with canonical projection π. If $x \in \overline{\mathcal{M}}$ is a critical point for $\bar{f} = f \circ \pi$, then the vertical space V_x is included in the kernel of* $\mathrm{Hess}\bar{f}(x)$*. In particular, if* $\dim \mathcal{M} < \dim \overline{\mathcal{M}}$ *then* $\mathrm{Hess}\bar{f}(x)$ *is not positive definite.*

Proof. Pick an arbitrary vertical vector $v \in \mathrm{V}_x$. Since the fiber of x is an embedded submanifold of $\overline{\mathcal{M}}$ with tangent space V_x at x, we can pick a smooth curve $\bar{c}$ on the fiber of x such that $\bar{c}(0) = x$ and $\bar{c}'(0) = v$. With $\bar{\nabla}$ and $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}$ denoting the Riemannian connection and induced covariant derivative on $\overline{\mathcal{M}}$, we have identities as in (5.17):

$$\mathrm{Hess}\bar{f}(x)[v] = \bar{\nabla}_v \mathrm{grad}\bar{f} = \frac{\bar{\mathrm{D}}}{\mathrm{d}t}\mathrm{grad}\bar{f}(\bar{c}(t))\bigg|_{t=0}.$$

By Proposition 9.39 , the fact that x is a critical point for $\bar{f}$ implies that $[x]$ is a critical point for f. Still using that same proposition, we also see that, for all t in the domain of $\bar{c}$,

$$\mathrm{grad}\bar{f}(\bar{c}(t)) = \mathrm{lift}_{\bar{c}(t)}(\mathrm{grad}f([\bar{c}(t)])) = \mathrm{lift}_{\bar{c}(t)}(\mathrm{grad}f([x])) = 0.$$

It follows that $\mathrm{Hess}\bar{f}(x)[v] = 0$. □

Thus, the standard theory does not predict fast local convergence for RGD on $\bar{f}$.

The good news is: the trivial eigenvalues of the Hessian of $\bar{f}$ associated to vertical directions do not appear in the spectrum of the Hessian of f on the quotient manifold (see Exercise 9.46). Thus, if the version of RGD we actually run on $\overline{\mathcal{M}}$ is equivalent to RGD on $\mathcal{M}$, we may apply the local convergence results of Section 4.6 to f rather than to $\bar{f}$. In many instances, the local minimizers of f *do* have the property that the Hessian there is positive definite, in which case we can claim (and indeed observe) fast local convergence.

As a closing remark: bear in mind that, in full generality, given a sequence $x_0, x_1, x_2, \ldots$ on $\overline{\mathcal{M}}$, it may happen that the sequence of equivalence classes $[x_0], [x_1], [x_2], \ldots$ converges to a limit point in $\mathcal{M}$, while $x_0, x_1, x_2, \ldots$ itself does *not* converge in $\overline{\mathcal{M}}$.

9.10 Connections

Let $\overline{\mathcal{M}}$ be a Riemannian manifold. Recall from Theorem 8.63 that $\overline{\mathcal{M}}$ is equipped with a uniquely defined Riemannian connection, here denoted by $\bar{\nabla}$. Likewise, a Riemannian quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ is equipped with its own uniquely defined Riemannian connection, ∇. Conveniently, due to the strong link between the Riemannian metric on $\overline{\mathcal{M}}$ and that on $\mathcal{M}$, their Riemannian connections are also tightly related. The main object of this section is to establish the formula:

$$\mathrm{lift}(\nabla_U V) = \mathrm{Proj}^{\mathrm{H}}\big(\bar{\nabla}_{\bar{U}} \bar{V}\big), \tag{9.43}$$

where $\mathrm{lift}\colon \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\overline{\mathcal{M}})$ extracts the horizontal lift of a vector field as in (9.21), $\bar{U} = \mathrm{lift}(U)$, $\bar{V} = \mathrm{lift}(V)$, and $\mathrm{Proj}^{\mathrm{H}}\colon \mathfrak{X}(\overline{\mathcal{M}}) \to \mathfrak{X}(\overline{\mathcal{M}})$ orthogonally projects each tangent vector of a vector field to the horizontal space at its base.

The proof of this statement is based on the Koszul formula (5.11), which we first encountered in the proof of Theorem 5.6. Recall that this formula completely characterizes the Riemannian connection in terms of the Riemannian metric and Lie brackets: for all $U, V, W \in \mathfrak{X}(\mathcal{M})$,

$$\begin{aligned} 2\left\langle \nabla_U V, W\right\rangle = U\langle V, W\rangle + V\langle W, U\rangle - W\langle U, V\rangle \\ - \langle U, [V, W]\rangle + \langle V, [W, U]\rangle + \langle W, [U, V]\rangle . \end{aligned}$$

To make progress, we must first understand how Lie brackets on the quotient manifold are related to Lie brackets of horizontal lifts.

Proposition 9.42. *For any two smooth vector fields $U, V \in \mathfrak{X}(\mathcal{M})$ and their horizontal lifts $\bar{U}, \bar{V} \in \mathfrak{X}(\overline{\mathcal{M}})$,*

$$\mathrm{lift}([U, V]) = \mathrm{Proj}^{\mathrm{H}}([\bar{U}, \bar{V}]). \tag{9.44}$$

Proof. From (9.26) and (9.33), recall that for all $U, V \in \mathfrak{X}(\mathcal{M})$ and $f \in \mathfrak{F}(\mathcal{M})$ and their lifts $\bar{U} = \mathrm{lift}(U)$, $\bar{V} = \mathrm{lift}(V)$ and $\bar{f} = f \circ \pi$:

$$(Vf)\circ\pi = \bar{V}\bar{f}, \qquad\qquad \langle U, V\rangle \circ \pi = \left\langle \bar{U}, \bar{V}\right\rangle. \tag{9.45}$$

Then, by definition of Lie brackets,

$$\begin{aligned} \mathrm{lift}([U, V])\bar{f} &= ([U, V]f)\circ\pi \\ &= (UVf)\circ\pi - (VUf)\circ\pi \\ &= \bar{U}\bar{V}\bar{f} - \bar{V}\bar{U}\bar{f} \\ &= [\bar{U}, \bar{V}]\bar{f} \\ &= \mathrm{Proj}^{\mathrm{H}}([\bar{U}, \bar{V}])\bar{f}, \end{aligned}$$

where the last equality holds because $\bar{f}$ is constant along vertical directions. The fact that this holds for all lifted functions $\bar{f}$ allows to conclude. Slightly more explicitly, using $\bar{V}\bar{f} = \left\langle \bar{V}, \mathrm{grad}\bar{f}\right\rangle$ twice, the above can be reformulated as:

$$\langle \mathrm{lift}([U, V]), \mathrm{grad}\bar{f}\rangle = \langle \mathrm{Proj}^{\mathrm{H}}([\bar{U}, \bar{V}]), \mathrm{grad}\bar{f}\rangle.$$

Then, consider for each point $x \in \overline{\mathcal{M}}$ a collection of $\dim \mathcal{M}$ functions f whose gradients at $[x]$ form a basis for the tangent space $\mathrm{T}_{[x]}\mathcal{M}$: the gradients of their lifts form a basis for the horizontal space H_x, which forces the horizontal parts of $\mathrm{lift}([U,V])$ and $\mathrm{Proj}^{\mathrm{H}}([\bar{U},\bar{V}])$ to coincide. Since both fields are horizontal, they are equal. $\square$

Let $\bar{U}, \bar{V}, \bar{W}$ denote the horizontal lifts of U, V and W. Using identities (9.44) and (9.45) several times we find that:

$$\begin{aligned} (U\langle V, W\rangle) \circ \pi &= \bar{U}(\langle V, W\rangle \circ \pi) = \bar{U}\left\langle \bar{V}, \bar{W}\right\rangle, \text{ and} \\ \langle U, [V,W]\rangle \circ \pi &= \left\langle \bar{U}, \mathrm{lift}([V,W])\right\rangle = \left\langle \bar{U}, [\bar{V},\bar{W}]\right\rangle, \end{aligned} \tag{9.46}$$

where in the last equality we used that $\bar{U}$ is horizontal. With these identities in hand, compare the Koszul formulas for both ∇ and $\bar{\nabla}$: this justifies the second equality in

$$\left\langle \mathrm{Proj}^{\mathrm{H}}(\bar{\nabla}_{\bar{U}}\bar{V}), \bar{W}\right\rangle = \left\langle \bar{\nabla}_{\bar{U}}\bar{V}, \bar{W}\right\rangle = \langle \nabla_U V, W\rangle \circ \pi = \left\langle \mathrm{lift}(\nabla_U V), \bar{W}\right\rangle,$$

while the first equality holds owing to horizontality of $\bar{W}$. Once more, since this holds for all lifted horizontal fields $\bar{W}$, we see that (9.43) holds, as announced. This discussion warrants the following theorem.

Theorem 9.43. *Let $\mathcal{M}$ be a Riemannian quotient manifold of $\overline{\mathcal{M}}$. The Riemannian connections ∇ on $\mathcal{M}$ and $\bar{\nabla}$ on $\overline{\mathcal{M}}$ are related by*

$$\mathrm{lift}(\nabla_U V) = \mathrm{Proj}^{\mathrm{H}}\left(\bar{\nabla}_{\bar{U}}\bar{V}\right)$$

for all $U, V \in \mathfrak{X}(\mathcal{M})$, with $\bar{U} = \mathrm{lift}(U)$ and $\bar{V} = \mathrm{lift}(V)$.

Compare this result to (8.28) for Riemannian submanifolds.

Exercise 9.44. *Show that*

$$\bar{\nabla}_{\bar{U}}\bar{V} = \mathrm{lift}(\nabla_U V) + \frac{1}{2}\mathrm{Proj}^{\mathrm{V}}([\bar{U},\bar{V}]),$$

where $\mathrm{Proj}^{\mathrm{V}} = \mathrm{Id} - \mathrm{Proj}^{\mathrm{H}}$ is the orthogonal projector to vertical spaces. Argue furthermore that $\mathrm{Proj}^{\mathrm{V}}([\bar{U},\bar{V}])$ at x depends only on $\bar{U}(x)$ and $\bar{V}(x)$.

9.11 Hessians

For a smooth function f on a Riemannian quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$, the Hessian of f is defined at $([x], \xi) \in \mathrm{T}\mathcal{M}$ by

$$\mathrm{Hess} f([x])[\xi] = \nabla_\xi \mathrm{grad} f. \tag{9.47}$$

For any vector field $V \in \mathfrak{X}(\mathcal{M})$, Theorem 9.43 tells us that

$$\mathrm{lift}_x(\nabla_\xi V) = \mathrm{Proj}^{\mathrm{H}}_x(\bar{\nabla}_u \bar{V}), \tag{9.48}$$

where $u = \mathrm{lift}_x(\xi)$ and $\bar{V} = \mathrm{lift}(V)$. Recall from Proposition 9.39 that

$$\mathrm{lift}(\mathrm{grad} f) = \mathrm{grad} \bar{f}, \tag{9.49}$$

with $\bar{f} = f \circ \pi$. Combining, we find that

$$\mathrm{lift}_x(\mathrm{Hess} f([x])[\xi]) = \mathrm{Proj}_x^{\mathrm{H}}\big(\bar{\nabla}_u \mathrm{grad} \bar{f}\big). \tag{9.50}$$

Finally, since $\bar{\nabla}_u \mathrm{grad} \bar{f} = \mathrm{Hess} \bar{f}(x)[u]$, we get the following result.

Proposition 9.45. *The Riemannian Hessian of f on a Riemannian quotient manifold is related to the Riemannian Hessian of the lifted function $\bar{f} = f \circ \pi$ on the total space as*

$$\mathrm{lift}_x(\mathrm{Hess} f([x])[\xi]) = \mathrm{Proj}_x^{\mathrm{H}}\big(\mathrm{Hess} \bar{f}(x)[u]\big), \tag{9.51}$$

for all $x \in \overline{\mathcal{M}}$ and $\xi \in \mathrm{T}_{[x]}\mathcal{M}$, with $u = \mathrm{lift}_x(\xi)$.

See Example 9.49 for an illustration.

We already knew from Proposition 9.6 that second-order criticality is preserved between $\mathcal{M}$ and $\overline{\mathcal{M}}$ for f and $\bar{f}$. This has implications for how their Hessians relate at critical points. It is an exercise to make this more precise for the Riemannian structures chosen here.

Exercise 9.46. *Let $f \colon \mathcal{M} \to \mathbb{R}$ be smooth on a Riemannian quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with canonical projection π, and let $\bar{f} = f \circ \pi$. We know that $x \in \overline{\mathcal{M}}$ is a first-order critical point for $\bar{f}$ if and only if $[x]$ is a first-order critical point for f. (See Proposition 9.6 or 9.39.)*

Show that if x is critical then the eigenvalues of $\mathrm{Hess} \bar{f}(x)$ are exactly the eigenvalues of $\mathrm{Hess} f([x])$ together with a set of $\dim \overline{\mathcal{M}} - \dim \mathcal{M}$ additional eigenvalues equal to zero. (Hint: use Lemma 9.41.)

9.12 A word about Riemannian Newton's method

We considered Newton's method on a general Riemannian manifold in Section 6.2. Applied to the minimization of a function f on a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with retraction R, the update equation is

$$[x_{k+1}] = \mathrm{R}_{[x_k]}(\xi_k), \tag{9.52}$$

where $\xi_k \in \mathrm{T}_{[x_k]}\mathcal{M}$ is the solution of the linear equation

$$\mathrm{Hess} f([x_k])[\xi_k] = -\mathrm{grad} f([x_k]), \tag{9.53}$$

which we assume to be unique. In the spirit of Section 9.9, we now discuss how to run (9.52) in practice.

Let us assume that $\mathcal{M}$ is a Riemannian quotient manifold with canonical projection π. We lift both sides of (9.53) to the horizontal space at x_k. With

$s_k = \mathrm{lift}_{x_k}(\xi_k)$, Propositions 9.39 and 9.45 tell us that the linear equation is equivalent to

$$\mathrm{Proj}^{\mathrm{H}}_{x_k}\big(\mathrm{Hess}\bar{f}(x_k)[s_k]\big) = -\mathrm{grad}\bar{f}(x_k), \tag{9.54}$$

where $\bar{f} = f \circ \pi$. This system is to be solved for $s_k \in \mathrm{H}_{x_k}$.

Although $\mathrm{Proj}^{\mathrm{H}}_{x_k}$ and $\mathrm{Hess}\bar{f}(x_k)$ are both symmetric maps on $\mathrm{T}_{x_k}\overline{\mathcal{M}}$, their composition is not necessarily symmetric. This is easily resolved: since s_k is horizontal, we may also rewrite the above as

$$\mathrm{Proj}^{\mathrm{H}}_{x_k}\big(\mathrm{Hess}\bar{f}(x_k)\big[\mathrm{Proj}^{\mathrm{H}}_{x_k}(s_k)\big]\big) = -\mathrm{grad}\bar{f}(x_k). \tag{9.55}$$

This is a linear system with symmetric linear map

$$\mathrm{Proj}^{\mathrm{H}}_{x_k} \circ \mathrm{Hess}\bar{f}(x_k) \circ \mathrm{Proj}^{\mathrm{H}}_{x_k}. \tag{9.56}$$

By construction, if (9.53) has a unique solution ξ_k, then (9.55) has a unique horizontal solution $s_k = \mathrm{lift}_{x_k}(\xi_k)$. (If we solve (9.55) in the whole tangent space $\mathrm{T}_{x_k}\overline{\mathcal{M}}$, then all solutions are of the form $s_k + v$ with $v \in \mathrm{V}_{x_k}$ arbitrary, and s_k is the solution of minimal norm.)

If the retraction R is related to a retraction $\overline{\mathrm{R}}$ on $\overline{\mathcal{M}}$ via (9.27), then continuing from (9.52) we see that

$$[x_{k+1}] = \mathrm{R}_{[x_k]}(\xi_k) = \big[\overline{\mathrm{R}}_{x_k}(\mathrm{lift}_{x_k}(\xi_k))\big] = \big[\overline{\mathrm{R}}_{x_k}(s_k)\big]\,. \tag{9.57}$$

In summary, to run Newton's method on f in practice, we may iterate

$$x_{k+1} = \overline{\mathrm{R}}_{x_k}(s_k) \tag{9.58}$$

with $s_k \in \mathrm{H}_{x_k}$ the horizontal solution of (9.55) (unique if and only if the solution of (9.53) is unique). It is an exercise to check that the conjugate gradients algorithm (CG) from Section 6.3 is well attuned to the computation of s_k.

In contrast, Newton's method on $\overline{\mathcal{M}}$ also iterates (9.58) but with s_k a solution of the following linear system over $\mathrm{T}_{x_k}\overline{\mathcal{M}}$ (if one exists):

$$\mathrm{Hess}\bar{f}(x_k)[s_k] = -\mathrm{grad}\bar{f}(x_k). \tag{9.59}$$

Such a solution may not be horizontal (and its horizontal part may not be a solution), or it may not be unique. Thus, running Newton's method in the total space is *not* equivalent to running it on the quotient manifold. What is more, if x_k converges to a critical point (which is desirable), Lemma 9.41 tells us that $\mathrm{Hess}\bar{f}(x_k)$ converges to a singular map. Thus, we must expect difficulties in solving the linear system on the total space. (However, see Exercise 9.48 for a numerical twist.)

The reasoning above extends to see how to run the Riemannian trust-region method on Riemannian quotient manifolds as well.

Exercise 9.47. *In light of eq.* (9.55)*, consider the linear system $Hs = b$ with*

$$H = \mathrm{Proj}^{\mathrm{H}}_x \circ \mathrm{Hess}\bar{f}(x) \circ \mathrm{Proj}^{\mathrm{H}}_x \qquad \textit{and} \qquad b = -\mathrm{grad}\bar{f}(x)$$

defined at some point $x \in \overline{\mathcal{M}}$. Show that the eigenvalues of H (a self-adjoint map on $\mathrm{T}_x\overline{\mathcal{M}}$) are those of $\mathrm{Hess} f([x])$ *together with an additional* $\dim \overline{\mathcal{M}} - \dim \mathcal{M}$ *trivial eigenvalues. In contrast with Lemma 9.41, show this even if x is not a critical point. Conclude that H is positive definite on* H_x *exactly if* $\mathrm{Hess} f([x])$ *is positive definite. In that scenario, discuss how the iterates of CG (Algorithm 6.2) behave when applied to the system $Hs = b$, especially minding horizontality. How many iterations need to be run at most?*

Exercise 9.48. *Continuing from Exercise 9.47, assume $[\tilde{x}]$ is a strict second-order critical point:* $\mathrm{grad} f([\tilde{x}]) = 0$ *and* $\mathrm{Hess} f([\tilde{x}]) \succ 0$. *Newton's method on f converges to $[\tilde{x}]$ if it ever gets close enough. What happens if we ignore the quotient structure and optimize $\bar{f} = f \circ \pi$ instead?*

Consider the Newton system for $\bar{f}$ at a point x near $\tilde{x}$ on the total space, ignoring the quotient structure:

$$\mathrm{Hess}\bar{f}(x)[s] = -\mathrm{grad}\bar{f}(x), \qquad s \in \mathrm{T}_x\overline{\mathcal{M}}. \tag{9.60}$$

From Lemma 9.41, we know that $\mathrm{Hess}\bar{f}(\tilde{x})$ *has a kernel, hence for x close to $\tilde{x}$ we expect this system to be ill conditioned. And indeed, solving* (9.60) *exactly with a standard algorithm (e.g., Matlab's backslash operator after representing the Hessian and the gradient in matrix and vector form) can lead to catastrophic failure when x is close to $\tilde{x}$.*

However, it is much more common to (try to) solve (9.60) *with CG. A typical observation then would be that roughly* $\dim \mathcal{M}$ *iterations of CG on* (9.60) *are fairly well behaved. The next iteration would break CG. Since the iterates of CG are increasingly better approximate solutions of* (9.60)*, if that happens, it is reasonable to return the best solution reached so far: that is what Matlab's implementation of CG does (`pcg`). As it turns out, that (approximate) solution is numerically close to the solution one would compute if working on the quotient manifold (as in Exercise 9.47).*

Explain this observation. (Hint: compare the Krylov space implicitly generated by CG on (9.60) *to the Krylov space that CG would generate for Newton's method on the quotient manifold, as per Exercise 9.47. It is helpful to use local frames as in Proposition 9.30 together with Lemma 9.41 at $\tilde{x}$.)*

Exercise 9.48 gives some explanation as to why, empirically, running the trust-region method with truncated CG as subproblem solver in the total space (ignoring the quotient) or on the quotient manifold (with matching retractions) often yields strikingly similar results, even though the superlinear convergence guarantees (Section 6.6) break in the total space due to Hessian singularity.

The use of CG (or another Krylov space-based solver) is important here: the solution of the Newton system in the total space at a point which is close to a strict (in the quotient) second-order critical point is not, in general, close to a lift of the Newton step in the quotient space; but numerically the CG algorithm finds an approximate solution to that linear system which happens to mimic the quotient approach.

Thus, the quotient formalism provides us with a basis to understand why running particular second-order optimization algorithms on the total space behaves far better than one might reasonably expect if ignoring the quotient. Moreover, it provides us with a clean alternative that resolves those numerical issues altogether: running the second-order methods on the quotient manifold, through horizontal lifts.

9.13 Total space embedded in a linear space

For all quotient manifolds described in Exercise 9.20, the total space $\overline{\mathcal{M}}$ is an embedded submanifold of a linear space $\mathcal{E}$ (a space of matrices). It is often convenient to make $\overline{\mathcal{M}}$ into a Riemannian submanifold of $\mathcal{E}$, then to make $\mathcal{M} = \overline{\mathcal{M}}/\sim$ into a Riemannian quotient manifold of $\overline{\mathcal{M}}$ (when possible). In this scenario, the geometric tools for $\mathcal{M}$ can be described directly in terms of objects in $\mathcal{E}$.

Consider a smooth function $\bar{\bar{f}} \colon \mathcal{E} \to \mathbb{R}$ (possibly only defined on a neighborhood of $\overline{\mathcal{M}}$). Its restriction, $\bar{f} = \bar{\bar{f}}|_{\overline{\mathcal{M}}}$, is smooth too. Since $\overline{\mathcal{M}}$ is a Riemannian submanifold of $\mathcal{E}$, we know from (3.39) that

$$\mathrm{grad}\bar{f}(x) = \mathrm{Proj}_x\left(\mathrm{grad}\bar{\bar{f}}(x)\right), \tag{9.61}$$

where Proj_x is the orthogonal projector from $\mathcal{E}$ to $\mathrm{T}_x\overline{\mathcal{M}}$.

If furthermore $\bar{f}$ is invariant under $\sim$ so that $\bar{f} = f \circ \pi$ for some smooth function f on the Riemannian quotient manifold $\mathcal{M}$, then Proposition 9.39 tells us that

$$\mathrm{lift}_x(\mathrm{grad}f([x])) = \mathrm{Proj}_x\left(\mathrm{grad}\bar{\bar{f}}(x)\right). \tag{9.62}$$

This notably shows that the right-hand side is a horizontal vector, even though we have only asked for $\bar{f}$ to be invariant: there is no such requirement for all of $\bar{\bar{f}}$, as the equivalence relation $\sim$ is not even formally defined outside of $\overline{\mathcal{M}}$. We exploit this observation to write also:

$$\mathrm{lift}_x(\mathrm{grad}f([x])) = \mathrm{Proj}_x^{\mathrm{H}}\left(\mathrm{grad}\bar{\bar{f}}(x)\right), \tag{9.63}$$

where $\mathrm{Proj}_x^{\mathrm{H}}$ is the orthogonal projector from $\mathcal{E}$ to the horizontal space H_x. To see this, note that $\mathrm{Proj}_x^{\mathrm{H}}(u) = u$ for all $u \in \mathrm{H}_x$, and $\mathrm{Proj}_x^{\mathrm{H}} \circ \mathrm{Proj}_x = \mathrm{Proj}_x^{\mathrm{H}}$ since $\mathrm{H}_x \subseteq \mathrm{T}_x\overline{\mathcal{M}}$, then apply to (9.62).

Similarly, we can express the Hessian of f in terms of $\bar{\bar{f}}$. Indeed, Proposition 9.45 states

$$\mathrm{lift}_x(\mathrm{Hess}f([x])[\xi]) = \mathrm{Proj}_x^{\mathrm{H}}\left(\mathrm{Hess}\bar{f}(x)[u]\right) \tag{9.64}$$

with $u = \mathrm{lift}_x(\xi)$. Moreover, we know from connections on Riemannian submanifolds (5.4) that

$$\mathrm{Hess}\bar{f}(x)[u] = \bar{\nabla}_u\mathrm{grad}\bar{f}(x) = \mathrm{Proj}_x\left(\mathrm{D}\bar{\bar{G}}(x)[u]\right), \tag{9.65}$$

where $\bar{\bar{G}}$ is a smooth extension of $\mathrm{grad}\bar{f}$ to a neighborhood of $\overline{\mathcal{M}}$ in $\mathcal{E}$. Note that to pick the extension $\bar{\bar{G}}$ we are free to entertain expressions for $\mathrm{Proj}_x\Big(\mathrm{grad}\bar{\bar{f}}(x)\Big)$ or $\mathrm{Proj}^{\mathrm{H}}_x\Big(\mathrm{grad}\bar{\bar{f}}(x)\Big)$ as they are both equal to $\mathrm{grad}\bar{f}(x)$ but one may lead to more convenient intermediate expressions than the other. Then,

$$\mathrm{lift}_x\big(\mathrm{Hess}f([x])[\xi]\big) = \mathrm{Proj}^{\mathrm{H}}_x\Big(\mathrm{D}\bar{\bar{G}}(x)[u]\Big). \tag{9.66}$$

These formulas are best illustrated through an example.

Example 9.49. *Let us go a few steps further in Example 9.40. Consider the Grassmann manifold* $\mathrm{Gr}(n,p) = \mathrm{St}(n,p)/\mathrm{O}(p)$ *as a Riemannian quotient manifold of* $\mathrm{St}(n,p)$*, itself a Riemannian submanifold of* $\mathbb{R}^{n\times p}$ *equipped with the usual trace inner product. The cost function* $\bar{\bar{f}}(X) = \frac{1}{2}\operatorname{Tr}(X^\top AX)$ *is smooth on* $\mathbb{R}^{n\times p}$*, hence its restriction* $\bar{f} = \bar{\bar{f}}|_{\mathrm{St}(n,p)}$ *is smooth too. Since* $\bar{f}$ *is invariant under* $\mathrm{O}(p)$*, we further find that* f *is smooth on* $\mathrm{Gr}(n,p)$*, with* $f([X]) = \bar{f}(X)$*. The Euclidean derivatives of* $\bar{\bar{f}}$ *are:*

$$\mathrm{grad}\bar{\bar{f}}(X) = AX \qquad \text{and} \qquad \mathrm{Hess}\bar{\bar{f}}(X)[U] = AU.$$

The horizontal spaces are given by $\mathrm{H}_X = \{U \in \mathbb{R}^{n\times p} : X^\top U = 0\}$*, and the corresponding orthogonal projectors are*

$$\mathrm{Proj}^{\mathrm{H}}_X(Z) = (I_n - XX^\top)Z = Z - X(X^\top Z). \tag{9.67}$$

Choosing to work with (9.63) *(rather than* (9.62)*) because* $\mathrm{Proj}^{\mathrm{H}}_X$ *is somewhat simpler than* Proj_X *(the projector to tangent spaces of the Stiefel manifold), we deduce that the lifted gradient of* f *is:*

$$\begin{aligned}\mathrm{lift}_X(\mathrm{grad}f([X])) &= (I_n - XX^\top)\mathrm{grad}\bar{\bar{f}}(X)\\ &= AX - X(X^\top AX).\end{aligned} \tag{9.68}$$

An obvious smooth extension to all of $\mathbb{R}^{n\times p}$ *is simply given by*

$$\bar{\bar{G}}(X) = (I_n - XX^\top)\mathrm{grad}\bar{\bar{f}}(X),$$

with directional derivatives

$$\mathrm{D}\bar{\bar{G}}(X)[U] = (I_n - XX^\top)\mathrm{Hess}\bar{\bar{f}}(X)[U] - (UX^\top + XU^\top)\mathrm{grad}\bar{\bar{f}}(X).$$

Then, we get the lifted Hessian via (9.66)*. For any* $U = \mathrm{lift}_X(\xi)$*, since* U *is horizontal, we get after some simplifications:*

$$\begin{aligned}\mathrm{lift}_X\big(\mathrm{Hess}f([X])[\xi]\big) &= (I_n - XX^\top)\mathrm{Hess}\bar{\bar{f}}(X)[U] - UX^\top\mathrm{grad}\bar{\bar{f}}(X)\\ &= AU - X(X^\top AU) - U(X^\top AX).\end{aligned} \tag{9.69}$$

We can also compute with the Hessian in a quadratic form:

$$\begin{aligned}\langle \xi, \mathrm{Hess}f([X])[\xi]\rangle_{[X]} &= \langle U, AU - X(X^\top AU) - U(X^\top AX)\rangle_X\\ &= \langle U, AU - U(X^\top AX)\rangle,\end{aligned}$$

where $\langle U, V\rangle = \mathrm{Tr}(U^\top V)$ *is the usual trace (Frobenius) inner product.*

Notice how intermediate formulas in (9.68) *and* (9.69) *provide convenient expressions directly in terms of the Euclidean gradient and Hessian of* $\bar{\bar{f}}$*. In Manopt, these formulas are implemented in* `grassmannfactory` *as* `egrad2rgrad` *and* `ehess2rhess`*.*

Example 9.50. *In the previous example, projections to horizontal spaces are more convenient than projections to tangent spaces of the total space. This is not always the case. For example, let* $\mathcal{E} = \mathbb{R}^{d\times n}$ *be the embedding space for*

$$\overline{\mathcal{M}} = \{X \in \mathbb{R}^{d\times n} : \det(XX^\top) \neq 0 \text{ and } X\mathbf{1} = 0\},$$

that is, rank-d *matrices whose columns sum to zero, and let* $\mathcal{M} = \overline{\mathcal{M}}/\mathrm{O}(d)$ *be defined by the equivalence classes* $[X] = \{QX : Q \in \mathrm{O}(d)\}$*. Equivalence classes are one-to-one with non-degenerate clouds of* n *labeled points in* $\mathbb{R}^d$ *up to rigid motion. Make* $\mathcal{M}$ *into a Riemannian quotient manifold of* $\overline{\mathcal{M}}$*, itself a Riemannian submanifold of* $\mathcal{E}$*.*

In this case, $\overline{\mathcal{M}}$ *is an open subset of an affine subspace of* $\mathcal{E}$*. Consequently, the tangent spaces* $\mathrm{T}_X\overline{\mathcal{M}}$ *are all the same:* Proj_X *is independent of* X*; let us denote it with* Proj*. It is more convenient then to use* (9.62) *for the gradient:*

$$\mathrm{lift}_X(\mathrm{grad} f([X])) = \mathrm{Proj}\Big(\mathrm{grad}\bar{\bar{f}}(X)\Big).$$

The right-hand side offers a suitable smooth extension $\bar{\bar{G}}$ *of* $\mathrm{grad}\bar{f}$*. It is easy to differentiate it since* Proj *is constant:* $\mathrm{D}\bar{\bar{G}}(X)[U] = \mathrm{Proj}\Big(\mathrm{Hess}\bar{\bar{f}}(X)[U]\Big)$*. We conclude via* (9.66) *that*

$$\mathrm{lift}_X\left(\mathrm{Hess} f([X])[\xi]\right) = \mathrm{Proj}_X^{\mathrm{H}}\Big(\mathrm{Hess}\bar{\bar{f}}(X)[U]\Big),$$

where $U = \mathrm{lift}_X(\xi)$ *and we used* $\mathrm{Proj}_X^{\mathrm{H}} \circ \mathrm{Proj} = \mathrm{Proj}_X^{\mathrm{H}}$*.*

Exercise 9.51. *Let* $A \in \mathbb{R}^{n\times n}$ *be a symmetric matrix. A subspace* V *is stable (or invariant) under* A *if* $v \in V \implies Av \in V$*. Show that any such subspace admits an orthonormal basis composed of eigenvectors of* A *(and vice versa). Based on Example 9.49, establish the following facts about the problem* $\min_{[X]\in\mathrm{Gr}(n,p)} \frac{1}{2}\mathrm{Tr}(X^\top AX)$ *(called Rayleigh quotient optimization):*

1. *The critical points are the subspaces of dimension* p *stable under* A*.*
2. *The global minimizers are the subspaces spanned by* p *orthonormal eigenvectors associated to* p *smallest eigenvalues of* A *(counting multiplicities).*
3. *The second-order critical points are the global minimizers.*

In particular, all local minimizers are second-order critical hence they are global minimizers. This is a well-known fact sometimes referred to as the hidden convexity of eigenvalue computation.

9.14 Horizontal curves and covariant derivatives

Our primary goal in this section is to understand the covariant derivative of smooth vector fields along smooth curves on a Riemannian quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$, following their discussion for general manifolds in Section 8.12. In so doing, we aim to relate the (uniquely defined) covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$ along a curve c on $\mathcal{M}$ to that of $\overline{\mathcal{M}}$ along a related curve $\bar{c}$, denoted by $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}$.

We can push any curve $\bar{c}$ from $\overline{\mathcal{M}}$ to a smooth curve c on $\mathcal{M}$, defined through $c = \pi \circ \bar{c}$. By the chain rule, their velocities are related as:

$$c'(t) = \mathrm{D}\pi(\bar{c}(t))[\bar{c}'(t)]. \tag{9.70}$$

Since $\mathrm{lift}_{\bar{c}(t)} \circ \mathrm{D}\pi(\bar{c}(t))$ is the orthogonal projector to the horizontal space $\mathrm{H}_{\bar{c}(t)}$, we can also write

$$\mathrm{lift}_{\bar{c}(t)}(c'(t)) = \mathrm{Proj}^{\mathrm{H}}_{\bar{c}(t)}(\bar{c}'(t)). \tag{9.71}$$

In particular, $\|c'(t)\|_{c(t)} = \|\mathrm{Proj}^{\mathrm{H}}_{\bar{c}(t)}\bar{c}'(t)\|_{\bar{c}(t)} \leq \|\bar{c}'(t)\|_{\bar{c}(t)}$: speed can only decrease in going to the quotient.

Expression (9.71) simplifies in a way that proves particularly useful for our purpose if the velocity of $\bar{c}$ is everywhere horizontal.

Definition 9.52. *A smooth curve $\bar{c}\colon I \to \overline{\mathcal{M}}$ is a* horizontal curve *if $\bar{c}'(t)$ is a horizontal vector for all t, that is, if $\bar{c}'(t) \in \mathrm{H}_{\bar{c}(t)}$ for all t.*

For a smooth vector field $Z \in \mathfrak{X}(c)$, we define its horizontal lift $\bar{Z}$ by $\bar{Z}(t) = \mathrm{lift}_{\bar{c}(t)}(Z(t))$, and we write $\bar{Z} = \mathrm{lift}(Z)$ for short. We use similar definitions for $\mathrm{Proj}^{\mathrm{H}}$ acting on vector fields of $\mathfrak{X}(\bar{c})$. It is an exercise to show that smoothness is preserved.

Theorem 9.53. *Given a horizontal curve $\bar{c}\colon I \to \overline{\mathcal{M}}$ and the corresponding smooth curve $c = \pi \circ \bar{c}$ on the Riemannian quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$, the covariant derivative of a vector field $Z \in \mathfrak{X}(c)$ is given by*

$$\frac{\mathrm{D}}{\mathrm{d}t}Z = \mathrm{D}\pi(\bar{c})\left[\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\bar{Z}\right], \tag{9.72}$$

where $\bar{Z} = \mathrm{lift}(Z)$ is the horizontal lift of Z to the curve $\bar{c}$.

Proof. First of all, this is well defined. Indeed, for any $Z \in \mathfrak{X}(c)$, the lift $\bar{Z} \in \mathfrak{X}(\bar{c})$ is uniquely defined, its covariant derivative $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\bar{Z}$ is indeed in $\mathfrak{X}(\bar{c})$, and pushing it through $\mathrm{D}\pi(\bar{c})$ produces a specific smooth vector field in $\mathfrak{X}(c)$. We need to prove that this vector field happens to be $\frac{\mathrm{D}}{\mathrm{d}t}Z$. To this end, we contemplate the three defining properties of $\frac{\mathrm{D}}{\mathrm{d}t}$ in Theorem 8.67.

The first property, $\mathbb{R}$-linearity in Z, follows easily from linearity of lift, $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}$ and $\mathrm{D}\pi(\bar{c})$. The second property, the Leibniz rule, does too for similar reasons. More work is needed to verify the chain rule: we must show that, for all $U \in \mathfrak{X}(\mathcal{M})$, the proposed formula (9.72) satisfies

$$\frac{\mathrm{D}}{\mathrm{d}t}(U \circ c) = \nabla_{c'} U.$$

Since $\mathrm{lift}(U \circ c) = \bar{U} \circ \bar{c}$, the right-hand side of (9.72) yields

$$\mathrm{D}\pi(\bar{c})\left[\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,\mathrm{lift}(U \circ c)\right] = \mathrm{D}\pi(\bar{c})\left[\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,(\bar{U} \circ \bar{c})\right] = \mathrm{D}\pi(\bar{c})\big[\bar{\nabla}_{\bar{c}'}\bar{U}\big]\,, \tag{9.73}$$

where we used the chain rule for $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}$. Importantly, we now use that $\bar{c}'$ is horizontal to invoke the pointwise formula for ∇ (9.48). More specifically, using that vertical vectors are in the kernel of $\mathrm{D}\pi(\bar{c})$ and $\bar{c}' = \mathrm{lift}(c')$ as in (9.71) owing to horizontality, we have

$$\mathrm{D}\pi(\bar{c})\big[\bar{\nabla}_{\bar{c}'}\bar{U}\big] = \mathrm{D}\pi(\bar{c})\big[\mathrm{Proj}^{\mathrm{H}}\big(\bar{\nabla}_{\bar{c}'}\bar{U}\big)\big] = \nabla_{c'}U.$$

This concludes the proof. $\square$

Under the same assumptions, the formula in Theorem 9.53 can be stated equivalently as:

$$\mathrm{lift}\left(\frac{\mathrm{D}}{\mathrm{d}t}Z\right) = \mathrm{Proj}^{\mathrm{H}}\left(\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\bar{Z}\right), \tag{9.74}$$

which is more informative regarding numerical representation. See Section 9.17 for a comment about the horizontality assumption in Theorem 9.53.

Exercise 9.54. *Consider a smooth curve $\bar{c} \colon I \to \overline{\mathcal{M}}$ and its projection to $\mathcal{M}$: $c = \pi \circ \bar{c}$. With Z a vector field along c, show that Z is smooth if and only if $\bar{Z} = \mathrm{lift}(Z)$ is smooth along $\bar{c}$. Furthermore, show that if $\bar{Z}$ is a (not necessarily horizontal) smooth vector field along $\bar{c}$, then $\mathrm{Proj}^{\mathrm{H}}(\bar{Z})$ is a (necessarily horizontal) smooth vector field along $\bar{c}$.*

9.15 Acceleration, geodesics and second-order retractions

The acceleration c'' of a smooth curve c on a quotient manifold $\mathcal{M}$ is defined—as it is in the general case—as the covariant derivative of its velocity. Owing to Theorem 9.53, if $\bar{c}$ is a horizontal curve related to c by $c = \pi \circ \bar{c}$, and if $\mathcal{M}$ is a Riemannian quotient manifold of $\overline{\mathcal{M}}$, then

$$c''(t) = \frac{\mathrm{D}}{\mathrm{d}t}c'(t) = \mathrm{D}\pi(\bar{c}(t))\left[\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\bar{c}'(t)\right] = \mathrm{D}\pi(\bar{c}(t))[\bar{c}''(t)]\,, \tag{9.75}$$

which we can also write as

$$\mathrm{lift}_{\bar{c}(t)}(c''(t)) = \mathrm{Proj}^{\mathrm{H}}_{\bar{c}(t)}(\bar{c}''(t))\,. \tag{9.76}$$

In particular, $\|c''(t)\|_{c(t)} = \|\mathrm{Proj}^{\mathrm{H}}_{\bar{c}(t)}\bar{c}''(t)\|_{\bar{c}(t)} \leq \|\bar{c}''(t)\|_{\bar{c}(t)}$: acceleration can only decrease in going to the quotient.

Recall that geodesics are curves with zero acceleration. A direct consequence of (9.75) is that *horizontal geodesics* (that is, horizontal curves which are also geodesics) on the total space descend to geodesics on the quotient manifold.

Corollary 9.55. *Let $\mathcal{M}$ be a Riemannian quotient manifold of $\overline{\mathcal{M}}$, with canonical projection π. If $\bar{c}\colon I \to \overline{\mathcal{M}}$ is a horizontal geodesic on $\overline{\mathcal{M}}$, then $c = \pi \circ \bar{c}$ is a geodesic on $\mathcal{M}$.*

In this last corollary, one can show that it is sufficient to have $\bar{c}$ be a geodesic with horizontal velocity at any single time, e.g., $\bar{c}'(0)$ horizontal, as then it is necessarily horizontal at all times. Anticipating the definition of completeness of a manifold (Section 10.1), this notably implies that $\mathcal{M}$ is complete if $\overline{\mathcal{M}}$ is complete [GHL04, Prop. 2.109]. A local converse to Corollary 9.55 also holds. The proof (omitted) relies on standard results about ordinary differential equations. See Section 9.17 for a discussion.

Proposition 9.56. *Let $c\colon I \to \overline{\mathcal{M}}/\sim$ be a smooth curve on a Riemannian quotient manifold such that $c'(t_0) \neq 0$, with I an open interval around t_0.*

1. *For any x_0 such that $c(t_0) = [x_0]$, there exists an open interval $J \subseteq I$ around t_0 and a unique, smooth curve $\bar{c}\colon J \to \overline{\mathcal{M}}$ such that $c|_J = \pi \circ \bar{c}$ (that is, $\bar{c}$ is a local lift of c), $\bar{c}$ is horizontal, and $\bar{c}(t_0) = x_0$.*
2. *On J, the curve c is a geodesic if and only if $\bar{c}$ is a geodesic.*

Example 9.57. *Consider the total space of full-rank matrices $\overline{\mathcal{M}} = \mathbb{R}^{d\times n}_d$ with $d \leq n$: an open submanifold of $\mathbb{R}^{d\times n}$. Consider also the quotient space $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with equivalence relation $X \sim Y \iff X^\top X = Y^\top Y$, that is, two clouds of n labeled points in $\mathbb{R}^d$ are equivalent if they have the same Gram matrix. Equivalence classes are of the form $[X] = \{QX : Q \in \mathrm{O}(d)\}$, that is, two clouds are equivalent if they are the same up to rotation and reflection. Use Theorem 9.18 to verify that $\mathcal{M}$ is a quotient manifold of $\overline{\mathcal{M}}$. Its points are in one-to-one correspondence with positive semidefinite matrices of size n and rank d. With the usual metric $\langle U, V\rangle_X = \mathrm{Tr}(U^\top V)$ on $\mathbb{R}^{d\times n}_d$, we can further turn $\mathcal{M}$ into a Riemannian quotient manifold (use Theorem 9.38).*

Given $X, Y \in \mathbb{R}^{d\times n}_d$, the straight line $\bar{c}(t) = (1-t)X + tY$ is a geodesic on $[0,1]$ provided it remains in $\mathbb{R}^{d\times n}_d$. Assuming this is the case, we may further ask: what does it take for $\bar{c}$ to be horizontal? Since the fiber of X is the submanifold $\{QX : Q \in \mathrm{O}(d)\}$, we find that the vertical spaces are $\mathrm{V}_X = \{\Omega X : \Omega \in \mathrm{Skew}(d)\}$, hence the horizontal spaces are given by $\mathrm{H}_X = \{U \in \mathbb{R}^{d\times n} : XU^\top = UX^\top\}$. Thus, $\bar{c}'(t) = Y - X$ belongs to $\mathrm{H}_{\bar{c}(t)}$ exactly if the following is symmetric:

$$\begin{aligned}\bar{c}(t)\bar{c}'(t)^\top &= ((1-t)X + tY)(Y - X)^\top \\ &= XY^\top - t(XY^\top + YX^\top) + tYY^\top - (1-t)XX^\top.\end{aligned}$$

This holds for all t exactly if $XY^\top$ is symmetric. If so, $\bar{c}$ is a horizontal geodesic and Corollary 9.55 states $c(t) = [(1-t)X + tY]$ is a geodesic on $\mathcal{M}$.

