NICOLAS BOUMAL

AN INTRODUCTION TO OPTIMIZATION ON SMOOTH MANIFOLDS

THIS VERSION COMPILED: MAY 25, 2020

Contents

Exponential maps: Retractions whose curves are geodesics

Preface

This book grew out of the lecture notes for MAT–APC 588, *Topics in Numerical Analysis: Optimization on Smooth Manifolds*, taught in the spring semesters of 2019 and 2020 at Princeton University.

Over the last few decades, optimization on manifolds has garnered increasing interest in research and engineering, recognized as a wide, beautiful and effective generalization of unconstrained optimization. Undoubtedly, this ascension was accelerated by the release[1] in 2008 of the ~~de facto reference book on the matter,~~ *Optimization algorithms on Riemannian manifolds* by Pierre-Antoine Absil, Robert Mahony and Rodolphe Sepulchre.

[1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008

Optimization problems on smooth manifolds arise often in engineering applications, including in machine learning, computer vision, signal processing, dynamical systems and scientific computing. Yet, engineering programs seldom include training in differential geometry. Furthermore, existing textbooks on this topic usually have a focus more aligned with the interests of mathematicians than with the needs of engineers and applied mathematicians.

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 most 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.

Readers are assumed to be comfortable with linear algebra and multivariable calculus. For important computational aspects, it is helpful to have notions of numerical linear algebra, for which I recommend the approachable textbook by Trefethen and Bau.[2] Central to the raison d'être of this book, there are no prerequisites in differential geometry or optimization.

[2] L.N. Trefethen and D. Bau. *Numerical linear algebra*. Society for Industrial and Applied Mathematics, 1997

Building on these expectations, the aim is to give full proofs and justification for all concepts that are introduced, at least for the case of 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.

While the exposition is original in several ways, and some results are hard to locate in the literature, few results are new. The main differential geometry references I used are the fantastic books by Lee,[3,4] O'Neill,[5] and Brickell and Clark.[6]

[3] J.M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer-Verlag New York, 2nd edition, 2012

[4] J.M. Lee. *Introduction to Riemannian Manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 2018

[5] B. O'Neill. *Semi-Riemannian geometry: with applications to relativity*, volume 103. Academic Press, 1983

[6] F. Brickell and R.S. Clark. *Differentiable manifolds: an introduction*. Van Nostrand Reinhold, 1970

A number of people offered decisive comments. I thank Pierre-Antoine Absil (who taught me most of this) and Rodolphe Sepulchre for their input at the early stages of planning for this book, as well as (in no particular order) Stephen McKeown, Eitan Levin, Chris Criscitiello, Razvan-Octavian Radu, Joe Kileel, Bart Vandereycken, Bamdev Mishra and Suvrit Sra, for numerous conversations that led to direct improvements. I am also indebted to the mathematics department at Princeton University for supporting me while I was writing. On a distinct note, I also thank Bil Kleb, Bill Wood and Kevin Godby for developing and freely distributing tufte-latex: the style files used here.

This book is still evolving. Your input on all aspects, however major or minor, is immensely welcome. Please feel free to contact me.

Nicolas Boumal
nboumal@math.princeton.edu

Princeton, New Jersey,
May 2020

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* $f: S \to \mathbb{R}$ 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. In Chapter 2, we discuss a few simple examples relevant to our discussion.

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

Imp.

Linear Space $\mathbb{R}^n$ endowed with inner product forms the Euclidean space.

$\Downarrow$

gives rise to the notion of gradient.

Beyond 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. Intuitively, 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: not unlike 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[1] find it hard to imagine that there exists such a thing as a third dimension, feeling free as birds in their own subspace.

A natural question then is: can we generalize the standard algorithms from unconstrained optimization to handle the broader class of optimization over smooth manifolds? The answer is yes, going back to the 70s,[2,3] the 80s[4] and the 90s,[5,6,7,8,9] and generating 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, turning it into a *Riemannian manifold*.

RM

A Riemannian structure is sufficient to define gradients and Hessians on the manifold, paving the way for optimization. There may exist more than one Riemannian structure on the same manifold, leading to different 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 a 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[10] 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

More than one
Riemannian Inner Product
Resulting in different rithmic performance.