What is the significance of the condition $XY^\top = YX^\top$? Consider the Euclidean distance between QX and Y in the total space, where $Q \in \mathrm{O}(d)$ remains *unspecified for now:*

$$\|QX - Y\|^2 = \|X\|^2 + \|Y\|^2 - 2\,\mathrm{Tr}(Q^\top Y X^\top).$$

It can be shown that this distance is minimized with respect to $Q \in \mathrm{O}(d)$ *if* $Q^\top Y X^\top$ *is symmetric and positive semidefinite. Specifically, if* $YX^\top = U\Sigma V^\top$ *is an SVD decomposition, then the minimum is attained by* $Q = UV^\top$ *(the polar factor of* $YX^\top$*). Replacing* X *by* QX *(which does not change its equivalence class) "aligns"* X *to* Y *in the sense that* $\|X - Y\|$ *is minimized.*

Assume X *and* Y *are aligned as described, so that* $XY^\top = YX^\top \succeq 0$*. Since* X *and* Y *have full rank, we see that*

$$\bar{c}(t)\bar{c}(t)^\top = (1-t)^2 XX^\top + t(1-t)(XY^\top + YX^\top) + t^2 YY^\top$$

is positive definite for all $t \in [0,1]$*. In other words: if* X *and* Y *are aligned, the straight line* $\bar{c}(t)$ *connecting them indeed remains in* $\mathbb{R}_d^{d\times n}$ *for* $t \in [0,1]$ *and it is horizontal. Since* $\bar{c}$ *is a horizontal geodesic,* c *is a geodesic too.*

Anticipating concepts of length and distance from Section 10.1, we claim that the length of $\bar{c}$ *on* $[0,1]$ *is* $\|Y - X\|$*, and that* c *has the same length as* $\bar{c}$*. Since no shorter curve connects the same end points, the Riemannian distance between the equivalence classes* $[X]$ *and* $[Y]$ *is nothing but the Euclidean distance between their best aligned representatives. See also Exercise 10.15.*

Massart and Absil discuss the geometry of $\mathcal{M}$ *in detail [MA20]. The points of* $\mathcal{M}$ *(that is, the equivalence classes of* $\overline{\mathcal{M}}$*) are one-to-one with the positive semidefinite matrices of size* n *and rank* d*. The latter form an embedded submanifold of the symmetric matrices of size* n*. The Riemannian metric for that manifold as constructed here on* $\mathcal{M}$ *is called the* Bures–Wasserstein metric *when* $d = n$ *(sometimes also when* $d \leq n$ *by extension). It is different from the Riemannian submanifold metric for that same manifold [VAV09].*

Another direct consequence of (9.75) is that second-order retractions (Definition 8.69) on $\overline{\mathcal{M}}$ which satisfy condition (9.28) and whose curves are horizontal yield second-order retractions on the quotient manifold.

Corollary 9.58. *Let* $\overline{\mathrm{R}}$ *be a retraction on* $\overline{\mathcal{M}}$ *such that* R *as defined by* (9.27) *is a retraction on the Riemannian quotient manifold* $\mathcal{M} = \overline{\mathcal{M}}/\sim$*. If* $\overline{\mathrm{R}}$ *is second order and its retraction curves,*

$$\bar{c}(t) = \overline{\mathrm{R}}_x(tu),$$

are horizontal for every x *and every* $u \in \mathrm{H}_x$*, then* R *is second order on* $\mathcal{M}$*.*

Proof. The retraction on the quotient manifold generates curves

$$c(t) = \mathrm{R}_{[x]}(t\xi) = \left[\overline{\mathrm{R}}_x(tu)\right]$$

with $u = \mathrm{lift}_x(\xi)$; hence, $c = \pi \circ \bar{c}$. Since $\bar{c}$ is horizontal, we may apply (9.75) and evaluate at $t = 0$. (This also makes it clear that $\bar{c}$ needs only be horizontal in a neighborhood of $t = 0$.) □

Example 9.59. *Recall the polar retraction* (7.24) *on* $\mathrm{St}(n,p)$*:*

$$\mathrm{R}_X(U) = (X+U)\left(I_p + U^\top U\right)^{-1/2}. \tag{9.77}$$

This is a second-order retraction. We already checked condition (9.28) *for it, so that it yields a retraction on the quotient manifold* $\mathrm{Gr}(n,p) = \mathrm{St}(n,p)/\mathrm{O}(p)$*. Furthermore, the retraction curves are horizontal. Indeed, for any* $U \in \mathrm{H}_X$ *(meaning* $X^\top U = 0$*), consider the curve*

$$\bar{c}(t) = \mathrm{R}_X(tU) = (X+tU)\left(I_p + t^2 U^\top U\right)^{-1/2}$$

and its velocity

$$\bar{c}'(t) = U\left(I_p + t^2U^\top U\right)^{-1/2} + (X+tU)\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(I_p + t^2U^\top U\right)^{-1/2}\right).$$

This curve is horizontal if, for all t*, the matrix*

$$\begin{aligned}\bar{c}(t)^\top\bar{c}'(t) = \left(I_p + t^2U^\top U\right)^{-1/2}(tU^\top U)\left(I_p + t^2U^\top U\right)^{-1/2} \\ + \left(I_p + t^2U^\top U\right)^{+1/2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(I_p + t^2U^\top U\right)^{-1/2}\right)\end{aligned}$$

is zero. Replace $U^\top U$ *with its eigendecomposition* $VDV^\top$*, with* $V \in \mathrm{O}(p)$ *and* D *diagonal: the right-hand side is diagonal in the basis* V*, and it is a simple exercise to conclude that it is indeed identically zero. As a result, the polar retraction is horizontal on* $\mathrm{St}(n,p)$ *and, when used as a retraction on the Grassmann manifold* $\mathrm{Gr}(n,p)$*, it is second order as well.*

9.16 Grassmann manifold: summary*

For convenience, this section collects the various tools we have constructed throughout the chapter to work on the Grassmann manifold $\mathrm{Gr}(n,p)$. Equivalent constructions appear early in [EAS98].

We view $\mathrm{Gr}(n,p)$ as a Riemannian quotient manifold of the Stiefel manifold $\mathrm{St}(n,p)$, itself a Riemannian submanifold of $\mathbb{R}^{n\times p}$ with the usual trace inner product $\langle A, B\rangle = \mathrm{Tr}(A^\top B)$. (See the end of this section for an embedded viewpoint.) The orthogonal group $\mathrm{O}(p)$ acts on $\mathrm{St}(n,p)$ as $(X,Q) \mapsto XQ$, so that the projection

$$\pi\colon \mathrm{St}(n,p) \to \mathrm{Gr}(n,p)\colon X \mapsto \pi(X) \triangleq [X] = \{XQ : Q \in \mathrm{O}(p)\}$$

is surjective and smooth from $\mathrm{St}(n,p)$ to $\mathrm{Gr}(n,p) = \mathrm{St}(n,p)/\mathrm{O}(p)$, and its differentials $\mathrm{D}\pi(X)$ are surjective. The dimension of $\mathrm{Gr}(n,p)$ is

$$\dim \mathrm{Gr}(n,p) = \dim \mathrm{St}(n,p) - \dim \mathrm{O}(p) = p(n-p).$$

A point $[X]$ on $\mathrm{Gr}(n,p)$ is represented by a matrix $X \in \mathrm{St}(n,p)$ (an arbitrary representative of the equivalence class $[X]$).

For an arbitrary manifold $\mathcal{M}$, a map $F\colon \mathrm{Gr}(n,p) \to \mathcal{M}$ is smooth if and only if $F \circ \pi \colon \mathrm{St}(n,p) \to \mathcal{M}$ is smooth. Likewise, a map $G \colon \mathcal{M} \to \mathrm{Gr}(n,p)$ is smooth at $x \in \mathcal{M}$ if and only if there exists a neighborhood $\mathcal{U}$ of x on $\mathcal{M}$ and a map $\bar{G} \colon \mathcal{U} \to \mathrm{St}(n,p)$ (smooth at x) such that $G|_{\mathcal{U}} = \pi \circ \bar{G}$. As usual, G is smooth if it is so at all points.

Given $X \in \mathrm{St}(n,p)$, the tangent space $\mathrm{T}_X\mathrm{St}(n,p)$ splits in two components, orthogonal for the inner product $\langle A, B\rangle_X = \mathrm{Tr}(A^\top B)$:

$$\mathrm{T}_X\mathrm{St}(n,p) = \mathrm{V}_X + \mathrm{H}_X, \qquad \begin{aligned} \mathrm{V}_X &= \{X\Omega : \Omega \in \mathrm{Skew}(p)\}, \\ \mathrm{H}_X &= \{U \in \mathbb{R}^{n\times p} : X^\top U = 0\}. \end{aligned}$$

The orthogonal projectors to V_X and H_X (both from $\mathbb{R}^{n\times p}$ and from $\mathrm{T}_X\mathrm{St}(n,p)$) take the form

$$\mathrm{Proj}_X^{\mathrm{V}}(Z) = X\,\mathrm{skew}(X^\top Z) = X\frac{X^\top Z - Z^\top X}{2}, \tag{9.78}$$

$$\mathrm{Proj}_X^{\mathrm{H}}(Z) = (I_n - XX^\top)Z = Z - X(X^\top Z). \tag{9.79}$$

The vertical space V_X is the tangent space to the fiber $\pi^{-1}(\pi(X))$. The horizontal space H_X is one-to-one with the tangent space of $\mathrm{Gr}(n,p)$ at $[X]$ via the differential of π; its inverse is the horizontal lift:

$$\begin{aligned} \mathrm{D}\pi(X)|_{\mathrm{H}_X} &\colon \mathrm{H}_X \to \mathrm{T}_{[X]}\mathrm{Gr}(n,p), \\ \mathrm{lift}_X = (\mathrm{D}\pi(X)|_{\mathrm{H}_X})^{-1} &\colon \mathrm{T}_{[X]}\mathrm{Gr}(n,p) \to \mathrm{H}_X. \end{aligned}$$

If $[X] \in \mathrm{Gr}(n,p)$ is represented by $X \in \mathrm{St}(n,p)$, then we represent a tangent vector $\xi \in \mathrm{T}_{[X]}\mathrm{Gr}(n,p)$ with the (unique) matrix $U \in \mathrm{H}_X$ such that $\mathrm{D}\pi(X)[U] = \xi$, or equivalently, such that $U = \mathrm{lift}_X(\xi)$. We could of course represent $[X]$ with a different matrix; for example, with XQ where $Q \in \mathrm{O}(p)$ is arbitrary. The horizontal lifts of ξ at X and at XQ are related by

$$\mathrm{lift}_{XQ}(\xi) = \mathrm{lift}_X(\xi)Q.$$

(See Example 9.26.)

Say $F \colon \mathrm{Gr}(n,p) \to \mathcal{M}$ is smooth, so that $\bar{F} = F \circ \pi \colon \mathrm{St}(n,p) \to \mathcal{M}$ is smooth. Still equivalently, there exists a smooth extension $\bar{\bar{F}}$ of $\bar{F}$ to a neighborhood $\mathcal{U}$ of $\mathrm{St}(n,p)$ in $\mathbb{R}^{n\times p}$ such that $\bar{F} = \bar{\bar{F}}|_{\mathrm{St}(n,p)}$. For all $(X, U) \in \mathrm{TSt}(n,p)$, we have

$$\mathrm{D}\bar{\bar{F}}(X)[U] = \mathrm{D}\bar{F}(X)[U] = \mathrm{D}F([X])[\mathrm{D}\pi(X)[U]]. \tag{9.80}$$

Stated the other way around, this means that if $([X], \xi) \in \mathrm{TGr}(n,p)$ is represented by (X, U) with $U = \mathrm{lift}_X(\xi)$, then

$$\mathrm{D}F([X])[\xi] = \mathrm{D}\bar{F}(X)[U] = \mathrm{D}\bar{\bar{F}}(X)[U]. \tag{9.81}$$

Now, say $G \colon \mathcal{M} \to \mathrm{Gr}(n,p)$ is smooth at $x \in \mathcal{M}$, that is, there exists a neighborhood $\mathcal{U}$ of x on $\mathcal{M}$ and a smooth map $\bar{G} \colon \mathcal{U} \to \mathrm{St}(n,p)$ such that $G|_{\mathcal{U}} = \pi \circ \bar{G}$. Then, for all $v \in \mathrm{T}_x\mathcal{M}$ we have

$$\mathrm{D}G(x)[v] = \mathrm{D}\pi(\bar{G}(x))[\mathrm{D}\bar{G}(x)[v]]. \tag{9.82}$$

Accordingly, the lift of $\mathrm{D}G(x)[v] \in \mathrm{T}_{G(x)}\mathrm{Gr}(n,p)$ to the horizontal space at $\bar{G}(x)$ is given by:

$$\mathrm{lift}_{\bar{G}(x)}(\mathrm{D}G(x)[v]) = \mathrm{Proj}^{\mathrm{H}}_{\bar{G}(x)}(\mathrm{D}\bar{G}(x)[v]). \tag{9.83}$$

We turn $\mathrm{Gr}(n,p)$ into a Riemannian manifold as a Riemannian quotient by pushing the metric from $\mathrm{St}(n,p)$ to $\mathrm{Gr}(n,p)$ through π. This turns $\mathrm{D}\pi(X)|_{\mathrm{H}_X}$ and lift_X into isometries. Explicitly, for all $\xi, \zeta \in \mathrm{T}_{[X]}\mathrm{Gr}(n,p)$, we may choose an arbitrary representative X of $[X]$ and lift ξ, ζ as $U = \mathrm{lift}_X(\xi)$ and $V = \mathrm{lift}_X(\zeta)$; then:

$$\langle \xi, \zeta \rangle_{[X]} = \langle U, V \rangle_X = \mathrm{Tr}(U^\top V).$$

This structure provides gradients. Explicitly, if $f\colon \mathrm{Gr}(n,p) \to \mathbb{R}$ is smooth, then $\bar{f} = f \circ \pi\colon \mathrm{St}(n,p) \to \mathbb{R}$ is smooth and it has a smooth extension $\bar{\bar{f}}$ to a neighborhood of $\mathrm{St}(n,p)$ in $\mathbb{R}^{n\times p}$. Their gradients are related as follows:

$$\mathrm{lift}_X(\mathrm{grad} f([X])) = \mathrm{grad}\bar{f}(X) = (I_n - XX^\top)\mathrm{grad}\bar{\bar{f}}(X). \tag{9.84}$$

Let us turn to the Riemannian connection on $\mathrm{Gr}(n,p)$. Say W is a smooth vector field on $\mathrm{Gr}(n,p)$, meaning $\bar{W} = \mathrm{lift}(W)$ is smooth on $\mathrm{St}(n,p)$. Recall that the Riemannian connection $\bar{\nabla}$ on $\mathrm{St}(n,p)$ is given by $\bar{\nabla}_U \bar{W} = \mathrm{Proj}_X(\mathrm{D}\bar{W}(X)[U])$ for all $U \in \mathrm{T}_X\mathrm{St}(n,p)$, where Proj_X is the orthogonal projector from $\mathbb{R}^{n\times p}$ to $\mathrm{T}_X\mathrm{St}(n,p)$ (and $\mathrm{D}\bar{W}(X)[U] = \mathrm{D}\bar{\bar{W}}(X)[U]$ using any smooth extension $\bar{\bar{W}}$ of $\bar{W}$ to a neighborhood of $\mathrm{St}(n,p)$ in $\mathbb{R}^{n\times p}$). Then, we reason from (9.43) that the Riemannian connection ∇ on $\mathrm{Gr}(n,p)$ satisfies:

$$\mathrm{lift}_X(\nabla_\xi W) = \mathrm{Proj}^{\mathrm{H}}_X\big(\bar{\nabla}_U \bar{W}\big) = \mathrm{Proj}^{\mathrm{H}}_X\big(\mathrm{D}\bar{W}(X)[U]\big), \tag{9.85}$$

where $U = \mathrm{lift}_X(\xi)$.

We can specialize the above discussion to Hessians, with $W = \mathrm{grad} f$ and $\bar{W} = \mathrm{lift}(\mathrm{grad} f) = \mathrm{grad}\bar{f}$, using $\bar{f} = f \circ \pi$. Then, with $\bar{\bar{f}}$ a smooth extension of $\bar{f}$ as above and with $U = \mathrm{lift}_X(\xi)$, we have

$$\begin{aligned}\mathrm{lift}_X(\mathrm{Hess} f([X])[\xi]) &= \mathrm{Proj}^{\mathrm{H}}_X\big(\mathrm{Hess}\bar{f}(X)[U]\big) \qquad (9.86)\\ &= (I_n - XX^\top)\mathrm{Hess}\bar{\bar{f}}(X)[U] - UX^\top\mathrm{grad}\bar{\bar{f}}(X).\end{aligned}$$

(To obtain the last expression, it is useful to refer to (7.29) for Hessians on $\mathrm{St}(n,p)$, noting that the vertical part of $\mathrm{grad}\bar{\bar{f}}(X)$ is zero.) Treating the Hessian as a quadratic map yields

$$\begin{aligned}\langle \xi, \mathrm{Hess} f([X])[\xi]\rangle_{[X]} &= \big\langle U, \mathrm{Hess}\bar{f}(X)[U]\big\rangle \\ &= \langle U, \mathrm{Hess}\bar{\bar{f}}(X)[U] - UX^\top\mathrm{grad}\bar{\bar{f}}(X)\rangle, \qquad (9.87)\end{aligned}$$

still with $U = \mathrm{lift}_X(\xi)$.

Several retractions on $\mathrm{St}(n,p)$ descend to well-defined retractions on $\mathrm{Gr}(n,p)$, including the QR-based (7.22) and the polar-based (7.24) retractions:

$$\mathrm{R}^{\mathrm{QR}}_{[X]}(\xi) = [\mathrm{qfactor}(X + \mathrm{lift}_X(\xi))]\,, \text{ and} \tag{9.88}$$

$$\mathrm{R}^{\mathrm{pol}}_{[X]}(\xi) = [\mathrm{pfactor}(X + \mathrm{lift}_X(\xi))]\,, \tag{9.89}$$

What is more, the polar retraction is second order. We build a transporter compatible with that retraction in Example 10.67. Expressions for the geodesics of $\mathrm{Gr}(n,p)$ (hence for its exponential map) and for parallel transport along geodesics appear in [EAS98, Thms 2.3, 2.4].

An embedded geometry for $\mathrm{Gr}(n,p)$.
It is instructive to know that the Riemannian quotient geometry detailed above can also be realized as a Riemannian submanifold of a Euclidean space [MS85, AMT13, SI14], [BH15, Def. 2.3, §4.2], [BZA20, LLY20]. Interestingly, working out efficient numerical tools to optimize on the Grassmannian with either perspective involves essentially the same steps: below, we only show equivalence of the geometries without numerical concerns.

Explicitly, consider the set

$$\mathcal{M} = \{P \in \mathbb{R}^{n\times n} : P = P^\top, P^2 = P \text{ and } \mathrm{Tr}(P) = p\}. \tag{9.90}$$

Each matrix in $\mathcal{M}$ is an orthogonal projector to a subspace of dimension p in $\mathbb{R}^n$. It can be shown that $\mathcal{M}$ is a smooth embedded submanifold of $\mathbb{R}^{n\times n}$ of dimension $p(n-p)$: that is the same dimension as $\mathrm{Gr}(n,p)$. It is easy to check that

$$\varphi\colon \mathrm{Gr}(n,p) \to \mathcal{M}\colon [X] \mapsto \varphi([X]) = XX^\top \tag{9.91}$$

is well defined. It is also smooth. Indeed, the map $\phi\colon \mathrm{St}(n,p) \to \mathbb{R}^{n\times n}$ defined by $\phi(X) = XX^\top$ is clearly smooth, and $\phi = \varphi \circ \pi$ where π is the quotient map of $\mathrm{Gr}(n,p)$. One can further check that φ is bijective, so that φ^{-1} is well defined.

Endow $\mathbb{R}^{n\times n}$ with the trace inner product $\langle A, B\rangle^{\mathbb{R}^{n\times n}} = \mathrm{Tr}(A^\top B)$, and let $\mathcal{M}$ be a Riemannian submanifold of $\mathbb{R}^{n\times n}$ with that metric: we write $\langle A, B\rangle^{\mathcal{M}}_P = \langle A, B\rangle^{\mathbb{R}^{n\times n}}$ for the inner product on $\mathrm{T}_P\mathcal{M}$. Consider two arbitrary tangent vectors $\xi, \zeta \in \mathrm{T}_{[X]}\mathrm{Gr}(n,p)$ and their horizontal lifts $U, V \in \mathrm{H}_X$. Compute $\mathrm{D}\phi(X)[U]$ in two different ways: directly as

$$\mathrm{D}\phi(X)[U] = UX^\top + XU^\top,$$

then also via the chain rule as

$$\mathrm{D}\phi(X)[U] = \mathrm{D}(\varphi\circ\pi)(X)[U] = \mathrm{D}\varphi(\pi(X))[\mathrm{D}\pi(X)[U]] = \mathrm{D}\varphi([X])[\xi].$$

Thus,

$$\mathrm{D}\varphi([X])[\xi] = UX^\top + XU^\top. \tag{9.92}$$

Owing to $X^\top X = I_p$ and $X^\top U = X^\top V = 0$, it follows that

$$\begin{aligned}\langle \mathrm{D}\varphi([X])[\xi], \mathrm{D}\varphi([X])[\zeta]\rangle^{\mathcal{M}}_{\varphi([X])} &= \langle UX^\top + XU^\top, VX^\top + XV^\top\rangle^{\mathbb{R}^{n\times n}} \\ &= 2\,\langle U, V\rangle^{\mathrm{St}(n,p)}_X \\ &= 2\,\langle \xi, \zeta\rangle^{\mathrm{Gr}(n,p)}_{[X]}, \end{aligned} \tag{9.93}$$

where $\langle\cdot,\cdot\rangle^{\mathrm{St}(n,p)}$ and $\langle\cdot,\cdot\rangle^{\mathrm{Gr}(n,p)}$ denote the Riemannian metrics on $\mathrm{St}(n,p)$ and $\mathrm{Gr}(n,p)$, respectively.

We have found that $\mathrm{D}\varphi([X])$ is (up to a factor $\sqrt{2}$) a linear isometry between $\mathrm{T}_{[X]}\mathrm{Gr}(n,p)$ and $\mathrm{T}_{\varphi([X])}\mathcal{M}$. The inverse function theorem then implies that φ is a local diffeomorphism around each point. Since φ is also invertible, it follows that φ is a global diffeomorphism: $\mathrm{Gr}(n,p)$ and $\mathcal{M}$ have the same smooth geometry. Moreover, their Riemannian metrics are identical up to a factor $\sqrt{2}$. Thus, the Riemannian geometries of $\mathrm{Gr}(n,p)$ and $\mathcal{M}$ are equivalent.

See [BZA20] for more about these and other representations of the Grassmann manifold, including discussions about equivalences and numerical aspects.

9.17 Notes and references

The main source for this chapter is the book by Absil et al. [AMS08, §3 and §5], which gives an original and concise treatment of quotient manifolds for Riemannian optimization, with an emphasis on generality and practicality. The main sources for differential geometric aspects and proofs are the differential geometry books by Brickell and Clark [BC70, §6] and Lee [Lee12, §4 and §21]. O'Neill provides useful results regarding connections on Riemannian submersions [O'N83, pp212–213], as do Gallot et al. [GHL04]. The recent book by Gallier and Quaintance [GQ20] offers an in depth look at the geometry of Lie groups and manifolds arising through group actions.

Here are further references for results above which included a proof: Theorem 9.21 follows [Lee12, Thm. 4.29]; Theorem 9.27 follows [Lee18, Prop. 2.25]; Theorem 9.33 follows [AMS08, Prop. 4.1.3]; Theorem 9.38 appears in [Lee18, Cor. 2.29]; Proposition 9.42 appears in [Lee18, p146]; Theorem 9.43 appears in [O'N83, Lem. 7.45]; and Corollary 9.55 appears in [O'N83, Lem. 7.46]. Moreover, Exercise 9.44 echoes [GHL04, Prop. 3.35], [dC92, Ex. 8.9] and [Lee18, p146].

Many results hold generally for the case where $\overline{\mathcal{M}}$ and $\mathcal{M}$ are smooth manifolds (not necessarily related by an equivalence relation) and $\pi\colon \overline{\mathcal{M}} \to \mathcal{M}$ is a submersion (it is smooth and its differentials are surjective) or a Riemannian submersion (its differentials are isometries once restricted to horizontal spaces). Reference books often state results at this level of generality. In contrast, we also require π to be surjective (it is a quotient map). For example, while $\pi\colon \mathbb{R}^2\backslash\{0\} \to \mathbb{R}^2$ (both with their usual Riemannian structures) defined by $\pi(x) = x$ is a Riemannian submersion, it is not a quotient map because it is not surjective: we exclude such cases. Certain results that do not hold for general submersions may hold for surjective submersions.

An advantage of Brickell and Clark's treatment is that they define smooth manifolds without topological restrictions (recall Section 8.2). As a result, the role of the two topological properties (Hausdorff and second countability) is apparent throughout their developments. This proves helpful here, considering the fact that certain quotient spaces fail to be quotient manifolds specifically because their quotient topologies fail to have these properties.

For a full characterization of when a quotient space is or is not a quotient manifold, see [AMS08, Prop. 3.4.2]. In this chapter, we presented a necessary condition (Proposition 9.3, which implies fibers of a quotient manifold must all be submanifolds of the same dimension), and a sufficient condition (Theorem 9.18, which states free, proper and smooth group actions yield quotient manifolds). For a partial converse to the latter, see [Lee12, Pb. 21-5].

In Definition 9.24, we summon a Riemannian structure on the total space to define horizontal spaces. Technically, that structure is not required: one could just as well define H_x to be any subspace of $\mathrm{T}_x\overline{\mathcal{M}}$ such that the direct sum of V_x and H_x coincides with $\mathrm{T}_x\overline{\mathcal{M}}$, and still obtain that the restriction of $\mathrm{D}\pi(x)$ to H_x is a bijection to $\mathrm{T}_{[x]}\mathcal{M}$. For practical purposes, it is then useful to arrange the choice of H_x to vary smoothly with x, leading to a *horizontal distribution*. However, this leaves a lot of freedom that we do not need. We opt for a more directive (and quite common) definition of horizontal space, while noting that other authors use the terminology in a broader sense [AMS08, §3.5.8].

When the group acting on the total space is not compact, it may be delicate to determine whether its action is proper. The following characterization may help in this regard [Lee12, Prop. 21.5].

Proposition 9.60. *Let $\mathcal{G}$ be a Lie group acting smoothly on a manifold $\overline{\mathcal{M}}$. The following are equivalent:*

1. *The action θ is proper.*
2. *If $x_0, x_1, x_2, \ldots$ is a sequence on $\overline{\mathcal{M}}$ and $g_0, g_1, g_2, \ldots$ is a sequence on $\mathcal{G}$ such that both $\{x_k\}_{k=0,1,2,\ldots}$ and $\{\theta(g_k, x_k)\}_{k=0,1,2,\ldots}$ converge, then a subsequence of $\{g_k\}_{k=0,1,2,\ldots}$ converges.*
3. *For every compact $K \subseteq \overline{\mathcal{M}}$, the set $\mathcal{G}_K = \{g \in \mathcal{G} : \theta(g, K) \cap K \neq \emptyset\}$ is compact.*

The smoothness criterion for vector fields on quotient manifolds given in Theorem 9.27 is an exercise in do Carmo's book [dC92, Ex. 8.9, p186] and is linked to the concept of π-related vector fields [Lee12, Pb. 8-18c]. The main difference for the latter is that Theorem 9.27 states results with respect to the special horizontal distribution we chose (emanating from the Riemannian metric on the total space), whereas results regarding π-related vector fields often focus on the horizontal distribution tied to charts in normal form (Proposition 9.5).

In the proof of Theorem 9.53, one may wonder what goes wrong if $\bar{c}$ is not horizontal. In that case, to proceed from (9.73) we separate $\bar{c}'(t)$ into its horizontal and vertical parts, and we use linearity:

$$\begin{aligned}\mathrm{D}\pi(\bar{c})\big[\bar{\nabla}_{\bar{c}'}\bar{U}\big] &= \mathrm{D}\pi(\bar{c})\big[\bar{\nabla}_{\mathrm{Proj}^{\mathrm{H}}(\bar{c}')}\bar{U}\big] + \mathrm{D}\pi(\bar{c})\big[\bar{\nabla}_{\mathrm{Proj}^{\mathrm{V}}(\bar{c}')}\bar{U}\big]\\ &= \nabla_{c'}U + \mathrm{D}\pi(\bar{c})\big[\bar{\nabla}_{\mathrm{Proj}^{\mathrm{V}}(\bar{c}')}\bar{U}\big]. \end{aligned} \tag{9.94}$$

For the horizontal part, the same argument as in the proof of Theorem 9.53 based on (9.48) and (9.71) still applies, yielding the first term. The second term though, does not vanish in general. This is because, in general, $\bar{U}$ is *not* "constant" along fibers: the lift of $U([x])$ at x need not be the "same" as its lift at $y \sim x$. (To make sense of the quoted terms, see the notion of parallel vector fields in Section 10.3.)

We verify this on an example. Consider $\mathrm{Gr}(n,p) = \mathrm{St}(n,p)/\mathrm{O}(p)$. We know a horizontally lifted vector field on $\mathrm{St}(n,p)$: take for example $\bar{U}(X) = \mathrm{grad}\bar{f}(X) = AX - X(X^\top AX)$, where $\bar{f}(X) = \frac{1}{2}\mathrm{Tr}(X^\top AX)$. Furthermore, any vertical vector at X is of the form $X\Omega$ for some $\Omega \in \mathrm{Skew}(p)$. Then, using the formula for the connection $\bar{\nabla}$ on $\mathrm{St}(n,p)$ (5.4),

$$\begin{aligned}\bar{\nabla}_{X\Omega}\bar{U} &= \mathrm{Proj}_X\big(AX\Omega - X\Omega(X^\top AX) - X(\Omega^\top X^\top AX + X^\top AX\Omega)\big)\\ &= (I_n - XX^\top)AX\Omega,\end{aligned}$$

where Proj_X is the orthogonal projector to $\mathrm{T}_X\mathrm{St}(n,p)$ (7.27). To our point, this vector is horizontal, and it can be nonzero, hence the vector $\mathrm{D}\pi(X)[\bar{\nabla}_{X\Omega}\bar{U}]$ can be nonzero.

It is useful to add a word about Proposition 9.56: this concerns the possibility of horizontally lifting curves $c\colon I \to \mathcal{M}$ from the quotient manifold to curves $\bar{c}$ on the total space. That this can be done locally (meaning: that we can obtain a horizontal lift $\bar{c}$ defined on an open interval around any $t_0 \in I$) is relatively direct, invoking standard results from ordinary differential equations (ODE) [Lee12, Thm. 9.12].

The argument goes like this: if $c'(t_0) \neq 0$, then there exists an interval $J \subseteq I$ around t_0 such that $c(J)$ is an embedded submanifold of $\mathcal{M}$. As a result, it is possible to extend the smooth vector field c' on $c(J)$ to a smooth vector field V defined on a neighborhood $\mathcal{U}$ of $c(J)$. It satisfies $V(c(t)) = c'(t)$ for all $t \in J$. Pick $x_0 \in \overline{\mathcal{M}}$ such that $c(t_0) = [x_0]$, and consider the ODE below whose unknown is the curve γ on $\mathcal{U}$:

$$\gamma'(t) = V(\gamma(t)), \qquad\qquad \gamma(t_0) = [x_0].$$

Clearly, $\gamma(t) = c(t)$ is a solution for $t \in J$. Since solutions of ODEs are unique, we deduce that $\gamma|_J = c|_J$. Now we turn to constructing horizontal lifts. Consider $\bar{V} = \mathrm{lift}(V)$: this is a smooth vector field on $\overline{\mathcal{U}} = \pi^{-1}(\mathcal{U})$. We can again write down an ODE:

$$\bar{c}'(t) = \bar{V}(\bar{c}(t)), \qquad\qquad \bar{c}(t_0) = x_0.$$

There exist an open interval J' around t_0 and a solution $\bar{c}\colon J' \to \overline{\mathcal{U}}$, smooth and unique. Clearly, $\bar{c}$ is horizontal because $\bar{c}'(t)$ is a horizontal vector by construction. If we project $\bar{c}$ to the quotient, then we get a curve $\gamma = \pi \circ \bar{c}$. Notice that $\gamma(t_0) = [x_0]$ and

$$\gamma'(t) = \mathrm{D}\pi(\bar{c}(t))[\bar{c}'(t)] = \mathrm{D}\pi(\bar{c}(t))[\bar{V}(\bar{c}(t))] = V([\bar{c}(t)]) = V(\gamma(t)).$$

Thus, γ satisfies the first ODE we considered, and we conclude that $\pi \circ \bar{c} = \gamma = c|_{J'}$. In words: $\bar{c}$ is a horizontal lift of c on the interval J'. Moreover, the lifted curve depends smoothly on the choice of representative $x_0 \in c(t_0)$ because solutions of smooth ODEs depend smoothly not only on time but also on initial conditions.

This argument and a few more elements form the basis of the proof of Proposition 9.56 presented by Gallot et al. [GHL04, Prop. 2.109], where the emphasis is on the case where c is a geodesic.

It is natural to ask: if c is defined on the interval I, can we not lift it to a horizontal curve $\bar{c}$ also defined on all of I? The argument above is not sufficient to reach this stronger conclusion, in part because it only uses the fact that π is a Riemannian submersion: it does not use the fact that π is surjective. Hence, as Gallot et al. point out, we might be in the case where $\pi \colon \mathbb{R}^2 \backslash \{0\} \to \mathbb{R}^2$ is the map $\pi(x) = x$ between the punctured plane and the plane, both equipped with their usual Riemannian metrics. This is indeed a Riemannian submersion, but it is not surjective (in particular, it is not a quotient map). It is clear that a geodesic (a straight line) through the origin in $\mathbb{R}^2$ cannot be lifted entirely to $\mathbb{R}^2 \backslash \{0\}$.

Thus, at the very least, we should require π to be surjective (which it is in the setting of quotient manifolds). Unfortunately, that is not sufficient. John M. Lee shares[3] a counter-example with $\overline{\mathcal{M}} = (-1, 1) \times \mathbb{R}$ and $\mathcal{M} = \mathrm{S}^1$ as Riemannian submanifolds of $\mathbb{R}^2$. Consider the map $\pi(x, y) = (\cos(2\pi x), \sin(2\pi x))$. This is indeed a surjective Riemannian submersion from $\overline{\mathcal{M}}$ to $\mathcal{M}$. Yet, consider the curve $c \colon (-2, 2) \to \mathcal{M}$ defined by $c(t) = (\cos(2\pi t), \sin(2\pi t))$. Its unique horizontal lift satisfying $\bar{c}(0) = (0, 0)$ is given by $\bar{c}(t) = (t, 0)$: evidently, $\bar{c}$ can only be extended up to $(-1, 1)$, not up to $(-2, 2)$.

Fortunately, there exist several sufficient conditions. One can read the following about that question in a classic book by Besse [Bes87, §9.E] (See Section 10.1 for a definition of complete manifolds.) Recall that if $\mathcal{M}$ is a Riemannian quotient manifold of $\overline{\mathcal{M}}$ then the canonical projection π is a surjective Riemannian submersion. Completeness is defined in Section 10.1. If $\overline{\mathcal{M}}$ is not connected, apply the claim to each complete connected component.

Proposition 9.61. *Let $\pi \colon \overline{\mathcal{M}} \to \mathcal{M}$ be a surjective Riemannian submersion. If $\overline{\mathcal{M}}$ is connected and complete, then $\mathcal{M}$ is also connected and complete, all fibers are complete (but not necessarily connected), and for every smooth curve $c \colon I \to \mathcal{M}$ with non-vanishing velocity and for every $x_0 \in c(t_0)$ there exists a*

[3] math.stackexchange.com/questions/3524475

unique horizontal lift $\bar{c}\colon I \to \overline{\mathcal{M}}$ *such that* $c = \pi \circ \bar{c}$ *and* $\bar{c}(t_0) = x_0$*. Moreover, c is a geodesic if and only if* $\bar{c}$ *is a geodesic.*

If π has this property, namely, that all curves with non-vanishing velocity can be lifted horizontally on their whole domain and made to pass through any representative of the equivalence class at some initial point, we say it is *Ehresmann-complete*. In this same reference, it is also noted that when this property holds, then π is a smooth *fiber bundle*. Furthermore, the property may hold without $\overline{\mathcal{M}}$ being complete.

Here is a special case of interest. If the quotient is obtained as per Theorem 9.18 through a smooth, free and proper Lie group action on a smooth manifold, then it also forms a fiber bundle [Lee12, Pb. 21-6]—in that case, the fiber bundle is also called a *principal G-bundle*, and the *standard fiber* is the Lie group itself. It can be shown that if (but not only if) the standard fiber is compact, then the fiber bundle is Ehresmann-complete [Mic08, pp204–206]. This is summarized as follows.

Proposition 9.62. *If* $\mathcal{G}$ *is a compact Lie group acting smoothly and freely on* $\overline{\mathcal{M}}$*, then* $\mathcal{M} = \overline{\mathcal{M}}/\mathcal{G}$ *is a quotient manifold, and it has the property that any smooth curve* $c\colon I \to \mathcal{M}$ *with non-vanishing velocity can be lifted to a unique horizontal curve* $\bar{c}\colon I \to \overline{\mathcal{M}}$ *passing through any* $x_0 \in c(t_0)$ *at* t_0*. Moreover, c is a geodesic if and only if* $\bar{c}$ *is a geodesic.*

See also [Mic08, Lem. 26.11] and [KN63, Prop. II.3.1, p69]. Thanks to P.-A. Absil, John M. Lee, Mario Lezcano-Casado and Estelle Massart for discussions on this topic.

Regarding Exercise 9.51: That second-order critical points are global optimizers is shown in [SI14, Prop. 3.4, Prop. 4.1] under the assumption that there is a gap between the pth and $(p+1)$st smallest eigenvalues of A. With some care, the eigengap assumption can be removed (this is part of the exercise). The other claims in the exercise are standard. It is also possible to control the spectrum of the Hessian at approximate critical points [LTW21].

We close with a proof of Proposition 9.6 which provides strong links between the salient points of optimization problems related through a map π (which may or may not be a quotient map).

Proof of Proposition 9.6. Let us verify each claim in turn:

1. The range of real values attained by $\bar{f}$ is $\bar{f}(\overline{\mathcal{M}}) = f(\pi(\overline{\mathcal{M}}))$, and $\pi(\overline{\mathcal{M}}) = \mathcal{M}$ since π is surjective. Thus, $\bar{f}(\overline{\mathcal{M}}) = f(\mathcal{M})$.
2. If x is a local minimizer of $\bar{f}$, there exists a neighborhood $\overline{\mathcal{U}}$ of x on $\overline{\mathcal{M}}$ such that $\bar{f}(x) = \inf \bar{f}(\overline{\mathcal{U}})$. Since π is open, the set $\mathcal{U} = \pi(\overline{\mathcal{U}})$ is open, and it contains $\pi(x)$. Thus, $\mathcal{U}$ is a neighborhood of $\pi(x)$. Also, $f(\pi(x)) = \bar{f}(x) = \inf \bar{f}(\overline{\mathcal{U}}) = \inf f(\pi(\overline{\mathcal{U}})) = \inf f(\mathcal{U})$, hence $\pi(x)$ is a local minimizer of f.

   If $\pi(x)$ is a local minimizer of f, there exists a neighborhood $\mathcal{U}$ of $\pi(x)$ on $\mathcal{M}$ such that $f(\pi(x)) = \inf f(\mathcal{U})$. Since π is continuous, the set $\overline{\mathcal{U}} = \pi^{-1}(\mathcal{U})$

is open, and it contains x. Thus, $\overline{\mathcal{U}}$ is a neighborhood of x. We also have $\bar{f}(\overline{\mathcal{U}}) = f(\pi(\pi^{-1}(\mathcal{U}))) \subseteq f(\mathcal{U})$ (with equality if π is surjective), hence $\bar{f}(x) = f(\pi(x)) = \inf f(\mathcal{U}) \leq \inf \bar{f}(\overline{\mathcal{U}})$. (In fact, equality holds because x is in $\overline{\mathcal{U}}$.) Therefore, x is a local minimizer of $\bar{f}$.

3. Recall Definitions 4.4 and 6.1 for first- and second-order critical points. Note that these do not require any Riemannian structure. For the following, we tacitly assume that f is once or twice differentiable, as needed.
   (a) By the chain rule, $\mathrm{D}\bar{f}(x) = \mathrm{D}f(\pi(x)) \circ \mathrm{D}\pi(x)$. If $\mathrm{D}f(\pi(x)) = 0$, then $\mathrm{D}\bar{f}(x) = 0$. The other way around, if $\mathrm{D}\bar{f}(x) = 0$, then $\mathrm{D}f(\pi(x))$ is zero on the image of $\mathrm{D}\pi(x)$. Since $\mathrm{D}\pi(x)$ is surjective, this implies that $\mathrm{D}f(\pi(x)) = 0$.
   (b) Assume $\pi(x)$ is second-order critical for f on $\mathcal{M}$. Pick an arbitrary smooth curve $\bar{c}$ on $\overline{\mathcal{M}}$ such that $\bar{c}(0) = x$. Let $c = \pi \circ \bar{c}$: this is a smooth curve on $\mathcal{M}$ such that $c(0) = \pi(x)$. Notice that $\bar{f} \circ \bar{c} = f \circ \pi \circ \bar{c} = f \circ c$. In particular, $(\bar{f} \circ \bar{c})'(0) = (f \circ c)'(0)$ and $(\bar{f} \circ \bar{c})''(0) = (f \circ c)''(0)$. Since $\pi(x)$ is second-order critical, we know that $(f \circ c)'(0) = 0$ and $(f \circ c)''(0) \geq 0$. Thus, $(\bar{f} \circ \bar{c})'(0) = 0$ and $(\bar{f} \circ \bar{c})''(0) \geq 0$, which implies that x is second-order critical for $\bar{f}$.
   For the other direction, it is convenient to use that $\mathcal{M}$ can always be endowed with a Riemannian structure [Lee12, Prop. 13.3] (it does not matter which one). Assume x is second-order critical for $\bar{f}$. For $v \in \mathrm{T}_x\overline{\mathcal{M}}$ (to be determined), we can always select a smooth curve $\bar{c}$ such that $\bar{c}(0) = x$ and $\bar{c}'(0) = v$. The curve $c = \pi \circ \bar{c}$ satisfies $c(0) = \pi(x)$ and $c'(0) = \mathrm{D}\pi(x)[v]$. Pick an arbitrary $w \in \mathrm{T}_{\pi(x)}\mathcal{M}$. Since $\mathrm{D}\pi(x)$ is surjective, we can now choose v such that $\mathrm{D}\pi(x)[v] = w$. From above, we know that $\pi(x)$ is (at least) first-order critical. Then, owing to the usual Taylor expansion of $f \circ c$ (5.25), we have
   $$\langle w, \mathrm{Hess}f(\pi(x))[w]\rangle_{\pi(x)} = (f \circ c)''(0),$$
   where $\mathrm{Hess}f$ is the Riemannian Hessian of f with respect to the arbitrarily chosen metric. Since w is arbitrary and $(f \circ c)''(0) = (\bar{f} \circ \bar{c})''(0) \geq 0$ (because x is second-order critical for $\bar{f}$), it follows that $\mathrm{Hess}f(\pi(x))$ is positive semidefinite, that is, $\pi(x)$ is second-order critical for f (Proposition 6.3).

Finally, if π is the projection of a quotient manifold, then it is surjective and smooth and its differentials are surjective by Definition 9.1. The map π is also open: see Exercise 9.9. $\square$

10 Additional tools

At times, it is useful to resort to some of the more advanced tools Riemannian geometry has to offer. We discuss some of these here in relation to optimization. The background on differential geometry given in Chapters 3 and 5 is often sufficient. We omit classical proofs, pointing to standard texts instead.

We start with the notion of Riemannian distance, which allows us to turn a (connected) Riemannian manifold into a metric space. It turns out that the associated metric space topology coincides with the manifold topology, and that shortest paths between pairs of points are geodesics. From there, we discuss the Riemannian exponential map: this is a retraction whose curves are geodesics. Then, we also give a formal and not-so-standard treatment of the inverse of the exponential map and, more generally, of the inverse of retractions.

Moving on, parallel transports allow us to move tangent vectors around, from tangent space to tangent space, isometrically. Combined with the exponential map, this tool makes it possible to define a notion of Lipschitz continuity for the gradient and the Hessian of a cost function on a Riemannian manifold. This leads to a sharp understanding of the regularity assumptions we made in Chapters 4 and 6 to control the worst-case behavior of optimization algorithms.

We follow up with the (also not-so-standard) notion of transporter—a poor man's version of parallel transport. This is useful to design certain algorithms. It also affords us a practical notion of finite difference approximation for the Hessian, which makes it possible to use second-order optimization algorithms without computing the Hessian.

In closing, we discuss covariant differentiation of tensor fields of any order.

As a notable omission, we do not discuss curvature at all: see for example [Lee18, Ch. 1, 7] for an introduction.

10.1 Distance, geodesics and completeness

Two points of $\mathcal{M}$ belong to the same *connected component* if there exists a continuous curve on $\mathcal{M}$ joining them. We say $\mathcal{M}$ is *connected* if it has a single connected component. A manifold has finitely many, or countably infinitely many connected components, owing to second-countability of the atlas topology (see Section 8.2).

Definition 10.1. *A* distance *on a set* $\mathcal{M}$ *is a function* $\mathrm{dist}\colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ *such that, for all* $x, y, z \in \mathcal{M}$,

1. $\mathrm{dist}(x, y) = \mathrm{dist}(y, x)$;
2. $\mathrm{dist}(x, y) \geq 0$, *and* $\mathrm{dist}(x, y) = 0$ *if and only* $x = y$; *and*
3. $\mathrm{dist}(x, z) \leq \mathrm{dist}(x, y) + \mathrm{dist}(y, z)$.

Equipped with a distance, $\mathcal{M}$ *is a* metric space.

A natural topology on a metric space is the *metric topology*, defined such that the functions $x \mapsto \mathrm{dist}(x, y)$ are continuous. Specifically, a subset $\mathcal{U} \subseteq \mathcal{M}$ is open if and only if, for every $x \in \mathcal{U}$, there exists a radius $r > 0$ such that the ball $\{y \in \mathcal{M} : \mathrm{dist}(x, y) < r\}$ is included in $\mathcal{U}$.