[1] E. Abbott. *Flatland: A Romance of Many Dimensions*. Seeley & Co., 1884

[2] D.G. Luenberger. The gradient projection method along geodesics. *Management Science*, 18(11):620–631, 1972

[3] 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

[4] D. Gabay. Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, 37(2):177–219, 1982

[5] C. Udrişte. *Convex functions and optimization methods on Riemannian manifolds*, volume 297 of *Mathematics and its applications*. Kluwer Academic Publishers, 1994

[6] S.T. Smith. Optimization techniques on Riemannian manifolds. *Fields Institute Communications*, 3(3):113–135, 1994

[7] U. Helmke and J.B. Moore. *Optimization and dynamical systems*. Springer Science & Business Media, 1996

[8] T. Rapcsák. *Smooth Nonlinear Optimization in $\mathbb{R}^n$*, volume 19 of *Nonconvex Optimization and Its Applications*. Springer, 1997

[9] 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

[10] 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.

Linear embedding space is useful for intuition.
but notions are intrinsic.

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 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. Quotient manifolds offer a beautiful way to harness symmetry, so common in applications. In their treatment, we follow Absil et al.[11]

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.

Little to no space is devoted to the use of existing software packages for optimization on manifolds. At the risk of revealing personal biases, see for example Manopt (Matlab), PyManopt (Python) and Manopt.jl (Julia). Such packages may significantly speed up the deployment of tools presented here in new situations.

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 modern applications Riemann himself—arguably—could have never envisioned. Through this book, my hope is to contribute to making this singularly satisfying success of mathematics approachable to a wider audience, with an eye to turn geometry into algorithms.

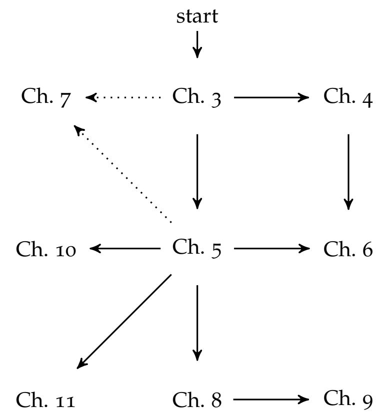
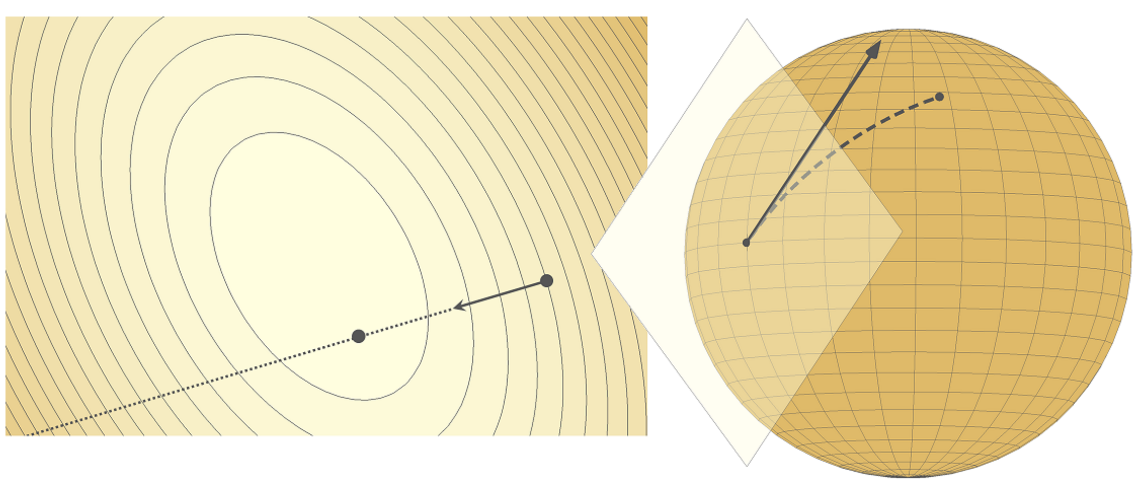


Figure 1.1: An arrow from A to B means it is preferable to read chapter A before reading chapter B. Chapter 7 is meant for on-and-off reading, starting with Chapter 3 and continuing into Chapter 5.

[11] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008

manopt.org
pymanopt.org
manoptjl.org



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.

2.1 *Logistic regression*

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 in terms of θ as follows:

For inner products, see Definition 3.1.

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

Read: the probability that y is 0 or 1, given x and θ.

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, then we can easily compute the probability that a new example x belongs to either class. The task is to learn θ from given, labeled examples.

If we have θ, *and* if the logistic model is accurate. . .

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*

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 clever way is to assume the following:[1]

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

[1] 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

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

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

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:

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

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. In particular, see the paper by Hand et al. (which is the source for this presentation of the problem) for a cost function robust against outliers; the resulting algorithm is called ShapeFit.

2.3 *Single extreme eigenvalue or singular value*

Let $\mathrm{Sym}(n)$ denote the set of real, symmetric matrices of size n. By the spectral theorem, a matrix $A \in \mathrm{Sym}(n)$ 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 $\mathbb{R}^n$: $\langle u, v\rangle = u_1v_1 + \cdots + u_nv_n$.

More generally, we may use any inner product on a finite dimensional linear space and consider a linear operator self-adjoint with respect to that inner product: see Definition 3.3.

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*,

We will rediscover the properties of the Rayleigh quotient through the prism of optimization on manifolds as a running example.

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

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.

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 form of the quotient, 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 optimization problem becomes:

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

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

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

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

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 notations, 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 over-complete 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 small coefficients. For the magnitudes of coefficients to be meaningful, we further require 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):

$$\min_{D\in\mathbb{R}^{d\times n}, C\in\mathbb{R}^{n\times k}} \|Y - DC\|^2 + \lambda\|C\|_0 \quad \text{subject to } \|b_1\| = \cdots = \|b_n\| = 1. \tag{2.1}$$

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 (2.1) 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 non-convex, so that finding a global optimum for dictionary learning as stated above is challenging: see work by Sun et al.[2] for some guarantees.

[2] 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

2.5 *Principal component analysis*

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 variations 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 the main direction of variation, called the *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 solve

$X = [x_1, \ldots, x_n]_{d\times n}.$

See (3.16) below for properties of Frobenius norms and inner products.

$$\max_{u\in\mathrm{S}^{d-1}} \sum_{i=1}^n \|uu^\top x_i\|^2 = \|uu^\top X\|^2 = \left\langle uu^\top X, uu^\top X\right\rangle = \left\langle XX^\top u, u\right\rangle.$$

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

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

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 both 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:

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

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 this problem:

$$\max_{U \in \mathrm{St}(d,k)} \sum_{i=1}^{k} \alpha_i \left\langle XX^\top u_i, u_i \right\rangle = \left\langle XX^\top U, UD \right\rangle, \tag{2.2}$$

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 (2.2), 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 to promote robustness against outliers or sparsity, for example to develop variants such as robust PCA [GZAL14, MT11] and sparse PCA [dBEG08, JNRS10]. 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 of singular value decomposition.

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,

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

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* $\sim$ on the Stiefel manifold:

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.

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

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 space*:

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

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 space is in one-to-one correspondence with the set of subspaces of dimension k in $\mathbb{R}^d$. The latter is called the *Grassmann manifold*:

The symbol $\equiv$ reads "is equivalent to", where context indicates in what sense.

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

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

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

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 the above, 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.

2.6 *Synchronization of rotations*

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.

A similar task 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. 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*

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 form the notion that certain users may share similar traits, so that their experiences may transpose well to one another. Similarly, 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.

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

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

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 biviariate 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.

2.8 *Gaussian mixture models*

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 $\Sigma_1, \ldots, \Sigma_d \in \operatorname{Sym}(d)^+$ and mixing probabilities $(w_1, \ldots, w_K) \in \Delta_+^{K-1}$ are to be determined. We use the following notation:

$$\operatorname{Sym}(d)^+ = \{X \in \operatorname{Sym}(d) : X \succ 0\}$$

for symmetric, positive definite matrices of size d, and

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

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:

$$\max_{\substack{\hat{\mu}_1,\ldots,\hat{\mu}_K \in \mathbb{R}^d, \\ \hat{\Sigma}_1,\ldots,\hat{\Sigma}_K \in \mathrm{Sym}(d)^+, \\ w \in \Delta_+^{K-1}}} \sum_{i=1}^n \log\left(\sum_{k=1}^K w_k \frac{1}{\sqrt{2\pi \det(\Sigma_k)}} e^{-\frac{(x-\mu_k)^\top \Sigma_k^{-1}(x-\mu_k)}{2}}\right).$$

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 (2.3) 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

$$\min_{X \in \mathrm{Sym}(n)} \langle C, X\rangle \qquad \text{subject to} \qquad \mathcal{A}(X) = b \text{ and } X \succeq 0, \tag{2.3}$$

where $\langle A, B\rangle = \mathrm{Tr}(A^\top B)$ and $\mathcal{A} \colon \mathrm{Sym}(n) \to \mathbb{R}^m$ is a linear operator 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,[3] 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

[3] S. Burer and R.D.C. Monteiro. Local minima and convergence in low-rank semidefinite programming. *Mathematical Programming*, 103(3):427–444, 2005

$$\min_{Y \in \mathbb{R}^{n\times p}} \langle CY, Y\rangle \qquad \text{subject to} \qquad \mathcal{A}(YY^\top) = b. \tag{2.4}$$

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*[4,5] assures 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.