In this section, we first state without proof that if $\mathcal{M}$ is a Riemannian manifold then its Riemannian metric induces a distance on (the connected components of) $\mathcal{M}$, and that the topology of $\mathcal{M}$ as a manifold is equivalent to the topology of $\mathcal{M}$ as a metric space equipped with that distance. Intuitively, $\mathrm{dist}(x, y)$ is the length of the shortest "reasonable" curve on $\mathcal{M}$ joining x and y, or the infimum over the lengths of such curves. To make this precise, we first need to discuss various types of curves on manifolds [Lee18, pp33–34].

Definition 10.2. *A* curve segment *on a manifold* $\mathcal{M}$ *is a continuous map* $c\colon [a, b] \to \mathcal{M}$, *where* $a \leq b$ *are real. A curve segment* $c\colon [a, b] \to \mathcal{M}$ *is:*

- smooth *if* c *can be extended to a smooth map* $\tilde{c}\colon I \to \mathcal{M}$ *on a neighborhood* I *of* $[a, b]$, *in which case* $c'(a)$ *and* $c'(b)$ *denote* $\tilde{c}'(a)$ *and* $\tilde{c}'(b)$, *respectively;*
- regular *if it is smooth and* $c'(t) \neq 0$ *for all* $t \in [a, b]$;
- piecewise smooth *(resp.,* piecewise regular*) if there exists a finite set of times* $a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b$ *such that the restrictions* $c|_{[t_{i-1}, t_i]}$ *are smooth (resp., regular) curve segments for* $i = 1, \ldots, k$.

In particular, piecewise regular curves are piecewise smooth. We say a curve segment $c\colon [a, b] \to \mathcal{M}$ connects x *to* y *if* $c(a) = x$ *and* $c(b) = y$.

Let $\mathcal{M}$ be a Riemannian manifold. Given a piecewise smooth curve segment $c\colon [a, b] \to \mathcal{M}$, we define the *length* of c as the integral of its *speed* $\|c'(t)\|_{c(t)}$:

$$L(c) = \int_a^b \|c'(t)\|_{c(t)} \, \mathrm{d}t. \tag{10.1}$$

While $c'(t_i)$ may be undefined, the speed of each curve segment $c|_{[t_{i-1}, t_i]}$ is smooth: the integral is computed by summing over these intervals.

The notion of length of a curve leads to a natural notion of distance on $\mathcal{M}$, called the *Riemannian distance*:

$$\mathrm{dist}(x, y) = \inf_c L(c), \tag{10.2}$$

where the infimum is taken over all piecewise regular curve segments on $\mathcal{M}$ which connect x to y. It is equivalent to take the infimum over all piecewise smooth curve segments. We have the following important result [Lee18, Thm. 2.55].

Theorem 10.3. *If $\mathcal{M}$ is connected (meaning there exists a curve segment connecting every pair of points), equation* (10.2) *defines a distance. Equipped with this distance, $\mathcal{M}$ is a metric space whose metric topology coincides with its atlas topology.*

If $\mathcal{M}$ is not connected, we may consider this result on the connected components of $\mathcal{M}$ separately. Sometimes, it helps to extend the definition to accept $\mathrm{dist}(x, y) = \infty$ when x, y belong to distinct connected components.

If the infimum in (10.2) is attained[1] for some curve segment c, we call c a *minimizing curve*. Remarkably, up to parameterization, these are geodesics (Definition 5.38) [Lee18, Thm. 6.4]. In other words, two competing generalizations of the notion of straight line from linear spaces to manifolds turn out to be equivalent: one based on shortest paths, one based on zero acceleration.

Theorem 10.4. *Every minimizing curve admits a constant-speed parameterization such that it is a geodesic, called a* minimizing geodesic.

This theorem admits a partial converse [Lee18, Thm. 6.15]. It could not have a full converse since, for example, two nearby points on a sphere can be connected through both a short and a long geodesic.

Theorem 10.5. *Every geodesic γ on $\mathcal{M}$ is* locally minimizing, *that is, every t in the domain of γ has a neighborhood I in the domain of γ such that, if $a, b \in I$ satisfy $a < b$, then the restriction $\gamma|_{[a,b]}$ is a minimizing curve.*

Equipped with a distance, we define a first notion of *completeness*. Recall that a sequence $x_0, x_1, x_2, \ldots$ is *Cauchy* if for every $\varepsilon > 0$ there exists an integer k such that, for all $m, n > k$, $\mathrm{dist}(x_m, x_n) < \varepsilon$.

Definition 10.6. *A connected Riemannian manifold is* metrically complete *if it is complete as a metric space equipped with the Riemannian distance, that is, if every Cauchy sequence on the manifold converges on the manifold.*

There exists another useful notion of completeness for manifolds.

Definition 10.7. *A Riemannian manifold is* geodesically complete *if every geodesic can be extended to a geodesic defined on the whole real line.*

The following theorem is an important classical result: see [Lee18, Thm. 6.19, Pb. 6-14] for a proof. It justifies omitting to specify whether we mean metric or geodesic completeness. The last part of the statement is the *Heine–Borel property*: recall Definition 8.26 for compact sets; a set S is *bounded* if $\sup_{x,y \in S} \mathrm{dist}(x, y)$ is finite.

[1] This is not always the case: think of $\mathcal{M} = \mathbb{R}^2 \backslash \{0\}$ as a Riemannian submanifold of $\mathbb{R}^2$, and connect x and $-x$.

Theorem 10.8 (Hopf–Rinow). *A connected Riemannian manifold $\mathcal{M}$ is metrically complete if and only if it is geodesically complete. Additionally, $\mathcal{M}$ is complete (in either sense) if and only if its compact subsets are exactly its closed and bounded subsets.*

For disconnected manifolds, *complete* refers to geodesic completeness, which is equivalent to metric completeness of each connected component. For example, the orthogonal group $\mathrm{O}(n)$, which has two connected components, is complete in this sense.

On a complete manifold, two points in the same connected component can always be connected by a (not necessarily unique) geodesic which attains the infimum in (10.2) [Lee18, Cor. 6.21].

Theorem 10.9. *If $\mathcal{M}$ is complete, then any two points x, y in the same connected component are connected by a minimizing geodesic segment $c\colon [0,1] \to \mathcal{M}$ such that $c(0) = x$, $c(1) = y$ and $\mathrm{dist}(x, y) = L(c)$.*

The converse of Theorem 10.9 does not hold: consider $\mathcal{M} = (0, 1)$ as a Riemannian submanifold of $\mathbb{R}$.

Example 10.10. *Compact Riemannian manifolds are complete.*

Example 10.11. *A finite-dimensional Euclidean space $\mathcal{E}$ is connected and complete. The unique minimizing geodesic from x to y is the line segment $t \mapsto (1-t)x + ty$ on $[0,1]$, and the Riemannian distance $\mathrm{dist}(x, y)$ is equal to the Euclidean distance $\|x - y\|$.*

Exercise 10.12. *Let $\mathcal{M}$ be a Riemanian submanifold of $\mathcal{N}$. Show that if $\mathcal{N}$ is complete and if $\mathcal{M}$ is a closed subset of $\mathcal{N}$, then $\mathcal{M}$ is complete. In particular, Riemannian submanifolds of Euclidean spaces which are closed are complete.*

Exercise 10.13. *Show that the length of a piecewise smooth curve segment $c\colon [a, b] \to \mathcal{M}$ is independent of parameterization, in that $L(c \circ h) = L(c)$ for any monotone, piecewise regular $h\colon [0, 1] \to [a, b]$ such that $h(0) = a$ and $h(1) = b$. Further show that if c is a piecewise regular curve segment then h can be chosen such that $c \circ h$ is piecewise regular with speed (whenever it is defined) equal to the constant $L(c)$.*

Exercise 10.14. *Show that the distance on a Riemannian product manifold $\mathcal{M} = \mathcal{M}_1 \times \cdots \times \mathcal{M}_n$ is given by*

$$\mathrm{dist}(x, y) = \sqrt{\sum_{i=1}^{n} \mathrm{dist}(x_i, y_i)^2},$$

where dist *denotes Riemannian distance on $\mathcal{M}$ and $\mathcal{M}_1, \ldots, \mathcal{M}_n$ alike, as indicated by context. Hint: the equality can be established through a pair of matching inequalities. For one side, Jensen's inequality may be helpful. For the other side, it may be helpful to use the results of Exercise 10.13.*

Exercise 10.15. *Let $\mathcal{M} = \overline{\mathcal{M}}/\sim$ be a Riemannian quotient manifold with quotient map $\pi\colon \overline{\mathcal{M}} \to \mathcal{M}$, as in Section 9.7. Let $\mathrm{dist}_{\mathcal{M}}$ and $\mathrm{dist}_{\overline{\mathcal{M}}}$ denote the Riemannian distances on $\mathcal{M}$ and $\overline{\mathcal{M}}$, respectively. Show that for all $[x], [y] \in \mathcal{M}$ the distances satisfy*

$$\mathrm{dist}_{\mathcal{M}}([x], [y]) \leq \inf_{x' \sim x, y' \sim y} \mathrm{dist}_{\overline{\mathcal{M}}}(x', y').$$

Moreover, show that if $\overline{\mathcal{M}}$ is connected and complete then the inequality holds with equality. Hint: for the second part, use Proposition 9.61.

10.2 Exponential and logarithmic maps

Using standard tools from the study of ordinary differential equations, one can show that on a Riemannian manifold, for every $(x, v) \in \mathrm{T}\mathcal{M}$, there exists a unique *maximal* geodesic [Lee18, Cor. 4.28]

$$\gamma_v\colon I \to \mathcal{M}, \qquad \text{with} \qquad \gamma_v(0) = x \qquad \text{and} \qquad \gamma_v'(0) = v.$$

Here, *maximal* refers to the fact that the interval I is as large as possible (this is *not* in contrast to the notion of minimizing geodesic we just defined in the previous section). We use these geodesics to define a special map.

Definition 10.16. *Consider the following subset of the tangent bundle:*

$$\mathcal{O} = \{(x, v) \in \mathrm{T}\mathcal{M} : \gamma_v \textit{ is defined on an interval containing } [0, 1]\}.$$

The exponential map $\mathrm{Exp}\colon \mathcal{O} \to \mathcal{M}$ *is defined by*

$$\mathrm{Exp}(x, v) = \mathrm{Exp}_x(v) = \gamma_v(1).$$

The restriction Exp_x *is defined on* $\mathcal{O}_x = \{v \in \mathrm{T}_x\mathcal{M} : (x, v) \in \mathcal{O}\}$.

For example, in a Euclidean space, $\mathrm{Exp}_x(v) = x + v$. By Definition 10.7, a manifold $\mathcal{M}$ is (geodesically) complete exactly if the domain of the exponential map is the whole tangent bundle $\mathrm{T}\mathcal{M}$.

Given $t \in \mathbb{R}$, it holds that $\gamma_{tv}(1) = \gamma_v(t)$ whenever either is defined. This allows us to express the exponential map as

$$\mathrm{Exp}_x(tv) = \gamma_v(t), \tag{10.3}$$

which is often more convenient. In particular, the domain of Exp_x is *star-shaped* around the origin in $\mathrm{T}_x\mathcal{M}$, that is,

$$v \in \mathcal{O}_x \implies tv \in \mathcal{O}_x \text{ for all } t \in [0, 1].$$

Conveniently, Exp is smooth [Lee18, Lem. 5.18, Prop. 5.19].

Proposition 10.17. *The exponential map is smooth on its domain $\mathcal{O}$, which is open in $\mathrm{T}\mathcal{M}$.*

The domain $\mathcal{O}$ contains all tangent space origins. We say that $\mathcal{O}$ is a neighborhood of the *zero section of the tangent bundle*:

$$\{(x, 0) \in \mathrm{T}\mathcal{M} : x \in \mathcal{M}\} \subset \mathcal{O}. \tag{10.4}$$

The exponential map is a retraction on its domain. More precisely:

Proposition 10.18. *The exponential map is a second-order retraction, with a possibly restricted domain $\mathcal{O} \subseteq \mathrm{T}\mathcal{M}$.*

Proof. Proposition 10.17 claims smoothness of Exp. By definition, for all (x, v) in $\mathrm{T}\mathcal{M}$ the curve $c(t) = \mathrm{Exp}_x(tv) = \gamma_v(t)$ satisfies $c(0) = x$ and $c'(0) = v$ so that Exp is a retraction. Finally, it is clear that this retraction is second order (Definition 5.42) since $\gamma_v''(t)$ is zero for all t, hence in particular $c''(0) = 0$. □

Given a point x and a (sufficiently short) tangent vector v, the exponential map produces a new point $y = \mathrm{Exp}_x(v)$. One may reasonably wonder whether, given the two points x, y, one can recover the tangent vector v. In what follows, we aim to understand to what extent the exponential map can be (smoothly) inverted.

A first observation, rooted in the inverse function theorem for manifolds (see Theorem 4.16), is that any retraction R at a point x is locally a diffeomorphism around the origin in the tangent space at x, because R_x is smooth and $\mathrm{DR}_x(0)$ is the identity (hence a fortiori invertible). This applies in particular to the exponential map. For the latter, the *injectivity radius* quantifies how large the local domains can be.

⋆ When we say a map F is a diffeomorphism on an open domain U, we mean that $F(U)$ is open and F is a diffeomorphism from U to $F(U)$.

Definition 10.19. *The* injectivity radius *of a Riemannian manifold $\mathcal{M}$ at a point x, denoted by* $\mathrm{inj}(x)$, *is the supremum over radii $r > 0$ such that* Exp_x *is defined and is a diffeomorphism on the open ball*

$$B(x, r) = \{v \in \mathrm{T}_x\mathcal{M} : \|v\|_x < r\}.$$

By the inverse function theorem, $\mathrm{inj}(x) > 0$.

Consider the ball $U = B(x, \mathrm{inj}(x))$ in the tangent space at x. Its image $\mathcal{U} = \mathrm{Exp}_x(U)$ is a neighborhood of x in $\mathcal{M}$. By definition, $\mathrm{Exp}_x : U \to \mathcal{U}$ is a diffeomorphism: it has a well-defined smooth inverse $\mathrm{Exp}_x^{-1} : \mathcal{U} \to U$. With these choices of domains, $v = \mathrm{Exp}_x^{-1}(y)$ is the unique shortest tangent vector at x such that $\mathrm{Exp}_x(v) = y$. Indeed, if there existed another vector $u \in \mathrm{T}_x\mathcal{M}$ such that $\mathrm{Exp}_x(u) = y$ and $\|u\|_x \leq \|v\|_x$, then u would be included in U, which would contradict invertibility. This motivates the following definition.[2]

[2] Note that Log_x may be discontinuous (think of x and y close to each other on a sphere from which we remove the midpoint $(x + y)/\|x + y\|$). Nevertheless, see below for smoothness properties on restricted domains (Corollary 10.25 in particular), and see also [Lee18, §10] for a related discussion of *cut locus* and *conjugate points*.

Definition 10.20. *For $x \in \mathcal{M}$, let* Log_x *denote the* logarithmic map at x,

$$\mathrm{Log}_x(y) = \arg\min_{v \in \mathcal{O}_x} \|v\|_x \textit{ subject to } \mathrm{Exp}_x(v) = y, \tag{10.5}$$

with domain such that this is uniquely defined.

For example, in a Euclidean space, $\mathrm{Log}_x(y) = y - x$ for all x, y. In particular, with domains $U = B(x, \mathrm{inj}(x))$ and $\mathcal{U} = \mathrm{Exp}_x(U)$ as above, the inverse of $\mathrm{Exp}_x : U \to \mathcal{U}$ is the (possibly restricted) map $\mathrm{Log}_x : \mathcal{U} \to U$, which is then a diffeomorphism. With different domain restrictions however, the inverse of Exp_x may be different from Log_x. This is illustrated in the following example.

Example 10.21. *On the sphere S^{n-1}, the exponential map is (Example 5.37)*

$$v \mapsto \mathrm{Exp}_x(v) = \cos(\|v\|)x + \frac{\sin(\|v\|)}{\|v\|}v.$$

This is smooth over the whole tangent bundle, with the usual smooth extension $\sin(t)/t = 1$ at $t = 0$. Given $x, y \in \mathrm{S}^{n-1}$, we seek an expression for $\mathrm{Exp}_x^{-1}(y)$. Since $x^\top x = 1$ and $x^\top v = 0$, considering $y = \mathrm{Exp}_x(v)$ as above, we deduce that $x^\top y = \cos(\|v\|)$. Thus, the following vector (which is readily computed given x and y) is parallel to v:

$$u \triangleq y - (x^\top y)x = \mathrm{Proj}_x(y) = \frac{\sin(\|v\|)}{\|v\|}v.$$

It has norm $|\sin(\|v\|)|$ and is parallel to v. Let us exclude the case $u = 0$ which is easily treated separately. Then, dividing u by its norm yields:

$$\frac{u}{\|u\|} = \mathrm{sign}(\sin(\|v\|))\frac{v}{\|v\|}.$$

If we restrict the domain of Exp_x to contain exactly those tangent vectors v whose norm is strictly less than π, then $\mathrm{sign}(\sin(\|v\|)) = 1$. Furthermore, the equation $x^\top y = \cos(\|v\|)$ then admits the unique solution $\|v\| = \arccos(x^\top y)$, where $\arccos : [-1, 1] \to [0, \pi]$ is the principal inverse of $\cos$. Overall, this yields the following expression:

$$y \mapsto \mathrm{Exp}_x^{-1}(y) = \arccos(x^\top y)\frac{u}{\|u\|}, \tag{10.6}$$

smooth over $\mathrm{S}^{n-1} \backslash \{-x\}$. Since the chosen domain for Exp_x is $B(x, \pi)$, the inverse is the logarithm: $\mathrm{Exp}_x^{-1} = \mathrm{Log}_x$. Also, $\mathrm{dist}(x, y) = \arccos(x^\top y)$.

The crucial point is that, in deriving this expression, we made the (somewhat arbitrary) choice of defining the domain of Exp_x in a specific way. This leads to a particular formula for the inverse, and a particular domain for Exp_x^{-1}. If we choose the domain of Exp_x differently, we may very well obtain a different formula for Exp_x^{-1} (not equal to Log_x) and a different domain for it as well. For example, on the circle S^1, we could decide that if y is ahead of x (counterclockwise) by an angle less than $\pi/2$, then $\mathrm{Exp}_x^{-1}(y)$ returns a vector of length *less than $\pi/2$, and otherwise it returns a vector of length less than $3\pi/2$, pointing in the clockwise direction.*

Before moving on to broader smoothness concerns, we quote useful relations between Exp, Log and dist [Lee18, Prop. 6.11].

Proposition 10.22. *If $\|v\|_x < \mathrm{inj}(x)$, the geodesic $c(t) = \mathrm{Exp}_x(tv)$ on the interval $[0, 1]$ is the minimizing curve connecting x to $y = \mathrm{Exp}_x(v)$, unique up to parameterization. In particular, $\mathrm{dist}(x, y) = \|v\|_x$, and $\mathrm{Log}_x(y) = v$.*

So far, we have fixed the point x, allowing us to claim that, on some domains, the map $y \mapsto \mathrm{Exp}_x^{-1}(y)$ is smooth in y. In order to discuss smoothness jointly in x and y, we need more work. We start with a general discussion valid for all retractions, and specialize to the exponential map later on.

Proposition 10.23. *Let $\mathcal{M}$ be a manifold with retraction R defined on a neighborhood $\mathcal{O}$ of the zero section of $\mathrm{T}\mathcal{M}$. Consider the following map:*

$$E : \mathcal{O} \to \mathcal{M} \times \mathcal{M} : (x, v) \mapsto E(x, v) = (x, \mathrm{R}_x(v)). \tag{10.7}$$

If $\mathcal{T} \subseteq \mathcal{O}$ is open in $\mathrm{T}\mathcal{M}$ such that, for all x, R_x is a diffeomorphism on $\mathcal{T}_x = \{v \in \mathrm{T}_x\mathcal{M} : (x, v) \in \mathcal{T}\}$, then $\mathcal{V} = E(\mathcal{T})$ is open in $\mathcal{M} \times \mathcal{M}$ and $E : \mathcal{T} \to \mathcal{V}$ is a diffeomorphism.

Proof. First, to see that $E : \mathcal{T} \to \mathcal{V}$ is invertible, consider any pairs $(x, v), (y, w) \in \mathcal{T}$ such that $E(x, v) = E(y, w)$. In other words, we have $(x, \mathrm{R}_x(v)) = (y, \mathrm{R}_y(w))$, so that $x = y$ and $v, w \in \mathcal{T}_x$. By assumption, R_x is injective on $\mathcal{T}_x$, hence we deduce from $\mathrm{R}_x(v) = \mathrm{R}_x(w)$ that $v = w$.

Second, to show that $\mathcal{V}$ is open and $E : \mathcal{T} \to \mathcal{V}$ is a diffeomorphism, it remains to check that the differential of E is invertible everywhere in $\mathcal{T}$ (the result then follows from applying the inverse function theorem at each point of $\mathcal{T}$, see Theorem 4.16). To this end, consider any $(x, v) \in \mathcal{T}$. Somewhat informally, the differential $\mathrm{D}E(x, v)$ is a block matrix of size two-by-two as follows. (To be formal, we should give a more precise description of the tangent space to $\mathrm{T}\mathcal{M}$ at (x, v). Here, it is identified with $\mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M}$.)

$$\mathrm{D}E(x, v) \simeq \begin{bmatrix} I & 0 \\ * & \mathrm{DR}_x(v) \end{bmatrix}.$$

Indeed, the differential of the first entry of $E(x, v) = (x, \mathrm{R}_x(v))$ with respect to x is the identity, and it is zero with respect to v. The second entry has some unspecified differential with respect to x, while its differential with respect to v is $\mathrm{DR}_x(v)$. Crucially, since v is in $\mathcal{T}_x$, we know by assumption that R_x is a diffeomorphism around v, hence $\mathrm{DR}_x(v)$ is invertible. We conclude that $\mathrm{D}E(x, v)$ is invertible for all $(x, v) \in \mathcal{T}$, as announced. $\square$

In particular, under the stated conditions, $(x, y) \mapsto (x, \mathrm{R}_x^{-1}(y))$ is a diffeomorphism from $\mathcal{V}$ to $\mathcal{T}$, meaning the inverse retraction can be defined smoothly jointly in x and y (with care when it comes to domains).

In this last proposition, the fact that $\mathcal{T}$ is open is crucial: this is what ties the domains $\mathcal{T}_x$ together. Without this assumption, we can still have an inverse, but not necessarily a smooth inverse. Furthermore, it is natural to want to include the tangent space origins in $\mathcal{T}$, that is, to make $\mathcal{T}$ a neighborhood of the zero section in $\mathrm{T}\mathcal{M}$. It is convenient to make this happen using a continuous function $\Delta\colon \mathcal{M} \to (0, \infty]$:

$$\mathcal{T} = \big\{(x, v) \in \mathrm{T}\mathcal{M} : \|v\|_x < \Delta(x)\big\}. \tag{10.8}$$

If R_x is defined and is a diffeomorphism on the open ball $B(x, \Delta(x))$ in $\mathrm{T}_x\mathcal{M}$ for all x, then Proposition 10.23 applies, and $\mathcal{V}$ contains the diagonal $\{(x, x) : x \in \mathcal{M}\}$.

Conveniently, for the exponential map, we can take Δ to be as large as one could possibly hope, namely: we can choose Δ to be the injectivity radius function. (This holds even if $\mathcal{M}$ is not connected or complete, see Section 10.8.)

Proposition 10.24. *On a Riemannian manifold $\mathcal{M}$, the injectivity radius function* $\mathrm{inj}\colon \mathcal{M} \to (0, \infty]$ *is continuous.*

Corollary 10.25. *The map $(x, v) \mapsto (x, \mathrm{Exp}_x(v))$ is a diffeomorphism from*

$$\mathcal{T} = \big\{(x, v) \in \mathrm{T}\mathcal{M} : \|v\|_x < \mathrm{inj}(x)\big\}$$

to

$$\mathcal{V} = \big\{(x, y) \in \mathcal{M} \times \mathcal{M} : \mathrm{dist}(x, y) < \mathrm{inj}(x)\big\}.$$

Its inverse is $(x, y) \mapsto (x, \mathrm{Log}_x(y))$, smooth from $\mathcal{V}$ to $\mathcal{T}$.

Under this corollary, we see that $(x, y) \mapsto \mathrm{Log}_x(y)$ is smooth jointly in x and y over some domain. This is apparent in Example 10.21 for the sphere, where $\mathrm{inj}(x) = \pi$ for all x.

More generally, we show that for any retraction there exists a positive and continuous function Δ which can be used to argue existence of a smooth inverse.

Proposition 10.26. *On a Riemannian manifold $\mathcal{M}$, consider the following open subsets of the tangent bundle $\mathrm{T}\mathcal{M}$:*

$$V_\delta(x) = \{(x', v') \in \mathrm{T}\mathcal{M} : \mathrm{dist}(x, x') < \delta \text{ and } \|v'\|_{x'} < \delta\}.$$

(In particular, x and x' must be in the same connected component.) Notice that $(x, 0)$ is in $V_\delta(x)$ for all $\delta > 0$. For any retraction R on $\mathcal{M}$ defined on a neighborhood $\mathcal{O}$ of the zero section in $\mathrm{T}\mathcal{M}$, define $\Delta\colon \mathcal{M} \to (0, \infty]$ by:

$$\Delta(x) = \sup\big\{\delta > 0 : V_\delta(x) \subseteq \mathcal{O} \text{ and } E \text{ is a diffeomorphism on } V_\delta(x)\big\},$$

where E is as defined in (10.7). Then, Δ is positive and continuous, and R_x is defined and is a diffeomorphism on $B(x, \Delta(x))$ for all x.

Proof. To see that $\Delta(x)$ is positive at every x, apply the inverse function theorem to the fact that the differential of E at $(x, 0)$ is invertible, since it is of the form $\mathrm{D}E(x, 0) \simeq \left[\begin{smallmatrix} I & 0 \\ * & I \end{smallmatrix}\right]$ (same as in the proof of Proposition 10.23).

It is sufficient to reason on each connected component of $\mathcal{M}$ separately, hence we may assume $\mathcal{M}$ is connected. If $\Delta(x) = \infty$ for some x, then $\mathcal{O} = \mathrm{T}\mathcal{M}$ and E is a diffeomorphism on that domain, so that $\Delta(x) = \infty$ for all x: this is compatible with the claim. We can now assume $\Delta(x)$ is finite for all x.

To see that Δ is continuous, we show that $\Delta(x) - \Delta(x') \leq \mathrm{dist}(x, x')$ for every two points $x, x' \in \mathcal{M}$. Then, switching the roles of x and x', we find $|\Delta(x) - \Delta(x')| \leq \mathrm{dist}(x, x')$, which shows Δ is continuous with respect to the Riemannian distance. This is equivalent to continuity with respect to the atlas topology by Theorem 10.3.

Pick any two points $x, x' \in \mathcal{M}$. If $\mathrm{dist}(x, x') \geq \Delta(x)$, the claim is clear. So assume $\mathrm{dist}(x, x') < \Delta(x)$, and define $\delta = \Delta(x) - \mathrm{dist}(x, x')$. We claim that

$$V_\delta(x') \subset V_{\Delta(x)}(x).$$

Indeed, pick any $(x'', v'') \in V_\delta(x')$. Then,

1. $\|v''\|_{x''} < \delta \leq \Delta(x)$, and
2. $\mathrm{dist}(x'', x) \leq \mathrm{dist}(x'', x') + \mathrm{dist}(x', x) < \delta + \mathrm{dist}(x', x) = \Delta(x)$.

We know that E is a diffeomorphism on $V_{\Delta(x)}(x)$. Thus, a fortiori, E is a diffeomorphism on $V_\delta(x')$. By definition of $\Delta(x')$, this implies $\Delta(x') \geq \delta = \Delta(x) - \mathrm{dist}(x, x')$, which is what we needed to show.

The conclusion about R_x follows from the fact that E is a diffeomorphism on each $V_{\Delta(x)}(x)$ (which covers $B(x, \Delta(x))$) and from the form of $\mathrm{D}E$, as in the proof of Proposition 10.23. □

For general retractions, we obtain the following corollary which notably means that, over some domain of $\mathcal{M} \times \mathcal{M}$ which contains all pairs (x, x), the map $(x, y) \mapsto \mathrm{R}_x^{-1}(y)$ can be defined smoothly jointly in x and y. There is no need to require that $\mathcal{M}$ be a Riemannian manifold because the existence of a Riemannian metric is guaranteed [Lee12, Prop. 13.3]: that is sufficient to apply Proposition 10.26.

Corollary 10.27. *For any retraction* R *on a manifold* $\mathcal{M}$ *there exists a neighborhood* $\mathcal{T}$ *of the zero section of the tangent bundle* $\mathrm{T}\mathcal{M}$ *on which*

$$(x, v) \mapsto (x, \mathrm{R}_x(v))$$

is a diffeomorphism; $\mathcal{T}$ *can be taken of the form* (10.8) *(with respect to an arbitrary Riemannian metric) with* $\Delta \colon \mathcal{M} \to (0, \infty]$ *continuous.*

We close with the notion of injectivity radius of a whole manifold. It may be zero, positive or infinite. The set of manifolds with positive injectivity radius is strictly included in the set of complete manifolds: Exercise 10.30. See [Lee18, Lem. 6.16] for a proof of Proposition 10.29.

Definition 10.28. *The injectivity radius* $\mathrm{inj}(\mathcal{M})$ *of a Riemannian manifold* $\mathcal{M}$ *is the infimum of* $\mathrm{inj}(x)$ *over* $x \in \mathcal{M}$.

Proposition 10.29. *For a compact Riemannian manifold,* $\mathrm{inj}(\mathcal{M}) \in (0, \infty)$.

A Euclidean space has infinite injectivity radius. The unit sphere S^{n-1} has injectivity radius π. Importantly, the manifold $\mathbb{R}_r^{m\times n}$ of matrices with fixed rank r embedded in $\mathbb{R}^{m\times n}$ (Section 7.5) has *zero* injectivity radius. This is in part due to the fact that there exist matrices in $\mathbb{R}_r^{m\times n}$ that are arbitrarily close to matrices of rank strictly less than r, as measured in the embedding space ($\mathbb{R}_r^{m\times n}$ is not complete).

Exercise 10.30. *Show that if* $\mathrm{inj}(\mathcal{M})$ *is positive then* $\mathcal{M}$ *is complete. The converse is not true: give an example of a complete, connected manifold whose injectivity radius is zero.*

Exercise 10.31. *Let* $\mathcal{K}$ *be any subset of a Riemannian manifold* $\mathcal{M}$ *and let* $r \colon \mathcal{K} \to \mathbb{R}^+$ *be continuous, with* $\mathbb{R}^+ = \{t \in \mathbb{R} : t \geq 0\}$. *Show that*

$$\mathcal{T} = \{(x, s) \in \mathrm{T}\mathcal{M} : x \in \mathcal{K} \text{ and } \|s\|_x \leq r(x)\}$$

is compact in $\mathrm{T}\mathcal{M}$ *if and only if* $\mathcal{K}$ *is compact in* $\mathcal{M}$.

Exercise 10.32. *Let* $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$ *be a Riemannian product manifold, as in Example 3.57. Let* $\mathcal{O}_1, \mathcal{O}_2$ *be the domains of the exponential maps on* $\mathcal{M}_1, \mathcal{M}_2$, *respectively. Show that the domain of the exponential map on* $\mathcal{M}$ *is* $\mathcal{O} = \mathcal{O}_1 \times \mathcal{O}_2$, *with the identification of tangent bundles on product manifolds as in* (3.31). *For* $(x_1, v_1) \in \mathcal{O}_1$ *and* $(x_2, v_2) \in \mathcal{O}_2$, *show that*

$$\mathrm{Exp}_x(v) = \left(\mathrm{Exp}_{x_1}(v_1), \mathrm{Exp}_{x_2}(v_2)\right),$$

where $x = (x_1, x_2)$ *and* $v = (v_1, v_2)$, *and* Exp *denotes the exponential map on each respective manifold. Hint: use Exercise 5.39.*

10.3 Parallel transport

Consider a manifold $\mathcal{M}$ equipped with a connection ∇, and a tangent vector $u \in \mathrm{T}_x\mathcal{M}$.[3] In several situations, it is desirable to somehow transport u from x to another point $y \in \mathcal{M}$. In so doing, we would like for u and its transported version to be related in some meaningful way.

The geometric tool of choice for this task is called *parallel transport* (or *parallel translation*). Let $c \colon I \to \mathcal{M}$ be a smooth curve such that

$$c(0) = x \qquad \text{and} \qquad c(1) = y.$$

Consider a smooth vector field $Z \in \mathfrak{X}(c)$ on this curve with $Z(0) = u$. If Z does

[3] In this section, $\mathcal{M}$ may or may not be Riemannian, and ∇ may or may not be the Riemannian connection.

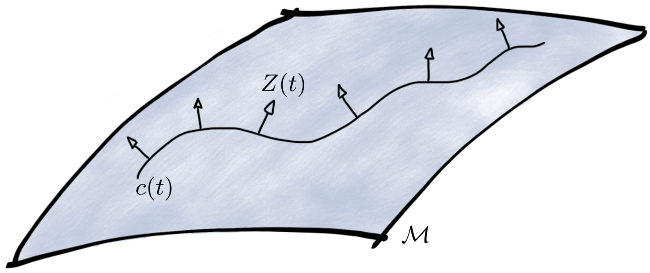


Figure 10.1 Parallel transports 'move' tangent vectors from one tangent space to another, along a specified curve. They can be used to compare or combine tangent vectors at different points by transporting them to a common tangent space.

not 'vary' too much, it is tempting to consider $Z(1)$ as a transport of u to y. One convenient way to formalize this is to require that Z be *parallel* with respect to the chosen connection ∇. Explicitly, using the covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$ induced by ∇ (Theorem 5.29), we require

$$Z \in \mathfrak{X}(c), \qquad Z(0) = u, \qquad \text{and} \qquad \frac{\mathrm{D}}{\mathrm{d}t} Z = 0. \tag{10.9}$$

Using standard tools from linear ordinary differential equations, one can show that such a vector field exists and is unique [Lee18, Thm. 4.32].

Definition 10.33. *A vector field $Z \in \mathfrak{X}(c)$ such that $\frac{\mathrm{D}}{\mathrm{d}t} Z = 0$ is* parallel.

Theorem 10.34. *On a manifold $\mathcal{M}$ with a connection and induced covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$, for any smooth curve $c\colon I \to \mathcal{M}$, $t_0 \in I$ and $u \in \mathrm{T}_{c(t_0)}\mathcal{M}$, there exists a unique parallel vector field $Z \in \mathfrak{X}(c)$ such that $Z(t_0) = u$.*

This justifies the following definition of parallel transport.

Definition 10.35. *Given a smooth curve c on $\mathcal{M}$, the* parallel transport *of tangent vectors at $c(t_0)$ to the tangent space at $c(t_1)$ along c is the map*

$$\mathrm{PT}^c_{t_1 \leftarrow t_0} \colon \mathrm{T}_{c(t_0)}\mathcal{M} \to \mathrm{T}_{c(t_1)}\mathcal{M}$$

defined by $\mathrm{PT}^c_{t_1 \leftarrow t_0}(u) = Z(t_1)$, where $Z \in \mathfrak{X}(c)$ is the unique parallel vector field such that $Z(t_0) = u$.

In particular, $t \mapsto \mathrm{PT}^c_{t \leftarrow t_0}(u)$ is a parallel vector field along c. Also, if Z is parallel along c, then $Z(t_1) = \mathrm{PT}^c_{t_1 \leftarrow t_0}(Z(t_0))$ for all t_0, t_1 in the domain of c.

On occasion, we may write $\mathrm{PT}^c_{y \leftarrow x}$ or even $\mathrm{PT}_{y \leftarrow x}$ when the times t_0, t_1 and the curve c such that $x = c(t_0)$ and $y = c(t_1)$ are clear from context, but beware: even if we use the Riemannian connection,

Parallel transport from x to y depends on the curve connecting x and y.

Indeed, think of a tangent vector at the equator pointing North. Transport it to the North pole via the shortest path. Alternatively, transport the same vector by first moving along the equator for some distance before going to the North pole: the results are different. This is, in fact, a crucial feature of Riemannian geometry, intimately related to the notion of curvature [Lee18, Ch. 7].

On a Riemannian manifold, when the curve c is not specified, one often implicitly means to move along the minimizing geodesic connecting x and y, assuming it exists and is unique.

Proposition 10.36. *The parallel transport operator* $\mathrm{PT}^c_{t_1 \leftarrow t_0}$ *is linear. Also,* $\mathrm{PT}^c_{t_2 \leftarrow t_1} \circ \mathrm{PT}^c_{t_1 \leftarrow t_0} = \mathrm{PT}^c_{t_2 \leftarrow t_0}$ *and* $\mathrm{PT}^c_{t \leftarrow t}$ *is the identity. In particular, the inverse of* $\mathrm{PT}^c_{t_1 \leftarrow t_0}$ *is* $\mathrm{PT}^c_{t_0 \leftarrow t_1}$. *If* $\mathcal{M}$ *is Riemannian and* ∇ *is compatible with the Riemannian metric, then*[4] *parallel transport is an isometry, that is,*

$$\forall u, v \in \mathrm{T}_{c(t_0)}\mathcal{M}, \qquad \langle u, v \rangle_{c(t_0)} = \left\langle \mathrm{PT}^c_{t_1 \leftarrow t_0}(u), \mathrm{PT}^c_{t_1 \leftarrow t_0}(v) \right\rangle_{c(t_1)}.$$

Stated differently, the adjoint and the inverse of $\mathrm{PT}^c_{t_1 \leftarrow t_0}$ *coincide.*

Proof. For linearity, consider $u, v \in \mathrm{T}_{c(t_0)}\mathcal{M}$ and $a, b \in \mathbb{R}$, arbitrary. By Theorem 10.34, there exist unique parallel vector fields $Z_u, Z_v \in \mathfrak{X}(c)$ such that $Z_u(t_0) = u$ and $Z_v(t_0) = v$. Since $Z = aZ_u + bZ_v \in \mathfrak{X}(c)$ is also parallel and $Z(t_0) = au + bv$, we conclude that Z is the unique parallel vector field used in the definition of

$$\begin{aligned} \mathrm{PT}^c_{t_1 \leftarrow t_0}(au + bv) = Z(t_1) &= aZ_u(t_1) + bZ_v(t_1) \\ &= a\mathrm{PT}^c_{t_1 \leftarrow t_0}(u) + b\mathrm{PT}^c_{t_1 \leftarrow t_0}(v), \end{aligned}$$

which shows linearity. The composition rule is clear, as is the fact that $\mathrm{PT}^c_{t \leftarrow t}$ is the identity. Then, the inverse follows by setting $t_2 = t_0$ in the composition rule.

To verify isometry, notice that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle Z_u(t), Z_v(t)\rangle_{c(t)} = \left\langle \frac{\mathrm{D}}{\mathrm{d}t} Z_u(t), Z_v(t) \right\rangle_{c(t)} + \left\langle Z_u(t), \frac{\mathrm{D}}{\mathrm{d}t} Z_v(t) \right\rangle_{c(t)} = 0,$$

using compatibility of the covariant derivative with the Riemannian metric and the fact that Z_u, Z_v are parallel. Thus, the inner product is constant along c. $\square$

One convenient tool afforded to us by parallel transports is the notion of *parallel frames* along a curve c. Consider an arbitrary basis $e_1, \ldots, e_d$ for the tangent space at $c(\bar{t})$, for an arbitrary $\bar{t}$ in the domain of definition of c. Construct the parallel vector fields

$$E_i(t) = \mathrm{PT}^c_{t \leftarrow \bar{t}}(e_i), \qquad i = 1, \ldots, d. \tag{10.10}$$

Since parallel transports are invertible, for all t, the vectors $E_i(t)$ form a basis for the tangent space at $c(t)$. (Also, if the manifold is Riemannian and we transport via a connection that is compatible with the metric, then orthonormality would

[4] The converse also holds: if parallel transport is an isometry, then ∇ is compatible with the metric [Lee18, Prop. 5.5].

be preserved.) As a result, for any $Y \in \mathfrak{X}(c)$ there exist unique, real functions α_i such that

$$Y(t) = \sum_{i=1}^{d} \alpha_i(t) E_i(t). \tag{10.11}$$

These functions are smooth since Y is smooth. Owing to linearity,

$$\mathrm{PT}^c_{t_1 \leftarrow t_0}(Y(t_0)) = \sum_{i=1}^{d} \alpha_i(t_0) E_i(t_1).$$

In particular, we see that this is smooth in both t_0 and t_1.

Using parallel frames, we can show that covariant derivatives admit a convenient expression in terms of parallel transports: transport the vector field to a common tangent space, differentiate in the usual way (in that fixed tangent space), then transport back.

Proposition 10.37. *Consider a smooth curve $c\colon I \to \mathcal{M}$. Given a vector field $Z \in \mathfrak{X}(c)$ and $t_0 \in I$, let $z\colon I \to \mathrm{T}_{c(t_0)}\mathcal{M}$ be $z(t) = \mathrm{PT}^c_{t_0 \leftarrow t} Z(t)$. Then,*

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \mathrm{PT}^c_{t \leftarrow t_0}\left(\frac{\mathrm{d}}{\mathrm{d}t} z(t)\right) = \lim_{\delta \to 0} \frac{\mathrm{PT}^c_{t \leftarrow t+\delta} Z(t+\delta) - Z(t)}{\delta}.$$

Proof. Transport any basis $e_1, \ldots, e_d$ of $\mathrm{T}_{c(t_0)}\mathcal{M}$ along c to form a frame $E_i(t) = \mathrm{PT}^c_{t \leftarrow t_0} e_i$. There exist unique, smooth, real functions $\alpha_1, \ldots, \alpha_d$ such that $Z(t) = \sum_{i=1}^{d} \alpha_i(t) E_i(t)$. Then, by the properties of covariant derivatives (Theorem 5.29) and $\frac{\mathrm{D}}{\mathrm{d}t} E_i = 0$,

$$\begin{aligned}
\frac{\mathrm{D}}{\mathrm{d}t} Z(t) &= \sum_{i=1}^{d} \alpha_i'(t) E_i(t) \\
&= \mathrm{PT}^c_{t \leftarrow t_0} \sum_{i=1}^{d} \alpha_i'(t) e_i \\
&= \mathrm{PT}^c_{t \leftarrow t_0} \frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{d} \alpha_i(t) e_i = \mathrm{PT}^c_{t \leftarrow t_0} \frac{\mathrm{d}}{\mathrm{d}t} z(t),
\end{aligned}$$

as announced. The important point is that z is a map between (fixed) linear spaces, hence why we can take a classical derivative $\frac{\mathrm{d}}{\mathrm{d}t}$. $\square$

Exercise 10.38. *On the sphere S^{n-1} as a Riemannian submanifold of $\mathbb{R}^n$ with the usual metric and connection, parallel transport along the geodesic $c(t) = \mathrm{Exp}_x(tv)$ admits the following explicit expression [QGA10a]:*

$$\mathrm{PT}^c_{t \leftarrow 0}(u) = \left(I_n + (\cos(t\|v\|) - 1)\frac{vv^\top}{\|v\|^2} - \sin(t\|v\|)\frac{xv^\top}{\|v\|}\right) u.$$

Verify this claim. (Recall Example 5.37 for the exponential.)

Exercise 10.39. *Let $c = (c_1, c_2)$ be a smooth curve on a Riemannian product manifold $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$. Verify that parallel transport along c is given in terms of parallel transports along c_1 and c_2 as*

$$\mathrm{PT}^{c}_{t_b \leftarrow t_a}(v) = \left(\mathrm{PT}^{c_1}_{t_b \leftarrow t_a}(v_1), \mathrm{PT}^{c_2}_{t_b \leftarrow t_a}(v_2)\right),$$

where $v = (v_1, v_2)$ is tangent to $\mathcal{M}$ at $c(t_a)$. Hint: recall Exercise 5.34.

10.4 Lipschitz conditions and Taylor expansions

One of the most convenient regularity assumptions one can make regarding the cost function f of an optimization problem is that it or its derivatives be Lipschitz continuous. Indeed, in the Euclidean case, it is well known that such properties lead to global bounds on the discrepancy between f and its Taylor expansions of various orders. These, in turn, ease worst-case iteration complexity analyses. Here, we consider definitions of Lipschitz continuity on Riemannian manifolds, and we derive Taylor bounds analogous to their Euclidean counterparts. In so doing, we are careful not to require the manifold to be complete.

In this section, ⋆

1. We implicitly assume all of our manifolds are Riemannian,
2. We usually omit subscripts for inner products and norms (writing $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ instead of $\langle \cdot, \cdot \rangle_x$ and $\|\cdot\|_x$), and
3. We state explicitly how many times we need maps to be (continuously) differentiable, thus not assuming smoothness (infinite differentiability) by default: recall Remark 8.6.

Let A, B be two metric spaces. A map $F \colon A \to B$ is *L-Lipschitz continuous* if $L \geq 0$ is such that

$$\forall x, y \in A, \qquad \mathrm{dist}_B(F(x), F(y)) \leq L \, \mathrm{dist}_A(x, y), \tag{10.12}$$

where dist_A, dist_B denote the distances on A and B. In particular:

Definition 10.40. *A function $f \colon \mathcal{M} \to \mathbb{R}$ on a connected manifold $\mathcal{M}$ is L-*Lipschitz continuous *if*

$$\forall x, y \in \mathcal{M}, \qquad |f(x) - f(y)| \leq L \, \mathrm{dist}(x, y), \tag{10.13}$$

where dist *is the Riemannian distance on $\mathcal{M}$. If $\mathcal{M}$ is disconnected, we require the condition to hold on each connected component separately.*

The definition above can be reformulated as we show below. This second formulation is more convenient to study iterates of optimization algorithms presented as $x_{k+1} = \mathrm{Exp}_{x_k}(s_k)$ for some $s_k \in \mathrm{T}_{x_k}\mathcal{M}$.

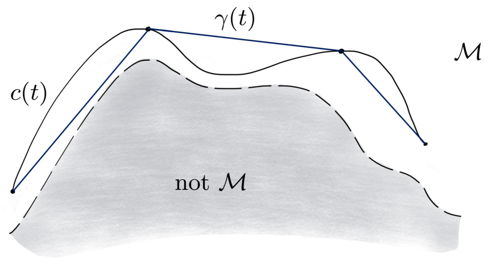


Figure 10.2 In Lemma 10.42, the curve γ is made of a finite number of minimizing geodesic segments, with endpoints on c.

Proposition 10.41. *A function $f\colon \mathcal{M} \to \mathbb{R}$ is L-Lipschitz continuous if and only if*

$$\forall (x, s) \in \mathcal{O}, \qquad |f(\mathrm{Exp}_x(s)) - f(x)| \leq L\|s\|, \tag{10.14}$$

where $\mathcal{O} \subseteq \mathrm{T}\mathcal{M}$ is the domain of the exponential map (Definition 10.16).

To prove this, we first introduce a lemma which states that any continuous curve c can be interpolated by a 'broken geodesic' γ. Also, if c is piecewise smooth it has a length and we have $L(\gamma) \leq L(c)$.

Lemma 10.42. *Given $c\colon [0, 1] \to \mathcal{M}$ continuous on a manifold $\mathcal{M}$, there exist a finite number of times $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = 1$ such that $\mathrm{dist}(c(t_i), c(t_{i+1})) < \mathrm{inj}(c(t_i))$ for $i = 0, \ldots, n-1$.*

These times define a piecewise regular curve $\gamma\colon [0, 1] \to \mathcal{M}$ satisfying $\gamma(t_i) = c(t_i)$ and such that $\gamma|_{[t_i, t_{i+1}]}$ is the minimizing geodesic connecting its endpoints. As such, there exist tangent vectors $s_0, \ldots, s_{n-1}$ such that $\gamma(t_{i+1}) = \mathrm{Exp}_{\gamma(t_i)}(s_i)$ and $\sum_{i=0}^{n-1} \|s_i\| = L(\gamma)$.

Proof. Consider the recursive routine `construct` with inputs a, b which proceeds as follows: if $\mathrm{dist}(c(a), c(b)) < \mathrm{inj}(c(a))$, return (a, b); if not, return the results of `construct`$(a, (a + b)/2)$ and `construct`$((a + b)/2, b)$ merged. We claim that `construct`$(0, 1)$ is an appropriate selection. Indeed, the routine terminates after a finite number of steps because $\mathrm{inj} \circ c$ is continuous and positive on the compact domain $[0, 1]$ so that it is bounded away from zero, and c is continuous so that for $\varepsilon > 0$ small enough we can have $\mathrm{dist}(c(t), c(t+\varepsilon))$ arbitrarily small. Furthermore, for all selected t_i, t_{i+1}, we have $\mathrm{dist}(c(t_i), c(t_{i+1})) < \mathrm{inj}(c(t_i))$. Hence, there exists a (unique) minimizing geodesic connecting $c(t_i)$ to $c(t_{i+1})$, for all i, and $\|s_i\| = L(\gamma|_{[t_i, t_{i+1}]})$. (We used Proposition 10.22.) ☐

Proof of Proposition 10.41. Under condition (10.13), the claim is clear:

$$\forall (x, s) \in \mathcal{O}, \qquad |f(\mathrm{Exp}_x(s)) - f(x)| \leq L\, \mathrm{dist}(\mathrm{Exp}_x(s), x) \leq L\|s\|$$

since $t \mapsto \mathrm{Exp}_x(ts)$ is a smooth curve defined on $[0, 1]$ with length $\|s\|$.

The other way around, let us assume condition (10.14) holds. If $\mathcal{M}$ is complete (that is, $\mathcal{O} = \mathrm{T}\mathcal{M}$) the claim is also clear: by Theorem 10.9, for all $x, y \in \mathcal{M}$ in the same connected component there exists s in $\mathrm{T}_x\mathcal{M}$ such that $y = \mathrm{Exp}_x(s)$ and $\|s\| = \mathrm{dist}(x, y)$.

If $\mathcal{M}$ is not complete, we proceed as follows: by definition of distance (10.2), for all $x, y \in \mathcal{M}$ in the same connected component and for all $\varepsilon > 0$, there exists a piecewise regular curve $c\colon [0,1] \to \mathcal{M}$ with length $L(c) \le \mathrm{dist}(x,y) + \varepsilon$ such that $c(0) = x$ and $c(1) = y$. Construct a broken geodesic γ as provided by Lemma 10.42: there exist times $0 = t_0 < \cdots < t_n = 1$ and tangent vectors $s_0, \ldots, s_{n-1}$ such that $\gamma(t_i) = c(t_i)$ and $\gamma(t_{i+1}) = \mathrm{Exp}_{\gamma(t_i)}(s_i)$ for all i, and $\sum_{i=0}^{n-1} \|s_i\| = L(\gamma) \le L(c)$. Then,

$$|f(x) - f(y)| \le \sum_{i=0}^{n-1} |f(\gamma(t_i)) - f(\gamma(t_{i+1}))| \overset{(10.14)}{\le} \sum_{i=0}^{n-1} L \cdot \|s_i\| \le L \cdot L(c).$$

This reasoning holds for all $\varepsilon > 0$, hence condition (10.13) follows. □

If f has a continuous gradient, then f is Lipschitz continuous exactly if its gradient is bounded.

Proposition 10.43. *If $f\colon \mathcal{M} \to \mathbb{R}$ has a continuous gradient, then f is L-Lipschitz continuous if and only if*

$$\forall x \in \mathcal{M}, \qquad \|\mathrm{grad} f(x)\| \le L. \tag{10.15}$$

Proof. For any $(x, s) \in \mathcal{O}$, consider $c(t) = \mathrm{Exp}_x(ts)$ for $t \in [0,1]$. Then,

$$f(c(1)) - f(c(0)) = \int_0^1 (f \circ c)'(t)\mathrm{d}t = \int_0^1 \langle \mathrm{grad} f(c(t)), c'(t) \rangle \, \mathrm{d}t.$$

Thus, if the gradient norm is bounded by L at all points along c,

$$|f(\mathrm{Exp}_x(s)) - f(x)| \le L \int_0^1 \|c'(t)\|\mathrm{d}t = L \cdot L(c) = L\|s\|.$$

This shows that (10.14) holds.

The other way around, for any $x \in \mathcal{M}$, assuming (10.14) holds and using that the domain of Exp_x is open around the origin, we have:

$$\begin{aligned}
\|\mathrm{grad} f(x)\| &= \max_{s \in \mathrm{T}_x\mathcal{M}, \|s\|=1} \langle \mathrm{grad} f(x), s \rangle \\
&= \max_{s \in \mathrm{T}_x\mathcal{M}, \|s\|=1} \mathrm{D}f(x)[s] \\
&= \max_{s \in \mathrm{T}_x\mathcal{M}, \|s\|=1} \lim_{t \to 0} \frac{f(\mathrm{Exp}_x(ts)) - f(x)}{t} \le L,
\end{aligned}$$

since $f(\mathrm{Exp}_x(ts)) - f(x) \le L\|ts\| = L|t|$. □

We now turn to defining Lipschitz continuity for the Riemannian gradient of a function f. Since $\mathrm{grad} f$ is a map from $\mathcal{M}$ to $\mathrm{T}\mathcal{M}$, to apply the general notion of Lipschitz continuity directly we would need to pick a distance on the tangent bundle. However, this would not lead to interesting notions for us. Indeed, the distance between $\mathrm{grad}f(x)$ and $\mathrm{grad}f(y)$ would necessarily have to be positive if $x \neq y$ since they would always be distinct points in $\mathrm{T}\mathcal{M}$. Contrast this to the Euclidean case $f : \mathcal{E} \to \mathbb{R}$, where it is natural to measure $\mathrm{grad}f(x) - \mathrm{grad}f(y)$ in the Euclidean metric, disregarding the base points. With this in mind, it is reasonable to resort to parallel transport (Section 10.3) to compare tangent vectors at distinct points. Since parallel transport is dependent on paths, this leaves some leeway in the definition.

The following definition is fairly common. Notice how the restriction by the injectivity radius allows us to choose a privileged path along which to transport (owing to Proposition 10.22). For our purpose, the notion below is particularly relevant with $V = \mathrm{grad}f$, in which case we would say f has an *L-Lipschitz continuous gradient.*

Definition 10.44. *A vector field V on a connected manifold $\mathcal{M}$ is L-*Lipschitz continuous *if, for all $x, y \in \mathcal{M}$ with* $\mathrm{dist}(x, y) < \mathrm{inj}(x)$*,*

$$\|\mathrm{PT}^{\gamma}_{0\leftarrow 1} V(y) - V(x)\| \leq L\, \mathrm{dist}(x, y), \tag{10.16}$$

where $\gamma : [0, 1] \to \mathcal{M}$ is the unique minimizing geodesic connecting x to y. If $\mathcal{M}$ is disconnected, we require the condition on each connected component.

Here too, we provide an equivalent definition in terms of the exponential map: this may be more convenient to analyze optimization algorithms, and has the added benefit of allowing the comparison of points which are further apart than the injectivity radius (but still connected by a geodesic).

Proposition 10.45. *A vector field V on a manifold $\mathcal{M}$ is L-Lipschitz continuous if and only if*

$$\forall (x, s) \in \mathcal{O}, \qquad \|P_s^{-1} V(\mathrm{Exp}_x(s)) - V(x)\| \leq L\|s\|, \tag{10.17}$$

where $\mathcal{O} \subseteq \mathrm{T}\mathcal{M}$ is the domain of Exp *and P_s denotes parallel transport along $\gamma(t) = \mathrm{Exp}_x(ts)$ from $t = 0$ to $t = 1$.*

Proof. For any $x, y \in \mathcal{M}$ such that $\mathrm{dist}(x, y) < \mathrm{inj}(x)$, there exists a unique $s \in \mathrm{T}_x\mathcal{M}$ such that $y = \mathrm{Exp}_x(s)$ and $\|s\| = \mathrm{dist}(x, y)$. Thus, if condition (10.17) holds, then (10.16) holds.

The other way around, for any $(x, s) \in \mathcal{O}$, consider the geodesic $\gamma(t) = \mathrm{Exp}_x(ts)$ defined over $[0, 1]$. It may or may not be minimizing. In any case, owing to Lemma 10.42, the interval $[0, 1]$ can be partitioned by $0 = t_0 < \cdots < t_n = 1$ such that $\mathrm{dist}(\gamma(t_i), \gamma(t_{i+1})) < \mathrm{inj}(\gamma(t_i))$. Since

$$\mathrm{PT}^{\gamma}_{0\leftarrow 1} = \mathrm{PT}^{\gamma}_{t_0\leftarrow t_{n-1}} \circ \mathrm{PT}^{\gamma}_{t_{n-1}\leftarrow t_n}$$

and since parallel transport is an isometry, we find that

$$\begin{aligned}\|\mathrm{PT}^{\gamma}_{t_0\leftarrow t_n}V(\gamma(t_n)) - V(x)\| &= \|\mathrm{PT}^{\gamma}_{t_{n-1}\leftarrow t_n}V(\gamma(t_n)) - \mathrm{PT}^{\gamma}_{t_{n-1}\leftarrow t_0}V(x)\| \\ &\leq \|\mathrm{PT}^{\gamma}_{t_{n-1}\leftarrow t_n}V(\gamma(t_n)) - V(\gamma(t_{n-1}))\| \\ &\quad + \|\mathrm{PT}^{\gamma}_{t_{n-1}\leftarrow t_0}V(x) - V(\gamma(t_{n-1}))\| \\ &\leq L\,\mathrm{dist}(\gamma(t_{n-1}), \gamma(t_n)) \\ &\quad + \|\mathrm{PT}^{\gamma}_{t_0\leftarrow t_{n-1}}V(\gamma(t_{n-1})) - V(x)\|,\end{aligned}$$

where in the last step we were able to use (10.16) since $\gamma|_{[t_{n-1},t_n]}$ is the unique minimizing geodesic connecting $\gamma(t_{n-1})$ and $\gamma(t_n)$. Repeat this argument on the right-most term $n-1$ times to see that

$$\|\mathrm{PT}^{\gamma}_{0\leftarrow 1}V(\mathrm{Exp}_x(s)) - V(x)\| \leq L\sum_{i=0}^{n-1}\mathrm{dist}(\gamma(t_i), \gamma(t_{i+1})) = L\cdot L(\gamma),$$

using that γ is a geodesic. To conclude, note that $L(\gamma) = \|s\|$ and that $\mathrm{PT}^{\gamma}_{0\leftarrow 1} = P_s^{-1}$, so that condition (10.17) holds. □

If the vector field V is continuously differentiable, then Lipschitz continuity of V is equivalent to boundedness of its covariant derivative. In turn, this makes it possible to compare the values of V at points connected by curves other than geodesics.

Proposition 10.46. *If V is a continuously differentiable vector field on a manifold $\mathcal{M}$, then it is L-Lipschitz continuous if and only if*

$$\forall (x, s) \in \mathrm{T}\mathcal{M}, \qquad \|\nabla_s V\| \leq L\|s\|, \tag{10.18}$$

where ∇ is the Riemannian connection. In that case, for any smooth curve $c\colon [0,1] \to \mathcal{M}$ connecting any x to any y, it holds that

$$\|\mathrm{PT}^{c}_{0\leftarrow 1}V(y) - V(x)\| \leq L\cdot L(c). \tag{10.19}$$

Proof. We first show that (10.18) implies (10.19). Since the latter itself implies (10.17), this also takes care of showing that (10.18) implies V is L-Lipschitz continuous. To this end, consider an orthonormal basis $e_1, \ldots, e_d \in \mathrm{T}_x\mathcal{M}$ and their parallel transports $E_i(t) = \mathrm{PT}^{c}_{t\leftarrow 0}(e_i)$. Then, $V(c(t)) = \sum_{i=1}^d v_i(t)E_i(t)$ for some continuously differentiable functions v_i, and

$$\sum_{i=1}^d v_i'(t)E_i(t) = \frac{\mathrm{D}}{\mathrm{d}t}(V\circ c)(t) = \nabla_{c'(t)}V.$$

Furthermore,

$$\begin{aligned}\mathrm{PT}^{c}_{0\leftarrow 1}V(c(1)) - V(c(0)) &= \sum_{i=1}^d (v_i(1) - v_i(0))e_i \\ &= \sum_{i=1}^d \left(\int_0^1 v_i'(t)\mathrm{d}t\right)e_i = \int_0^1 \mathrm{PT}^{c}_{0\leftarrow t}\big(\nabla_{c'(t)}V\big)\,\mathrm{d}t.\end{aligned}$$

Consequently, using that parallel transports are isometric,

$$\|\mathrm{PT}_{0\leftarrow 1}^{c} V(y) - V(x)\| \leq \int_0^1 \|\nabla_{c'(t)} V\| \mathrm{d}t \overset{(10.18)}{\leq} L \int_0^1 \|c'(t)\| \mathrm{d}t = L \cdot L(c).$$

Now for the other direction: assume V is L-Lipschitz continuous. For any $x \in \mathcal{M}$, using that the domain of Exp_x is open around the origin, we know that for all $s \in \mathrm{T}_x\mathcal{M}$ the smooth curve $c(t) = \mathrm{Exp}_x(ts)$ is defined around $t = 0$. Then, by Proposition 10.37,

$$\nabla_s V = \frac{\mathrm{D}}{\mathrm{d}t} V(c(t))\bigg|_{t=0} = \lim_{t \to 0} \frac{\mathrm{PT}_{0\leftarrow t}^{c} V(c(t)) - V(c(0))}{t}.$$

By (10.17), the norm of the numerator is bounded by $L\|ts\|$, which concludes the proof. ☐

Corollary 10.47. *If $f : \mathcal{M} \to \mathbb{R}$ is twice continuously differentiable on a manifold $\mathcal{M}$, then* $\mathrm{grad} f$ *is L-Lipschitz continuous if and only if* $\mathrm{Hess} f(x)$ *has operator norm bounded by L for all x, that is, if for all x we have*

$$\|\mathrm{Hess} f(x)\| = \max_{\substack{s \in \mathrm{T}_x\mathcal{M} \\ \|s\|=1}} \|\mathrm{Hess} f(x)[s]\| \leq L.$$

Let us summarize these findings.

Corollary 10.48. *For a vector field V on a manifold $\mathcal{M}$, these are equivalent:*

1. *V is L-Lipschitz continuous.*
2. *For all x, y in the same component with* $\mathrm{dist}(x, y) < \mathrm{inj}(x)$*, it holds that* $\|\mathrm{PT}_{0\leftarrow 1}^{\gamma} V(y) - V(x)\| \leq L \,\mathrm{dist}(x, y)$ *with γ the unique minimizing geodesic connecting x to y.*
3. *For all (x, s) in the domain of* Exp, $\|P_s^{-1} V(\mathrm{Exp}_x(s)) - V(x)\| \leq L\|s\|$ *where P_s is parallel transport along $c(t) = \mathrm{Exp}_x(ts)$ from $t = 0$ to $t = 1$.*

If V is continuously differentiable, the above are equivalent to the following:

1 *For all smooth $c : [0, 1] \to \mathcal{M}$,* $\|\mathrm{PT}_{0\leftarrow 1}^{c} V(c(1)) - V(c(0))\| \leq L \cdot L(c)$.
2 *For all $(x, s) \in \mathrm{T}\mathcal{M}$,* $\|\nabla_s V\| \leq L\|s\|$.

Particularizing the above to $V = \mathrm{grad} f$ provides a good understanding of functions f with Lipschitz continuous gradients.

Going one degree higher, we now define and (begin to) discuss functions with a *Lipschitz continuous Hessian.* The Hessian of f associates to each x a linear map $\mathrm{Hess} f(x)$ from $\mathrm{T}_x\mathcal{M}$ to itself. The following definition applies.

Definition 10.49. *For each $x \in \mathcal{M}$, let $H(x) : \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ be linear. If $\mathcal{M}$ is connected, we say H is* L-Lipschitz continuous *if for all $x, y \in \mathcal{M}$ such that* $\mathrm{dist}(x, y) < \mathrm{inj}(x)$ *we have*

$$\|\mathrm{PT}_{0\leftarrow 1}^{\gamma} \circ H(y) \circ \mathrm{PT}_{1\leftarrow 0}^{\gamma} - H(x)\| \leq L \,\mathrm{dist}(x, y), \tag{10.20}$$

where $\|\cdot\|$ denotes the operator norm with respect to the Riemannian metric, *and* $\gamma\colon [0,1] \to \mathcal{M}$ *is the unique minimizing geodesic connecting* x *to* y*. If* $\mathcal{M}$ *is disconnected, we require the condition on each connected component.*

Recall the operator norm is $\|H(x)\| = \max_{s\in \mathrm{T}_x\mathcal{M}, \|s\|=1} \|H(x)[s]\|$. The proof of the following proposition is left as an exercise.

Proposition 10.50. *For each* $x \in \mathcal{M}$*, let* $H(x)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ *be linear. The map* H *is* L*-Lipschitz continuous if and only if*

$$\forall (x,s) \in \mathcal{O}, \qquad \|P_s^{-1} \circ H(\mathrm{Exp}_x(s)) \circ P_s - H(x)\| \leq L\|s\|, \tag{10.21}$$

where $\mathcal{O} \subseteq \mathrm{T}\mathcal{M}$ *is the domain of* Exp *and* P_s *denotes parallel transport along* $\gamma(t) = \mathrm{Exp}_x(ts)$ *from* $t=0$ *to* $t=1$*.*

In Section 10.7, we define what it means for a map H as above to be differentiable, and we define its covariant derivative ∇H. Then, as a particular case of Proposition 10.83 we get the following claim, analogous to Proposition 10.46 above.

Proposition 10.51. *For each* $x \in \mathcal{M}$*, let* $H(x)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ *be linear. If* H *is continuously differentiable, it is* L*-Lipschitz continuous if and only if*

$$\forall (x,s) \in \mathrm{T}\mathcal{M}, \qquad \|\nabla_s H\| \leq L\|s\|, \tag{10.22}$$

where ∇ *is the Riemannian connection and* $\nabla_s H\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ *is linear. In that case, for any smooth curve* $c\colon [0,1] \to \mathcal{M}$ *connecting* x *to* y*, we have*

$$\|\mathrm{PT}^c_{0\leftarrow 1} \circ H(y) \circ \mathrm{PT}^c_{1\leftarrow 0} - H(x)\| \leq L \cdot L(c). \tag{10.23}$$

Corollary 10.52. *If* $f\colon \mathcal{M} \to \mathbb{R}$ *is three times continuously differentiable on a manifold* $\mathcal{M}$*, then* $\mathrm{Hess} f$ *is* L*-Lipschitz continuous if and only if*

$$\forall (x,s) \in \mathrm{T}\mathcal{M}, \qquad \|\nabla_s \mathrm{Hess} f\| \leq L\|s\|,$$

where $\nabla_s \mathrm{Hess} f$ *is a self-adjoint linear map on* $\mathrm{T}_x\mathcal{M}$ *defined by* (10.49)*.*

A summary of the same kind as Corollary 10.48 holds here as well.

In Section 10.7, we show how to cast f, $\mathrm{grad} f$ and $\mathrm{Hess} f$ as tensor fields of order zero, one and two respectively. We show how the Riemannian connection can be used to differentiate tensor fields in general, and we discuss Lipschitz continuity at that level of generality. This provides for the missing details in our brief discussion of Lipschitz continuous Hessians, and indicates how to deal with derivatives of arbitrary order.

We can now derive some of the most useful consequences of Lipschitz continuity, namely, bounds on the difference between a function f (or its derivatives) and corresponding Taylor expansions.

We use the following notation often: given (x,s) in the domain $\mathcal{O}$ of the exponential map, let $\gamma(t) = \mathrm{Exp}_x(ts)$ be the corresponding geodesic (defined in particular on the interval $[0,1]$); then, we let

$$P_{ts} = \mathrm{PT}^{\gamma}_{t\leftarrow 0} \tag{10.24}$$

denote parallel transport from x to $\mathrm{Exp}_x(ts)$ along γ. Since γ is a geodesic, its velocity vector field is parallel and we have

$$\gamma'(t) = P_{ts}\gamma'(0) = P_{ts}s. \tag{10.25}$$

This will be helpful a number of times.

Proposition 10.53. *Let $f\colon \mathcal{M} \to \mathbb{R}$ be continuously differentiable on a manifold $\mathcal{M}$. Let $\gamma(t) = \mathrm{Exp}_x(ts)$ be defined on $[0, 1]$ and assume there exists $L \geq 0$ such that, for all $t \in [0, 1]$,*

$$\left\|P_{ts}^{-1}\mathrm{grad}f(\gamma(t)) - \mathrm{grad}f(x)\right\| \leq L\|ts\|.$$

Then, the following inequality holds:

$$|f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad}f(x)\rangle| \leq \frac{L}{2}\|s\|^2.$$

Proof. Consider the real function $f \circ \gamma$ on $[0, 1]$; we have:

$$\begin{aligned} f(\gamma(1)) &= f(\gamma(0)) + \int_0^1 (f\circ\gamma)'(t)\mathrm{d}t \\ &= f(x) + \int_0^1 \langle \mathrm{grad}f(\gamma(t)), \gamma'(t)\rangle \,\mathrm{d}t \\ &= f(x) + \int_0^1 \left\langle P_{ts}^{-1}\mathrm{grad}f(\gamma(t)), s\right\rangle \mathrm{d}t, \end{aligned}$$

where on the last line we used $\gamma'(t) = P_{ts}s$ (10.25) and the fact that P_{ts} is an isometry, so that its adjoint with respect to the Riemannian metric is equal to its inverse. Moving $f(x)$ to the left-hand side and subtracting $\langle \mathrm{grad}f(x), s\rangle$ on both sides, we get

$$f(\mathrm{Exp}_x(s)) - f(x) - \langle \mathrm{grad}f(x), s\rangle = \int_0^1 \left\langle P_{ts}^{-1}\mathrm{grad}f(\gamma(t)) - \mathrm{grad}f(x), s\right\rangle \mathrm{d}t.$$

Using Cauchy–Schwarz and our main assumption, it follows that

$$|f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad}f(x)\rangle| \leq \int_0^1 tL\|s\|^2\mathrm{d}t = \frac{L}{2}\|s\|^2,$$

as announced. ☐

The following corollary shows that the regularity assumptions A4.3 (p60) and A6.6 (p136) hold for the exponential retraction over its whole domain provided $\mathrm{grad}f$ is L-Lipschitz (Definition 10.44). See also Exercise 10.58.

Corollary 10.54. *If $f\colon \mathcal{M} \to \mathbb{R}$ has L-Lipschitz continuous gradient, then*

$$|f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad}f(x)\rangle| \leq \frac{L}{2}\|s\|^2$$

for all (x, s) in the domain of the exponential map.

Proposition 10.55. *Let $f\colon \mathcal{M} \to \mathbb{R}$ be twice continuously differentiable on a manifold $\mathcal{M}$. Let $\gamma(t) = \mathrm{Exp}_x(ts)$ be defined on $[0,1]$ and assume there exists $L \geq 0$ such that, for all $t \in [0,1]$,*

$$\left\|P_{ts}^{-1} \circ \mathrm{Hess} f(\gamma(t)) \circ P_{ts} - \mathrm{Hess} f(x)\right\| \leq L\|ts\|.$$

Then, the two following inequalities hold:

$$\left| f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle - \frac{1}{2}\langle s, \mathrm{Hess} f(x)[s]\rangle \right| \leq \frac{L}{6}\|s\|^3,$$

$$\left\|P_s^{-1}\mathrm{grad} f(\mathrm{Exp}_x(s)) - \mathrm{grad} f(x) - \mathrm{Hess} f(x)[s]\right\| \leq \frac{L}{2}\|s\|^2.$$

Proof. The proof is in three steps.

Step 1: a preliminary computation. Pick an arbitrary basis $e_1, \ldots, e_d$ for $\mathrm{T}_x\mathcal{M}$ and define the parallel vector fields $E_i(t) = P_{ts}e_i$ along $\gamma(t)$. The vectors $E_1(t), \ldots, E_d(t)$ form a basis for $\mathrm{T}_{\gamma(t)}\mathcal{M}$ for each $t \in [0,1]$. As a result, we can express the gradient of f along $\gamma(t)$ in these bases,

$$\mathrm{grad} f(\gamma(t)) = \sum_{i=1}^{d} \alpha_i(t) E_i(t), \tag{10.26}$$

with $\alpha_1(t), \ldots, \alpha_d(t)$ differentiable. Using the Riemannian connection ∇ and associated covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$, we find on the one hand that

$$\frac{\mathrm{D}}{\mathrm{d}t}\mathrm{grad} f(\gamma(t)) = \nabla_{\gamma'(t)}\mathrm{grad} f = \mathrm{Hess} f(\gamma(t))[\gamma'(t)],$$

and on the other hand that

$$\frac{\mathrm{D}}{\mathrm{d}t}\sum_{i=1}^{d}\alpha_i(t)E_i(t) = \sum_{i=1}^{d}\alpha_i'(t)E_i(t) = P_{ts}\sum_{i=1}^{d}\alpha_i'(t)e_i.$$

Combining with $\gamma'(t) = P_{ts}s$ (10.25), we deduce that

$$\sum_{i=1}^{d}\alpha_i'(t)e_i = \left(P_{ts}^{-1} \circ \mathrm{Hess} f(\gamma(t)) \circ P_{ts}\right)[s].$$

Going back to (10.26), we also see that

$$G(t) \triangleq P_{ts}^{-1}\mathrm{grad} f(\gamma(t)) = \sum_{i=1}^{d}\alpha_i(t)e_i$$

is a map from (a subset of) $\mathbb{R}$ to $\mathrm{T}_x\mathcal{M}$—two linear spaces—so that we can differentiate it in the usual way:

$$G'(t) = \sum_{i=1}^{d}\alpha_i'(t)e_i.$$

Overall, we conclude that

$$G'(t) = \frac{\mathrm{d}}{\mathrm{d}t}P_{ts}^{-1}\mathrm{grad} f(\gamma(t)) = \left(P_{ts}^{-1} \circ \mathrm{Hess} f(\gamma(t)) \circ P_{ts}\right)[s]. \tag{10.27}$$

This comes in handy in the next step.

Step 2: Taylor expansion of the gradient. Since G' is continuous,

$$P_{ts}^{-1}\mathrm{grad}f(\gamma(t)) = G(t) = G(0) + \int_0^t G'(\tau)\mathrm{d}\tau$$
$$= \mathrm{grad}f(x) + \int_0^t \left(P_{\tau s}^{-1} \circ \mathrm{Hess}f(\gamma(\tau)) \circ P_{\tau s}\right)[s]\mathrm{d}\tau.$$

Moving $\mathrm{grad}f(x)$ to the left-hand side and subtracting $\mathrm{Hess}f(x)[ts]$ on both sides, we find

$$P_{ts}^{-1}\mathrm{grad}f(\gamma(t)) - \mathrm{grad}f(x) - \mathrm{Hess}f(x)[ts]$$
$$= \int_0^t \left(P_{\tau s}^{-1} \circ \mathrm{Hess}f(\gamma(\tau)) \circ P_{\tau s} - \mathrm{Hess}f(x)\right)[s]\mathrm{d}\tau.$$

Using the main assumption on $\mathrm{Hess}f$ along γ, it follows that

$$\left\|P_{ts}^{-1}\mathrm{grad}f(\gamma(t)) - \mathrm{grad}f(x) - \mathrm{Hess}f(x)[ts]\right\|$$
$$\leq \int_0^t \tau L\|s\|^2\mathrm{d}\tau = \frac{L}{2}\|ts\|^2. \quad (10.28)$$

For $t = 1$, this is one of the announced inequalities.

Step 3: Taylor expansion of the function value. With the same start as in the proof of Proposition 10.53 and subtracting the term $\frac{1}{2}\langle s, \mathrm{Hess}f(x)[s]\rangle$ on both sides, we get

$$f(\mathrm{Exp}_x(s)) - f(x) - \langle \mathrm{grad}f(x), s\rangle - \frac{1}{2}\langle s, \mathrm{Hess}f(x)[s]\rangle$$
$$= \int_0^1 \left\langle P_{ts}^{-1}\mathrm{grad}f(\gamma(t)) - \mathrm{grad}f(x) - \mathrm{Hess}f(x)[ts], s\right\rangle \mathrm{d}t.$$

Using (10.28) and Cauchy–Schwarz, it follows that

$$\left|f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad}f(x)\rangle - \frac{1}{2}\langle s, \mathrm{Hess}f(x)[s]\rangle\right|$$
$$\leq \int_0^1 t^2\frac{L}{2}\|s\|^3\mathrm{d}t = \frac{L}{6}\|s\|^3,$$

as announced. □

The following corollary shows that the regularity assumption A6.7 (p136) holds for the exponential retraction over its whole domain provided $\mathrm{Hess}f$ is L-Lipschitz (Definition 10.49). See also Exercise 10.87.

Corollary 10.56. *If $f \colon \mathcal{M} \to \mathbb{R}$ has L-Lipschitz continuous Hessian, then*

$$\left|f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad}f(x)\rangle - \frac{1}{2}\langle s, \mathrm{Hess}f(x)[s]\rangle\right| \leq \frac{L}{6}\|s\|^3,$$
$$\text{and } \left\|P_s^{-1}\mathrm{grad}f(\mathrm{Exp}_x(s)) - \mathrm{grad}f(x) - \mathrm{Hess}f(x)[s]\right\| \leq \frac{L}{2}\|s\|^2,$$

for all (x, s) in the domain of the exponential map.

To close this section, the following statement from [LKB21] provides Lipschitz-type bounds for pullbacks through arbitrary retractions so long as we restrict our attention to a compact subset of the tangent bundle. The proof (omitted) merely uses the fact that if $f \circ \mathrm{R}$ is sufficiently many times continuously differentiable then its derivatives are bounded on any compact set. In turn, that can be used to bound truncation errors on Taylor expansions of $f \circ \mathrm{R}_x$ in $\mathrm{T}_x\mathcal{M}$ uniformly in x over a compact set. This is convenient to verify typical regularity assumptions such as A4.3, A6.6 and A6.7, though it should be noted that the constants L_1, L_2 below exist merely owing to the compactness-and-continuity argument: this provides little insight (let alone control) over those constants.

Lemma 10.57. *Consider a retraction* R *on* $\mathcal{M}$*, a compact subset* $\mathcal{K} \subseteq \mathcal{M}$ *and a continuous, nonnegative function* $r \colon \mathcal{K} \to \mathbb{R}$*. The set*

$$\mathcal{T} = \{(x, s) \in \mathrm{T}\mathcal{M} : x \in \mathcal{K} \text{ and } \|s\| \leq r(x)\}$$

is compact in the tangent bundle $\mathrm{T}\mathcal{M}$ *(Exercise 10.31). Assume* $f \colon \mathcal{M} \to \mathbb{R}$ *is twice continuously differentiable. There exists a constant* L_1 *such that, for all* $(x, s) \in \mathcal{T}$*, with* $\hat{f}_x = f \circ \mathrm{R}_x$*, we have*

$$\begin{aligned} |f(\mathrm{R}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle| &\leq \frac{L_1}{2}\|s\|^2, \\ \left\|\mathrm{grad}\hat{f}_x(s) - \mathrm{grad}\hat{f}_x(0)\right\| &\leq L_1\|s\|, \end{aligned}$$

and $\|\mathrm{Hess}\hat{f}_x(0)\| \leq L_1$ *for all* $x \in \mathcal{K}$*. If additionally* f *is three times continuously differentiable, then there exists a constant* L_2 *such that, for all* $(x, s) \in \mathcal{T}$*,*

$$\begin{aligned} \left|f(\mathrm{R}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle - \frac{1}{2}\langle s, \mathrm{Hess}\hat{f}_x(0)[s]\rangle\right| &\leq \frac{L_2}{6}\|s\|^3, \\ \left\|\mathrm{grad}\hat{f}_x(s) - \mathrm{grad}\hat{f}_x(0) - \mathrm{Hess}\hat{f}_x(0)[s]\right\| &\leq \frac{L_2}{2}\|s\|^2, \\ \left\|\mathrm{Hess}\hat{f}_x(s) - \mathrm{Hess}\hat{f}_x(0)\right\| &\leq L_2\|s\|. \end{aligned}$$

(Recall $\mathrm{grad}\hat{f}_x(0) = \mathrm{grad} f(x)$ *and, if the retraction is second order,* $\mathrm{Hess}\hat{f}_x(0) = \mathrm{Hess} f(x)$*.)*

The exercise below has clear implications for the regularity assumptions A4.3 (p60) and A6.6 (p136). See also Exercise 10.87.

Exercise 10.58. *Let* $f \colon \mathcal{M} \to \mathbb{R}$ *be twice continuously differentiable on a manifold* $\mathcal{M}$ *equipped with a retraction* R*. Assume we have*

$$|f(\mathrm{R}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle| \leq \frac{L}{2}\|s\|^2 \tag{10.29}$$

for all (x, s) in a neighborhood of the zero section in the tangent bundle. With $\hat{f}_x = f \circ \mathrm{R}_x$*, show that* $\|\mathrm{Hess}\hat{f}_x(0)\| \leq L$*. Deduce that if* R *is second order then the* *inequalities* (10.29) *hold only if* $\mathrm{grad} f$ *is L-Lipschitz continuous. With* $\mathrm{R} = \mathrm{Exp}$ *in particular, verify that the three following claims are equivalent:*

1. *Inequalities* (10.29) *hold in a neighborhood of the zero section in* $\mathrm{T}\mathcal{M}$*;*
2. $\mathrm{grad} f$ *is L-Lipschitz continuous;*
3. *Inequalities* (10.29) *hold over the whole domain of* Exp*.*

Exercise 10.59. *Give a proof of Proposition 10.50, for example by adapting that of Proposition 10.45.*

10.5 Transporters

The strong properties of parallel transports (Section 10.3) make them great for theoretical purposes, and in some cases they can even be computed via explicit expressions. In general though, computing parallel transports involves numerically solving ordinary differential equations, which is typically too expensive in practice. Furthermore, we may want to dispense with the need to choose a curve connecting x and y explicitly to transport vectors from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_y\mathcal{M}$, as this may add to the computational burden (e.g., require computing $\mathrm{Log}_x(y)$ if we mean to transport along minimizing geodesics).

As an alternative, we define a poor man's version of parallel transports called *transporters*.[5] There is no need for a Riemannian structure or connection. Informally, for x and y close enough to one another, we aim to define linear maps of the form

$$\mathrm{T}_{y \leftarrow x} \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_y\mathcal{M},$$

with $\mathrm{T}_{x \leftarrow x}$ in particular being the identity map. If $\mathcal{M}$ is an embedded submanifold of a Euclidean space, we present a simple transporter based on orthogonal projections to tangent spaces in Proposition 10.66.

It is natural and convenient to ask that these maps vary smoothly with respect to x and y. One indirect way to make sense of this statement would be to require that the map $((x, u), y) \mapsto \mathrm{T}_{y \leftarrow x} u$ be smooth from (an open submanifold of) $\mathrm{T}\mathcal{M} \times \mathcal{M}$ to $\mathrm{T}\mathcal{M}$. However, it is more instructive (and eventually more comfortable) to endow the set of linear maps between tangent spaces of two manifolds with a smooth structure. (Here, the two manifolds are the same.) Once this is done, we can formalize the notion of smoothness for a map $(x, y) \mapsto \mathrm{T}_{y \leftarrow x}$. This is in direct analogy with how we defined the tangent bundle $\mathrm{T}\mathcal{M}$ as a disjoint union of tangent spaces, associating a linear space $\mathrm{T}_x\mathcal{M}$ to each point $x \in \mathcal{M}$. Here, we associate to each pair $(x, y) \in \mathcal{M} \times \mathcal{N}$ the linear space of linear maps from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_y\mathcal{N}$. The proof is an exercise.

[5] This is different from the notion of *vector transport* as defined in [AMS08, §8.1]: we connect both concepts at the end of this section.

Proposition 10.60. *For manifolds $\mathcal{M}$ and $\mathcal{N}$ of dimensions m and n, the disjoint union of linear maps from the tangent spaces of $\mathcal{M}$ to those of $\mathcal{N}$,*

$$\mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{N}) = \{(x, y, \mathcal{L}) \,|\, x \in \mathcal{M}, y \in \mathcal{N} \text{ and } \mathcal{L} \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_y\mathcal{N} \text{ is linear}\},$$

is itself a manifold, with charts as follows: for any pair of charts $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ of $\mathcal{M}$ and $\mathcal{N}$ respectively, pick local frames on $\mathcal{U}$ and $\mathcal{V}$ as in Proposition 8.51; then,

$$\Phi(x, y, \mathcal{L}) = (\varphi(x), \psi(y), \mathrm{Mat}(\mathcal{L})) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n\times m}$$

is a chart on $\pi^{-1}(\mathcal{U} \times \mathcal{V})$, where $\mathrm{Mat}(\mathcal{L})$ is the matrix that represents $\mathcal{L}$ with respect to the bases of $\mathrm{T}_x\mathcal{M}$ and $\mathrm{T}_y\mathcal{N}$ provided by the local frames, and $\pi(x, y, \mathcal{L}) = (x, y)$ is the projector from $\mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{N})$ to $\mathcal{M} \times \mathcal{N}$.

The manifold $\mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{N})$ is a *vector bundle* of $\mathcal{M} \times \mathcal{N}$ in that it (smoothly) attaches a linear space to each point of that manifold. Maps such as transporters defined below have the property that they map (x, y) to $(x, y, \mathcal{L})$ for some $\mathcal{L}$: these are called *sections* of the vector bundle. In the same way, vector fields are called sections of the tangent bundle.

Definition 10.61. *Given a manifold $\mathcal{M}$, let $\mathcal{V}$ be open in $\mathcal{M} \times \mathcal{M}$ such that $(x, x) \in \mathcal{V}$ for all $x \in \mathcal{M}$. A* transporter *on $\mathcal{V}$ is a smooth map*

$$\mathrm{T} \colon \mathcal{V} \to \mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{M}) \colon (x, y) \mapsto \mathrm{T}_{y\leftarrow x}$$

such that $\mathrm{T}_{y\leftarrow x}$ is linear from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_y\mathcal{M}$ and $\mathrm{T}_{x\leftarrow x}$ is the identity.

In this definition, smoothness of T is understood with $\mathcal{V}$ as an open submanifold of the product manifold $\mathcal{M}\times\mathcal{M}$ and $\mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{M})$ equipped with the smooth structure of Proposition 10.60. Formally, this means that for any pair $(\bar{x}, \bar{y}) \in \mathcal{V}$ and local frames defined on neighborhoods $\mathcal{U}_{\bar{x}}$ and $\mathcal{U}_{\bar{y}}$, the matrix that represents $\mathrm{T}_{y\leftarrow x}$ with respect to these local frames varies smoothly with (x, y) in $\mathcal{U}_{\bar{x}} \times \mathcal{U}_{\bar{y}}$. We detail this in the proof of the next proposition, which shows that inverting the linear maps of a transporter yields a transporter.

Proposition 10.62. *For a transporter T on $\mathcal{V}$, let $\mathcal{V}'$ be the set of pairs $(x, y) \in \mathcal{V}$ such that $\mathrm{T}_{x\leftarrow y}$ is invertible. Then, the maps*

$$\mathrm{T}'_{y\leftarrow x} = (\mathrm{T}_{x\leftarrow y})^{-1} \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_y\mathcal{M}$$

define a transporter T' on $\mathcal{V}'$.

Proof. For all $x \in \mathcal{M}$, since $\mathrm{T}_{x\leftarrow x}$ is the identity map, clearly $(x, x) \in \mathcal{V}'$ and $\mathrm{T}'_{x\leftarrow x}$ is itself the identity. Likewise, for all $(x, y) \in \mathcal{V}'$, it is clear that $\mathrm{T}'_{y\leftarrow x}$ is linear. It remains to argue that $\mathcal{V}'$ is open in $\mathcal{M}\times\mathcal{M}$ and that T' is smooth from $\mathcal{V}'$ to $\mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{M})$.

To this end, consider an arbitrary pair $(\bar{x}, \bar{y}) \in \mathcal{V}'$ and let $U_1, \ldots, U_d$ be a local frame on a neighborhood $\mathcal{U}_{\bar{x}}$ of $\bar{x}$ with $d = \dim \mathcal{M}$—see Proposition 8.51. Likewise, let $W_1, \ldots, W_d$ be a local frame on a neighborhood $\mathcal{U}_{\bar{y}}$ of $\bar{y}$. If need be, reduce $\mathcal{U}_{\bar{x}}$ and $\mathcal{U}_{\bar{y}}$ to smaller neighborhoods of $\bar{x}$ and $\bar{y}$ so that $\mathcal{U}_{\bar{x}} \times \mathcal{U}_{\bar{y}} \subseteq \mathcal{V}$. (We can do this because $\mathcal{V}$ is open in $\mathcal{M} \times \mathcal{M}$, hence it is a union of products of open sets in $\mathcal{M}$. One of these products, say $\tilde{\mathcal{U}} \times \hat{\mathcal{U}}$, contains $(\bar{x}, \bar{y})$, as otherwise it would not be in $\mathcal{V}$; thus: $\bar{x} \in \tilde{\mathcal{U}}$ and $\bar{y} \in \hat{\mathcal{U}}$. Replace $\mathcal{U}_{\bar{x}}$ by its intersection with $\tilde{\mathcal{U}}$, and similarly for $\mathcal{U}_{\bar{y}}$: now $\mathcal{U}_{\bar{x}} \times \mathcal{U}_{\bar{y}} \subseteq \mathcal{V}$ is a neighborhood of $(\bar{x}, \bar{y})$ and the local frames are well defined.) Since T is smooth, the matrix $G(x, y)$ in $\mathbb{R}^{d \times d}$ that represents $\mathrm{T}_{x \leftarrow y}$ with respect to the bases $U_1(x), \ldots, U_d(x)$ and $W_1(y), \ldots, W_d(y)$ varies smoothly with (x, y) in $\mathcal{U}_{\bar{x}} \times \mathcal{U}_{\bar{y}}$. In particular, the function $(x, y) \mapsto \det G(x, y)$ is smooth on this domain, so that the subset of $\mathcal{U}_{\bar{x}} \times \mathcal{U}_{\bar{y}}$ over which $\det G(x, y) \neq 0$ (that is, over which $\mathrm{T}_{x \leftarrow y}$ is invertible) is open, and it contains $(\bar{x}, \bar{y})$. In other words: this subset is a neighborhood of $(\bar{x}, \bar{y})$ in $\mathcal{V}'$. Since each point in $\mathcal{V}'$ admits such a neighborhood, we find that $\mathcal{V}'$ is open. Furthermore, the matrix that represents $\mathrm{T}'_{y \leftarrow x}$ is simply $G(x, y)^{-1}$. This is a smooth function of (x, y) on the open set where the inverse is well defined, confirming that T' is smooth from $\mathcal{V}'$ to $\mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{M})$. □

With similar developments, we also get the following result once we equip the manifold with a Riemannian metric.