[4] 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

[5] A.I. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete & Computational Geometry*, 13(1):189–202, 1995

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:

$$\mathcal{M} = \{Y \in \mathbb{R}^{n \times p} : \mathcal{A}(YY^\top) = b\}. \tag{2.5}$$

Assume the map $h(Y) = \mathcal{A}(YY^\top)$ has the property that its differential at any 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:[6]

[6] 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

> *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.*

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 in general p cannot be taken much lower than that threshold [WW18], 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

what is meant by smooth space & smooth function on this space?

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 $f \colon \mathcal{M} \to \mathbb{R}$ 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.

When we write $F \colon A \to B$, we mean that the domain of F is all of A. When it is not, we state so explicitly.

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

$$\mathrm{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}, \tag{3.2}$$

where $\|x\| = \sqrt{x^\top x}$ is the Euclidean norm on $\mathbb{R}^n$. Intuitively, we think of S^{n-1} as a smooth nonlinear space in $\mathbb{R}^n$. Our definitions below are compatible with this intuition, and we call S^{n-1} an *embedded submanifold* of $\mathbb{R}^n$.

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

$x_1^2 + x_2^2 = 1$

$2x_1, 2x_2$

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

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

Regarding smoothness of functions, we may expect that any function $f \colon \mathrm{S}^{n-1} \to \mathbb{R}$ obtained by restriction to S^{n-1} of a smooth function on $\mathbb{R}^n$ (smooth in the usual sense for functions from $\mathbb{R}^n$ to $\mathbb{R}$) 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, $\mathcal{E} = \mathbb{R}^n$ 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:

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

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. Context tells which is meant.

where the $\alpha_k > 0$ are appropriately chosen step-sizes and $\mathrm{grad} f : \mathcal{E} \to \mathcal{E}$ is the gradient of f. Under mild assumptions, the limit 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

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

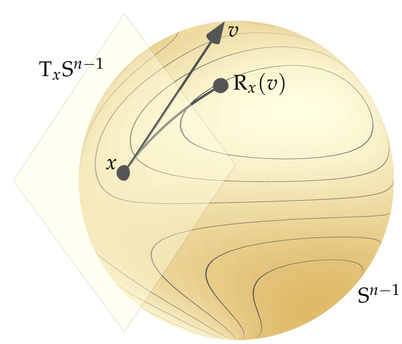
with some smooth function f on the sphere. The most pressing point is ~~to find an alternative for the implicit use of linearity in (3.5).~~ Indeed, above, both x_k and $\mathrm{grad} f(x_k)$ are elements of $\mathcal{E}$. Since $\mathcal{E}$ is a linear space, they can be combined through linear combination.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

Retraction.

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

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

Surely, if the purpose is to remain on S^{n-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 $\mathrm{T}_x\mathrm{S}^{n-1}$, we only consider moving along directions in $\mathrm{T}_x\mathrm{S}^{n-1}$. To this end, we introduce the concept of *retraction* at x: a map $\mathrm{R}_x : \mathrm{T}_x\mathrm{S}^{n-1} \to \mathrm{S}^{n-1}$ 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:



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

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

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

Directional Derivative

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

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

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

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

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

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

and the canonical basis $e_1, \ldots, e_n \in \mathbb{R}^n$ (the columns of the identity matrix), the ith coordinate of $\operatorname{grad} f(x) \in \mathbb{R}^n$ if given by

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

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

$T_x S^{n-1}$ - linear space.

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

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

Cool.

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

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

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

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

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

where $\operatorname{Proj}_x \colon \mathbb{R}^n \to \mathrm{T}_x\mathrm{S}^{n-1}$ is the orthogonal projector from $\mathbb{R}^n$ to $\mathrm{T}_x\mathrm{S}^{n-1}$ (orthogonal with respect to the inner product on $\mathbb{R}^n$.) The functions f and $\bar{f}$ often have the same analytical expression. For example, $f(x) = x^\top A x$ (for some matrix $A \in \mathbb{R}^{n \times n}$) is smooth on S^{n-1} because $\bar{f}(x) = x^\top A x$ is smooth on $\mathbb{R}^n$ and they coincide on S^{n-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^{n-1}: given $x_0 \in \mathrm{S}^{n-1}$,

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

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

f: $\mathbb{R}^n \to \mathbb{R}$
Derivative is Df^T
$Df^T(v) = Df^T v$.
matrix applied to vector.