Proposition 10.63. *Let $\mathcal{M}$ be a Riemannian manifold and let T be a transporter for $\mathcal{M}$ on $\mathcal{V}$. Then, T' defined by the maps*

$$\mathrm{T}'_{y \leftarrow x} = (\mathrm{T}_{x \leftarrow y})^* \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_y\mathcal{M}$$

is a transporter on $\mathcal{V}$. (As always, the superscript $*$ denotes the adjoint, here with respect to the Riemannian metric at x and y.)

Proof sketch. Compared to Proposition 10.62 (and using the same notation), an extra step in the proof is to show that, using local frames, the Riemannian metric can be represented as a smooth map from x to $M(x)$: a symmetric, positive definite matrix of size d which allows us to write $\langle u, v \rangle_x = \bar{u}^\top M(x) \bar{v}$ with $\bar{u}, \bar{v} \in \mathbb{R}^d$ denoting the coordinate vectors of u, v in the same local frame. Upon doing so, it is straightforward to show that the matrix which represents $(\mathrm{T}_{x \leftarrow y})^*$ is $M(y)^{-1} G(x, y)^\top M(x)$, which is indeed smooth in (x, y). □

Upon choosing a smoothly varying collection of curves that uniquely connect pairs of nearby points on $\mathcal{M}$, it is easy to construct a transporter from parallel transport along those curves (with respect to some connection). One way of choosing such families of curves is via a retraction.

Conveniently, the differentials of a retraction also provide a transporter. This is because, with $y = \mathrm{R}_x(v)$, by perturbing v in $\mathrm{T}_x\mathcal{M}$ we perturb $\mathrm{R}_x(v)$ away from y, thus producing a tangent vector in $\mathrm{T}_y\mathcal{M}$. That is a good alternative when parallel transports are out of reach.

Proposition 10.64. *For a retraction R on a manifold $\mathcal{M}$, let $\mathcal{T}$ be a neighborhood of the zero section of $\mathrm{T}\mathcal{M}$ such that $E(x, v) = (x, \mathrm{R}_x(v))$ is a diffeomorphism from $\mathcal{T}$ to $\mathcal{V} = E(\mathcal{T})$—such neighborhoods exist by Corollary 10.27. For* *our purpose, this means* $(x, y) \mapsto (x, \mathrm{R}_x^{-1}(y))$ *is a diffeomorphism from* $\mathcal{V}$ *to* $\mathcal{T}$*, yielding a smooth choice of curves joining pairs* (x, y)*.*

1. *Assume* $\mathcal{T}_x = \{v \in \mathrm{T}_x\mathcal{M} : (x, v) \in \mathcal{T}\}$ *is star-shaped around the origin for all* x*. Parallel transport along retraction curves defines a transporter on* $\mathcal{V}$ *via* $\mathrm{T}_{y \leftarrow x} = \mathrm{PT}^c_{1 \leftarrow 0}$*, where* $c(t) = \mathrm{R}_x(tv)$ *and* $v = \mathrm{R}_x^{-1}(y)$*.*
2. *The differentials of the retraction define a transporter on* $\mathcal{V}$ *via* $\mathrm{T}_{y \leftarrow x} = \mathrm{DR}_x(v)$*, where* $v = \mathrm{R}_x^{-1}(y)$*.*

Proof sketch. The domain $\mathcal{V} \subseteq \mathcal{M} \times \mathcal{M}$ is open and indeed contains all pairs (x, x) since $E(x, 0) = (x, x)$. For both proposed transporters, it is clear that $\mathrm{T}_{x \leftarrow x}$ is the identity and that $\mathrm{T}_{y \leftarrow x}$ is a linear map from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_y\mathcal{M}$. Smoothness for parallel transport can be argued with tools from ordinary differential equations (it takes some work). Smoothness for the retraction-based transport follows by composition of smooth maps since $\mathrm{T}_{y \leftarrow x} = \mathrm{DR}_x(\mathrm{R}_x^{-1}(y))$. □

Example 10.65. *Transporters can be used to transport linear maps between certain tangent spaces to other tangent spaces. This is useful notably in defining a Riemannian version of the famous BFGS algorithm. For example, if* $\mathcal{A}$ *is a linear map from* $\mathrm{T}_x\mathcal{M}$ *to* $\mathrm{T}_x\mathcal{M}$*, then we may transport it to a linear map from* $\mathrm{T}_y\mathcal{M}$ *to* $\mathrm{T}_y\mathcal{M}$ *in at least three ways using a transporter* T*:*

$$\mathrm{T}_{y \leftarrow x} \circ \mathcal{A} \circ \mathrm{T}_{x \leftarrow y}, \qquad (\mathrm{T}_{x \leftarrow y})^* \circ \mathcal{A} \circ \mathrm{T}_{x \leftarrow y}, \qquad (\mathrm{T}_{x \leftarrow y})^{-1} \circ \mathcal{A} \circ \mathrm{T}_{x \leftarrow y}.$$

If $\mathcal{A}$ *is self-adjoint, then so is the second operator. If the transporter is obtained through parallel transport as in Proposition 10.64 and the curve connecting* x *to* y *is the same as the curve connecting* y *to* x *(for example, if we use unique minimizing geodesics), then all three operators are equal: see Proposition 10.36.*

For manifolds embedded in Euclidean spaces, an especially convenient transporter is given by orthogonal projectors to the tangent spaces. In contrast to Proposition 10.64, it does not involve retractions.

Proposition 10.66. *Let* $\mathcal{M}$ *be an embedded submanifold of a Euclidean space* $\mathcal{E}$*. For all* $x, y \in \mathcal{M}$*, exploiting the fact that both* $\mathrm{T}_x\mathcal{M}$ *and* $\mathrm{T}_y\mathcal{M}$ *are subspaces of* $\mathcal{E}$*, define the linear maps*

$$\mathrm{T}_{y \leftarrow x} = \mathrm{Proj}_y\big|_{\mathrm{T}_x\mathcal{M}},$$

where Proj_y *is the orthogonal projector from* $\mathcal{E}$ *to* $\mathrm{T}_y\mathcal{M}$*, here restricted to* $\mathrm{T}_x\mathcal{M}$*. This is a transporter on all of* $\mathcal{M} \times \mathcal{M}$*.*

Proof. By design, $\mathrm{T}_{x \leftarrow x}$ is the identity and $\mathrm{T}_{y \leftarrow x}$ is linear from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_y\mathcal{M}$. Moreover, T is smooth as can be deduced by an argument along the same lines as in Exercise 3.66. □

For a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$, in the same way that we discussed conditions for a retraction $\overline{\mathrm{R}}$ on the total space $\overline{\mathcal{M}}$ to induce a retraction R on the quotient manifold $\mathcal{M}$, it is tempting to derive a transporter T on $\mathcal{M}$ from a transporter $\overline{\mathrm{T}}$ on $\overline{\mathcal{M}}$. We show through an example how this can be done.

Example 10.67. *Consider the Grassmann manifold* $\mathrm{Gr}(n,p) = \mathrm{St}(n,p)/\sim$ *(Section 9.16). Equip the total space with the polar retraction* (7.24), $\overline{\mathrm{R}}_X(V) = \mathrm{pfactor}(X+V)$, *and with the projection transporter,*

$$\overline{\mathrm{T}}_{Y\leftarrow X} = \mathrm{Proj}_Y^{\mathrm{St}}|_{\mathrm{T}_X\mathrm{St}(n,p)},$$

the orthogonal projector from $\mathbb{R}^{n\times p}$ *to* $\mathrm{T}_Y\mathrm{St}(n,p)$ *restricted to* $\mathrm{T}_X\mathrm{St}(n,p)$. *This transporter is defined globally on* $\mathrm{St}(n,p)\times\mathrm{St}(n,p)$. *Our tentative transporter on* $\mathrm{Gr}(n,p)$ *is:*

$$\mathrm{T}_{[Y]\leftarrow[X]}(\xi) = \mathrm{D}\pi(Y)\left[\overline{\mathrm{T}}_{Y\leftarrow X}(\mathrm{lift}_X(\xi))\right], \tag{10.30}$$

where X *is an arbitrary representative of* $[X]$, *and* Y *is a representative of* $[Y]$ *such that* $Y = \overline{\mathrm{R}}_X(V)$ *for some* $V \in \mathrm{H}_X$, *assuming one exists. When such a choice of* Y *and* V *exists, it is unique. Indeed, consider the map*

$$\begin{aligned} E\colon \mathrm{TGr}(n,p) &\to \mathrm{Gr}(n,p)\times\mathrm{Gr}(n,p) \\ : ([X],\xi) &\mapsto E([X],\xi) = ([X],[\overline{\mathrm{R}}_X(V)]), \end{aligned} \tag{10.31}$$

where $V = \mathrm{lift}_X(\xi)$. *In Exercise 10.71, we find that* E *from* $\mathrm{TGr}(n,p)$ *to* $\mathcal{V} = E(\mathrm{TGr}(n,p))$ *is smoothly invertible. In other words: if* $[Y]$ *can be reached from* $[X]$ *through retraction, it is so by a unique tangent vector* ξ*; the latter has a specific horizontal lift* V *once we choose a specific representative* X. *Furthermore, since* E^{-1} *is continuous,* $\mathcal{V}$ *is open. Finally,* $\mathcal{V}$ *contains all pairs of the form* $([X],[X])$. *This set* $\mathcal{V}$ *is meant to be the domain of* T.

Now restricting our discussion to $\mathcal{V}$, *We rewrite* (10.30) *equivalently as*

$$\begin{aligned} \mathrm{lift}_Y\big(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\big) &= \mathrm{Proj}_Y^{\mathrm{H}}\left(\overline{\mathrm{T}}_{Y\leftarrow X}(\mathrm{lift}_X(\xi))\right) \\ &= \mathrm{Proj}_Y^{\mathrm{H}}\left(\mathrm{lift}_X(\xi)\right), \end{aligned} \tag{10.32}$$

where we used that $\mathrm{Proj}_Y^{\mathrm{H}}\circ\mathrm{Proj}_Y^{\mathrm{St}} = \mathrm{Proj}_Y^{\mathrm{H}}$. *We must check (a) that* $\mathrm{T}_{[Y]\leftarrow[X]}$ *is well defined, and (b) that it defines a transporter, both on* $\mathcal{V}$.

For (a), we must check that the right-hand side of (10.30) *does not depend on our choice of representatives* X *and* Y. *To this end, consider* (10.32). *Recall from Example 9.26 that if we choose the representative* XQ *instead of* X *for* $[X]$ *with some arbitrary* $Q \in \mathrm{O}(p)$, *then* $\mathrm{lift}_{XQ}(\xi) = \mathrm{lift}_X(\xi)Q$. *The representative* Y *also changes as a result. Indeed, given* $V \in \mathrm{H}_X$ *such that* $\overline{\mathrm{R}}_X(V) = Y$, *we know that* $VQ \in \mathrm{H}_{XQ}$ *is such that* $\overline{\mathrm{R}}_{XQ}(VQ) = YQ$ *(this is specific to the polar retraction by* (9.9)*), and this is the only horizontal vector at* XQ *that maps to* $[Y]$. *Since* $\mathrm{Proj}_Y^{\mathrm{H}} = I_n - YY^\top = \mathrm{Proj}_{YQ}^{\mathrm{H}}$, *we find that*

$$\begin{aligned} \mathrm{lift}_{YQ}\big(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\big) &= \mathrm{Proj}_{YQ}^{\mathrm{H}}\left(\mathrm{lift}_{XQ}(\xi)\right) \\ &= (I_n - YY^\top)\,\mathrm{lift}_X(\xi)Q \\ &= \mathrm{lift}_Y\big(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\big)\,Q. \end{aligned}$$

This confirms that the lifted vectors correspond to each other in the appropriate way, that is, the result $\mathrm{T}_{[Y]\leftarrow[X]}(\xi)$ *does not depend on our choice of representative* X.

Regarding (b), it is clear that $\mathrm{T}_{[Y]\leftarrow[X]}$ *is a linear map from* $\mathrm{T}_{[X]}\mathrm{Gr}(n,p)$ *to* $\mathrm{T}_{[Y]}\mathrm{Gr}(n,p)$*, as a composition of linear maps. Likewise,* T *is smooth as a composition of smooth maps (this also follows from Exercise 10.71, which shows that given* $([X],[Y]) \in \mathcal{V}$*, for any choice of representative* X*, there is a smooth choice of* Y *and* V *(horizontal) such that* $Y = \overline{\mathrm{R}}_X(V)$*). It is easy to see that* $\mathrm{T}_{[X]\leftarrow[X]}$ *is the identity. Finally, we already checked that* $\mathcal{V}$ *is an appropriate domain for a transporter.*

How do we use this transporter in practice? If we are simply given two representatives X *and* Y *and the lift* U *of* ξ *at* X*, then before applying* (10.32) *we must replace* Y *by* YQ*, for the unique* Q *such that there exists* $V \in \mathrm{H}_X$ *with* $\overline{\mathrm{R}}_X(V) = YQ$*. This can be done if and only if* $X^\top Y$ *is invertible. Explicitly, one can reason from Exercise 10.71 that* Q *is nothing but the polar factor of* $X^\top Y$*. Then, we can follow this procedure:*

1. *Compute* $Q \in \mathrm{O}(p)$ *via SVD, as* $Q = \tilde{U}\tilde{V}^\top$ *with* $\tilde{U}\tilde{\Sigma}\tilde{V}^\top = X^\top Y$*;*
2. *By* (10.32)*,* $\mathrm{lift}_{YQ}\big(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\big) = \mathrm{Proj}^{\mathrm{H}}_{YQ}(\mathrm{lift}_X(\xi)) = U - Y(Y^\top U)$*;*
3. *Finally,* $\mathrm{lift}_Y\big(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\big) = (U - Y(Y^\top U))Q^\top$*.*

Often times though, Y *is a point that was generated by retraction of some horizontal vector from* X*. If that retraction is the polar retraction, then using this transporter is straightforward:* $X^\top Y$ *is symmetric and positive definite, hence its polar factor is* $Q = I_p$*, and it is sufficient to compute* $U - Y(Y^\top U)$*.*

In closing, we connect the notion of transporter (used in [HGA15, §4.3]) to that of vector transport (favored in [AMS08, Def. 8.1.1]).

Definition 10.68. *A* vector transport *on a manifold* $\mathcal{M}$ *is a smooth map*

$$(x,u,v) \mapsto \mathrm{VT}_{(x,u)}(v)$$

from the Whitney sum *(which can be endowed with a smooth structure)*

$$\mathrm{T}\mathcal{M} \oplus \mathrm{T}\mathcal{M} = \{(x,u,v) : x \in \mathcal{M} \text{ and } u,v \in \mathrm{T}_x\mathcal{M}\}$$

to $\mathrm{T}\mathcal{M}$*, satisfying the following for some retraction* R *on* $\mathcal{M}$*:*

1. $\mathrm{VT}_{(x,u)}$ *is a linear map from the tangent space at* x *to the tangent space at* $\mathrm{R}_x(u)$ *for all* $(x,u) \in \mathrm{T}\mathcal{M}$*; and*
2. $\mathrm{VT}_{(x,0)}$ *is the identity on* $\mathrm{T}_x\mathcal{M}$ *for all* $x \in \mathcal{M}$*.*

Equivalently, we can define a vector transport associated to a retraction R as a smooth map $\mathrm{VT}\colon \mathrm{T}\mathcal{M} \to \mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{M})$ such that $\mathrm{VT}_{(x,u)}$ is a linear map from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_{\mathrm{R}_x(u)}\mathcal{M}$ and $\mathrm{VT}_{(x,0)}$ is the identity on $\mathrm{T}_x\mathcal{M}$. From this perspective, it is clear that a transporter T and a retraction R can be combined to define a vector transport through $\mathrm{VT}_{(x,u)} = \mathrm{T}_{\mathrm{R}_x(u)\leftarrow x}$. However, not all vector transports are of this form because in general we could have $\mathrm{VT}_{(x,u)} \neq \mathrm{VT}_{(x,w)}$ even if $\mathrm{R}_x(u) = \mathrm{R}_x(w)$, which the transporter construction does not allow. The other way around, a vector transport with associated retraction R can be used to define a transporter if we first restrict the domain such that $(x, u) \mapsto (x, \mathrm{R}_x(u))$ admits a smooth inverse (see Corollary 10.27).

Exercise 10.69. *Give a proof of Proposition 10.60.*

Exercise 10.70. *Give a proof of Proposition 10.63.*

Exercise 10.71. *With notation as in Example 10.67, show that E (10.31) is invertible. Furthermore, show that given two arbitrary representatives X and Y of $[X]$ and $[\overline{\mathrm{R}}_X(V)]$ (respectively), V is given by*

$$V = Y(X^\top Y)^{-1} - X, \tag{10.33}$$

and deduce that the inverse of E is smooth. From this formula, it is also apparent that $[Y]$ can be reached from $[X]$ if and only if $X^\top Y$ is invertible. Compare with Exercise 7.2.

Exercise 10.72. *For the orthogonal group ($\mathcal{M} = \mathrm{O}(n)$) or the group of rotations ($\mathcal{M} = \mathrm{SO}(n)$) as a Riemannian submanifold of $\mathbb{R}^{n\times n}$ (see Section 7.4), it is natural to consider the following transporter:*

$$\mathrm{T}_{Y\leftarrow X}(U) = YX^\top U, \tag{10.34}$$

where $X, Y \in \mathcal{M}$ are orthogonal matrices of size n and $U \in \mathrm{T}_X\mathcal{M}$ is such that $X^\top U$ is skew-symmetric. Show that this is indeed a transporter and that it is isometric. Then, show that this is not *parallel transport along geodesics (see Exercise 7.3). If we represent tangent vectors $U = X\Omega$ simply as their skew-symmetric part Ω, then this transporter requires no computations.*

Exercise 10.73. *Let R be a retraction on a Riemannian manifold $\mathcal{M}$. The differentiated retraction plays a special role as a link between the Riemannian gradient and Hessian of $f \colon \mathcal{M} \to \mathbb{R}$ and the (classical) gradients and Hessians of the pullbacks $\hat{f} = f \circ \mathrm{R}_x \colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}$.*

Prove the following identities [ABBC20, §6]: if f is differentiable, then

$$\mathrm{grad}\hat{f}(s) = T_s^* \mathrm{grad} f(\mathrm{R}_x(s)), \tag{10.35}$$

where $T_s = \mathrm{DR}_x(s)$ is a linear map from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_{\mathrm{R}_x(s)}\mathcal{M}$, and T_s^* is its adjoint. If f is twice differentiable, then

$$\mathrm{Hess}\hat{f}(s) = T_s^* \circ \mathrm{Hess} f(\mathrm{R}_x(s)) \circ T_s + W_s, \tag{10.36}$$

with W_s a self-adjoint linear map on $\mathrm{T}_x\mathcal{M}$ defined by

$$\langle \dot{s}, W_s(\dot{s}) \rangle_x = \langle \mathrm{grad} f(\mathrm{R}_x(s)), c''(0) \rangle_{\mathrm{R}_x(s)}, \tag{10.37}$$

where $c''(0) = \frac{\mathrm{D}}{\mathrm{d}t} c'(0)$ is the initial intrinsic acceleration of the smooth curve $c(t) = \mathrm{R}_x(s + t\dot{s})$. Argue that W_s is indeed linear and self-adjoint.

Check that these formulas generalize Propositions 8.59 and 8.71 (as well as their embedded counter-parts, Propositions 3.59 and 5.45).

As a comment: For all $u, v \in \mathrm{T}_x\mathcal{M}$*, we can use* (10.37) *to compute* $\langle u, W_s(v)\rangle_x$ *owing to the polarization identity:*

$$\langle u, W_s(v)\rangle_x = \frac{1}{4}\big(\langle u+v, W_s(u+v)\rangle_x - \langle u-v, W_s(u-v)\rangle_x \big).$$

This is why (10.37) *fully determines* W_s*.*

10.6 Finite difference approximation of the Hessian

In order to minimize a smooth function $f\colon \mathcal{M} \to \mathbb{R}$ on a Riemannian manifold, several optimization algorithms (notably those in Chapter 6) require computation of the Riemannian Hessian applied to a vector: $\mathrm{Hess}f(x)[u]$. Since obtaining an explicit expression for the Hessian may be tedious,[6] it is natural to explore avenues to approximate it numerically. To this end, we consider *finite difference approximations*.

For any smooth curve $c\colon I \to \mathcal{M}$ such that $c(0) = x$ and $c'(0) = u$, it holds that

$$\mathrm{Hess}f(x)[u] = \nabla_u \mathrm{grad}f = \frac{\mathrm{D}}{\mathrm{d}t}(\mathrm{grad}f \circ c)(0). \tag{10.38}$$

Using Proposition 10.37, we can further rewrite the right-hand side in terms of parallel transport along c:

$$\mathrm{Hess}f(x)[u] = \lim_{t\to 0} \frac{\mathrm{PT}^c_{0\leftarrow t}(\mathrm{grad}f(c(t))) - \mathrm{grad}f(x)}{t}. \tag{10.39}$$

This suggests the approximation

$$\mathrm{Hess}f(x)[u] \approx \frac{\mathrm{PT}^c_{0\leftarrow \bar{t}}(\mathrm{grad}f(c(\bar{t}))) - \mathrm{grad}f(x)}{\bar{t}} \tag{10.40}$$

for some well-chosen $\bar{t} > 0$: small enough to be close to the limit (see Corollary 10.56 to quantify this error), large enough to avoid numerical issues. Of course, we could also use higher-order finite differences.

In light of Section 10.5, we may ask: is it legitimate to replace the parallel transport in (10.40) with a transporter? We already verified this for a special case in Example 5.32, where we considered a Riemannian submanifold $\mathcal{M}$ of a Euclidean space $\mathcal{E}$ with the transporter obtained by orthogonal projection to tangent spaces. In this section, we consider the general setting.

With a transporter T on a Riemannian manifold $\mathcal{M}$, we contemplate the following candidate approximation for the Hessian:

$$\mathrm{Hess}f(x)[u] \approx \frac{\mathrm{T}_{x\leftarrow c(\bar{t})}(\mathrm{grad}f(c(\bar{t}))) - \mathrm{grad}f(x)}{\bar{t}}. \tag{10.41}$$

Implementing this formula takes little effort compared to the hassle of deriving formulas for the Hessian by hand. For example, in the Manopt toolbox,

[6] See also Section 4.7 for a word regarding automatic differentiation.

the default behavior when the Hessian is needed but unavailable is to fall back on (10.41) with $c(\bar{t}) = \mathrm{R}_x(\bar{t}u)$ and $\bar{t} > 0$ set such that $\|\bar{t}u\|_x = 2^{-14}$. This costs one retraction, one gradient evaluation (assuming $\mathrm{grad} f(x)$ is available), and one call to a transporter. Moreover, this satisfies radial linearity as required in A6.1 (p134). To justify (10.41), we generalize Proposition 10.37.

Proposition 10.74. *Let $c\colon I \to \mathcal{M}$ be a smooth curve on a Riemannian manifold equipped with a transporter* T*. For a fixed $t_0 \in I$, let $v_1, \ldots, v_d$ form a basis of $\mathrm{T}_{c(t_0)}\mathcal{M}$ and define the vector fields*

$$V_i(t) = \left(\mathrm{T}_{c(t_0) \leftarrow c(t)}\right)^{-1}(v_i). \tag{10.42}$$

Given a vector field $Z \in \mathfrak{X}(c)$, it holds that

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t_0) = \lim_{\delta \to 0} \frac{\mathrm{T}_{c(t_0) \leftarrow c(t_0+\delta)} Z(t_0 + \delta) - Z(t_0)}{\delta} + \sum_{i=1}^{d} \alpha_i(t_0) \frac{\mathrm{D}}{\mathrm{d}t} V_i(t_0),$$

where $\alpha_1(t_0), \ldots, \alpha_d(t_0)$ are the coefficients of $Z(t_0)$ in the basis $v_1, \ldots, v_d$.

Proof. The vector fields $V_1, \ldots, V_d$ play a role similar to parallel frames. By Proposition 10.62, these vector fields depend smoothly on t in a neighborhood I_0 of t_0. Furthermore, I_0 can be chosen small enough so that $V_1(t), \ldots, V_d(t)$ form a basis of $\mathrm{T}_{c(t)}\mathcal{M}$ for each $t \in I_0$. Hence, there exists a unique set of smooth functions $\alpha_i\colon I_0 \to \mathbb{R}$ such that

$$Z(t) = \sum_{i=1}^{d} \alpha_i(t) V_i(t).$$

On the one hand, using properties of covariant derivatives,

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \sum_{i=1}^{d} \alpha_i'(t) V_i(t) + \alpha_i(t) \frac{\mathrm{D}}{\mathrm{d}t} V_i(t).$$

On the other hand, defining G as

$$G(t) = \mathrm{T}_{c(t_0) \leftarrow c(t)}(Z(t)) = \sum_{i=1}^{d} \alpha_i(t) v_i,$$

we find that

$$G'(t) = \sum_{i=1}^{d} \alpha_i'(t) v_i.$$

Combining both findings at t_0 using $V_i(t_0) = v_i$, it follows that

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t_0) = G'(t_0) + \sum_{i=1}^{d} \alpha_i(t_0) \frac{\mathrm{D}}{\mathrm{d}t} V_i(t_0).$$

Since G is a map between (open subsets of) linear spaces, we can write $G'(t_0)$ as a limit in the usual way. $\square$

Applying this to (10.38) yields a corollary relevant to formula (10.41).

Corollary 10.75. *For any smooth curve c on a Riemannian manifold $\mathcal{M}$ such that $c(0) = x$ and $c'(0) = u$, and for any transporter* T, orthonormal *basis $v_1, \ldots, v_d$ of $\mathrm{T}_x\mathcal{M}$ and associated vector fields $V_1, \ldots, V_d$ defined by*

$$V_i(t) = \left(\mathrm{T}_{x \leftarrow c(t)}\right)^{-1}(v_i), \tag{10.43}$$

it holds that

$$\mathrm{Hess} f(x)[u] = \lim_{t \to 0} \frac{\mathrm{T}_{x \leftarrow c(t)}(\mathrm{grad} f(c(t))) - \mathrm{grad} f(x)}{t} + \sum_{i=1}^{d} \langle \mathrm{grad} f(x), v_i \rangle_x \frac{\mathrm{D}}{\mathrm{d}t} V_i(0). \tag{10.44}$$

Thus we see that the approximation (10.41) is justified at or near a critical point. This is typically sufficient to obtain good performance with second-order optimization algorithms such as RTR.

The approximation is also justified at a general point x if the vectors $\frac{\mathrm{D}}{\mathrm{d}t} V_i(0)$ vanish. This is of course the case if we use parallel transport, recovering (10.39). Likewise, circling back to Example 5.32 for the case where $\mathcal{M}$ is a Riemannian submanifold of a Euclidean space $\mathcal{E}$ and the transporter is taken to be simply orthogonal projection to tangent spaces (Proposition 10.66), we also get the favorable simplification. As a reminder, this yields the particularly convenient formula

$$\mathrm{Hess} f(x)[u] = \lim_{t \to 0} \frac{\mathrm{Proj}_x(\mathrm{grad} f(c(t))) - \mathrm{grad} f(x)}{t}, \tag{10.45}$$

where Proj_x is the orthogonal projector from $\mathcal{E}$ to $\mathrm{T}_x\mathcal{M}$, $c(t)$ satisfies $c(0) = x$ and $c'(0) = u$, and $\mathrm{grad} f(c(t))$ is interpreted as a vector in $\mathcal{E}$.

10.7 Tensor fields and their covariant differentiation

Given a Riemannian manifold $\mathcal{M}$, we can think of a smooth vector field $U \in \mathfrak{X}(\mathcal{M})$ as a map from $\mathfrak{X}(\mathcal{M})$ to the set of smooth real-valued functions $\mathfrak{F}(\mathcal{M})$ as follows:

$$V \mapsto U(V) = \langle U, V \rangle.$$

This map is $\mathfrak{F}(\mathcal{M})$-linear in its argument, meaning

$$U(fV + gW) = fU(V) + gU(W)$$

for all $V, W \in \mathfrak{X}(\mathcal{M})$ and $f, g \in \mathfrak{F}(\mathcal{M})$. Likewise, we can think of the Riemannian metric itself as a map from $\mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M})$ to $\mathfrak{F}(\mathcal{M})$:

$$(U, V) \mapsto \langle U, V \rangle.$$

This mapping is $\mathfrak{F}(\mathcal{M})$-linear in each of its arguments. These two maps are examples of *tensor fields*, respectively of order one and two.

Definition 10.76. *A* smooth tensor field T of order k *on a manifold* $\mathcal{M}$ *is a map*

$$T\colon \mathfrak{X}(\mathcal{M}) \times \cdots \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{F}(\mathcal{M})$$

which is $\mathfrak{F}(\mathcal{M})$*-linear in each one of its* k *inputs. The set of such objects is denoted by* $\mathfrak{X}^k(\mathcal{M})$*. If the ordering of the inputs is irrelevant, we say* T *is a* symmetric (smooth) tensor field*. (See also Remark 10.84.)*

In our examples above, vector fields are identified with tensor fields of order one, while the Riemannian metric is a symmetric tensor field of order two. As a non-example, notice that the Riemannian connection ∇, conceived of as a map

$$\nabla\colon \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{F}(\mathcal{M})\colon (U, V, W) \mapsto \langle \nabla_U V, W\rangle\,,$$

is *not* a tensor field, because it is only $\mathbb{R}$-linear in V, not $\mathfrak{F}(\mathcal{M})$-linear. Indeed, $\nabla_U(fV) = f\nabla_U V + (Uf)V$ for $f \in \mathfrak{F}(\mathcal{M})$.

Importantly, tensor fields are pointwise objects, in that they associate to each point $x \in \mathcal{M}$ a well-defined multilinear map (i.e., a *tensor*) on the tangent space $\mathrm{T}_x\mathcal{M}$—hence the name tensor *field*. In order to see this, consider a tensor field T of order k and a local frame $W_1, \ldots, W_d$ on a neighborhood $\mathcal{U}$ of x (Section 3.9). Then, the input vector fields $U_1, \ldots, U_k \in \mathfrak{X}(\mathcal{M})$ can each be expanded in the local frame as

$$U_i|_{\mathcal{U}} = f_{i,1}W_1 + \cdots + f_{i,d}W_d = \sum_{j=1}^{d} f_{i,j}W_j,$$

where the $f_{i,j}$ are smooth functions on $\mathcal{U}$. Working only on the domain $\mathcal{U}$ and using the linearity properties of tensor fields, we find[7]

$$T(U_1, \ldots, U_k) = \sum_{j_1=1}^{d} \cdots \sum_{j_k=1}^{d} f_{1,j_1} \cdots f_{k,j_k} T(W_{j_1}, \ldots, W_{j_k}).$$

Evaluating this function at $x' \in \mathcal{U}$, the result depends on $U_1, \ldots, U_k$ only through the values of the $f_{i,j}$ at x', that is: $T(U_1, \ldots, U_k)(x')$ depends on the vector fields only through $U_1(x'), \ldots, U_k(x')$. Moreover, the dependence is linear in each one.

This offers a useful perspective on tensor fields of order k: they associate to each point x of $\mathcal{M}$ a k-linear function on the tangent space at that point, namely,

$$T(x)\colon \mathrm{T}_x\mathcal{M} \times \cdots \times \mathrm{T}_x\mathcal{M} \to \mathbb{R}.$$

This function is defined by

$$T(x)(u_1, \ldots, u_k) = T(U_1, \ldots, U_k)(x),$$

[7] Via bump functions as in Section 5.6, we can extend the vector fields from $\mathcal{U}$ to all of $\mathcal{M}$, which is necessary to apply T. To be formal, we should then also argue why the conclusions we reach based on these special vector fields generalize—see [Lee12, Lem. 12.24].

where the $U_i \in \mathfrak{X}(\mathcal{M})$ are arbitrary so long as $U_i(x) = u_i$.

Continuing with our examples, for a vector field $U \in \mathfrak{X}(\mathcal{M})$, the notation $u = U(x)$ normally refers to a tangent vector at x, while if we think of U as a tensor field, then $U(x)$ denotes the linear function $v \mapsto \langle u, v\rangle_x$ on the tangent space at x. The Riemannian metric is a tensor field of order two; let us call it G. Then, $G(x)$ is a bilinear function on $\mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M}$ such that $G(x)(u, v) = \langle u, v\rangle_x$.

The map $x \mapsto T(x)$ is smooth, in a sense we now make precise. Similarly to Proposition 10.60, we can define a *tensor bundle* of order k over $\mathcal{M}$ as:

$$\begin{aligned} \mathrm{T}^k\mathrm{T}\mathcal{M} &= \left\{(x, L) : x \in \mathcal{M} \text{ and } L \in \mathrm{T}^k\mathrm{T}_x\mathcal{M}\right\}, \text{ where} \\ \mathrm{T}^k\mathrm{T}_x\mathcal{M} &= \left\{k\text{-linear functions from } (\mathrm{T}_x\mathcal{M})^k \text{ to } \mathbb{R}\right\}. \end{aligned} \tag{10.46}$$

Each tensor bundle can be endowed with a natural smooth manifold structure such that $\pi \colon \mathrm{T}^k\mathrm{T}\mathcal{M} \to \mathcal{M}$ defined by $\pi(x, L) = x$ is smooth. This is identical to how we equipped the tangent bundle $\mathrm{T}\mathcal{M}$ with such a smooth structure. Then, any section of $\mathrm{T}^k\mathrm{T}\mathcal{M}$, that is, any map T from $\mathcal{M}$ to $\mathrm{T}^k\mathrm{T}\mathcal{M}$ such that $\pi(T(x)) = x$ is called a *tensor field* of order k. This is commonly taken as the definition of a tensor field. *Smooth* tensor fields as we defined them above are exactly the smooth sections as defined here [Lee12, Prop. 12.19, Lem. 12.24].

By convention, $\mathrm{T}^0\mathrm{T}_x\mathcal{M} = \mathbb{R}$, so that $\mathrm{T}^0\mathrm{T}\mathcal{M} = \mathcal{M} \times \mathbb{R}$. Notice that $\mathrm{T}^1\mathrm{T}_x\mathcal{M}$ can be identified with $\mathrm{T}_x\mathcal{M}$, and that $\mathrm{T}^2\mathrm{T}_x\mathcal{M}$ can be identified with the set of linear maps from $\mathrm{T}_x\mathcal{M}$ into itself. Thus, $\mathrm{T}^1\mathrm{T}\mathcal{M}$ can be identified with $\mathrm{T}\mathcal{M}$ itself, and $\mathrm{T}^2\mathrm{T}\mathcal{M}$ can be identified as:

$$\mathrm{T}^2\mathrm{T}\mathcal{M} \equiv \{(x, L) : x \in \mathcal{M} \text{ and } L \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M} \text{ is linear}\}. \tag{10.47}$$

Now that we think of smooth tensor fields as smooth maps on manifolds, it is natural to ask what happens if we differentiate them. In Chapter 5, we introduced the notion of connection ∇ on the tangent bundle $\mathrm{T}\mathcal{M}$. One can formalize the idea that ∇ induces a connection on any tensor bundle, unique once we require certain natural properties [Lee18, Prop. 4.15]. This gives meaning to the notation $\nabla_V T$ for $V \in \mathfrak{X}(\mathcal{M})$. Omitting quite a few details, we give an opportunistic construction of this object.

Recall how Vf is the derivative of a real function f against a vector field V. Since $T(U_1, \ldots, U_k)$ is a smooth function on $\mathcal{M}$, we can differentiate it against any smooth vector field V to obtain $VT(U_1, \ldots, U_k)$, also a smooth function on $\mathcal{M}$. The definition below is crafted to secure a natural chain rule for this differentiation.

Definition 10.77. *Given a smooth tensor field T of order k on a manifold $\mathcal{M}$ with a connection ∇, the* total covariant derivative ∇T *of T is a smooth tensor field of order $k+1$ on $\mathcal{M}$, defined for all $U_1, \ldots, U_k, V \in \mathfrak{X}(\mathcal{M})$ by*

$$\begin{aligned} \nabla T(U_1, \ldots, U_k, V) &= VT(U_1, \ldots, U_k) \\ &\quad - \sum_{i=1}^{k} T(U_1, \ldots, U_{i-1}, \nabla_V U_i, U_{i+1}, \ldots, U_k). \end{aligned} \tag{10.48}$$

(If $\nabla T = 0$ we call T parallel.*) We also let $\nabla_V T$ be a smooth tensor field of order k (symmetric if T is symmetric), defined by*

$$(\nabla_V T)(U_1, \ldots, U_k) = \nabla T(U_1, \ldots, U_k, V).$$

(It is an exercise to check that these are indeed tensor fields.)

As an example, it is instructive to see how gradients and Hessians of scalar fields fit into the framework of covariant differentiation of tensor fields.

Example 10.78. *Let $\mathcal{M}$ be a Riemannian manifold with its Riemannian connection ∇. A smooth function $f: \mathcal{M} \to \mathbb{R}$ is a tensor field of order zero. Differentiating f as a tensor field using Definition 10.77, we find that ∇f is a tensor field of order one defined by*

$$\nabla f(U) = Uf = \mathrm{D}f(U) = \langle \mathrm{grad} f, U \rangle .$$

In other words, ∇f is the differential $\mathrm{D}f$, which we identify with the gradient vector field through the Riemannian metric. We now differentiate ∇f to produce $\nabla(\nabla f) = \nabla^2 f$: a tensor field of order two defined by

$$\begin{aligned} \nabla^2 f(U, V) &= V\nabla f(U) - \nabla f(\nabla_V U) \\ &= V\langle \mathrm{grad} f, U \rangle - \langle \mathrm{grad} f, \nabla_V U \rangle \\ &= \langle \nabla_V \mathrm{grad} f, U \rangle \\ &= \langle \mathrm{Hess} f(V), U \rangle . \end{aligned}$$

In other words, $\nabla^2 f$ and the Riemannian Hessian $\mathrm{Hess} f$ *are identified through the Riemannian metric. This also shows that $\nabla^2 f$ is symmetric. Going one step further, we differentiate $\nabla^2 f$ to produce $\nabla^3 f = \nabla(\nabla^2 f)$: a tensor field of order three defined by*

$$\begin{aligned} \nabla^3 f(U, V, W) &= W\nabla^2 f(U, V) \\ &\qquad - \nabla^2 f(\nabla_W U, V) - \nabla^2 f(U, \nabla_W V) \\ &= W\langle \mathrm{Hess} f(V), U \rangle \\ &\qquad - \langle \mathrm{Hess} f(V), \nabla_W U \rangle - \langle \mathrm{Hess} f(\nabla_W V), U \rangle \\ &= \langle \nabla_W (\mathrm{Hess} f(V)) - \mathrm{Hess} f(\nabla_W V), U \rangle . \end{aligned}$$

Notice that $\nabla^3 f$ is symmetric in its first two inputs, but not necessarily in its third input (see also Exercise 10.87). Based on the above, for a given $W \in \mathfrak{X}(\mathcal{M})$ it is useful to introduce $\nabla_W \mathrm{Hess} f : \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$: an operator of the same type as the Riemannian Hessian itself given by

$$\nabla_W \mathrm{Hess} f(V) = \nabla_W (\mathrm{Hess} f(V)) - \mathrm{Hess} f(\nabla_W V). \tag{10.49}$$

The smooth tensor field $(U, V) \mapsto \langle \nabla_W \mathrm{Hess} f(V), U \rangle = \nabla^3 f(U, V, W)$ is symmetric in U and V.

Since tensor fields are pointwise objects, we can make sense of the notation

$\nabla_v T$ for $v \in \mathrm{T}_x\mathcal{M}$ as follows: ∇T is a $(k+1)$-tensor field on $\mathcal{M}$, so that $(\nabla T)(x)$ is a $(k+1)$-linear map on $\mathrm{T}_x\mathcal{M}$; fixing the last input to be v, we are left with

$$(\nabla_v T)(u_1, \ldots, u_k) = \big((\nabla T)(x)\big)(u_1, \ldots u_k, v). \tag{10.50}$$

Thus, $\nabla_v T$ is a k-linear map on $\mathrm{T}_x\mathcal{M}$; if T is symmetric, so is $\nabla_v T$. In particular, $\nabla_v \mathrm{Hess} f$ is self-adjoint on $\mathrm{T}_x\mathcal{M}$ for $v \in \mathrm{T}_x\mathcal{M}$.

Given a curve c on $\mathcal{M}$, we defined the notion of covariant derivative of a vector field along c in Section 5.7. This extends to tensors, in direct analogy with Theorem 5.29.

Definition 10.79. *Let $c\colon I \to \mathcal{M}$ be a smooth curve on a manifold $\mathcal{M}$. A* smooth tensor field Z of order k along c *is a map*

$$Z\colon \mathfrak{X}(c) \times \cdots \times \mathfrak{X}(c) \to \mathfrak{F}(I)$$

which is $\mathfrak{F}(I)$-linear in each one of its k inputs. The set of such objects is denoted by $\mathfrak{X}^k(c)$.

Here too, we can reason that tensor fields are pointwise objects, in that $Z(t)$ is a k-linear map from $(\mathrm{T}_{c(t)}\mathcal{M})^k$ to $\mathbb{R}$: see [Lee18, Thm. 4.24, Prop. 5.15].

Theorem 10.80. *Let $c\colon I \to \mathcal{M}$ be a smooth curve on a manifold with a connection ∇. There exists a unique operator $\frac{\mathrm{D}}{\mathrm{d}t}\colon \mathfrak{X}^k(c) \to \mathfrak{X}^k(c)$ satisfying these properties for all $Y, Z \in \mathfrak{X}^k(c)$, $T \in \mathfrak{X}^k(\mathcal{M})$, $g \in \mathfrak{F}(I)$ and $a, b \in \mathbb{R}$:*

1. *$\mathbb{R}$-linearity: $\frac{\mathrm{D}}{\mathrm{d}t}(aY + bZ) = a\frac{\mathrm{D}}{\mathrm{d}t}Y + b\frac{\mathrm{D}}{\mathrm{d}t}Z$;*
2. *Leibniz rule: $\frac{\mathrm{D}}{\mathrm{d}t}(gZ) = g'Z + g\frac{\mathrm{D}}{\mathrm{d}t}Z$;*
3. *Chain rule: $\big(\frac{\mathrm{D}}{\mathrm{d}t}(T \circ c)\big)(t) = \nabla_{c'(t)}T$ for all $t \in I$.*

This operator is called the induced covariant derivative*.*

In the statement above, we understand $\nabla_{c'(t)}T$ through (10.50).

As a result of Definition 10.77, we also have the following chain rule: given

$Z \in \mathfrak{X}^k(c)$ and $U_1, \ldots, U_k \in \mathfrak{X}(c)$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\big(Z(U_1, \ldots, U_k)\big) &= \frac{\mathrm{d}}{\mathrm{d}t}\big(Z(t)(U_1(t), \ldots, U_k(t))\big) \\
&= \left(\frac{\mathrm{D}}{\mathrm{d}t} Z(t)\right)(U_1(t), \ldots, U_k(t)) \\
&\quad + Z(t)\left(\frac{\mathrm{D}}{\mathrm{d}t} U_1(t), U_2(t), \ldots, U_k(t)\right) + \cdots \\
&\quad + Z(t)\left(U_1(t), \ldots, U_{k-1}(t), \frac{\mathrm{D}}{\mathrm{d}t} U_k(t)\right) \\
&= \left(\frac{\mathrm{D}}{\mathrm{d}t} Z\right)(U_1, \ldots, U_k) \\
&\quad + Z\left(\frac{\mathrm{D}}{\mathrm{d}t} U_1, U_2, \ldots, U_k\right) + \cdots \\
&\quad + Z\left(U_1, \ldots, U_{k-1}, \frac{\mathrm{D}}{\mathrm{d}t} U_k\right).
\end{aligned} \tag{10.51}$$

Example 10.81. *If $c\colon I \to \mathcal{M}$ is a geodesic on the Riemannian manifold $\mathcal{M}$ with Riemannian connection ∇ and $f\colon \mathcal{M} \to \mathbb{R}$ is smooth, we have:*

$$\begin{aligned}
(f \circ c)' &= \nabla_{c'} f = (\nabla f \circ c)(c'), \\
(f \circ c)'' &= (\nabla_{c'} \nabla f)(c') + (\nabla f \circ c)(c'') = (\nabla^2 f \circ c)(c', c'), \\
(f \circ c)''' &= (\nabla_{c'} \nabla^2 f)(c', c') + (\nabla^2 f \circ c)(c'', c') + (\nabla^2 f \circ c)(c', c'') \\
&= (\nabla^3 f \circ c)(c', c', c').
\end{aligned}$$

In particular, if $c(0) = x$ and $c'(0) = u$, then at $t = 0$ it follows that:

$$\begin{aligned}
(f \circ c)'(0) &= \nabla f(x)(u) = \langle \mathrm{grad} f(x), u \rangle_x, \\
(f \circ c)''(0) &= \nabla^2 f(x)(u, u) = \langle \mathrm{Hess} f(x)[u], u \rangle_x, \\
(f \circ c)'''(0) &= \nabla^3 f(x)(u, u, u) = \langle (\nabla_u \mathrm{Hess} f)[u], u \rangle_x,
\end{aligned}$$

with $\nabla_u \mathrm{Hess} f$ as defined through (10.49) *and* (10.50)*.*

Circling back to Section 10.4, let us now discuss Lipschitz continuity of tensor fields. As afforded by Remark 8.6, here we do not require tensor fields to be smooth (that is, infinitely differentiable). We understand the differentiability properties of tensor fields as maps between manifolds, as outlined around (10.46). We start with a definition.

Definition 10.82. *A tensor field T of order k on a Riemannian manifold $\mathcal{M}$ with its Riemannian connection is L-*Lipschitz continuous *if for all (x, s) in the domain of the exponential map and for all $u_1, \ldots, u_k \in \mathrm{T}_x\mathcal{M}$ we have*

$$|T(\mathrm{Exp}_x(s))(P_s u_1, \ldots, P_s u_k) - T(x)(u_1, \ldots, u_k)| \le L\|s\|_x \|u_1\|_x \cdots \|u_k\|_x,$$

where P_s is parallel transport along $\gamma(t) = \mathrm{Exp}_x(ts)$ from $t = 0$ to $t = 1$.

It is an exercise to show that this definition is compatible with the ones we introduced earlier, for example for f, $\mathrm{grad} f$ and $\mathrm{Hess} f$.

Assume T is differentiable. Let c be a smooth curve on $\mathcal{M}$ satisfying $c(0) = x$ and $c'(0) = s$. Parallel transport $u_1, \ldots, u_k \in \mathrm{T}_x\mathcal{M}$ along c to form $U_i(t) = \mathrm{PT}^c_{t\leftarrow 0} u_i$ for $i = 1, \ldots, k$. Then, owing to $\frac{\mathrm{D}}{\mathrm{d}t} U_i = 0$ for all i we have (using (10.51) and the chain rule in Theorem 10.80):

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\Big((T \circ c)(U_1, \ldots, U_k)\Big) &= \left(\frac{\mathrm{D}}{\mathrm{d}t}(T \circ c)\right)(U_1, \ldots, U_k) \\ &= (\nabla_{c'} T)(U_1, \ldots, U_k) \\ &= (\nabla T \circ c)(U_1, \ldots, U_k, c'). \end{aligned} \tag{10.52}$$

In particular, at $t = 0$ this says:

$$\nabla T(x)(u_1, \ldots, u_k, s) = (\nabla_s T)(u_1, \ldots, u_k) = \frac{\mathrm{d}}{\mathrm{d}t}(T \circ c)(U_1, \ldots, U_k)\Big|_{t=0}. \tag{10.53}$$

In light of these identities, we have the following simple observations.

Proposition 10.83. *Let T be a tensor field of order k on $\mathcal{M}$ (Riemannian).*

1. *If T is differentiable and L-Lipschitz continuous, then ∇T is bounded by L, that is, for all $x \in \mathcal{M}$ and for all $u_1, \ldots u_k, s \in \mathrm{T}_x\mathcal{M}$,*
$$|\nabla T(x)(u_1, \ldots, u_k, s)| \;=\; |(\nabla_s T)(u_1, \ldots, u_k)| \;\le\; L\|s\|_x \|u_1\|_x \cdots \|u_k\|_x.$$
2. *If T is continuously differentiable and ∇T is bounded by L, then T is L-Lipschitz continuous.*

Proof. For the first claim, use (10.53), continuity of the absolute value function and Definition 10.82 with $c(t) = \mathrm{Exp}_x(ts)$ to see that

$$\begin{aligned} |\nabla T(x)(u_1, \ldots, u_k, s)| &= |(\nabla_s T)(u_1, \ldots, u_k)| \\ &= \lim_{t\to 0} \frac{|T(\mathrm{Exp}_x(ts))(P_{ts}u_1, \ldots, P_{ts}u_k) - T(x)(u_1, \ldots, u_k)|}{|t|} \\ &\le \lim_{t\to 0} \frac{1}{|t|} L\|ts\|_x \|u_1\|_x \cdots \|u_k\|_x = L\|s\|_x \|u_1\|_x \cdots \|u_k\|_x, \end{aligned}$$

as announced.

For the second claim, a bit of standard calculus and (10.52) (including notation there) yield

$$\begin{aligned} & T(c(1))(U_1(1), \ldots, U_k(1)) - T(x)(u_1, \ldots, u_k) \\ &\quad = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\Big(T(c(t))(U_1(t), \ldots, U_k(t))\Big)\mathrm{d}t \\ &\quad = \int_0^1 (\nabla T \circ c)(U_1, \ldots, U_k, c')(t)\mathrm{d}t. \end{aligned}$$

Therefore, using that ∇T is bounded by L,

$$
\begin{aligned}
&|T(c(1))(U_1(1), \ldots, U_k(1)) - T(x)(u_1, \ldots, u_k)| \\
&\quad \le \int_0^1 L\|U_1(t)\|_{c(t)} \cdots \|U_k(t)\|_{c(t)} \|c'(t)\|_{c(t)} \mathrm{d}t \\
&\quad = L\|u_1\|_x \cdots \|u_k\|_x \cdot L(c),
\end{aligned}
$$

where $L(c)$ is the length of the curve c over $[0, 1]$. The claim follows in particular by letting $c(t) = \mathrm{Exp}_x(ts)$, in which case $L(c) = \|s\|_x$. □

Remark 10.84. *The Riemannian metric establishes a one-to-one correspondence between the tangent vector $u \in \mathrm{T}_x\mathcal{M}$ and the linear map $v \mapsto \langle u, v\rangle_x$ from $\mathrm{T}_x\mathcal{M}$ to $\mathbb{R}$. In the absence of a metric, we distinguish between these two types of objects, respectively called vectors and* covectors. *When doing so, it is useful to define tensor fields as maps that transform vector fields and/or covector fields into scalar fields, leading to the notions of* covariant, contravariant and mixed tensor fields. *These terms are likely to come up in discussions of tensor fields on Riemannian manifolds as well, because they are often familiar to readers from non-Riemannian smooth geometry. See [Lee12, Ch. 12] and [Lee18, Ch. 4, Ch. 5] for details.*

Exercise 10.85. *Check that ∇T as provided by Definition 10.77 is indeed a tensor field of order $k+1$, and that if T is symmetric, then $\nabla_V T$ is a symmetric tensor field of order k.*

Exercise 10.86. *Let $\mathcal{M}$ be a Riemannian manifold and let G be the smooth tensor field of order two defined by $G(U, V) = \langle U, V\rangle$. Check that a connection ∇ on $\mathcal{M}$ is compatible with the metric if and only if G is parallel. (See also [Lee18, Prop. 5.5] for further characterizations of compatibility between the metric and the connection.)*

Exercise 10.87. *Let $f : \mathcal{M} \to \mathbb{R}$ be three times continuously differentiable on a Riemannian manifold $\mathcal{M}$ equipped with a retraction* R. *Assume*

$$
\left| f(\mathrm{R}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle_x - \frac{1}{2}\langle s, \mathrm{Hess} f(x)[s]\rangle_x \right| \le \frac{L}{6}\|s\|_x^3 \tag{10.54}
$$

for all (x, s) in a neighborhood of the zero section in the tangent bundle. (Compare with regularity assumption A6.7 on p136; see also Exercise 10.58.) Further assume R *is a* third-order retraction*, which we define to be a second-order retraction for which all curves $c(t) = \mathrm{R}_x(ts)$ with $(x, s) \in \mathrm{T}\mathcal{M}$ obey $c'''(0) = 0$, where $c''' = \frac{\mathrm{D}}{\mathrm{d}t}c''$. In particular, the exponential map is a third-order retraction (see also Exercise 10.88). Show for all $(x, s) \in \mathrm{T}\mathcal{M}$ that*

$$
|\langle (\nabla_s \mathrm{Hess} f)[s], s\rangle_x| = |\nabla^3 f(x)(s, s, s)| \le L\|s\|_x^3. \tag{10.55}
$$

In other words: the symmetric part of $\nabla^3 f(x)$ *is bounded by L.*

Conversely, show that if (10.55) *holds for all $(x, s) \in \mathrm{T}\mathcal{M}$, then* (10.54) *holds with* $\mathrm{R} = \mathrm{Exp}$ *for all (x, s) in the domain of* Exp.

We state here without details that $\nabla^3 f(x)$ may not be symmetric when x is not a critical point of f and also when $\mathcal{M}$ is not flat at x (this follows from the Ricci identity *applied to ∇f). Therefore, $\nabla^3 f(x)$ may not be bounded by L even if its symmetric part is. Consequently, the developments here do not allow us to conclude as to the Lipschitz properties of* Hess$f$ *(which would have otherwise followed from Corollary 10.52 or Proposition 10.83).*

Exercise 10.88. *Consider a Riemannian submanifold $\mathcal{M}$ of a Euclidean space $\mathcal{E}$. Given $(x, v) \in \mathrm{T}\mathcal{M}$, let $c\colon I \to \mathcal{M}$ be a smooth curve on $\mathcal{M}$ such that $c(0) = x$ and $x + tv - c(t)$ is orthogonal to $\mathrm{T}_{c(t)}\mathcal{M}$ for all t: this is the case if $c(t)$ is a curve obtained through metric projection retraction (see Section 5.12). Show that $c'''(0) = -2\mathcal{W}\big(v, \frac{\mathrm{d}}{\mathrm{d}t}c'(0)\big)$, where $\mathcal{W}$ is the Weingarten map* (5.38)*. In general, this is nonzero. Thus, we do not expect the metric projection retraction to be third order in general. Hint: use* (5.18) *to express $c'''(0)$ in terms of extrinsic derivatives of c, and simplify that expression by computing one more derivative of g in the proof of Proposition 5.55.*

Exercise 10.89. *Given a vector field V on $\mathcal{M}$ (Riemannian), let $T = \langle V, \cdot \rangle$ be the associated tensor field of order one. Show that T is L-Lipschitz continuous in the sense of Definition 10.82 if and only if V is L-Lipschitz continuous in the sense of Definition 10.44. Likewise, for each $x \in \mathcal{M}$, let $H(x)$ denote a linear map from $\mathrm{T}_x\mathcal{M}$ into itself, and let T denote the associated tensor field of order two defined through $T(x)(u, v) = \langle H(x)(u), v \rangle_x$. Show that T is L-Lipschitz continuous in the sense of Definition 10.82 if and only if H is L-Lipschitz continuous in the sense of Definition 10.49.*

10.8 Notes and references

Considering geodesics as length-minimizing curves, it is possible to generalize the concept of geodesic to arbitrary metric spaces, specifically, without the need for a smooth or Riemannian structure. See for example the monograph by Bačák [Bac14] for an introduction to *geodesic metric spaces*, and applications in convex analysis and optimization on Hadamard spaces.

For a connected manifold, there always exists a Riemannian metric which makes it complete [NO61].

Propositions 10.23 and 10.26 (and their proofs and corollaries) were designed with Eitan Levin and also discussed with Stephen McKeown. They are inspired by the proof of the Tubular Neighborhood Theorem in [Lee18, Thm. 5.25] and [Pet06, Prop. 5.18]. They notably imply that the injectivity radius function $\mathrm{inj}\colon \mathcal{M} \to \mathbb{R}$ is lower-bounded by a positive, continuous function, with (one could argue) fewer technicalities than are required to prove inj itself is continuous.

Many Riemannian geometry textbooks restrict their discussion of the injectivity radius to connected and complete manifolds. For disconnected manifolds, we defined *complete* to mean geodesically complete, i.e., each connected component is metrically complete. A function is continuous on a disconnected set if it is continuous on each connected component. Thus, continuity of inj on connected and complete manifolds implies continuity on complete manifolds. It is easy to find a published proof that the function inj: $\mathcal{M} \to \mathbb{R}$ is continuous if $\mathcal{M}$ is (connected and) complete (see for example [Lee18, Prop. 10.37]), but it has proven difficult to locate one that applies to incomplete manifolds as well. The claim appears without proof in [Cha06, Thm. III.2.3]. Following a discussion, Stephen McKeown provided[8] a proof that inj is lower-semicontinuous, then John M. Lee added a proof that inj is upper-semicontinuous, together confirming that it is continuous: see Lemmas 10.90 and 10.91 below. Both proofs rely on continuity in the complete case. They are rearranged to highlight commonalities.

Our main motivation to study continuity of inj is to reach Corollary 10.25, stating that the map $(x, y) \mapsto \mathrm{Log}_x(y)$ is smooth over the specified domain. O'Neill makes a similar statement: pick an *open* set $S \subseteq \mathcal{M}$; that set is deemed *convex* (by [O'N83, Def. 5.5]) if, for all $x \in S$, the map Exp_x is a diffeomorphism from some neighborhood of the origin in $\mathrm{T}_x\mathcal{M}$ to S. Then, $(x, v) \mapsto (x, \mathrm{Exp}_x(v))$ is a diffeomorphism from the appropriate set in $\mathrm{T}S$ to $S \times S$ [O'N83, Lem. 5.9]. This shows the map $L\colon S \times S \to \mathrm{T}S$ such that $L(x, y) \in \mathrm{T}_x\mathcal{M}$ is the initial velocity of the (unique) geodesic $\gamma\colon [0, 1] \to S$ connecting x to y is smooth: L is also a kind of inverse for the exponential (though not necessarily the same as Log).

The tool of choice to differentiate the exponential map (and the logaritmic map) is Jacobi fields [Lee18, Prop. 10.10]. Some examples of this are worked out in the context of optimization on manifolds in [CB20] and [LC20].

Parallel transporting a tangent vector u at x to all the points in a normal neighborhood [Lee18, p131] of x along geodesics through x contained in that neighborhood results in a smooth vector field. This is a well-known fact; details of the argument appear notably in [LB20, Lem. A.1].

The Riemannian notion of Lipschitz continuous gradient (and also Lipschitz continuous vector field) appears in [dCN95, Def. 3.1, p79] and [dCNdLO98, Def. 4.1], with a definition equivalent to the characterization we give in Proposition 10.45. This may be their first occurrence in an optimization context. There too, the motivation is to derive inequalities such as the ones in Proposition 10.53. Such inequalities appear often in optimization papers, see also [AMS08, §7.4], [SFF19, App. A] and [ABBC20], among many others. Lipschitz continuous Hessians appear in an optimization context in [FS02, Def. 2.2], in line with our characterization in Proposition 10.50. A general definition of Lipschitz continuous maps in tangent bundles of any order (covering tensors fields of any order) appears in [RW12], in the preliminaries on geometry. A general notion of 'fundamental theorem of calculus' for tensor fields on Riemannian manifolds, based on parallel transport, is spelled out in [ABM08, eq. (2.3)].

As an alternative to Definition 10.44, one could also endow the tangent bundle

[8] mathoverflow.net/questions/335032

with a metric space structure, so that we can then apply the standard notion of Lipschitz continuity to vector fields as maps between two metric spaces. A canonical choice would follow from the *Sasaki metric* [GHL04, §2.B.6]. See [dOF20] for a comparison of the two concepts.

The notion of vector transport appears in [AMS08, §8]. The related notion of transporter is introduced in [QGA10a] with reference to a linear structure space, and further developed in [HGA15] for general manifolds. The constructions of transporters from other transporters via inversions and adjoints are natural extensions.

Corollary 10.75 handles finite differences of the Riemannian Hessian using an arbitrary transporter (Definition 10.61). An analogous result for vector transports (Definition 10.68) appears in [AMS08, Lem. 8.2.2].

In Exercise 10.73, we contemplate the role of the initial acceleration of curves of the form $c(t) = \mathrm{R}_x(s+t\dot{s})$. Consider the special case where R is the exponential map. If $s = 0$, then c is a geodesic so that $c''(0) = 0$; but for $s \neq 0$ we expect $c''(0) \neq 0$ in general. The tangent vector $c''(0)$ and its norm are tightly related to curvature of the manifold. See [CB20, LC20] for a discussion of that vector and its effects on the Lipschitzness of pullbacks and their derivatives.

In Section 10.7, we follow do Carmo [dC92, §4.5], in that we rely on the Riemannian metric to avoid the need to distinguish between vectors and covectors, and we build up the differentiation of tensor fields by quoting the desired chain rule directly, bypassing many technical steps. This simplifies the discussion without loss of generality.

We close this section with the proofs of continuity of the injectivity radius. In these proofs, we do not need to worry about infinite values. Indeed, if $\mathrm{inj}(x) = \infty$ at some point x, then Exp_x is defined on all of $\mathrm{T}_x\mathcal{M}$. Thus, the connected component of x is complete [Lee18, Cor. 6.20], and it follows by [Lee18, Prop. 10.37] that inj is continuous on that component. (More specifically: inj is infinite at all points in that component.)

Lemma 10.90. *The injectivity radius function* $\mathrm{inj}\colon \mathcal{M} \to (0, \infty]$ *is lower-semi-continuous.*

Proof by Stephen McKeown. For contradiction, assume inj is not lower-semi-continuous at some point $x \in \mathcal{M}$. Then, there exists a sequence of points $x_0, x_1, x_2, \ldots$ on $\mathcal{M}$ such that

$$\lim_{k\to\infty} x_k = x \qquad \text{and} \qquad \forall k,\ \mathrm{inj}(x_k) \leq r < R = \mathrm{inj}(x).$$

Define $\varepsilon = \frac{R-r}{3}$ and $r' = r+\varepsilon$, $r'' = r+2\varepsilon$ so that $r < r' < r'' < R$. By definition of R, the exponential map Exp_x induces a diffeomorphism $\varphi\colon B(x, R) \to B^d(R)$: from the open geodesic ball on $\mathcal{M}$ centered at x with radius R to the open Euclidean ball in $\mathbb{R}^d$ centered at the origin with radius R, where $d = \dim \mathcal{M}$. (ϕ is constructed from the inverse of Exp_x followed by a linear isometry from $\mathrm{T}_x\mathcal{M}$ to $\mathbb{R}^d$.) Let g denote the Riemannian metric on $\mathcal{M}$, let $\tilde{g}$ denote the pushforward of $g|_{B(x,R)}$ to $B^d(R)$ (through φ), and let g_0 denote the Euclidean metric on $\mathbb{R}^d$. Consider a smooth bump function $\chi\colon \mathbb{R}^d \to [0,1]$ whose value is 1 on the closed ball $\bar{B}^d(r'')$ and with support in $\bar{B}^d(R)$ [Lee12, Prop. 2.25]. Then,

$$\hat{g} = \chi\tilde{g} + (1-\chi)g_0$$

is a Riemannian metric on $\mathbb{R}^d$ such that $\hat{\mathcal{M}} = (\mathbb{R}^d, \hat{g})$ is a (connected) complete Riemannian manifold [Lee18, Pb. 6-10].

Consequently, the injectivity radius function $\hat{\mathrm{inj}}\colon \hat{\mathcal{M}} \to (0,\infty]$ is continuous [Lee18, Prop. 10.37]. Furthermore, since the metrics g and $\hat{g}$ agree (through φ) on $\bar{B}(x,r'')$ and $\bar{B}^d(r'')$, we deduce that for all $y \in \mathcal{M}$ and $\rho > 0$ it holds that

$$B(y,\rho) \subseteq B(x,r'') \qquad \Longrightarrow \qquad \begin{cases} \hat{\mathrm{inj}}(\hat{y}) = \mathrm{inj}(y) & \text{if } \mathrm{inj}(y) < \rho, \\ \hat{\mathrm{inj}}(\hat{y}) \geq \rho & \text{otherwise,} \end{cases}$$

where $\hat{y} \in \mathbb{R}^d$ is the image of y through φ.

We use this fact in two ways:

1. $B(x,r'') \subseteq B(x,r'')$ and $\mathrm{inj}(x) = R > r''$, hence $\hat{\mathrm{inj}}(\hat{x}) \geq r''$, and
2. There exists k_0 large enough such that, for all $k \geq k_0$, $\mathrm{dist}(x_k, x) < \varepsilon$, so that $B(x_k, r') \subset B(x, r'')$. Moreover, $\mathrm{inj}(x_k) \leq r < r'$, so that $\hat{\mathrm{inj}}(\hat{x}_k) = \mathrm{inj}(x_k) \leq r$ for all $k \geq k_0$.

Together with the fact that $\hat{\mathrm{inj}}$ is continuous, these yield:

$$r < r'' \leq \hat{\mathrm{inj}}(\hat{x}) = \lim_{k\to\infty} \hat{\mathrm{inj}}(\hat{x}_k) \leq r,$$

a contradiction. □

Lemma 10.91. *The injectivity radius function* $\mathrm{inj}\colon \mathcal{M} \to (0,\infty]$ *is upper-semicontinuous.*

Proof by John M. Lee. The proof parallels that of lower-semicontinuity. For contradiction, assume inj is not upper-semicontinuous at some point $x \in \mathcal{M}$. Then, there exists a sequence of points $x_0, x_1, x_2, \ldots$ on $\mathcal{M}$ such that

$$\lim_{k\to\infty} x_k = x \qquad \text{and} \qquad \forall k,\ \mathrm{inj}(x_k) \geq R > r = \mathrm{inj}(x).$$

Define $\varepsilon = \frac{R-r}{3}$ and $r' = r + \varepsilon$, $r'' = r + 2\varepsilon$ so that $r < r' < r'' < R$. Because the injectivity radius at x is smaller than at the points in the sequence, it is not enough to consider Exp_x to setup a diffeomorphism with a ball in $\mathbb{R}^d$. Instead, we pick a special point in the sequence to act as a center. Let k_0 be large enough so that $\mathrm{dist}(x_k, x) < \varepsilon/2 < \varepsilon$ for all $k \geq k_0$. By triangular inequality, we also have $\mathrm{dist}(x_k, x_{k_0}) < \varepsilon$ for all $k \geq k_0$. Now, use $\mathrm{Exp}_{x_{k_0}}$ to setup a diffeomorphism $\varphi\colon B(x_{k_0}, R) \to B^d(R)$: we can do this since $\mathrm{inj}(x_{k_0}) \geq R$. Let g denote the Riemannian metric on $\mathcal{M}$, let $\tilde{g}$ denote the pushforward of $g|_{B(x_{k_0},R)}$ to $B^d(R)$ (through φ), and let g_0 denote the Euclidean metric on $\mathbb{R}^d$. Consider a smooth bump function $\chi \colon \mathbb{R}^d \to [0,1]$ whose value is 1 on the closed ball $\bar{B}^d(r'')$ and with support in $\bar{B}^d(R)$. Then,

$$\hat{g} = \chi\tilde{g} + (1-\chi)g_0$$

is a Riemannian metric on $\mathbb{R}^d$ such that $\hat{\mathcal{M}} = (\mathbb{R}^d, \hat{g})$ is a (connected) complete Riemannian manifold.

Consequently, the injectivity radius function $\hat{\mathrm{inj}} \colon \hat{\mathcal{M}} \to (0,\infty]$ is continuous. Furthermore, since the metrics g and $\hat{g}$ agree (through φ) on $\bar{B}(x_{k_0}, r'')$ and $\bar{B}^d(r'')$, we deduce that for all $y \in \mathcal{M}$ and $\rho > 0$ it holds that

$$B(y,\rho) \subseteq B(x_{k_0}, r'') \qquad \Longrightarrow \qquad \begin{cases} \hat{\mathrm{inj}}(\hat{y}) = \mathrm{inj}(y) & \text{if } \mathrm{inj}(y) < \rho, \\ \hat{\mathrm{inj}}(\hat{y}) \geq \rho & \text{otherwise,} \end{cases}$$

where $\hat{y} \in \mathbb{R}^d$ is the image of y through φ.

We use this fact in two ways:

1. $B(x, r') \subset B(x_{k_0}, r'')$ and $\mathrm{inj}(x) = r < r'$, hence $\hat{\mathrm{inj}}(\hat{x}) = \mathrm{inj}(x) = r$, and
2. For all $k \geq k_0$, $B(x_k, r') \subset B(x_{k_0}, r'')$ and $\mathrm{inj}(x_k) \geq R > r'$, thus $\hat{\mathrm{inj}}(\hat{x}_k) \geq r'$.

Together with the fact that $\hat{\mathrm{inj}}$ is continuous, these yield:

$$r = \mathrm{inj}(x) = \hat{\mathrm{inj}}(\hat{x}) = \lim_{k\to\infty} \hat{\mathrm{inj}}(\hat{x}_k) \geq r' > r,$$

a contradiction. $\square$

11 Geodesic convexity

In this chapter, we discuss elementary notions of convexity for optimization on manifolds. In so doing, we resort to notions of Riemannian distance, geodesics and completeness as covered in Section 10.1. At times, we also use the exponential map introduced in Section 10.2 and the concept of Lipschitz continuity from Section 10.4.

The study of convexity on Riemannian manifolds, called *geodesic convexity*, predates optimization on manifolds. In the context of optimization, it attracted a lot of attention as soon as the 70s. Excellent reference books on this topic include one by Udrişte [Udr94] and another by Rapcsák [Rap97].

There is some variation in how geodesically convex sets are defined by different authors. This is partly because the needs for convexity may differ depending on usage. We favor the permissive definition of Rapcsák [Rap97, §6] and relate it to two other popular definitions in Section 11.3.

All three definitions turn out to be equivalent for complete, simply connected Riemannian manifolds with nonpositive curvature. Those are called *Cartan–Hadamard manifolds.* They provide the most favorable playground for geodesic convexity, including Euclidean spaces, hyperbolic spaces (Section 7.6), the positive orthant $\mathbb{R}^n_+$ (Section 11.6) and the set of positive definite matrices $\mathrm{Sym}(n)^+$ (Section 11.7), all with the appropriate Riemannian metrics.

Applications of geodesically convex optimization notably include covariance matrix estimation [Wie12, NSAY$^+$19], Gaussian mixture modeling [HS15, HS19], matrix square root computation [Sra16], metric learning [ZHS16], statistics and averaging on manifolds [Moa03, Moa05, Fle13], a whole class of optimization problems called geometric programming [BKVH07], operator scaling in relation to the Brascamp–Lieb constant [Vis18, AZGL$^+$18], integrative PCA and matrix normal models [TA21, FORW21] and Tyler-M estimation [FM20].

Zhang and Sra analyze a collection of algorithms specifically designed for geodesically convex optimization [ZS16], providing worst-case iteration complexity results. We discuss gradient descent in Section 11.5. See also Section 11.8 for references to literature about the possibility of accelerating gradient descent on manifolds.

11.1 Convex sets and functions in linear spaces

Recall that a subset S of a linear space $\mathcal{E}$ is a *convex set* if for all x, y in S the line segment $t \mapsto (1-t)x + ty$ for $t \in [0,1]$ is in S. Furthermore,[1] $f\colon S \to \mathbb{R}$ is a *convex function* if S is convex and for all $x, y \in S$ we have:

$$\forall t \in [0,1], \qquad f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

Likewise, f is *strictly convex* if for $x \neq y$ we have

$$\forall t \in (0,1), \qquad f((1-t)x + ty) < (1-t)f(x) + tf(y).$$

If $\mathcal{E}$ is a Euclidean space with norm $\|\cdot\|$, we say f is *μ-strongly convex* for some $\mu > 0$ if $x \mapsto f(x) - \frac{\mu}{2}\|x\|^2$ is convex, or equivalently, if for all $x, y \in S$ and $t \in [0,1]$ it holds that:

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) - \frac{t(1-t)\mu}{2}\|x-y\|^2. \tag{11.1}$$

In optimization, our main reason to care about convex functions is that their local minimizers (if any exist) are global minimizers.

An equivalent way of defining convex functions is to define convexity for one-dimensional functions first. Then, $f\colon S \to \mathbb{R}$ is convex if and only if f is convex when restricted to all line segments in the convex set S, that is: for all x, y distinct in S, the composition $f \circ c\colon [0,1] \to \mathbb{R}$ is convex with $c(t) = (1-t)x + ty$. A similar statement holds for strict and strong convexity.

We adopt the latter perspective in the next section to generalize beyond linear spaces. To that end, a few basic facts about one-dimensional convex functions come in handy.

Lemma 11.1. *Let $g\colon I \to \mathbb{R}$ be defined on a connected set $I \subseteq \mathbb{R}$.*

1. *If g is convex, then g is continuous in the interior of I, denoted* $\operatorname{int} I$.
2. *If g is differentiable on I:*[2]
   (a) *g is convex if and only if $g(y) \geq g(x) + (y-x)g'(x)$ for all $x, y \in I$.*
   (b) *g is strictly convex if and only if $g(y) > g(x) + (y-x)g'(x)$ for all $x, y \in I$ distinct.*
   (c) *g is μ-strongly convex if and only if $g(y) \geq g(x) + (y-x)g'(x) + \frac{\mu}{2}(y-x)^2$ for all $x, y \in I$.*
3. *If g is continuously differentiable on I and twice differentiable on* $\operatorname{int} I$*:*
   (a) *g is convex if and only if $g''(x) \geq 0$ for all $x \in \operatorname{int} I$.*
   (b) *g is strictly convex if (but not only if) $g''(x) > 0$ for all $x \in \operatorname{int} I$.*
   (c) *g is μ-strongly convex if and only if $g''(x) \geq \mu$ for all $x \in \operatorname{int} I$.*

Proof. For the first point, see [HUL01, p15]. A proof of the second and third points follows for convenience. Note that it is allowed for I not to be open.

[1] With some care, we may allow f to take on infinite values [Roc70, HUL01].
[2] If I is not open, we mean differentiable in the sense that there exists an extension $\bar{g}$ of g differentiable on a neighborhood of I. Then, $g'(x) \triangleq \bar{g}'(x)$ for all $x \in I$.

2. (a) Assume the inequalities hold. Then, for $x, y \in I$ and $t \in [0,1]$ arbitrary, define $z = (1-t)x + ty$. Both of the following inequalities hold:

$$g(x) \geq g(z) + (x-z)g'(z), \qquad g(y) \geq g(z) + (y-z)g'(z).$$

Add them up with weights $1-t$ and t, respectively:

$$\begin{aligned}(1-t)g(x) + tg(y) &\geq g(z) + \big((1-t)(x-z) + t(y-z)\big)g'(z) \\ &= g(z) \\ &= g((1-t)x + ty).\end{aligned}$$

This shows g is convex. The other way around, if g is convex, then for all $x, y \in I$ and $t \in (0,1]$ we have

$$\begin{aligned}g(x + t(y-x)) &= g((1-t)x + ty) \\ &\leq (1-t)g(x) + tg(y) = g(x) + t(g(y) - g(x)).\end{aligned}$$

Move $g(x)$ to the left-hand side and divide by t to find:

$$g(y) \geq g(x) + \frac{g(x + t(y-x)) - g(x)}{t}.$$

Since this holds for all x, y, t as prescribed and since g is differentiable at x, we can take the limit for $t \to 0$ and conclude that the sought inequalities hold.

(b) Assume the strict inequalities hold. Then, for all $x, y \in I$ distinct and for all $t \in (0,1)$, define $z = (1-t)x + ty$; we have:

$$g(x) > g(z) + (x-z)g'(z), \qquad g(y) > g(z) + (y-z)g'(z).$$

Multiply by $1-t$ and t respectively, and add them up:

$$(1-t)g(x) + tg(y) > g(z) = g((1-t)x + ty),$$

which shows g is strictly convex. The other way around, assume g is strictly convex: it lies strictly below its chords, that is, for all x, y distinct in I,

$$\forall t \in (0,1), \qquad g((1-t)x + ty) < (1-t)g(x) + tg(y).$$

Since g is convex, it also lies above its first-order approximations:

$$\forall t \in [0,1], \qquad g(x + t(y-x)) \geq g(x) + t(y-x)g'(x).$$

The left-hand sides coincide, so that combining we find:

$$\forall t \in (0,1), \qquad (1-t)g(x) + tg(y) > g(x) + t(y-x)g'(x).$$

Subtract $g(x)$ on both sides and divide by t to conclude.

(c) By definition, g is μ-strongly convex if and only if $h(x) = g(x) - \frac{\mu}{2}x^2$ is convex, and we just showed that the latter is convex if and only if $h(y) \geq h(x) + (y-x)h'(x)$ for all $x, y \in I$, which is equivalent to the claim.

3. Taylor's theorem applies to g: for all x, y distinct in I, there exists z strictly between x and y such that

$$g(y) = g(x) + (y - x)g'(x) + \frac{1}{2}(y - x)^2 g''(z). \tag{11.2}$$

   (a) If $g''(z) \geq 0$ for all $z \in \operatorname{int} I$, then $g(y) \geq g(x) + (y - x)g'(x)$ for all $x, y \in I$ by (11.2), hence g is convex. The other way around, if g is convex, then for all x, y in I we have:

$$g(y) \geq g(x) + (y - x)g'(x) \qquad \text{and} \qquad g(x) \geq g(y) + (x - y)g'(y).$$

   Rearrange and combine to find

$$(y - x)g'(y) \geq g(y) - g(x) \geq (y - x)g'(x).$$

   We deduce that $y \geq x$ implies $g'(y) \geq g'(x)$, that is: g' is nondecreasing on I. For all $x \in \operatorname{int} I$, consider the following limit, where y goes to x while remaining in I:

$$0 \leq \lim_{y \to x} \frac{g'(y) - g'(x)}{y - x} = g''(x).$$

   This shows $g''(x) \geq 0$ for all x in $\operatorname{int} I$. (The same argument also shows that if g is twice differentiable on I and I has any boundary points, then g'' is also nonnegative on those points.)

   (b) If $g''(z) > 0$ for all $z \in \operatorname{int} I$, then $g(y) > g(x) + (y - x)g'(x)$ for all $x, y \in I$ distinct by (11.2), hence g is strictly convex. The converse is not true: $g(x) = x^4$ is smooth and strictly convex on $\mathbb{R}$, yet $g''(0) = 0$.

   (c) By definition, g is μ-strongly convex if and only if $h(x) = g(x) - \frac{\mu}{2}x^2$ is convex, and we just showed that the latter is convex if and only if $h''(x) \geq 0$ for all $x \in I$, which is equivalent to the claim. $\square$

11.2 Geodesically convex sets and functions

In this section, we present a classical generalization of convexity to Riemannian manifolds. The main idea is to use geodesic segments instead of line segments. There are, however, a number of subtly different ways one can do this, due to the fact that, in contrast to line segments in Euclidean spaces, geodesics connecting pairs of points may not exist, may not be unique, and may not be minimizing. See Section 11.3 for a discussion of popular alternative definitions.

Definition 11.2. *A subset S of a Riemannian manifold $\mathcal{M}$ is* geodesically convex *if, for every $x, y \in S$, there exists a geodesic segment $c\colon [0, 1] \to \mathcal{M}$ such that $c(0) = x$, $c(1) = y$ and $c(t)$ is in S for all $t \in [0, 1]$.*

In this definition, c is a geodesic for $\mathcal{M}$, not necessarily for S (which may or may not be a manifold). In particular, singletons and the empty set are geodesically convex.

If $\mathcal{M}$ is a Euclidean space, then a subset is convex in the usual sense if and only if it is geodesically convex, because the only geodesic connecting x to y (up to reparameterization) is $c(t) = (1-t)x + ty$.

By Theorem 10.9, a connected and complete Riemannian manifold is geodesically convex. This includes spheres, the Stiefel manifold $\mathrm{St}(n,p)$ for $p < n$, and the group of rotations $\mathrm{SO}(n)$—but we will soon see that such compact manifolds are not interesting for convexity unless we restrict our attention to subsets. More interestingly, any hemisphere of S^{n-1}, open or closed, is a geodesically convex subset of S^{n-1}. The hyperbolic space discussed in Section 7.6 is connected and complete hence geodesically convex. Likewise, the manifold of positive real numbers, $\mathbb{R}_+ = \{x > 0\}$, equipped with the metric $\langle u, v \rangle_x = \frac{uv}{x^2}$, is connected and complete, hence geodesically convex. We consider two generalizations of the latter in Sections 11.6 and 11.7 to handle $\mathbb{R}^n_+$ and $\mathrm{Sym}(n)^+$, that is, entrywise positive vectors and positive definite matrices respectively.

For a subset S of a manifold $\mathcal{M}$, we say a curve c on $\mathcal{M}$ *connects x to y in S* if it is continuous, $c(0) = x$, $c(1) = y$ and $c(t)$ is in S for all $t \in [0,1]$.

In a geodesically convex set S, any two points are connected in S by at least one geodesic segment c. Composing a function $f : S \to \mathbb{R}$ with c yields a real function on $[0,1]$. If all of these compositions are convex in the usual sense, we say f is convex in a geometric sense. Note that we do not require f to be smooth or even continuous.

Definition 11.3. *A function $f : S \to \mathbb{R}$ is* geodesically (strictly) convex *if S is geodesically convex and $f \circ c : [0,1] \to \mathbb{R}$ is (strictly) convex for each geodesic segment $c : [0,1] \to \mathcal{M}$ whose image is in S (with $c(0) \neq c(1)$).*

In the above definition, we are tacitly referring to the Riemannian structure on $\mathcal{M}$ for which $S \subseteq \mathcal{M}$ is geodesically convex and for which the curves c are geodesics. Here too, if $\mathcal{M}$ is a Euclidean space, we recover the standard notion of (strictly) convex function.

In other words, for S a geodesically convex set, we say $f : S \to \mathbb{R}$ is geodesically convex if for all $x, y \in S$ and all geodesics c connecting x to y in S the function $f \circ c : [0,1] \to \mathbb{R}$ is convex, that is,

$$\forall t \in [0,1], \qquad f(c(t)) \leq (1-t) f(x) + t f(y). \tag{11.3}$$

If additionally whenever $x \neq y$ we have

$$\forall t \in (0,1), \qquad f(c(t)) < (1-t) f(x) + t f(y), \tag{11.4}$$

then we say f is geodesically strictly convex.

Definition 11.4. *We say $f : S \to \mathbb{R}$ is* geodesically (strictly) concave *if $-f$ is geodesically (strictly) convex, and f is* geodesically linear *if it is both geodesically convex and concave.*

We also extend the notion of strong convexity, in analogy with (11.1). Recall that the length of a curve segment was defined in Section 10.1.

Definition 11.5. *A function* $f: S \to \mathbb{R}$ *is* geodesically μ-strongly convex *for some* $\mu > 0$ *if the set* S *is geodesically convex and for each geodesic segment* $c: [0,1] \to \mathcal{M}$ *whose image is in* S *we have*

$$f(c(t)) \leq (1-t)f(c(0)) + tf(c(1)) - \frac{t(1-t)\mu}{2} L(c)^2,$$

where $L(c) = \|c'(0)\|_{c(0)}$ *is the length of the geodesic segment. This is equivalent to the requirement that* $f \circ c: [0,1] \to \mathbb{R}$ *be* $\mu L(c)^2$*-strongly convex in the usual sense.*

Clearly, geodesic strong convexity implies geodesic strict convexity.

Same as for standard convexity in linear spaces, geodesic convexity ensures that local minimizers, if they exist, are global minimizers.

Theorem 11.6. *If* $f: S \to \mathbb{R}$ *is geodesically convex, then any local minimizer is a global minimizer.*

Proof. For contradiction, assume $x \in S$ is a local minimizer that is not a global minimizer. Then, there exists $y \in S$ such that $f(y) < f(x)$. There also exists a geodesic c connecting $c(0) = x$ to $c(1) = y$ in S such that, for all $t \in (0,1]$,

$$f(c(t)) \leq (1-t)f(x) + tf(y) = f(x) + t(f(y) - f(x)) < f(x),$$

which contradicts the claim that x is a local minimizer. □

Strict convexity yields uniqueness of minimizers, when they exist.

Theorem 11.7. *If* $f: S \to \mathbb{R}$ *is geodesically strictly convex, then it admits at most one local minimizer, which is necessarily the global minimizer.*

Proof. From Theorem 11.6, we know that any local minimizer is a global minimizer. Assume for contradiction that there exist two distinct global minimizers, x and y, so that $f(x) = f(y) = f_\star$. There exists a geodesic c connecting them in S such that, for $t \in (0,1)$,

$$f(c(t)) < (1-t)f(x) + tf(y) = f_\star,$$

which contradicts global optimality of x and y. □

The sublevel sets of geodesically convex functions are geodesically convex.[3] Moreover, the intersection of such sublevel sets is also geodesically convex. However, the intersection of arbitrary geodesically convex sets is not necessarily geodesically convex: see Section 11.3.

Proposition 11.8. *Let* $i \in I$ *index an arbitrary collection of geodesically convex functions* $f_i: S \to \mathbb{R}$ *and scalars* $\alpha_i \in \mathbb{R}$*. Define the sublevel sets*

$$S_i = \{x \in S : f_i(x) \leq \alpha_i\}.$$

[3] The converse does not hold: a function on a geodesically convex set is called *geodesically quasiconvex* if all of its sublevel sets are geodesically convex [Rap97, Lem. 13.1.1].

Their intersection $S' = \cap_{i \in I} S_i$ is geodesically convex. In particular, the sublevel sets of one geodesically convex function f are geodesically convex sets, and the set of global minimizers of f is geodesically convex.

Proof. The claim is clear if S' is empty. Assume it is not. Pick an arbitrary pair of points $x, y \in S'$, and an arbitrary geodesic $c \colon [0,1] \to \mathcal{M}$ connecting x to y in S: there exists at least one because x, y are in S and S is geodesically convex. For every $i \in I$ and for all $t \in [0,1]$, it holds that

$$f_i(c(t)) \leq (1-t)f_i(x) + tf_i(y) \leq (1-t)\alpha_i + t\alpha_i = \alpha_i,$$

where we used the fact that x and y belong to S_i. We conclude that $c(t)$ belongs to each S_i for all $t \in [0,1]$, so that c is in fact a geodesic connecting x and y in S'. Thus, S' is geodesically convex. $\square$

Here is a take-away from Proposition 11.8. Let S be a geodesically convex set on $\mathcal{M}$. If $f, f_1, \ldots, f_m$ are geodesically convex functions on S and $g_1, \ldots, g_p$ are geodesically linear functions on S (Definition 11.4), then for arbitrary reals $\alpha_1, \ldots, \alpha_m$ and $\beta_1, \ldots, \beta_p$ we call

$$\min_{x \in S} f(x) \qquad \text{subject to} \qquad \begin{aligned} & f_i(x) \leq \alpha_i \text{ for } i = 1, \ldots, m, \\ & g_j(x) = \beta_j \text{ for } j = 1, \ldots, p \end{aligned} \tag{11.5}$$

a *geodesically convex program*. Since the constraint $g_j(x) = \beta_j$ is equivalent to the two constraints $g_j(x) \leq \beta_j$ and $-g_j(x) \leq -\beta_j$, and since both g_j and $-g_j$ are geodesically convex, it follows that the set S' of points which satisfy all constraints in (11.5) is geodesically convex. Therefore, any local minimizer of $f|_{S'}$ is a global minimizer of $f|_{S'}$.

A connected, complete Riemannian manifold is a geodesically convex set. Its *interior*[4] is the whole manifold itself. Consider that observation with the following fact [Rap97, Thm. 6.1.8], [Udr94, Thm. 3.6].

Proposition 11.9. *If $f \colon S \to \mathbb{R}$ is geodesically convex, then f is continuous on the interior of S.*

Since compact manifolds are complete, and continuous functions on compact sets attain their maximum, we have the following corollary.

Corollary 11.10. *If $\mathcal{M}$ is a connected, compact Riemannian manifold and $f \colon \mathcal{M} \to \mathbb{R}$ is geodesically convex, then f is constant.*

The take-away is that on compact manifolds geodesic convexity is only interesting on subsets of a connected component.

When a geodesically convex function admits a *maximizer* (it may admit none, one or many), this maximizer typically occurs on the boundary of the geodesically convex domain. Indeed, a maximizer occurs 'inside' the domain only in

[4] The interior of a subset S of a manifold $\mathcal{M}$ is the union of all subsets of S open in $\mathcal{M}$.

uninteresting situations. We formalize this below, with a definition and a proposition. The definition is an extension from the classical case [Roc70, Thm. 6.4]. It is helpful to picture a two-dimensional triangle or disk in $\mathbb{R}^3$.

Definition 11.11. *Let S be a geodesically convex set on a Riemannian manifold $\mathcal{M}$. The* relative interior *of S, denoted by* $\operatorname{relint} S$, *is the set of points $x \in S$ with the following property: for all $y \in S$, all geodesics $c\colon [0,1] \to \mathcal{M}$ connecting $x = c(0)$ to $y = c(1)$ in S can be extended to the domain $[-\varepsilon, 1]$ for some $\varepsilon > 0$, and still be geodesics of $\mathcal{M}$ with image in S.*

Proposition 11.12. *Let $f\colon S \to \mathbb{R}$ be geodesically convex. If f attains its maximum at a point x in the relative interior of S, then f is constant on S.*

Proof. Pick an arbitrary point $y \in S$. Our goal is to show $f(y) = f(x)$. Consider any geodesic $c\colon [0,1] \to \mathcal{M}$ connecting $x = c(0)$ to $y = c(1)$ in S. Since x is in the relative interior of S, we can extend the domain of c to $[-\varepsilon, 1]$ for some $\varepsilon > 0$, and it is still a geodesic in S. Let $z = c(-\varepsilon)$. Since f is geodesically convex on S, $f \circ c$ is convex and we deduce:

$$f(x) \leq \frac{1}{1+\varepsilon} f(z) + \frac{\varepsilon}{1+\varepsilon} f(y).$$

Multiply by $1 + \varepsilon$; since x is a maximizer, $f(z) \leq f(x)$ and we find:

$$\varepsilon f(x) \leq \varepsilon f(y).$$

Since ε is positive, we deduce $f(x) \leq f(y)$. But x is a maximizer hence $f(x) \geq f(y)$. It follows that $f(x) = f(y)$, as announced. $\square$

Exercise 11.13. *Assume f and g are geodesically convex on the set S. Show that $x \mapsto \max(f(x), g(x))$ is geodesically convex on S. Further show that $x \mapsto \alpha f(x) + \beta g(x)$ is geodesically convex on S for all $\alpha, \beta \geq 0$.*

Exercise 11.14. *Let $f\colon S \to \mathbb{R}$ be geodesically convex. Show that if $h\colon \mathbb{R} \to \mathbb{R}$ is nondecreasing and convex, then $h \circ f$ is geodesically convex on S.*

Exercise 11.15. *Let S_1 be a geodesically convex set on a Riemannian manifold $\mathcal{M}_1$, and similarly for S_2 on $\mathcal{M}_2$. Verify that $S_1 \times S_2$ is geodesically convex on the Riemannian product manifold $\mathcal{M}_1 \times \mathcal{M}_2$.*

11.3 Alternative definitions of geodesically convex sets*

Definition 11.2 for a geodesically convex set S is the one preferred by Rapcsák, well suited for optimization purposes [Rap91, Def. 6.1.1]. It is rather permissive: it merely requires that every pair of points $x, y \in S$ be connected by some geodesic segment in the set. It does not require all geodesic segments connecting x and y to stay in S, nor does it require uniqueness of such a segment, nor that there exist a minimizing geodesic segment connecting x and y and that this one stay in S: all properties we have in Euclidean spaces.