3.1 Euclidean space

$\mathcal{E}$ - linear space (vector space over $\mathbb{R}$)

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. Frequently used examples include $\mathbb{R}^n$ (column vectors of length n), $\mathbb{R}^{n \times p}$ (matrices of size $n \times p$), $\mathrm{Sym}(n)$ (real, symmetric matrices of size n) and $\mathrm{Skew}(n)$ (real, skew-symmetric matrices of size n).

A *basis* for $\mathcal{E}$ is a set of vectors (elements of $\mathcal{E}$) $e_1, \ldots, e_d$ such that any vector $x \in \mathcal{E}$ can be expressed as a unique linear combination

Differential

$x = a_1e_1 + \cdots + a_d e_d$. All bases have the same number of elements, called the *dimension* of $\mathcal{E}$ ($\dim \mathcal{E}$): it is always finite in our treatment. Each basis induces a one-to-one mapping between $\mathcal{E}$ and $\mathbb{R}^d$: we write $\mathcal{E} \equiv \mathbb{R}^d$. Moreover, $\mathcal{E}$ inherits the usual *topology* of $\mathbb{R}^d$ in this way. A *neighborhood* of x in $\mathcal{E}$ is an open subset of $\mathcal{E}$ which contains x.

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. A subset is *closed* if its complement is open. A subset may be open, closed, both, or neither. Notions such as continuity and compactness are defined just based on a topology. More on this in Section 8.2.

For two linear spaces $\mathcal{E}$ and $\mathcal{E}'$ of dimensions d and d' respectively, using the identifications $\mathcal{E} \equiv \mathbb{R}^d$ and $\mathcal{E}' \equiv \mathbb{R}^{d'}$, a function $F \colon \mathcal{E} \to \mathcal{E}'$ is *smooth* if and only if it is smooth (infinitely differentiable) in the usual sense for a function from $\mathbb{R}^d$ to $\mathbb{R}^{d'}$. Then, the *differential* of F at x is a linear map $\mathrm{D}F(x) \colon \mathcal{E} \to \mathcal{E}'$ defined by:

$$\mathrm{D}F(x)[v] = \lim_{t \to 0} \frac{F(x+tv) - F(x)}{t} = \frac{\mathrm{d}}{\mathrm{d}t} F(x+tv)\bigg|_{t=0}. \tag{3.12}$$

Directional Derivative in the direction of v.

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

It is often useful to equip $\mathcal{E}$ with a (real) *inner product*.

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

1. *Symmetry: $\langle u, v \rangle = \langle v, u \rangle$;*
2. *Linearity: $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$; and*
3. *Positive definiteness: $\langle u, u \rangle \geq 0$ with $\langle u, u \rangle = 0 \iff u = 0$.*

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*:*

$$\|u\| = \sqrt{\langle u, u \rangle}. \tag{3.13}$$

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

$$\langle u, v \rangle = u^\top v = \sum_i u_i v_i, \qquad \|u\| = \sqrt{\sum_i u_i^2}. \tag{3.14}$$

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

$$\langle U, V \rangle = \mathrm{Tr}(U^\top V) = \sum_{ij} U_{ij} V_{ij}, \qquad \|U\| = \sqrt{\sum_{ij} U_{ij}^2}, \tag{3.15}$$

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 this inner product, with matrices U, V, W, A, B of compatible sizes:

$$\begin{aligned} \langle U, V \rangle &= \langle U^\top, V^\top \rangle, & \langle UA, V \rangle &= \langle U, VA^\top \rangle, \\ \langle BU, V \rangle &= \langle U, B^\top V \rangle, & \langle U \odot W, V \rangle &= \langle U, V \odot W \rangle, \end{aligned} \tag{3.16}$$

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:

$$\langle u, v \rangle = \Re\{u^* v\} = \Re\left\{\sum_k \bar{u}_k v_k\right\}, \qquad \|u\| = \sqrt{\sum_k |u_k|^2}, \tag{3.17}$$

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 $\mathbb{R}^{2n}$, with real and imaginary parts 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 standard inner product and norm:

$$\langle U, V \rangle = \Re\{\mathrm{Tr}(U^* V)\} = \Re\left\{\sum_{k\ell} \bar{U}_{k\ell} V_{k\ell}\right\}, \quad \|U\| = \sqrt{\sum_{k\ell} |U_{k\ell}|^2}. \tag{3.18}$$

We have similarly useful identities in the complex case:

$$\begin{aligned} \langle U, V \rangle &= \langle U^*, V^* \rangle, & \langle UA, V \rangle &= \langle U, VA^* \rangle, \\ \langle BU, V \rangle &= \langle U, B^* V \rangle, & \langle U \odot W, V \rangle &= \langle U, V \odot \bar{W} \rangle. \end{aligned} \tag{3.19}$$

We close with the definition of adjoint of a linear operator (or linear function, or linear map). This is often useful in deriving gradients of functions—more on this in Section 4.7.