This permissive definition still allows us to establish most optimization results we may desire, but it does have some undesirable effects. For example, the intersection of two geodesically convex sets may fail to be geodesically convex (notwithstanding Proposition 11.8). Indeed, let $\mathcal{M} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$ be the unit circle as a Riemannian submanifold of $\mathbb{R}^2$, and consider $S_1 = \{x \in \mathcal{M} : x_1 \geq 0\}$ and $S_2 = \{x \in \mathcal{M} : x_1 \leq 0\}$. Clearly, S_1 and S_2 are geodesically convex but their intersection $S_1 \cap S_2 = \{(0, 1), (0, -1)\}$ is not.

A common way to restrict Definition 11.2 is to require all geodesic segments connecting points in S to stay in S. Udrişte [Udr94, Def. 1.3] and Sakai [Sak96, Def. IV.5.1] call this total convexity, up to the following minor points: Udrişte tacitly requires $\mathcal{M}$ to be complete (instead, we here require existence of at least one geodesic segment connecting each pair of points in S), and Sakai requires S to be non-empty (we do not).

Definition 11.16. *A subset S of a Riemannian manifold $\mathcal{M}$ is* geodesically totally convex *if, for every $x, y \in S$, there is at least one geodesic segment $c\colon [0, 1] \to \mathcal{M}$ such that $c(0) = x$ and $c(1) = y$, and, for all such segments, $c(t)$ is in S for all $t \in [0, 1]$.*

Another way to restrict Definition 11.2 is to require each pair of points in S to be connected by a unique minimizing geodesic segment, and for that segment to stay in S. (Recall Theorem 10.4 for minimizing geodesics.) Lee calls such sets geodesically convex [Lee18, p166], whereas Sakai calls them strongly convex [Sak96, Def. IV.5.1]. We use the latter name.

Definition 11.17. *A subset S of a Riemannian manifold $\mathcal{M}$ is* geodesically strongly convex *if, for every $x, y \in S$, there exists a unique minimizing geodesic segment $c\colon [0, 1] \to \mathcal{M}$ such that $c(0) = x$ and $c(1) = y$; and $c(t)$ is in S for all $t \in [0, 1]$.*

More verbosely, Definition 11.17 requires the following: given $x, y \in S$ arbitrary, consider all geodesic segments $c\colon [0, 1] \to \mathcal{M}$ connecting $c(0) = x$ to $c(1) = y$ in $\mathcal{M}$; we must have that exactly one of those segments is minimizing, and moreover that this minimizing geodesic segment lies entirely in S.

If a set S is geodesically totally convex or geodesically strongly convex, then it is also geodesically convex. It is an exercise to show that (a) neither converse is true, and (b) total convexity does not imply strong convexity, nor the other way around.

Notwithstanding, for special manifolds all three notions of convexity are equivalent. The following result applies to Cartan–Hadamard manifolds.

Theorem 11.18. *Assume $\mathcal{M}$ is a complete Riemannian manifold such that each pair of points $x, y \in \mathcal{M}$ is connected by a unique geodesic segment. Then, the notions of geodesic convexity, geodesic total convexity and geodesic strong convexity are equivalent and can be stated as: $S \subseteq \mathcal{M}$ is geodesically convex if for all x, y in S the geodesic segment connecting them stays in S.*

Proof. By assumption, $\mathcal{M}$ is connected and complete. Theorem 10.9 provides that each pair of points x, y is connected by a minimizing geodesic segment. Still by assumption, no other geodesic segment connects x and y. Thus, the geodesic segment is minimizing. □

For each point x on a Riemannian manifold, there exists a positive $r_0 > 0$ such that every geodesic ball of radius $r \leq r_0$ centered at x is geodesically strongly convex [Lee18, Thm. 6.17].

The notion of geodesically convex function $f : S \to \mathbb{R}$ as in Definition 11.3 extends verbatim with S a geodesically totally or strongly convex set, and we still call these functions geodesically convex.

Exercise 11.19. *Let $\mathcal{M}$ be the unit sphere as a Riemannian submanifold of $\mathbb{R}^n$. Show that if $S \subseteq \mathcal{M}$ is geodesically totally convex, then either $S = \emptyset$ or $S = \mathcal{M}$. Further argue that the spherical cap $S_\alpha = \{x \in \mathcal{M} : x_1 \geq \alpha\}$ is*

1. *geodesically convex for all $\alpha \in \mathbb{R}$,*
2. *geodesically totally convex if and only if $\alpha \notin (-1, 1]$, and*
3. *geodesically strongly convex if and only if $\alpha > 0$.*

Deduce that, while geodesic total convexity and geodesic strong convexity both imply geodesic convexity, no other implications hold among these three notions in general.

Exercise 11.20. *Show that (unlike geodesic convexity) the properties of geodesic total convexity and geodesic strong convexity are closed under intersection.*

11.4 Differentiable geodesically convex functions

For functions which have a gradient or Hessian, geodesic convexity can be characterized in practical ways through inequalities involving derivatives at a base point. (Recall Remark 8.6 defining maps which are k times differentiable, as opposed to smooth.)

We start with a statement using gradients. On a technical note, recall from Section 10.2 that a geodesic segment $c : [0, 1] \to \mathcal{M}$ admits a unique extension to a maximally large open interval containing $[0, 1]$: this is how we make sense of $c'(0)$ and $c'(1)$.

Theorem 11.21. *Let S be a geodesically convex set on a Riemannian manifold $\mathcal{M}$ and let $f : \mathcal{M} \to \mathbb{R}$ be differentiable in a neighborhood of S. Then, $f|_S : S \to \mathbb{R}$ is geodesically convex if and only if for all geodesic segments $c : [0, 1] \to \mathcal{M}$ contained in S we have (letting $x = c(0)$):*

$$\forall t \in [0, 1], \qquad f(c(t)) \geq f(x) + t \left\langle \mathrm{grad} f(x), c'(0) \right\rangle_x . \tag{11.6}$$

Moreover, $f|_S$ is geodesically μ-strongly convex for some $\mu > 0$ if and only if

$$\forall t \in [0,1], \qquad f(c(t)) \geq f(x) + t \left\langle \mathrm{grad} f(x), c'(0) \right\rangle_x + t^2 \frac{\mu}{2} L(c)^2. \qquad (11.7)$$

Finally, $f|_S$ is geodesically strictly convex if and only if, whenever $c'(0) \neq 0$,

$$\forall t \in (0,1], \qquad f(c(t)) > f(x) + t \left\langle \mathrm{grad} f(x), c'(0) \right\rangle_x. \qquad (11.8)$$

Proof. By definition, $f|_S$ is geodesically (strictly) convex if and only if, for all $x, y \in S$ and all geodesics c connecting x to y in S, the composition $f \circ c$ is (strictly) convex from $[0,1]$ to $\mathbb{R}$. By extending the domain of c somewhat, we see that $f \circ c$ is differentiable on an open interval which contains $[0,1]$: this allows us to call upon Lemma 11.1.

First, $f \circ c$ is convex if and only if for all $s, t \in [0,1]$:

$$f(c(t)) \geq f(c(s)) + (t-s)(f \circ c)'(s).$$

Since f is differentiable in a neighborhood of S, we have

$$(f \circ c)'(s) = \mathrm{D} f(c(s))[c'(s)] = \left\langle \mathrm{grad} f(c(s)), c'(s) \right\rangle_{c(s)}.$$

Combine and set $s = 0$ to conclude that if $f|_S$ is geodesically convex then the inequalities (11.6) hold. The other way around, if the inequalities (11.6) hold, then (by reparameterization of c) we conclude that $f \circ c$ is convex for all c as prescribed, hence $f|_S$ is geodesically convex. The proof for strong convexity is similar.

Second, assuming $c'(0) \neq 0$, we have that $f \circ c$ is strictly convex if and only if for all s, t distinct in $[0,1]$:

$$f(c(t)) > f(c(s)) + (t-s)(f \circ c)'(s).$$

Again, using differentiability of f and setting $s = 0$, it follows that $f \circ c$ is strictly convex if and only if inequality (11.8) holds. Conclude similarly to the first part. □

In Section 11.5, we use the inequalities provided by geodesic strong convexity together with inequalities that hold if the gradient of f is Lipschitz continuous to analyze Riemannian gradient descent.

The following corollary is of particular importance to optimization. Note that we need the geodesically convex domain to be open. Indeed, it is possible for a global minimizer to have nonzero gradient if it lies on the boundary of S. (See also Exercise 11.26.)

Corollary 11.22. *If f is differentiable and geodesically convex on an open geodesically convex set, then x is a global minimizer of f if and only if* $\mathrm{grad} f(x) = 0$.

Proof. If $\mathrm{grad} f(x) = 0$, then Theorem 11.21 shows $f(x) \leq f(y)$ for all y in the domain of f (this does not require the domain to be open). The other way around, since the domain of f is open, it is in particular an open submanifold of

$\mathcal{M}$ and we can apply Proposition 4.5 to conclude that if x is a global minimizer, then $\mathrm{grad} f(x) = 0$. □

The next theorem provides a characterization of convexity based on second-order derivatives, also with the requirement that the domain be open.

Theorem 11.23. *Let $f: S \to \mathbb{R}$ be twice differentiable on an open geodesically convex set S. The function f is*

1. *Geodesically convex if and only if* $\mathrm{Hess} f(x) \succeq 0$*;*
2. *Geodesically μ-strongly convex if and only if* $\mathrm{Hess} f(x) \succeq \mu \,\mathrm{Id}$*;*
3. *Geodesically strictly convex if (but not only if)* $\mathrm{Hess} f(x) \succ 0$*,*

all understood to hold for all $x \in S$.

Proof. Similarly to the proof of Theorem 11.21, we start with the fact that f is geodesically convex if and only if $f \circ c: [0, 1] \to \mathbb{R}$ is convex for all geodesic segments $c: [0, 1] \to \mathcal{M}$ whose image lies in S. Calling upon Lemma 11.1, we find that this is the case if and only if, for all such geodesics, it holds that

$$\forall t \in (0, 1), \qquad\qquad (f \circ c)''(t) \geq 0.$$

Since f is twice differentiable everywhere in S, we get that

$$(f \circ c)''(t) = \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathrm{grad} f(c(t)), c'(t) \rangle_{c(t)} = \langle \mathrm{Hess} f(c(t))[c'(t)], c'(t) \rangle_{c(t)},$$

where we also used that $c''(t) = 0$ since c is a geodesic.

If $\mathrm{Hess} f(x)$ is positive semidefinite for all x in S, then $(f \circ c)''(t) \geq 0$ for all c as prescribed and $t \in (0, 1)$, so that f is geodesically convex. The other way around, if f is geodesically convex, it follows that

$$\langle \mathrm{Hess} f(c(0))[c'(0)], c'(0) \rangle_{c(0)} \geq 0$$

for all admissible c (where we particularized to $t = 0$). For all $x \in S$ and sufficiently small $v \in \mathrm{T}_x\mathcal{M}$, the geodesic c with $c(0) = x$ and $c'(0) = v$ remains in S for $t \in [0, 1]$ since S is open in $\mathcal{M}$. Thus, for all such x and v, we deduce that $\langle \mathrm{Hess} f(x)[v], v \rangle_x \geq 0$, which confirms $\mathrm{Hess} f(x)$ is positive semidefinite, and this holds at all points $x \in S$.

The same proof applies for strong convexity, either using or showing that $(f \circ c)''(t) \geq \mu L(c)^2$ for all admissible c and $t \in [0, 1]$, and recalling that $L(c) = \|c'(t)\|_{c(t)}$ since c is a geodesic defined over $[0, 1]$.

If $\mathrm{Hess} f(x)$ is positive definite at all $x \in S$, then $(f \circ c)''(t) > 0$ whenever $c'(0) \neq 0$, which confirms f is geodesically strictly convex. The converse is not true because it also does not hold in the Euclidean case: consider $f(x) = x^4$ on $S = (-1, 1) \subset \mathbb{R}$. □

Example 11.24. *Let $\mathcal{M}$ be a compact Riemannian manifold. If $f: \mathcal{M} \to \mathbb{R}$ has positive semidefinite Hessian at all points, then f is geodesically convex on each connected component of $\mathcal{M}$ by Theorem 11.23. Since $\mathcal{M}$ is compact, it follows* *from Corollary 11.10 that f is constant on each connected component. Therefore, the Hessian of f is in fact zero everywhere.*

Example 11.25. *Given a differentiable function $f\colon \mathcal{M} \to \mathbb{R}$ on a manifold $\mathcal{M}$, it is natural to wonder whether there exists a Riemannian metric for $\mathcal{M}$ such that f is geodesically convex. If such a metric exists, then (a) the domain of f is geodesically convex with that metric, and (b) the critical points of f are its global minimizers (by Corollary 11.22 since $\mathcal{M}$ is open regardless of the metric). Since the notions of criticality (Definition 4.4) and global optimality are independent of the Riemannian metric, it follows that no Riemannian metric makes f geodesically convex if f has a suboptimal critical point. A similar reasoning holds if f (not necessarily differentiable) has a suboptimal local minimizer (by Theorem 11.6) [Vis18]. Of course, it may still be possible to choose a metric such that f is geodesically convex on a submanifold of $\mathcal{M}$.*

Exercise 11.26. *Let S be a geodesically convex set in $\mathcal{M}$, not necessarily open. Define the* cone of feasible directions *K_x of S at x to be the set of vectors $c'(0)$ for all possible geodesic segments c in S satisfying $c(0) = x$. Let $f\colon \mathcal{M} \to \mathbb{R}$ be differentiable in a neighborhood of S and geodesically convex on S. Show that $x_\star \in S$ is a global minimizer of $f|_S$ if and only if*

$$\langle \mathrm{grad} f(x_\star), v\rangle_{x_\star} \geq 0$$

for all v in $K_{x_\star}$. (The closure of K_x is called the tangent cone *to S at x.)*

11.5 Geodesic strong convexity and Lipschitz continuous gradients

Recall Definition 10.16 for the exponential map $\mathrm{Exp}\colon \mathrm{T}\mathcal{M} \to \mathcal{M}$. If S is geodesically convex and x, y are two points in S, then there exists a geodesic segment in S connecting x to y. In terms of Exp, this can be stated as: there exists a tangent vector $v \in \mathrm{T}_x\mathcal{M}$ such that the curve $c(t) = \mathrm{Exp}_x(tv)$ stays in S for all $t \in [0, 1]$ with $c(0) = x$ and $c(1) = y$. Notice that the length of that geodesic segment satisfies $\mathrm{dist}(x, y) \leq L(c) = \|v\|_x$ (Section 10.1).

Thus, Theorem 11.21 provides that, if $f\colon \mathcal{M} \to \mathbb{R}$ is differentiable in a neighborhood of S and if $f|_S$ is geodesically convex, then given $x \in S$ and $v \in \mathrm{T}_x\mathcal{M}$ such that $c(t) = \mathrm{Exp}_x(tv)$ is in S for all $t \in [0, 1]$, we have

$$\forall t \in [0,1], \qquad f(\mathrm{Exp}_x(tv)) \geq f(x) + t\,\langle \mathrm{grad} f(x), v\rangle_x, \tag{11.9}$$

and the inequality is strict for $t \in (0, 1]$ if $f|_S$ is geodesically strictly convex. Furthermore, if $f|_S$ is geodesically μ-strongly convex, then

$$\forall t \in [0,1], \qquad f(\mathrm{Exp}_x(tv)) \geq f(x) + t\,\langle \mathrm{grad} f(x), v\rangle_x + t^2\frac{\mu}{2}\|v\|_x^2. \tag{11.10}$$

These convenient inequalities should be compared with the corresponding one we have if the gradient of f is L-Lipschitz continuous (Proposition 10.53):

$$\forall t \in [0,1], \qquad f(\mathrm{Exp}_x(tv)) \leq f(x) + t\,\langle \mathrm{grad} f(x), v\rangle_x + t^2 \frac{L}{2}\|v\|_x^2. \tag{11.11}$$

When both of the latter inequalities hold, we can obtain strong guarantees for optimization algorithms. To illustrate this, we first work through a couple of facts about geodesically strongly convex functions. For starters, strong convexity ensures existence and uniqueness of a minimizer. (Recall Theorem 10.8 regarding complete manifolds.)

Lemma 11.27. *Let S be a non-empty, closed and geodesically convex set in a complete manifold $\mathcal{M}$. Assume $f\colon \mathcal{M} \to \mathbb{R}$ is differentiable in a neighborhood of S. If $f|_S$ is geodesically μ-strongly convex with $\mu > 0$, then the sublevel sets of $f|_S$ are compact and $f|_S$ has exactly one global minimizer.*

Proof. Let $x_0 \in S$ be arbitrary. We first argue that the sublevel set $S_0 = \{x \in S : f(x) \leq f(x_0)\}$ is compact. Since f is continuous around S and S is closed, S_0 is closed. Since $\mathcal{M}$ is complete, it remains to show that S_0 is bounded in $\mathcal{M}$. For contradiction, assume that this is not the case. Then, there exists a sequence $x_1, x_2, x_3, \ldots$ in S_0 such that $\lim_{k\to\infty} \mathrm{dist}(x_0, x_k) = \infty$. Each x_k is in S and S is geodesically convex hence there exists $v_k \in \mathrm{T}_{x_0}\mathcal{M}$ such that $c(t) = \mathrm{Exp}_{x_0}(tv_k)$ remains in S for $t \in [0,1]$ and $c(1) = x_k$. Then, we have by (11.10) that

$$f(x_k) \geq f(x_0) + \langle \mathrm{grad} f(x_0), v_k\rangle_{x_0} + \frac{\mu}{2}\|v_k\|_{x_0}^2.$$

Since $\mathrm{dist}(x_0, x_k)$ goes to infinity and $\mathrm{dist}(x_0, x_k) \leq L(c) = \|v_k\|_{x_0}$, we have that $\|v_k\|_{x_0}$ goes to infinity. Thus, for all k,

$$f(x_k) \geq f(x_0) - \|\mathrm{grad} f(x_0)\|_{x_0}\|v_k\|_{x_0} + \frac{\mu}{2}\|v_k\|_{x_0}^2.$$

The right-hand side goes to infinity with $k \to \infty$, hence so does $f(x_k)$. This is incompatible with $f(x_k) \leq f(x_0)$ for all k, hence S_0 is compact.

Since $f|_{S_0}$ is continuous, it attains its minimum at some point $x_\star$ in S_0. Thus, for all $x \in S$, we either have $x \notin S_0$ in which case $f(x) > f(x_0) \geq f(x_\star)$, or we have $x \in S_0$ in which case $f(x) \geq f(x_\star)$. Therefore, $x_\star$ is also a minimizer for $f|_S$.

Since geodesic strong convexity implies geodesic strict convexity, it follows from Theorem 11.7 that $x_\star$ is the only minimizer of $f|_S$. ☐

In the same setting as the previous lemma, we find that the norm of the gradient of a geodesically strongly convex function at some point x provides crisp information about the optimality gap at x.

Lemma 11.28. *Let S be a non-empty, closed and geodesically convex set in a complete manifold $\mathcal{M}$. Assume $f\colon \mathcal{M} \to \mathbb{R}$ is differentiable in a neighborhood of*

S. *If* $f|_S$ *is geodesically* μ*-strongly convex with* $\mu > 0$*, then it satisfies a* Polyak–Łojasiewicz inequality*:*

$$\forall x \in S, \qquad f(x) - f(x_\star) \le \frac{1}{2\mu}\|\mathrm{grad} f(x)\|_x^2 \tag{11.12}$$

where $x_\star$ *is the minimizer of* $f|_S$*.*

Proof. The minimizer $x_\star$ of $f|_S$ exists and is unique by Lemma 11.27. Fix $x \in S$ arbitrary. Both x and $x_\star$ are in S which is geodesically convex, hence there exists $v_x \in \mathrm{T}_x\mathcal{M}$ such that $x_\star = \mathrm{Exp}_x(v_x)$ and $t \mapsto \mathrm{Exp}_x(tv_x)$ remains in S for all $t \in [0, 1]$. Therefore, geodesic μ-strong convexity (11.10) provides:

$$\begin{aligned} f(x_\star) = f(\mathrm{Exp}_x(v_x)) &\ge f(x) + \langle \mathrm{grad} f(x), v_x\rangle_x + \frac{\mu}{2}\|v_x\|_x^2 \\ &\ge \inf_{v \in \mathrm{T}_x\mathcal{M}} f(x) + \langle \mathrm{grad} f(x), v\rangle_x + \frac{\mu}{2}\|v\|_x^2 \\ &= f(x) - \frac{1}{2\mu}\|\mathrm{grad} f(x)\|_x^2, \end{aligned}$$

where the infimum is attained by $v = -\frac{1}{\mu}\mathrm{grad} f(x)$ (the critical point of the quadratic in v). Rearrange to conclude. □

The two last results provide sufficient context to study a simple version of Riemannian gradient descent applied to a function which is geodesically strongly convex and has a Lipschitz continuous gradient. See Section 11.8 for further references.

Theorem 11.29. *Let* $f : \mathcal{M} \to \mathbb{R}$ *be differentiable and geodesically convex on a complete manifold* $\mathcal{M}$*. Given* $x_0 \in \mathcal{M}$*, consider the sublevel set* $S_0 = \{x \in \mathcal{M} : f(x) \le f(x_0)\}$*. Assume* f *has* L*-Lipschitz continuous gradient on a neighborhood of* S_0 *and* $f|_{S_0}$ *is geodesically* μ*-strongly convex with* $\mu > 0$*. Consider gradient descent with exponential retraction and constant step-size* $1/L$ *initialized at* x_0*, namely,*

$$x_{k+1} = \mathrm{Exp}_{x_k}\left(-\frac{1}{L}\mathrm{grad} f(x_k)\right), \qquad k = 0, 1, 2, \ldots$$

The function f *has a unique minimizer* $x_\star$ *and the iterates converge to it at least linearly. More precisely, with* $\kappa = L/\mu \ge 1$ *(the* condition number *of* $f|_{S_0}$*), the whole sequence stays in* S_0 *and we have*

$$f(x_k) - f(x_\star) \le \left(1 - \frac{1}{\kappa}\right)^k (f(x_0) - f(x_\star)) \text{ and} \tag{11.13}$$

$$\mathrm{dist}(x_k, x_\star) \le \sqrt{1 - \frac{1}{\kappa}}^{\,k} \sqrt{\kappa}\, \mathrm{dist}(x_0, x_\star) \tag{11.14}$$

for all $k \ge 0$*. (Note that* $\sqrt{1 - \frac{1}{\kappa}} \le 1 - \frac{1}{2\kappa}$*.)*

Proof. By construction, S_0 is non-empty. It is also closed since f is continuous, and geodesically convex since f is geodesically convex. Thus, Lemma 11.27 provides that $f|_{S_0}$ has a unique minimizer $x_\star \in S_0$, and it is clear that $x_\star$ is also the unique minimizer of f on $\mathcal{M}$ since $x \notin S_0 \implies f(x) > f(x_0) \geq f(x_\star)$.

Next, we argue by induction that all x_k are in S_0. Of course, x_0 is in S_0. Assume x_k is in S_0. We know $\mathrm{grad} f$ is L-Lipschitz continuous on a neighborhood $\mathcal{U}$ of S_0. Consider the curve $c(t) = \mathrm{Exp}_{x_k}(-t\mathrm{grad} f(x_k))$. Notice that $c(0)$ is in $\mathcal{U}$. Let I denote the largest interval around $t = 0$ such that $c(t)$ is in $\mathcal{U}$ for all $t \in I$. This interval is open since $c^{-1}(\mathcal{U})$ is open. Then, L-Lipschitz continuity of the gradient provides:

$$\forall t \in I, \qquad f(c(t)) \leq f(x_k) - t\left(1 - t\frac{L}{2}\right)\|\mathrm{grad} f(x_k)\|_{x_k}^2.$$

We want to show that I contains $[0, 2/L)$. To this end, let $\bar{t} = \sup I$ be the first (positive) time such that $c(\bar{t})$ leaves $\mathcal{U}$. If $\bar{t}$ is infinite, we have nothing to do. Assume $\bar{t}$ is finite. The above inequality holds for all $0 \leq t < \bar{t}$. By continuity, it must also hold for $t = \bar{t}$. For contradiction, assume $\bar{t} < 2/L$. Then, $f(c(\bar{t})) \leq f(x_k) \leq f(x_0)$ because $\bar{t}\left(1 - \bar{t}\frac{L}{2}\right) > 0$. This implies that $c(\bar{t})$ is in S_0, a contradiction. Hence, I contains the interval $[0, 2/L)$. In particular, it contains $1/L$. Since $x_{k+1} = c(1/L)$, we deduce that

$$f(x_{k+1}) \leq f(x_k) - \frac{1}{2L}\|\mathrm{grad} f(x_k)\|_{x_k}^2, \tag{11.15}$$

confirming that the whole sequence remains in S_0.

Subtract $f(x_\star)$ on both sides of (11.15) to find

$$f(x_{k+1}) - f(x_\star) \leq f(x_k) - f(x_\star) - \frac{1}{2L}\|\mathrm{grad} f(x_k)\|_{x_k}^2 \tag{11.16}$$

for all k. Lemma 11.28 bounds the gradient norm at $x_k \in S_0$ as:

$$\|\mathrm{grad} f(x_k)\|_{x_k}^2 \geq 2\mu(f(x_k) - f(x_\star)). \tag{11.17}$$

Combining the latter two inequalities, it follows that

$$f(x_{k+1}) - f(x_\star) \leq \left(1 - \frac{\mu}{L}\right)(f(x_k) - f(x_\star)) \tag{11.18}$$

for all k. It is clear when comparing (11.10) and (11.11) that $L \geq \mu$, hence $\kappa = \frac{L}{\mu} \geq 1$ and we can conclude for the sequence $(f(x_k))_{k=0,1,2\ldots}$.

Since $x_\star$ and each x_k are in S_0 which is geodesically convex, there exists $v_k \in \mathrm{T}_{x_\star}\mathcal{M}$ such that the curve $c(t) = \mathrm{Exp}_{x_\star}(tv_k)$ connects $c(0) = x_\star$ to $c(1) = x_k$ while remaining in S_0 for all $t \in [0, 1]$. Then, geodesic strong convexity (11.10) provides

$$f(x_k) \geq f(x_\star) + \langle \mathrm{grad} f(x_\star), v_k \rangle_{x_\star} + \frac{\mu}{2}\|v_k\|_{x_\star}^2.$$

Since $x_\star$ is the minimizer of f on $\mathcal{M}$, we know that $\mathrm{grad} f(x_\star) = 0$. Moreover,

$\mathrm{dist}(x_k, x_\star) \leq L(c) = \|v_k\|_{x_\star}$. Thus,

$$\mathrm{dist}(x_k, x_\star)^2 \leq \frac{2}{\mu}(f(x_k) - f(x_\star)) \tag{11.19}$$

for all k. Combine with the bound on $f(x_k) - f(x_\star)$ to deduce:

$$\mathrm{dist}(x_k, x_\star) \leq \sqrt{\frac{2(f(x_0) - f(x_\star))}{\mu}} \sqrt{1 - \frac{1}{\kappa}}^k. \tag{11.20}$$

Now consider x_0 and $x_\star$. They are in the same connected component of $\mathcal{M}$ since all iterates x_k are in the same connected component (any two consecutive iterates are connected by a geodesic segment) and the sequence converges to $x_\star$. Hence, x_0 and $x_\star$ are connected by a minimizing geodesic γ of $\mathcal{M}$ such that $\gamma(0) = x_\star$, $\gamma(1) = x_0$ and $L(\gamma) = \mathrm{dist}(x_0, x_\star)$ since $\mathcal{M}$ is complete (Theorem 10.9). It is easy to see that $\gamma(t)$ is in S_0 for all $t \in [0, 1]$ because f is geodesically convex on all of $\mathcal{M}$ hence

$$\begin{aligned} \forall t \in [0,1], \qquad f(\gamma(t)) &\leq (1-t) f(\gamma(0)) + t f(\gamma(1)) \\ &= f(x_0) - (1-t)(f(x_0) - f(x_\star)) \leq f(x_0). \end{aligned}$$

Therefore, Lipschitz continuity of the gradient (11.11) provides:

$$f(x_0) \leq f(x_\star) + \frac{L}{2} \mathrm{dist}(x_0, x_\star)^2, \tag{11.21}$$

where we used $\mathrm{grad} f(x_\star) = 0$. Plug this into (11.20) to conclude. □

11.6 Example: Positive reals and geometric programming

As usual, let $\mathbb{R}^n$ denote the Euclidean space with metric $\langle u, v \rangle = u^\top v$. The positive orthant

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_1, \ldots, x_n > 0\} \tag{11.22}$$

is a convex subset of $\mathbb{R}^n$, in the usual sense. Being an open set, it is also an open submanifold of $\mathbb{R}^n$. Its tangent spaces are all identified with $\mathbb{R}^n$.

We can make $\mathbb{R}^n_+$ into a Riemannian submanifold of $\mathbb{R}^n$ using the Euclidean metric. Geodesic convexity on that manifold is equivalent to convexity in the usual sense: this is not particularly interesting. Furthermore, this manifold is not complete—its geodesics are the straight lines of $\mathbb{R}^n$: they cease to exist when they leave $\mathbb{R}^n_+$.

We can endow $\mathbb{R}^n_+$ with a different Riemannian metric so as to make it complete. This leads to a different notion of geodesic convexity on $\mathbb{R}^n_+$. The key is to establish a diffeomorphism between $\mathbb{R}^n$ and $\mathbb{R}^n_+$, and to pullback the Riemannian geometry of $\mathbb{R}^n$ to $\mathbb{R}^n_+$ through that diffeomorphism.

To this end, consider the map $\varphi \colon \mathbb{R}^n_+ \to \mathbb{R}^n$:

$$\varphi(x) = \log(x) = (\log(x_1), \ldots, \log(x_n))^\top. \tag{11.23}$$

This is a diffeomorphism between the manifolds $\mathbb{R}^n_+$ and $\mathbb{R}^n$ because it is smooth and its inverse $\varphi^{-1} \colon \mathbb{R}^n \to \mathbb{R}^n_+$ is smooth too:

$$\varphi^{-1}(y) = \exp(y) = (e^{y_1}, \ldots, e^{y_n})^\top. \tag{11.24}$$

Note also the following expressions for the differential of φ at $x \in \mathbb{R}^n_+$ and its inverse (both are maps from $\mathbb{R}^n$ to $\mathbb{R}^n$):

$$\mathrm{D}\varphi(x)[u] = \left(\frac{u_1}{x_1}, \ldots, \frac{u_n}{x_n}\right)^\top, \qquad (\mathrm{D}\varphi(x))^{-1}[z] \qquad = (x_1 z_1, \ldots, x_n z_n)^\top.$$

They will come in handy.

Equipped with this diffeomorphism, we can define a Riemannian metric $\langle \cdot, \cdot \rangle^+$ on $\mathbb{R}^n_+$ as follows: $\mathrm{D}\varphi(x)$ is an invertible linear map from $\mathrm{T}_x\mathbb{R}^n_+$ to $\mathrm{T}_{\varphi(x)}\mathbb{R}^n$, and we define the inner product on $\mathrm{T}_x\mathbb{R}^n_+$ so as to make this map an isometry, that is:[5]

$$\langle u, v \rangle^+_x \triangleq \langle \mathrm{D}\varphi(x)[u], \mathrm{D}\varphi(x)[v] \rangle = \sum_{i=1}^n \frac{u_i v_i}{x_i^2}. \tag{11.25}$$

(Notice how the metric at x is given by the Euclidean Hessian of the *log-barrier function* $x \mapsto -\sum_{i=1}^n \log(x_i)$.)

Since $\mathbb{R}^n_+$ now has two distinct Riemannian geometries, we let

$$\mathcal{M} = (\mathbb{R}^n_+, \langle \cdot, \cdot \rangle^+) \tag{11.26}$$

denote the Riemannian manifold obtained with the pullback metric, to avoid ambiguity. This is implemented in Manopt as `positivefactory`.

It is an exercise to show that the geodesics of $\mathcal{M}$ are exactly the images of geodesics of $\mathbb{R}^n$ through φ^{-1}, that is: all geodesics of $\mathcal{M}$ are of the form

$$c(t) = \varphi^{-1}(y + tz) = \exp(y + tz) = \left(e^{y_1 + tz_1}, \ldots, e^{y_n + tz_n}\right), \tag{11.27}$$

for some $y, z \in \mathbb{R}^n$. These are defined for all t, hence $\mathcal{M}$ is complete. (Intuitively, as we near the missing boundary of $\mathbb{R}^n_+$, that is, as some x_i nears zero, the metric's $1/x_i^2$ scaling distorts lengths, making the boundary seem infinitely far away.) Moreover, for any two points $x, x' \in \mathcal{M}$, there exists a unique geodesic $c \colon [0, 1] \to \mathcal{M}$ (necessarily minimizing) connecting them:

$$c(t) = \exp(\log(x) + t(\log(x') - \log(x))). \tag{11.28}$$

We are now in a good position to study geodesic convexity on $\mathcal{M}$.

Proposition 11.30. *A set $S \subseteq \mathbb{R}^n_+$ is geodesically convex on $\mathcal{M}$ if and only if $C = \log(S)$ is convex in $\mathbb{R}^n$.*

[5] Compare this with the metric we imposed on the relative interior of the simplex in Exercise 3.65, namely: $\langle u, v \rangle_x = \sum_{i=1}^n \frac{u_i v_i}{x_i}$. That one is a pullback from the usual metric on the positive orthant of the unit sphere (up to scaling); it is not complete.

Proof. Assume S is geodesically convex. For any two points $y, y' \in C$, let $x = \varphi^{-1}(y)$ and $x' = \varphi^{-1}(y')$ be the corresponding points in S. Since S is geodesically convex, the geodesic (11.28) is included in S for $t \in [0,1]$. Hence, C contains $\varphi(c(t)) = \log(x) + t(\log(x') - \log(x)) = y + t(y' - y)$ for $t \in [0,1]$: this is the line segment connecting y to y', hence C is convex. The proof is similar in the other direction. $\square$

Proposition 11.31. *Let S be geodesically convex on $\mathcal{M}$. Then, $f\colon S \to \mathbb{R}$ is geodesically (strictly) convex on $\mathcal{M}$ if and only if the function*

$$g\colon \log(S) \to \mathbb{R}\colon y \mapsto g(y) = f(\exp(y))$$

is (strictly) convex in $\mathbb{R}^n$.

Proof. By definition, f is geodesically convex if and only if for all $x, x' \in S$ and $t \in [0,1]$ it holds that

$$f(c(t)) \leq (1-t)f(x) + tf(x') = (1-t)g(y) + tg(y'),$$

where $x = \exp(y)$, $x' = \exp(y')$, and $c(t)$ is the geodesic uniquely specified by (11.28). Conclude with the observation that

$$f(c(t)) = f(\exp(\log(x) + t(\log(x') - \log(x)))) = g((1-t)y + ty').$$

(The argument is the same for geodesic strict convexity.) $\square$

Let us consider an example. The function $f\colon \mathbb{R}^n_+ \to \mathbb{R}$ defined by

$$f(x) = x_1^{a_1} \cdots x_n^{a_n} \tag{11.29}$$

with some $a \in \mathbb{R}^n$ is (usually) not convex on $\mathbb{R}^n_+$, but it is geodesically convex on $\mathcal{M}$. Indeed, $S = \mathbb{R}^n_+$ is geodesically convex (since $\mathcal{M}$ is connected and complete), and

$$g(y) = f(\exp(y)) = (e^{y_1})^{a_1} \cdots (e^{y_n})^{a_n} = e^{a^\top y}$$

is convex on all of $\mathbb{R}^n$ because it is the composition of a linear (hence convex) function of y with a convex, nondecreasing function (see also Exercise 11.14).

With this example, we can identify a whole class of geodesically convex functions on $\mathcal{M}$, based on the observation that nonnegative linear combinations of geodesically convex functions are geodesically convex (see Exercise 11.13).

Definition 11.32. *A* posynomial *is a function $f\colon \mathbb{R}^n_+ \to \mathbb{R}$ of the form*

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} \cdots x_n^{a_{nk}},$$

where $c_1, \ldots, c_K$ are nonnegative and the exponents a_{ik} are arbitrary. All posynomials are geodesically convex on $\mathcal{M}$. If $K = 1$, f is called a monomial.

By Proposition 11.8, this implies that sets of the form

$$\{x \in \mathbb{R}^n_+ : f(x) \leq \alpha\}$$

are geodesically convex in $\mathcal{M}$ for any posynomial f and $\alpha \in \mathbb{R}$.

We can say even more about monomials. Given $f(x) = cx_1^{a_1} \cdots x_n^{a_n}$ with $c > 0$, the function $\log f$ is well defined on $\mathbb{R}^n_+$. Moreover, $\log f$ is *geodesically linear* on $\mathcal{M}$ (Definition 11.4). Indeed,

$$\log(f(\exp(y))) = \log(c) + a^\top y$$

is affine. By Proposition 11.31, this implies both $\log f$ and $-\log f$ are geodesically convex, as announced. Consequently, sets of the form

$$\{x \in \mathbb{R}^n_+ : \log f(x) = \log \beta\} = \{x \in \mathbb{R}^n_+ : f(x) = \beta\}$$

are geodesically convex in $\mathcal{M}$ for any monomial f and $\beta > 0$.

Overall, we reach the conclusion that problems of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n_+} f(x) \qquad \text{subject to} \qquad & f_i(x) \leq 1, \qquad i = 1, \ldots, m, \\ & g_j(x) = 1, \qquad j = 1, \ldots, p, \end{aligned} \tag{11.30}$$

are geodesically convex whenever $f, f_1, \ldots, f_m$ are posynomials and $g_1, \ldots, g_p$ are monomials. These optimization problems are known as *geometric programs*: see the tutorials by Peterson [Pet76] and by Boyd, Kim, Vandenberghe and Hassibi [BKVH07] for the more standard construction of this class of problems and a list of applications. This is also discussed under the lens of geodesic convexity in [Rap97, Ch. 10].

By construction, $\mathcal{M}$ and $\mathbb{R}^n$ are not only diffeomorphic but also isometric: essentially, they are the same Riemannian manifold. Thus, the notion of geodesic convexity on $\mathcal{M}$ is not meaningfully different from classical Euclidean convexity in $\mathbb{R}^n$ (though it is different from classical convexity in $\mathbb{R}^n_+$). The next section presents a more interesting example.

Exercise 11.33. *Let $\mathcal{M}$ and $\tilde{\mathcal{M}}$ be two Riemannian manifolds with $\varphi\colon \mathcal{M} \to \tilde{\mathcal{M}}$ a diffeomorphism such that $\mathrm{D}\varphi(x)$ is an isometry for all $x \in \mathcal{M}$: $\langle u, v \rangle_x = \langle \mathrm{D}\varphi(x)[u], \mathrm{D}\varphi(x)[v] \rangle_{\varphi(x)}$ (that is, φ is a* Riemannian isometry*). In the context of this section, $\tilde{\mathcal{M}}$ is the Euclidean space $\mathbb{R}^n$, $\mathcal{M}$ is $\mathbb{R}^n_+$ with the metric* (11.25) *and φ is given by* (11.23).

Let ∇ and $\tilde{\nabla}$ denote the Riemannian connections on $\mathcal{M}$ and $\tilde{\mathcal{M}}$ respectively. Show that they are related by

$$\nabla_u V = \mathrm{D}\varphi(x)^{-1}\left[\tilde{\nabla}_{\mathrm{D}\varphi(x)[u]} \tilde{V}\right]$$

for all $u \in \mathrm{T}_x\mathcal{M}$ and $V \in \mathfrak{X}(\mathcal{M})$ with $\tilde{V} \circ \varphi = \mathrm{D}\varphi \circ V$. From there, deduce an expression for $\frac{\mathrm{D}}{\mathrm{d}t}$ on $\mathcal{M}$ in terms of the covariant derivative $\frac{\tilde{\mathrm{D}}}{\mathrm{d}t}$ on $\tilde{\mathcal{M}}$, and conclude that c is a geodesic on $\mathcal{M}$ if and only if $\varphi \circ c$ is a geodesic on $\tilde{\mathcal{M}}$.

Explicitly, with Exp *and* $\tilde{\text{Exp}}$ *the exponential maps on* $\mathcal{M}$ *and* $\tilde{\mathcal{M}}$ *respectively, establish the formula*

$$\mathrm{Exp}_x(u) = \varphi^{-1}\Big(\tilde{\mathrm{Exp}}_{\varphi(x)}(\mathrm{D}\varphi(x)[u])\Big)$$

for all $(x, u) \in \mathrm{T}\mathcal{M}$. *Use this to verify* (11.28) *as well as the fact that* $S \subseteq \mathcal{M}$ *is geodesically convex if and only if* $\varphi(S) \subseteq \tilde{\mathcal{M}}$ *is geodesically convex, and likewise for geodesic convexity of* $f\colon S \to \mathbb{R}$ *and* $f \circ \varphi^{-1}\colon \varphi(S) \to \mathbb{R}$.

11.7 Example: Positive definite matrices

Consider the set of symmetric, positive definite matrices of size n:

$$\mathrm{Sym}(n)^+ = \{X \in \mathrm{Sym}(n) : X \succ 0\}. \tag{11.31}$$

This is a convex set in the Euclidean space $\mathrm{Sym}(n)$ of symmetric matrices of size n, with the inner product $\langle U, V\rangle = \mathrm{Tr}(U^\top V) = \mathrm{Tr}(UV)$. It is an open submanifold; its tangent spaces are identified with $\mathrm{Sym}(n)$.

In analogy with $\mathbb{R}^n_+$, we aim to endow $\mathrm{Sym}(n)^+$ with a Riemannian structure, ideally one that makes it complete. There are at least two ways of doing this. In both cases, for $n = 1$ we recover the same Riemannian geometry as we constructed for $\mathbb{R}^1_+$ in the previous section.

One way is to construct a diffeomorphism between $\mathrm{Sym}(n)^+$ and a complete manifold, just like log provided a diffeomorphism from $\mathbb{R}^n_+$ to $\mathbb{R}^n$. Here, we can define $\varphi\colon \mathrm{Sym}(n)^+ \to \mathrm{Sym}(n)$ to be the principal matrix logarithm,[6]

$$\varphi(X) = \log(X). \tag{11.32}$$

Its inverse is the matrix exponential $\varphi^{-1}(Y) = \exp(Y)$. Both are smooth on the specified domains, hence φ is indeed a diffeomorphism. Based on this observation, we can pullback the Euclidean metric from $\mathrm{Sym}(n)$ to $\mathrm{Sym}(n)^+$ in order to define the following inner product on $\mathrm{T}_X\mathrm{Sym}(n)^+ = \mathrm{Sym}(n)$:

$$\langle U, V\rangle_X^{\log} \triangleq \langle \mathrm{D}\log(X)[U], \mathrm{D}\log(X)[V]\rangle\,. \tag{11.33}$$

This is the *Log-Euclidean metric* studied in detail by Arsigny et al. [AFPA07]. For the same reasons as in the previous section, we can easily describe its geodesics and geodesic convexity (Exercise 11.33):

- The unique (and minimizing) geodesic connecting $X, X' \in \mathrm{Sym}(n)^+$ with respect to the Log-Euclidean metric is

$$c(t) = \exp(\log(X) + t(\log(X') - \log(X))). \tag{11.34}$$

- A set $S \subseteq \mathrm{Sym}(n)^+$ is geodesically convex in that metric if and only if $\log(S)$ is convex in $\mathrm{Sym}(n)$.

[6] See Section 4.7 for questions related to the computation of matrix functions and their differentials.

- Given such a geodesically convex set S, a function $f\colon S \to \mathbb{R}$ is geodesically (strictly) convex if and only if $f \circ \exp$ is (strictly) convex on $\mathrm{Sym}(n)$.

Another—and by some measures, more common—metric on $\mathrm{Sym}(n)^+$ is the so-called *affine invariant metric*. On the tangent space $\mathrm{T}_X\mathrm{Sym}(n)^+$, it is defined as follows:

$$\langle U, V\rangle_X^{\mathrm{aff}} = \left\langle X^{-1/2}UX^{-1/2}, X^{-1/2}VX^{-1/2}\right\rangle = \mathrm{Tr}(X^{-1}UX^{-1}V). \tag{11.35}$$

The central expression ensures that the inputs to $\langle\cdot,\cdot\rangle$ are symmetric matrices. The metric at X matches the Hessian of the log-barrier $X \mapsto -\log(\det(X))$. This is implemented in Manopt as `sympositivedefinitefactory`.

This metric is named after the following property: for all $M \in \mathbb{R}^{n\times n}$ invertible, it holds that $MXM^\top$ is positive definite, and:

$$\left\langle MUM^\top, MVM^\top\right\rangle_{MXM^\top}^{\mathrm{aff}} = \langle U, V\rangle_X^{\mathrm{aff}}. \tag{11.36}$$

One concrete consequence is that if $c\colon [0,1] \to \mathrm{Sym}(n)^+$ is a smooth curve, then the length of c is equal to the length of the other curve $t \mapsto Mc(t)M^\top$ because their speeds are equal for all t. Likewise, the length of the curve $t \mapsto c(t)^{-1}$ is equal to that of c. One can show that the geodesic such that $c(0) = X$ and $c'(0) = V$ is given by [Bha07, Thm. 6.1.6], [Vis18, Ex. 4.9]:

$$\mathrm{Exp}_X(tV) = c(t) = X^{1/2}\exp\left(tX^{-1/2}VX^{-1/2}\right)X^{1/2}. \tag{11.37}$$

This is defined for all t, thus the manifold is complete. Moreover, the manifold is Cartan–Hadamard which makes it well suited for applications of geodesically convex optimization. In order to ensure $c(1) = X'$ (another positive definite matrix), set V to be

$$\mathrm{Log}_X(X') = X^{1/2}\log(X^{-1/2}X'X^{-1/2})X^{1/2}. \tag{11.38}$$

This provides the initial velocity at X of the unique geodesic segment connecting X and X'. It follows that

$$\mathrm{dist}(X, X')^2 = \langle \mathrm{Log}_X(X'), \mathrm{Log}_X(X')\rangle_X^{\mathrm{aff}} = \|\log(X^{-1/2}X'X^{-1/2})\|_{\mathrm{F}}^2, \tag{11.39}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm. With some care, it is possible to express Exp, Log and dist without any matrix square roots, but matrix inverses, exponentials and logarithms are still necessary.

To solve optimization problems over $\mathrm{Sym}(n)^+$ it is helpful to compute gradients and Hessians. Let $\bar f\colon \mathrm{Sym}(n) \to \mathbb{R}$ be a function over the space of symmetric matrices with the usual metric from $\mathbb{R}^{n\times n}$. Assume f is smooth on the open set $\mathrm{Sym}(n)^+$. Further let $f = \bar f|_{\mathrm{Sym}(n)^+}$ formally denote the restriction of $\bar f$ to the manifold of positive definite matrices equipped with the affine invariant metric. The gradients and Hessians of f and $\bar f$ are related as follows for all

$(X, V) \in \mathrm{TSym}(n)^+$:

$$\mathrm{grad} f(X) = X \mathrm{grad} \bar{f}(X) X, \tag{11.40}$$
$$\mathrm{Hess} f(X)[V] = X \mathrm{Hess} \bar{f}(X)[V] X + \frac{V \mathrm{grad} \bar{f}(X) X + X \mathrm{grad} \bar{f}(X) V}{2}.$$

If $\bar{f}$ is defined over all of $\mathbb{R}^{n \times n}$, then it is necessary to replace $\mathrm{grad} \bar{f}(X)$ and $\mathrm{Hess} \bar{f}(X)[V]$ by their symmetric parts. These formulas are derived from [SH15, §3] where expressions also appear for the Riemannian connection and parallel transport on $\mathrm{Sym}(n)^+$.

Example 11.34. *Let* $\bar{f} \colon \mathrm{Sym}(n) \to \mathbb{R}$ *be defined by* $\bar{f}(X) = \log(\det(X))$*, and let* $f = \bar{f}|_{\mathrm{Sym}(n)^+}$ *be its restriction to positive definite matrices with the affine invariant metric. From Example 4.28 we know that* $\mathrm{grad} \bar{f}(X) = X^{-1}$*, hence also that* $\mathrm{Hess} \bar{f}(X)[V] = -X^{-1} V X^{-1}$*. It thus follows from* (11.40) *that*

$$\mathrm{grad} f(X) = X \qquad \textit{and} \qquad \mathrm{Hess} f(X)[V] = 0$$

for all $(X, V) \in \mathrm{TSym}(n)^+$*. In particular,* $\mathrm{Hess} f(X)$ *is both positive and negative semidefinite for all* X*. It follows from Theorem 11.23 and Definition 11.4 that* f *is geodesically linear on* $\mathrm{Sym}(n)^+$*.*

Bhatia [Bha07, Ch. 6] and Moakher [Moa05] (among others) provide a discussion of the affine invariant geometry of positive definite matrices. Moakher as well as Sra and Hosseini [SH15] discuss geodesic convexity on $\mathrm{Sym}(n)^+$ endowed with the affine invariant geometry, with applications. See [NSAY$^+$19] for an overview of reasons to use the affine invariant metric when positive definite matrices represent zero-mean Gaussian distributions (in comparison with other possible structures), and for an application of geodesic convexity to robust distribution estimation.

11.8 Notes and references

Udrişte and Rapcsák wrote a large number of papers on the subject of Riemannian convexity through the late 70s, 80s and 90s: see the many references in [Rap91, Rap97] and [Udr94]. Several of the results discussed in this chapter (and more) can be found in those works. Other useful resources include [dCNdLO98], [Moa05], [SH15], [ZS16] to name a few.

When discussing convexity of a function in $\mathbb{R}^n$, one usually allows f to take on infinite values. It is also habitual to allow f to be nondifferentiable, in which case one resorts to *subgradients* instead of gradients. This can be generalized to Riemannian manifolds; see for example [FO98, ZS16, GH16, BFM17]. Another classical tool in the study of convex functions is the Fenchel dual: see [BHSL$^+$21] for a discussion of that notion on Riemannian manifolds.

Propositions 11.9 and 11.12 are akin to [Roc70, Thm. 10.1, Thm. 32.1] in

$\mathbb{R}^n$. Euclidean versions of Theorems 11.21 and 11.23 are classical, see for example [HUL01, Thm. B.4.1.1, p110, Thm. B.4.3.1, p115].

Some references for Exercise 11.33 regarding Riemannian isometries are [Lee18, Lem. 4.37, Prop. 4.38, Prop. 5.13] and [FLP20, §4].

On a Cartan–Hadamard manifold, given any point y, the function $f(x) = \frac{1}{2}\operatorname{dist}(x, y)^2$ is geodesically 1-strongly convex on the whole manifold [Lee18, Lem. 12.15]. In particular, any geodesic ball centered at y is geodesically convex since it is a sublevel set of f. More detailed information about the Hessian of the distance and the squared distance functions on complete manifolds with bounded curvature can be found in [Sak96, pp153–154].

One could also consider a notion of *retraction convexity* [Hua13, Def. 4.3.1]. Given a retraction R on a manifold $\mathcal{M}$, a set $S \subseteq \mathcal{M}$ is *retraction convex* if for all $x, y \in S$ there exists $v \in \mathrm{T}_x\mathcal{M}$ such that $c(t) = \mathrm{R}_x(tv)$ satisfies $c(0) = x$, $c(1) = y$ and $c([0, 1]) \subseteq S$. A function $f : S \to \mathbb{R}$ is retraction convex if f composed with all retraction curves in S is convex. For the exponential retraction, this reduces to the notion of geodesic convexity defined in this chapter. Retraction convexity is referenced notably in [TFBJ18] and [KSM18].

There is a link between geodesic convexity and barrier functions for interior point methods. Quiroz and Oliveira [QO04] for example study $\mathbb{R}^n_+$ with a general family of diagonal Riemannian metrics, and show applications to the design and analysis of interior point methods for linear programming.

See [Tak11, MMP18] and references therein for discussions of Riemannian geometries on $\mathrm{Sym}(n)^+$ related to the Wasserstein distance between probability distributions, particularized to Gaussian distributions with positive definite covariance matrices.

Section 11.5 provides some simple results regarding Riemannian gradient descent applied to a geodesically strongly convex function with Lipschitz continuous gradient, based on discussions with Chris Criscitiello. The main claim is Theorem 11.29. The proof relies on a Polyak–Łojasiewicz inequality built in Lemma 11.28. This is a direct extension from the Euclidean case [KNS16] to the Riemannian case. Such extensions also appear in various forms and for variations of the setting here in [ZS16], [CMRS20, Thm. 4] and [CB21].

Perhaps the most famous algorithm for convex optimization in $\mathbb{R}^n$ is the accelerated gradient method (also known as the fast gradient method or Nesterov's gradient method), for which a version of Theorem 11.29 holds in $\mathbb{R}^n$ with κ replaced by $\sqrt{\kappa}$. There is interest in determining whether that algorithm has a sensible analog on Riemannian manifolds. Recent work on this topic includes a discussion of the difficulties of the task [ZS18], with positive takes involving continuous-time perspectives [AOBL20a], methods based on estimate sequences [AS20, AOBL20b] and methods based on geodesic maps [MR20], but also negative takes (impossibility results) in [HM21, CB21], all applying to subtly different settings.

Bibliography

[Abb84] E. Abbott. *Flatland: A Romance of Many Dimensions.* Seeley & Co., 1884. [see p2]

[ABBC20] N. Agarwal, N. Boumal, B. Bullins, and C. Cartis. Adaptive regularization with cubics on manifolds. *Mathematical Programming*, 188(1):85–134, 2020. [see p154, 292, 304]

[ABG07] P.-A. Absil, C. G. Baker, and K. A. Gallivan. Trust-region methods on Riemannian manifolds. *Foundations of Computational Mathematics*, 7(3):303–330, 2007. [see p146, 148, 152]

[ABM08] F. Alvarez, J. Bolte, and J. Munier. A unifying local convergence result for Newton's method in Riemannian manifolds. *Foundations of Computational Mathematics*, 8(2):197–226, Apr 2008. [see p304]

[ADM+02] R.L. Adler, J.P. Dedieu, J.Y. Margulies, M. Martens, and M. Shub. Newton's method on Riemannian manifolds and a geometric model for the human spine. *IMA Journal of Numerical Analysis*, 22(3):359–390, 2002. [see p152]

[AFPA07] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.*, 29(1):328–347, 2007. [see p328]

[AK06] P.-A. Absil and K. Kurdyka. On the stable equilibrium points of gradient systems. *Systems & Control Letters*, 55(7):573–577, July 2006. [see p83]

[AM12] P.-A. Absil and J. Malick. Projection-like retractions on matrix manifolds. *SIAM Journal on Optimization*, 22(1):135–158, 2012. [see p116, 119, 162]

[AMH09] A. Al-Mohy and N. Higham. Computing the fréchet derivative of the matrix exponential, with an application to condition number estimation. *SIAM Journal on Matrix Analysis and Applications*, 30(4):1639–1657, 2009. [see p76]

[AMS08] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds.* Princeton University Press, Princeton, NJ, 2008. [see pxii, 51, 69, 76, 81, 82, 83, 118, 142, 148, 152, 153, 161, 163, 181, 211, 234, 255, 256, 286, 291, 304, 305]

[AMT13] P.-A. Absil, R. Mahony, and J. Trumpf. An extrinsic look at the Riemannian Hessian. In Frank Nielsen and Frédéric Barbaresco, editors, *Geometric Science of Information*, volume 8085 of *Lecture Notes in Computer Science*, pages 361–368. Springer Berlin Heidelberg, 2013. [see p119, 254]

[AO15] P.-A. Absil and I.V. Oseledets. Low-rank retractions: a survey and new results. *Computational Optimization and Applications*, 62(1):5–29, 2015. [see p170]

[AOBL20a] F. Alimisis, A. Orvieto, G. Bécigneul, and A. Lucchi. A continuous-time perspective for modeling acceleration in Riemannian optimization. In S. Chiappa and R. Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 1297–1307. PMLR, 2020. [see p331]

[AOBL20b] F. Alimisis, A. Orvieto, G. Bécigneul, and A. Lucchi. Practical accelerated optimization on Riemannian manifolds. *arXiv 2002.04144*, 2020. [see p331]

[AS20] K. Ahn and S. Sra. From Nesterov's estimate sequence to Riemannian acceleration. In *Conference on Learning Theory*, pages 84–118. PMLR, 2020. [see p331]

[ASS+09] S. Agarwal, N. Snavely, I. Simon, S.M. Seitz, and R. Szeliski. Building Rome in a day. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 72–79. IEEE, 2009. [see p12]

[AZGL+18] Z. Allen-Zhu, A. Garg, Y. Li, R. Oliveira, and A. Wigderson. Operator scaling via geodesically convex optimization, invariant theory and polynomial identity testing. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 172–181, 2018. [see p308]

[Bac14] M. Bacák. *Convex analysis and optimization in Hadamard spaces*, volume 22 of *De Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter GmbH & Co KG, 2014. [see p303]

[BAC18] N. Boumal, P.-A. Absil, and C. Cartis. Global rates of convergence for nonconvex optimization on manifolds. *IMA Journal of Numerical Analysis*, 39(1):1–33, February 2018. [see p82, 153]

[BAJN20] G.O. Berger, P.-A. Absil, R.M. Jungers, and Y. Nesterov. On the quality of first-order approximation of functions with Hölder continuous gradient. *Journal of Optimization Theory and Applications*, 185(1):17–33, 2020. [see p82]

[Bar95] A.I. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete & Computational Geometry*, 13(1):189–202, 1995. [see p16]

[BBJN18] S. Bhojanapalli, N. Boumal, P. Jain, and P. Netrapalli. Smoothed analysis for low-rank solutions to semidefinite programs in quadratic penalty form. In S. Bubeck, V. Perchet, and P. Rigollet, editors, *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 3243–3270. PMLR, 06–09 Jul 2018. [see p17]

[BBV16] A.S. Bandeira, N. Boumal, and V. Voroninski. On the low-rank approach for semidefinite programs arising in synchronization and community detection. In *Proceedings of The 29th Conference on Learning Theory, COLT 2016, New York, NY, June 23–26*, 2016. [see p17]

[BC70] F. Brickell and R.S. Clark. *Differentiable manifolds: an introduction.* Van Nostrand Reinhold, 1970. [see pxiii, 2, 183, 190, 191, 192, 211, 255]

[Ber21] R. Bergmann. Manopt.jl. `https://manoptjl.org`, 2021. [see p155]

[Bes87] A.L. Besse. *Einstein Manifolds.* Springer Berlin Heidelberg, 1987. [see p258]

[BFM17] G.C. Bento, O.P. Ferreira, and J.G. Melo. Iteration-complexity of gradient, subgradient and proximal point methods on Riemannian manifolds. *Journal of Optimization Theory and Applications*, 173(2):548–562, 2017. [see p82, 330]

[BH15] B.G. Bodmann and J. Haas. Frame potentials and the geometry of frames. *Journal of Fourier Analysis and Applications*, 21(6):1344–1383, May 2015. [see p83, 254]

[BH19] R. Bergmann and R. Herzog. Intrinsic formulation of KKT conditions and constraint qualifications on smooth manifolds. *SIAM Journal on Optimization*, 29(4):2423–2444, 2019. [see p82, 152]

[Bha07] R. Bhatia. *Positive definite matrices.* Princeton University Press, 2007. [see p329, 330]

[BHSL+21] R. Bergmann, R. Herzog, M. Silva Louzeiro, D. Tenbrinck, and J. Vidal-Núñez. Fenchel duality theory and a primal-dual algorithm on Riemannian manifolds. *Foundations of Computational Mathematics*, 21(6):1465–1504, 2021. [see p330]

[BKVH07] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi. A tutorial on geometric programming. *Optimization and Engineering*, 8(1):67–127, 2007. [see p308, 327]

[BM03] S. Burer and R.D.C. Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003. [see p17]

[BM05] S. Burer and R.D.C. Monteiro. Local minima and convergence in low-rank semidefinite programming. *Mathematical Programming*, 103(3):427–444, 2005. [see p16, 17]

[BM06] I. Brace and J.H. Manton. An improved BFGS-on-manifold algorithm for computing weighted low rank approximations. In *Proceedings of the 17h International Symposium on Mathematical Theory of Networks and Systems*, pages 1735–1738, 2006. [see p82]

[BMAS14] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. Manopt, a Matlab toolbox for optimization on manifolds. *Journal of Machine Learning Research*, 15(42):1455–1459, 2014. [see p155]

[Bon13] S. Bonnabel. Stochastic gradient descent on Riemannian manifolds. *Automatic Control, IEEE Transactions on*, 58(9):2217–2229, 2013. [see p82]

[BV21] P. Breiding and N. Vannieuwenhoven. The condition number of Riemannian approximation problems. *SIAM Journal on Optimization*, 31(1):1049–1077, 2021. [see p119]

[BVB16] N. Boumal, V. Voroninski, and A.S. Bandeira. The non-convex Burer–Monteiro approach works on smooth semidefinite programs. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 2757–2765. Curran Associates, Inc., 2016. [see p17]

[BVB19] N. Boumal, V. Voroninski, and A.S. Bandeira. Deterministic guarantees for Burer-Monteiro factorizations of smooth semidefinite programs. *Communications on Pure and Applied Mathematics*, 73(3):581–608, 2019. [see p16, 17]

[BZ05] J. Borwein and Q. Zhu. *Techniques of Variational Analysis*. CMS Books in Mathematics. Springer-Verlag, 2005. [see p53]

[BZA20] T. Bendokat, R. Zimmermann, and P.-A. Absil. A Grassmann manifold handbook: Basic geometry and computational aspects. *arXiv preprint arXiv:2011.13699*, 2020. [see p76, 254, 255]

[CB20] C. Criscitiello and N. Boumal. An accelerated first-order method for non-convex optimization on manifolds. *arXiv 2008.02252*, 2020. [see p304, 305]

[CB21] C. Criscitiello and N. Boumal. Negative curvature obstructs acceleration for geodesically convex optimization, even with exact first-order oracles. *arXiv 2111.13263*, 2021. [see p331]

[CGT00] A.R. Conn, N.I.M. Gould, and P.L. Toint. *Trust-region methods*. MPS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2000. [see p145, 146, 147, 148, 152]

[CGT11a] C. Cartis, N.I.M. Gould, and P. Toint. Adaptive cubic regularisation methods for unconstrained optimization. Part II: worst-case function- and derivative-evaluation complexity. *Mathematical Programming*, 130:295–319, 2011. [see p153]

[CGT11b] C. Cartis, N.I.M. Gould, and P.L. Toint. Adaptive cubic regularisation methods for unconstrained optimization. Part I: motivation, convergence and numerical results. *Mathematical Programming*, 127(2):245–295, 2011. [see p153]

[CGT12] C. Cartis, N.I.M. Gould, and P.L. Toint. Complexity bounds for second-order optimality in unconstrained optimization. *Journal of Complexity*, 28(1):93–108, 2012. [see p153]

[Cha06] I. Chavel. *Riemannian geometry: a modern introduction*, volume 108 of *Cambridge Tracts in Mathematics.* Cambridge University Press, 2006. [see p304]

[Cif19] D. Cifuentes. On the Burer–Monteiro method for general semidefinite programs. *arXiv 1904.07147*, 2019. [see p17]

[CLR18] F.E. Curtis, Z. Lubberts, and D.P. Robinson. Concise complexity analyses for trust region methods. *Optimization Letters*, 12(8):1713–1724, June 2018. [see p153]

[CM19] D. Cifuentes and A. Moitra. Polynomial time guarantees for the Burer-Monteiro method. *arXiv 1912.01745*, 2019. [see p17]

[CMRS20] S. Chewi, T. Maunu, P. Rigollet, and A.J. Stromme. Gradient descent algorithms for Bures–Wasserstein barycenters. In *Conference on Learning Theory (COLT)*, pages 1276–1304. PMLR, 2020. [see p331]

[CMSZ20] S. Chen, S. Ma, A.M.C. So, and T. Zhang. Proximal gradient method for nonsmooth optimization over the Stiefel manifold. *SIAM Journal on Optimization*, 30(1):210–239, 2020. [see p82]

[dBEG08] A. d'Aspremont, F. Bach, and L. El Ghaoui. Optimal solutions for sparse principal component analysis. *The Journal of Machine Learning Research*, 9:1269–1294, 2008. [see p11]

[dC92] M.P. do Carmo. *Riemannian geometry.* Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty. [see p255, 256, 305]

[dCN95] J.X. da Cruz Neto. *Métodos Geodésicos na Programaçao Matemática.* PhD thesis, Ph.D. Thesis, COPPE/UFRJ, Rio de Janeiro, Brazil, 1995. [see p304]

[dCNdLO98] J.X. da Cruz Neto, L.L. de Lima, and P.R. Oliveira. Geodesic algorithms in Riemannian geometry. *Balkan Journal of Geometry and Its Applications*, 3(2):89–100, 1998. [see p82, 304, 330]

[DE99] L. Dieci and T. Eirola. On smooth decompositions of matrices. *SIAM Journal on Matrix Analysis and Applications*, 20(3):800–819, 1999. [see p76]

[Deh95] J. Dehaene. *Continuous-time matrix algorithms, systolic algorithms and adaptive neural networks.* PhD thesis, Katholieke universiteit Leuven, 1995. [see p76]

[DH94] E. Dudek and K. Holly. Nonlinear orthogonal projection. *Annales Polonici Mathematici*, 59(1):1–31, 1994. [see p119]

[DH19] A. Douik and B. Hassibi. Manifold optimization over the set of doubly stochastic matrices: A second-order geometry. *IEEE Transactions on Signal Processing*, 67(22):5761–5774, November 2019. [see p181]

[DMV99] J. Dehaene, M. Moonen, and J. Vandewalle. Analysis of a class of continuous-time algorithms for principal component analysis and subspace tracking. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 46(3):364–372, 1999. [see p76]

[dOF20] F.R. de Oliveira and O.P. Ferreira. Newton method for finding a singularity of a special class of locally Lipschitz continuous vector fields on Riemannian manifolds. *Journal of Optimization Theory and Applications*, 185(2):522–539, 2020. [see p305]

[EAS98] A. Edelman, T.A. Arias, and S.T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM journal on Matrix Analysis and Applications*, 20(2):303–353, 1998. [see p2, 82, 163, 181, 251, 254]

[FCPJ04] P.T. Fletcher, Lu; C., S.M. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23(8):995–1005, August 2004. [see p181]

[Fep17] F. Feppon. Riemannian geometry of matrix manifolds for Lagrangian uncertainty quantification of stochastic fluid flows. Master's thesis, Massachusetts Institute of Technology, 2017. [see p76]

[FL19] F. Feppon and P.F.J. Lermusiaux. The extrinsic geometry of dynamical systems tracking nonlinear matrix projections. *SIAM Journal on Matrix Analysis and Applications*, 40(2):814–844, 2019. [see p76]

[Fle13] T.P. Fletcher. Geodesic regression and the theory of least squares on Riemannian manifolds. *International Journal of Computer Vision*, 105(2):171–185, Nov 2013. [see p308]

[FLP20] O.P. Ferreira, M.S. Louzeiro, and L.F. Prudente. Iteration-complexity and asymptotic analysis of steepest descent method for multiobjective optimization on Riemannian manifolds. *Journal of Optimization Theory and Applications*, 184:507–533, December 2020. [see p331]

[FM20] C. Franks and A. Moitra. Rigorous guarantees for Tyler's M-estimator via quantum expansion. In J. Abernethy and S. Agarwal, editors, *Proceedings of Thirty Third Conference on Learning Theory (COLT)*, volume 125 of *Proceedings of Machine Learning Research*, pages 1601–1632. PMLR, 09–12 Jul 2020. [see p308]

[FO98] O.P. Ferreira and P.R. Oliveira. Subgradient algorithm on Riemannian manifolds. *Journal of Optimization Theory and Applications*, 97(1):93–104, Apr 1998. [see p330]

[FORW21] C. Franks, R. Oliveira, A. Ramachandran, and M. Walter. Near optimal sample complexity for matrix and tensor normal models via geodesic convexity. *preprint arXiv:2110.07583*, 2021. [see p308]

[FS02] O.P. Ferreira and B.F. Svaiter. Kantorovich's theorem on Newton's method in Riemannian manifolds. *Journal of Complexity*, 18(1):304–329, 2002. [see p304]

[Gab82] D. Gabay. Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, 37(2):177–219, 1982. [see p2, 81, 82]

[GH16] P. Grohs and S. Hosseini. ε-subgradient algorithms for locally Lipschitz functions on Riemannian manifolds. *Advances in Computational Mathematics*, 42(2):333–360, 2016. [see p330]

[GHL04] S. Gallot, D. Hulin, and J. LaFontaine. *Riemannian geometry*. Springer Verlag, 2004. [see p249, 255, 258, 305]

[GQ20] J. Gallier and J. Quaintance. *Differential Geometry and Lie Groups*. Springer International Publishing, 2020. [see p255]

[Gri81] A. Griewank. The modification of Newton's method for unconstrained optimization by bounding cubic terms. Technical Report Technical report NA/12, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, 1981. [see p153]

[GW08] A. Griewank and A. Walther. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. Society for Industrial and Applied Mathematics, 2 edition, 2008. [see p79]

[GZAL14] J. Goes, T. Zhang, R. Arora, and G. Lerman. Robust stochastic principal component analysis. In *Artificial Intelligence and Statistics*, pages 266–274, 2014. [see p11]

[HAG16] W. Huang, P.-A. Absil, and K.A. Gallivan. *A Riemannian BFGS Method for Nonconvex Optimization Problems*, pages 627–634. Springer International Publishing, Cham, 2016. [see p82]

[HGA15] W. Huang, K.A. Gallivan, and P.-A. Absil. A Broyden class of quasi-Newton methods for Riemannian optimization. *SIAM Journal on Optimization*, 25(3):1660–1685, 2015. [see p82, 291, 305]

[Hig08] N. Higham. *Functions of Matrices*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008. [see p75, 76, 79]

[HLV18] P. Hand, C. Lee, and V. Voroninski. ShapeFit: Exact location recovery from corrupted pairwise directions. *Communications on Pure and Applied Mathematics*, 71(1):3–50, 2018. [see p6, 7]

[HM96] U. Helmke and J.B. Moore. *Optimization and dynamical systems*. Springer Science & Business Media, 1996. [see p2, 82]

[HM21] L. Hamilton and A. Moitra. A no-go theorem for acceleration in the hyperbolic plane. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021. [see p331]

[HS15] R. Hosseini and S. Sra. Matrix manifold optimization for Gaussian mixtures. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 910–918. Curran Associates, Inc., 2015. [see p15, 308]

[HS18] G. Heidel and V. Schulz. A Riemannian trust-region method for low-rank tensor completion. *Numerical Linear Algebra with Applications*, 25(6):e2175, 2018. [see p181]

[HS19] R. Hosseini and S. Sra. An alternative to EM for Gaussian mixture models: Batch and stochastic Riemannian optimization. *Mathematical Programming*, pages 1–37, 2019. [see p308]

[HS20] R. Hosseini and S. Sra. *Recent Advances in Stochastic Riemannian Optimization*, pages 527–554. Springer International Publishing, 2020. [see p82]

[HU17] S. Hosseini and A. Uschmajew. A Riemannian gradient sampling algorithm for nonsmooth optimization on manifolds. *SIAM Journal on Optimization*, 27(1):173–189, 2017. [see p82]

[Hua13] Wen Huang. *Optimization algorithms on Riemannian manifolds with applications.* PhD thesis, Florida State University, 2013. [see p331]

[HUL01] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis.* Grundlehren Text Editions. Springer-Verlag Berlin Heidelberg, 1 edition, 2001. [see p309, 331]

[JBAS10] M. Journée, F. Bach, P.-A. Absil, and R. Sepulchre. Low-rank optimization on the cone of positive semidefinite matrices. *SIAM Journal on Optimization*, 20(5):2327–2351, 2010. [see p17, 181]

[JD15] B. Jiang and Y.-H. Dai. A framework of constraint preserving update schemes for optimization on Stiefel manifold. *Mathematical Programming*, 153(2):535–575, 2015. [see p162]

[JMM19] P. Jawanpuria, M. Meghwanshi, and B. Mishra. Low-rank approximations of hyperbolic embeddings. *2019 IEEE 58th Conference on Decision and Control (CDC)*, Dec 2019. [see p181]

[JNRS10] M. Journée, Y. Nesterov, P. Richtárik, and R. Sepulchre. Generalized power method for sparse principal component analysis. *The Journal of Machine Learning Research*, 11:517–553, 2010. [see p11]

[KGB16] A. Kovnatsky, K. Glashoff, and M.M. Bronstein. *MADMM: A Generic Algorithm for Non-smooth Optimization on Manifolds*, pages 680–696. Springer International Publishing, Cham, 2016. [see p82]

[KMU+20] V. Khrulkov, L. Mirvakhabova, E. Ustinova, I. Oseledets, and V. Lempitsky. Hyperbolic image embeddings. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. [see p181]

[KN63] S. Kobayashi and K. Nomizu. *Foundations of differential geometry, Vol. I.* John Wiley and Sons, 1963. [see p259]

[KNS16] H. Karimi, J. Nutini, and M. Schmidt. Linear convergence of gradient and proximal-gradient methods under the Polyak–łojasiewicz condition. In *Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, pages 795–811. Springer International Publishing, 2016. [see p331]

[KSM18] H. Kasai, H. Sato, and B. Mishra. Riemannian stochastic recursive gradient algorithm. In J. Dy and A. Krause, editors, *Proceedings of* *the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 2516–2524, Stockholmsmassan, Stockholm Sweden, 10–15 Jul 2018. PMLR. [see p82, 331]

[KSV14] D. Kressner, M. Steinlechner, and B. Vandereycken. Low-rank tensor completion by Riemannian optimization. *BIT Numerical Mathematics*, 54(2):447–468, Jun 2014. [see p181]

[Lag07] C. Lageman. Pointwise convergence of gradient-like systems. *Mathematische Nachrichten*, 280(13–14):1543–1558, October 2007. [see p83]

[LB20] C. Liu and N. Boumal. Simple algorithms for optimization on Riemannian manifolds with constraints. *Applied Mathematics and Optimization*, 82(3):949–981, 2020. [see p182, 304]

[LC20] M. Lezcano-Casado. Curvature-dependant global convergence rates for optimization on manifolds of bounded geometry. *arXiv preprint arXiv:2008.02517*, 2020. [see p304, 305]

[Lee12] J.M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer-Verlag New York, 2nd edition, 2012. [see pxiii, 22, 28, 51, 52, 53, 67, 99, 100, 118, 119, 190, 191, 192, 198, 199, 208, 209, 210, 211, 212, 218, 219, 222, 223, 224, 255, 256, 257, 259, 260, 270, 296, 297, 302, 306]

[Lee18] J.M. Lee. *Introduction to Riemannian Manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 2018. [see pxiii, 51, 52, 118, 119, 175, 209, 210, 211, 212, 255, 261, 262, 263, 264, 265, 266, 268, 270, 272, 273, 297, 299, 302, 303, 304, 305, 306, 316, 317, 331]

[Lev20] E. Levin. Towards optimization on varieties. Undergraduate senior thesis, Princeton University, 2020. [see p181]

[Lic79] A. Lichnewsky. Une méthode de gradient conjugué sur des variétés: Application à certains problèmes de valeurs propres non linéaires. *Numerical Functional Analysis and Optimization*, 1(5):515–560, 1979. [see p2, 82]

[LKB21] E. Levin, J. Kileel, and N. Boumal. Finding stationary points on bounded-rank matrices: A geometric hurdle and a smooth remedy. *arXiv 2107.03877*, 2021. [see p15, 153, 181, 285]

[LLY20] Z. Lai, L.-H. Lim, and K. Ye. Simpler Grassmannian optimization. *arXiv preprint arXiv:2009.13502*, 2020. [see p254]

[Łoj65] S. Łojasiewicz. Ensembles semi-analytiques. *Lecture Notes IHES (Bures-sur-Yvette)*, 1965. [see p83]

[LTW21] S. Li, G. Tang, and M.B. Wakin. Landscape correspondence of empirical and population risks in the eigendecomposition problem. *arXiv 2106.06574*, 2021. [see p259]

[Lue72] D.G. Luenberger. The gradient projection method along geodesics. *Management Science*, 18(11):620–631, 1972. [see p2, 81, 82, 180]

[MA20] E. Massart and P.-A. Absil. Quotient geometry with simple geodesics for the manifold of fixed-rank positive-semidefinite matrices. *SIAM Journal on Matrix Analysis and Applications*, 41(1):171–198, 2020. [see p181, 250]

[Man02] J.H. Manton. Optimization algorithms exploiting unitary constraints. *IEEE Transactions on Signal Processing*, 50(3):635–650, March 2002. [see p152]

[Mat96] R. Mathias. A chain rule for matrix functions and applications. *SIAM Journal on Matrix Analysis and Applications*, 17(3):610–620, 1996. [see p75]

[Mey11] Gilles Meyer. *Geometric optimization algorithms for linear regression on fixed-rank matrices.* PhD thesis, Université de Liège, Belgique, 2011. [see p15]

[Mic08] P.W. Michor. *Topics in differential geometry*, volume 93. American Mathematical Society, 2008. [see p259]

[Mis14] Bamdev Mishra. *A Riemannian approach to large-scale constrained least-squares with symmetries.* PhD thesis, Université de Liège, Belgique, 2014. [see p15]

[MMP18] L. Malagò, L. Montrucchio, and G. Pistone. Wasserstein Riemannian geometry of positive definite matrices. 2018. [see p331]

[Moa03] M. Moakher. Means and averaging in the group of rotations. *SIAM Journal on Matrix Analysis and Applications*, 24(1):1–16, 2003. [see p308]

[Moa05] M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.*, 26(3):735–747, March 2005. [see p308, 330]

[MR20] D. Martínez-Rubio. Global Riemannian acceleration in hyperbolic and spherical spaces. *arXiv preprint arXiv:2012.03618*, 2020. [see p331]

[MS85] A. Machado and I. Salavessa. Grassman manifolds as subsets of Euclidean spaces. In Cordero, editor, *Differential Geometry, Proc. 5th Int. Colloq.*, volume 131 of *Research Notes in Mathematics*, pages 85–102, Boston, MA, 1985. Pitman. [see p254]

[MS16] B. Mishra and R. Sepulchre. Riemannian preconditioning. *SIAM Journal on Optimization*, 26(1):635–660, 2016. [see p147]

[MS20] S. Marsland and S. Sommer. Riemannian geometry on shapes and diffeomorphisms: Statistics via actions of the diffeomorphism group. In X. Pennec, S. Sommer, and T. Fletcher, editors, *Riemannian Geometric Statistics in Medical Image Analysis*, pages 135–167. Academic Press, 2020. [see p181]

[MT11] M. McCoy and J.A. Tropp. Two proposals for robust PCA using semidefinite programming. *Electronic Journal of Statistics*, 5:1123–1160, 2011. [see p11]

[MV13] B. Mishra and B. Vandereycken. A Riemannian approach to low-rank algebraic Riccati equations. *arXiv preprint arXiv:1312.4883*, 2013. [see p14]

[MZL19] T. Maunu, T. Zhang, and G. Lerman. A well-tempered landscape for non-convex robust subspace recovery. *Journal of Machine Learning Research*, 20(37):1–59, 2019. [see p11]

[NK17] M. Nickel and D. Kiela. Poincaré embeddings for learning hierarchical representations. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 6338–6347. Curran Associates, Inc., 2017. [see p181]

[NNSS20] S. Neumayer, M. Nimmer, S. Setzer, and G. Steidl. On the rotational invariant L1-norm PCA. *Linear Algebra and its Applications*, 587:243–270, February 2020. [see p11]

[NO61] K. Nomizu and H. Ozeki. The existence of complete Riemannian metrics. *Proceedings of the American Mathematical Society*, 12(6):889–891, 1961. [see p303]

[Nof17] V. Noferini. A formula for the Fréchet derivative of a generalized matrix function. *SIAM Journal on Matrix Analysis and Applications*, 38(2):434–457, 2017. [see p75]

[NP06] Y. Nesterov and B.T. Polyak. Cubic regularization of Newton method and its global performance. *Mathematical Programming*, 108(1):177–205, 2006. [see p153]

[NSAY+19] V.A. Nguyen, S. Shafieezadeh-Abadeh, M.-C. Yue, D. Kuhn, and W. Wiesemann. Calculating optimistic likelihoods using (geodesically) convex optimization. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. [see p308, 330]

[NW06] J. Nocedal and S. Wright. *Numerical optimization.* Springer Series in Operations Research and Financial Engineering. Springer Science & Business Media, 2 edition, 2006. [see p65, 122, 146, 152]

[O'N83] B. O'Neill. *Semi-Riemannian geometry: with applications to relativity*, volume 103. Academic Press, 1983. [see pxiii, 51, 52, 53, 118, 119, 176, 177, 211, 234, 255, 304]

[Pat98] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of operations research*, 23(2):339–358, 1998. [see p16]

[Pea94] B.A. Pearlmutter. Fast exact multiplication by the Hessian. *Neural Computation*, 6:147–160, 1994. [see p79]

[Pet76] Elmor L. Peterson. Geometric programming. *SIAM Review*, 18(1):1–51, January 1976. [see p327]

[Pet06] P. Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics.* Springer, 2 edition, 2006. [see p303]

[PJB18] T. Pumir, S. Jelassi, and N. Boumal. Smoothed analysis of the low-rank approach for smooth semidefinite programs. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 2283–2292. Curran Associates, Inc., 2018. [see p17]

[QGA10a] C. Qi, K.A. Gallivan, and P.-A. Absil. An efficient BFGS algorithm for Riemannian optimization. In *Proceedings of the 19th International Symposium on Mathematical Theory of Network and Systems (MTNS 2010)*, volume 1, pages 2221–2227, 2010. [see p274, 305]

[QGA10b] C. Qi, K.A. Gallivan, and P.A. Absil. Riemannian BFGS algorithm with applications. *Recent Advances in Optimization and its Applications in Engineering*, pages 183–192, 2010. [see p82]

[Qi11] C. Qi. *Numerical optimization methods on Riemannian manifolds*. PhD thesis, Florida State University, Tallahassee, FL, 2011. [see p154]

[QO04] E.A. Quiroz and P.R. Oliveira. New results on linear optimization through diagonal metrics and Riemannian geometry tools. Technical report, Technical Report ES-654/04, PESC COPPE, Federal University of Rio de Janeiro, 2004. [see p331]

[Rap91] T. Rapcsák. Geodesic convexity in nonlinear optimization. *Journal of Optimization Theory and Applications*, 69(1):169–183, Apr 1991. [see p315, 330]

[Rap97] T. Rapcsák. *Smooth Nonlinear Optimization in $\mathbb{R}^n$*, volume 19 of *Nonconvex Optimization and Its Applications*. Springer, 1997. [see p2, 82, 308, 313, 314, 327, 330]

[RDTEL21] D.M. Rosen, K.J. Doherty, A. Terán Espinoza, and J.J. Leonard. Advances in inference and representation for simultaneous localization and mapping. *Annual Review of Control, Robotics, and Autonomous Systems*, 4:215–242, 2021. [see p13]

[Roc70] R.T. Rockafellar. *Convex analysis*. Princeton University Press, Princeton, NJ, 1970. [see p309, 315, 330]

[RW12] W. Ring and B. Wirth. Optimization methods on Riemannian manifolds and their application to shape space. *SIAM Journal on Optimization*, 22(2):596–627, 2012. [see p82, 304]

[Sak96] T. Sakai. *Riemannian geometry*, volume 149. American Mathematical Society, 1996. [see p316, 331]

[Sat16] H. Sato. A Dai–Yuan-type Riemannian conjugate gradient method with the weak Wolfe conditions. *Computational Optimization and Applications*, 64(1):101–118, 2016. [see p82]

[Sat21] H. Sato. *Riemannian Optimization and Its Applications*. Springer International Publishing, 2021. [see p82]

[Sch66] P.H. Schönemann. A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, 31(1):1–10, mar 1966. [see p162]

[SFF19] Y. Sun, N. Flammarion, and M. Fazel. Escaping from saddle points on Riemannian manifolds. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 7276–7286. Curran Associates, Inc., 2019. [see p304]

[SH15] S. Sra and R. Hosseini. Conic geometric optimization on the manifold of positive definite matrices. *SIAM Journal on Optimization*, 25(1):713–739, 2015. [see p330]

[SI14] H. Sato and T. Iwai. Optimization algorithms on the Grassmann manifold with application to matrix eigenvalue problems. *Japan Journal of Industrial and Applied Mathematics*, 31(2):355–400, April 2014. [see p254, 259]

[SI15] H. Sato and T. Iwai. A new, globally convergent Riemannian conjugate gradient method. *Optimization*, 64(4):1011–1031, 2015. [see p82]

[SKM19] H. Sato, H. Kasai, and B. Mishra. Riemannian stochastic variance reduced gradient algorithm with retraction and vector transport. *SIAM Journal on Optimization*, 29(2):1444–1472, 2019. [see p82]

[Smi94] S.T. Smith. Optimization techniques on Riemannian manifolds. *Fields Institute Communications*, 3(3):113–135, 1994. [see p2, 81, 82]

[SN22] S. Sun and J. Nocedal. A trust region method for the optimization of noisy functions. *arXiv:2201.00973*, 2022. [see p145]

[SQW17] J. Sun, Q. Qu, and J. Wright. Complete dictionary recovery over the sphere II: Recovery by Riemannian trust-region method. *IEEE Transactions on Information Theory*, 63(2):885–914, Feb 2017. [see p10]

[Sra16] S. Sra. On the matrix square root via geometric optimization. *Electronic Journal of Linear Algebra*, 31(1):433–443, Jul 2016. [see p308]

[TA21] Tiffany M. Tang and Genevera I. Allen. Integrated principal components analysis. *Journal of Machine Learning Research*, 22(198):1–71, 2021. [see p308]

[Tak11] A. Takatsu. Wasserstein geometry of Gaussian measures. *Osaka Journal of Mathematics*, 48(4):1005–1026, 2011. [see p331]

[TB97] L.N. Trefethen and D. Bau. *Numerical linear algebra.* Society for Industrial and Applied Mathematics, 1997. [see pxi, 126, 130, 152]

[TD14] R. Tron and K. Daniilidis. On the quotient representation for the essential manifold. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, June 2014. [see p181]

[TFBJ18] N. Tripuraneni, N. Flammarion, F. Bach, and M.I. Jordan. Averaging stochastic gradient descent on Riemannian manifolds. In

Proceedings of The 31st Conference on Learning Theory, COLT, 2018. [see p331]

[TG21] N. Trendafilov and M. Gallo. *Multivariate Data Analysis on Matrix Manifolds (with Manopt)*. Springer International Publishing, 2021. [see p12]

[TKW16] J. Townsend, N. Koep, and S. Weichwald. PyManopt: a Python toolbox for optimization on manifolds using automatic differentiation. *Journal of Machine Learning Research*, 17(137):1–5, 2016. [see p155]

[Udr94] C. Udrişte. *Convex functions and optimization methods on Riemannian manifolds*, volume 297 of *Mathematics and its applications*. Kluwer Academic Publishers, 1994. [see p2, 82, 308, 314, 316, 330]

[UV13] A. Uschmajew and B. Vandereycken. The geometry of algorithms using hierarchical tensors. *Linear Algebra and its Applications*, 439(1):133–166, July 2013. [see p181]

[UV20] A. Uschmajew and B. Vandereycken. Geometric methods on low-rank matrix and tensor manifolds. In *Handbook of Variational Methods for Nonlinear Geometric Data*, pages 261–313. Springer International Publishing, 2020. [see p181]

[Van10] B. Vandereycken. Riemannian and multilevel optimization for rank-constrained matrix problems. *Faculty of Engineering, Katholieke Universiteit Leuven*, 2010. [see p14]

[Van13] B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013. [see p14, 181]

[Vav91] S.A. Vavasis. *Nonlinear optimization: complexity issues*. Oxford University Press, Inc., 1991. [see p147]

[VAV09] B. Vandereycken, P.-A. Absil, and S. Vandewalle. Embedded geometry of the set of symmetric positive semidefinite matrices of fixed rank. In *SSP09*, pages 389–392, 2009. [see p181, 250]

[Vis18] N.K. Vishnoi. Geodesic convex optimization: Differentiation on manifolds, geodesics, and convexity. *arXiv 1806.06373*, 2018. [see p308, 320, 329]

[Wie12] A. Wiesel. Geodesic convexity and covariance estimation. *IEEE Transactions on Signal Processing*, 60(12):6182–6189, Dec 2012. [see p308]

[WW20] I. Waldspurger and A. Waters. Rank optimality for the Burer–Monteiro factorization. *SIAM Journal on Optimization*, 30(3):2577–2602, 2020. [see p17]

[WY13] Z. Wen and W. Yin. A feasible method for optimization with orthogonality constraints. *Mathematical Programming*, 142(1–2):397–434, 2013. [see p162]

[YZS14] W.H. Yang, L.-H. Zhang, and R. Song. Optimality conditions for the nonlinear programming problems on Riemannian manifolds. *Pacific Journal of Optimization*, 10(2):415–434, 2014. [see p82, 152]

[ZHS16] P.H. Zadeh, R. Hosseini, and S. Sra. Geometric mean metric learning. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning*, ICML, pages 2464–2471. JMLR.org, 2016. [see p308]

[ZRS16] H. Zhang, S.J. Reddi, and S. Sra. Riemannian SVRG: Fast stochastic optimization on Riemannian manifolds. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 4592–4600. Curran Associates, Inc., 2016. [see p82]

[ZS16] H. Zhang and S. Sra. First-order methods for geodesically convex optimization. In *Conference on Learning Theory*, pages 1617–1638, 2016. [see p82, 308, 330, 331]

[ZS18] H. Zhang and S. Sra. An estimate sequence for geodesically convex optimization. In S. Bubeck, V. Perchet, and P. Rigollet, editors, *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 1703–1723. PMLR, 06–09 Jul 2018. [see p331]

[ZZ18] J. Zhang and S. Zhang. A cubic regularized Newton's method over Riemannian manifolds. *arXiv preprint arXiv:1805.05565*, 2018. [see p154]