Definition 3.3. *Let $\mathcal{E}$ and $\mathcal{E}'$ be two Euclidean spaces, with inner products $\langle \cdot, \cdot \rangle_a$ and $\langle \cdot, \cdot \rangle_b$ respectively. Let $\mathcal{A} \colon \mathcal{E} \to \mathcal{E}'$ be a linear operator. The* adjoint *of $\mathcal{A}$ is a linear operator $\mathcal{A}^* \colon \mathcal{E}' \to \mathcal{E}$ defined by this property:*

$$\forall u \in \mathcal{E}, v \in \mathcal{E}', \qquad \langle \mathcal{A}(u), v \rangle_b = \langle u, \mathcal{A}^*(v) \rangle_a. \quad \text{h OK.}$$

In particular, if $\mathcal{A}$ maps $\mathcal{E}$ to $\mathcal{E}$ equipped with an inner product $\langle \cdot, \cdot \rangle$ and ou

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

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

Properties laid out above in (3.16) and (3.19) notably define the adjoints of left- or right-multiplication by a matrix. These illustrate why the transpose (or Hermitian conjugate-transpose) of a matrix is often called the adjoint of that matrix, implicitly referring to the standard inner product. For the same reason, self-adjoint operators are sometimes called *symmetric* or *Hermitian*.

We conclude with a reminder of the spectral properties of self-adjoint operators.

Theorem 3.4 (spectral theorem)**.** *Let $\mathcal{A}$ be a self-adjoint linear operator on the Euclidean space $\mathcal{E}$ equipped with the inner product $\langle \cdot, \cdot \rangle$. Then, $\mathcal{A}$ admits an orthonormal basis of eigenvectors $v_1, \ldots, v_d \in \mathcal{E}$ associated to real eigenvalues $\lambda_1, \ldots, \lambda_d$; that is, with $d = \dim \mathcal{E}$:*

$$\forall i, j, \quad \mathcal{A}(v_i) = \lambda_i v_i \qquad \text{and} \qquad \langle v_i, v_j \rangle = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Definition 3.5. *Let $\mathcal{A}$ be a self-adjoint linear operator on the Euclidean space $\mathcal{E}$ equipped with the inner product $\langle \cdot, \cdot \rangle$. We say $\mathcal{A}$ 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.*

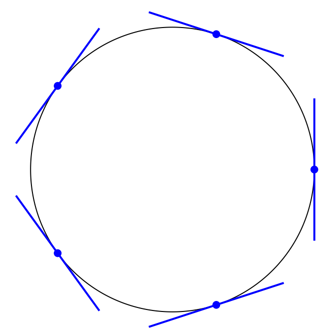
ok .

Smoothness- linearization possible .

3.2 *Embedded submanifolds of Euclidean space*

One reasonable notion of smoothness for a set $\mathcal{M}$ in $\mathcal{E}$ captures the idea that a smooth set can be linearized in some meaningful sense around each point. This readily includes (open subsets of) linear spaces. As a more interesting example, consider the sphere S^{n-1}. This is the set of vectors $x \in \mathbb{R}^n$ satisfying

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



Circle: $x_1^2 + x_2^2 - 1 = 0$

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 this Taylor expansion of h:

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

If x is in S^{n-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^{n-1} for small t. This motivates us to consider the subspace $\ker \mathrm{D}h(x)$ as a linearization of S^{n-1} around x. Since

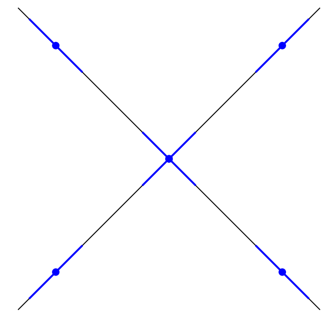
ker $\mathcal{A}$ denotes the *kernel* or *null space* of a linear operator $\mathcal{A}$.

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

$v \in$ Ker of $h(x)$

the kernel of $\mathrm{D}h(x)$ is the subspace orthogonal to x in $\mathbb{R}^n$. 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 the set is smooth and can be linearized by the kernels of Dh. However, this is not the case. Indeed, consider the following example in $\mathbb{R}^2$:



Cross: $x_1^2 - x_2^2 = 0$

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

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

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

At $x = 0$, D$h(0) = [0, 0]$, whose kernel is all of $\mathbb{R}^2$. The reason a kink could arise at the origin is the rank drop: D$h(x)$ has (full) rank one at all x in $\mathcal{X}$, except at the origin where the rank drops to zero.

ok. only,

These observations form the basis for the following definition. Since a set $\mathcal{M}$ may be equivalently defined by many different functions h, and since it may not be practical to define all of $\mathcal{M}$ with a single function h, we phrase the definition locally and in terms of the set $\mathcal{M}$ itself (as opposed to phrasing it with respect to a specific function h).

Definition 3.6. *Let $\mathcal{M}$ be a subset of a linear space $\mathcal{E}$. We say $\mathcal{M}$ is a (smooth)* embedded submanifold *of $\mathcal{E}$ if either of the following holds:*

About terminology: $\mathcal{M}$ is a *sub*manifold of the (linear) manifold $\mathcal{E}$. Submanifolds are manifolds too. 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.

1. *$\mathcal{M}$ is an open subset of $\mathcal{E}$. Then, we also call $\mathcal{M}$ an* open submanifold. *If $\mathcal{M} = \mathcal{E}$, we also call it a* linear manifold.
2. *For a fixed integer $k \geq 1$ 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) rank D$h(x) = k$.

   Such a function h is called a local defining function *for $\mathcal{M}$ at x.*

We call $\mathcal{E}$ the embedding space *or the* ambient space *of $\mathcal{M}$.*

Condition 2a above can be stated equivalently as:

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

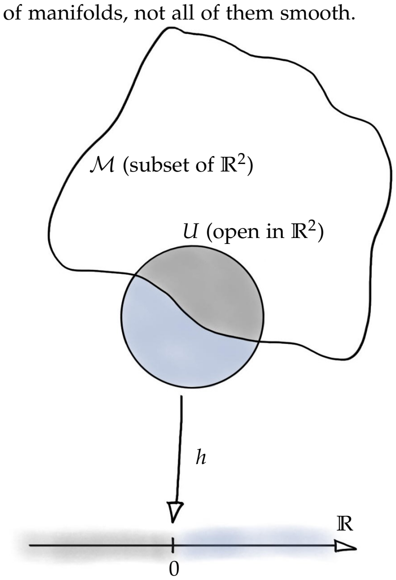
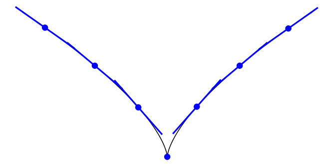


Figure 3.1: Local defining functions: h smoothly maps an open set U (a disk) to $\mathbb{R}^k$ ($k = 1$); exactly the intersection of U with the set $\mathcal{M}$ (a curve) is mapped to 0.

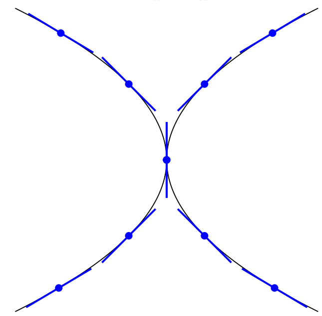
It is an exercise to verify that this definition excludes various types of pathologies such as the cross ($x_1^2 = x_2^2$), the cusp ($x_1^2 = x_2^3$) and the double parabola ($x_1^2 = x_2^4$).

Definition 3.6 is not normally taken to be the definition of an embedded submanifold of $\mathcal{E}$, but it is equivalent to the standard definition:

see Sections 3.10 and 8.14. Differential geometry defines a broader class of smooth sets called *(smooth) manifolds*. Embedded submanifolds are manifolds. When the statements we make hold true for manifolds in general, we use that word to signal it. The hope is that this entry point will prove natural to build all the tools and intuitions we need for optimization, in particular because many of the manifolds we encounter in applications are presented to us as constrained sets within a linear space.



Cusp: $x_1^2 - x_2^3 = 0$



Double parabola: $x_1^2 - x_2^4 = 0$

Let us take some time to build further support for our definition of smooth sets. In order to understand the local geometry of a set around a point, we may want 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 on $\mathcal{M}$ which pass through a given point, and record their velocities as they do so.

Definition 3.7. *Let $\mathcal{M}$ be a subset of $\mathcal{E}$. For all $x \in \mathcal{M}$, define:*

$$\mathrm{T}_x\mathcal{M} = \left\{ c'(0) \mid c\colon I \to \mathcal{M} \text{ is smooth around } 0 \text{ and } c(0) = x \right\}. \tag{3.20}$$

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.

Here, I is an open interval containing $t = 0$, and c is smooth in the usual sense as a map from (an open subset of) $\mathbb{R}$ to $\mathcal{E}$—two linear spaces.

Note that $\mathrm{T}_x\mathcal{M}$ is a subset of $\mathcal{E}$, but it is not necessarily a linear subspace. 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. On the other hand, for the cross $\mathcal{X}$, the set $\mathrm{T}_x\mathcal{M}$ at $x = 0$ is actually equal to $\mathcal{X}$, which is not a linear subspace. (Note though that while the cusp and double parabola are not embedded submanifolds, their sets $\mathrm{T}_x\mathcal{M}$ are linear subspaces of $\mathbb{R}^2$.)

Theorem 3.8. *Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$. Consider $x \in \mathcal{M}$ and the set $\mathrm{T}_x\mathcal{M}$ (3.20). 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.*

The main tool we need to prove this theorem is the standard *inverse function theorem*, stated here without proof.

Theorem 3.9 (Inverse function theorem)**.** *Suppose $U \subseteq \mathcal{E}$ and $V \subseteq \mathcal{E}'$ 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 $F|_{U'}\colon U' \to V'$ is a diffeomorphism; that is: it is smooth and invertible, and its inverse is smooth. Then, restricting F to U', we have $\mathrm{D}F^{-1}(F(x)) = (\mathrm{D}F(x))^{-1}$.*

[Lee12, Thm. C.34]

For the last claim, differentiate the identity $x = F^{-1}(F(x))$ on both sides with respect to x along v; by the chain rule,

$$\begin{aligned} v &= \mathrm{D}(F^{-1} \circ F)(x)[v] \\ &= \mathrm{D}F^{-1}(F(x))[\mathrm{D}F(x)[v]], \end{aligned}$$

i.e., $\mathrm{D}F^{-1}(F(x)) \circ \mathrm{D}F(x)$ is the identity.

Equipped with this tool, we proceed to prove our main theorem.

Proof of Theorem 3.8. 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 $c'(0) = v$ is in $\mathrm{T}_x\mathcal{M}$. This shows $\mathcal{E}$ is included in $\mathrm{T}_x\mathcal{M}$, so that the two coincide.

Considering the other case, let $h\colon U \to \mathbb{R}^k$ be any local defining function for $\mathcal{M}$ at 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. 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, which means 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. 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 $F\colon U \to \mathbb{R}^d$ (where $U \subseteq \mathcal{E}$ is the domain of h) defined by:

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

where $y_1, \ldots, y_d$ denote the coordinates of $y \in \mathcal{E}$. We aim to apply the inverse function theorem to F at x. Since F is evidently smooth, we only need to check the differential of F at x. Considering this one row at a time, we find

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

where I_{d-k} is the identity matrix of size $d-k$ and 0 here denotes a zero matrix of size $(d-k) \times k$. This matrix is invertible, with inverse given by:

$$(\mathrm{D}F(x))^{-1} = \begin{bmatrix} I_{d-k} & 0 \\ -B^{-1}A & B^{-1} \end{bmatrix}. \tag{3.21}$$

(Indeed, their product is I_d.) Hence, the inverse function theorem asserts we may reduce U to a possibly smaller neighborhood of x so that F (now restricted to that new neighborhood) has a smooth inverse, and:

$$\mathrm{D}F^{-1}(F(x)) = (\mathrm{D}F(x))^{-1}. \tag{3.22}$$

We use F^{-1} to construct smooth curves on $\mathcal{M}$ that go through x. Specifically, pick an arbitrary $u \in \mathbb{R}^{d-k}$ and let

$$c(t) = F^{-1}\left(F(x) + t\begin{bmatrix} u \\ 0 \end{bmatrix}\right). \tag{3.23}$$

(Here, 0 denotes a zero vector of size k.) This is indeed smooth for t in some interval around 0 since F^{-1} is smooth in a neighborhood of $F(x)$. Clearly, $c(0) = x$. Furthermore,

$$F(c(t)) = F(x) + t\begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + tu_1 \\ \vdots \\ x_{d-k} + tu_{d-k} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Since the last k coordinates of $F(c(t))$ correspond to $h(c(t))$, we find that $h(c(t)) = 0$, so 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 (3.23),

$$c'(t) = \mathrm{D}F^{-1}\left(F(x) + t\begin{bmatrix} u \\ 0 \end{bmatrix}\right)\begin{bmatrix} u \\ 0 \end{bmatrix}.$$

In particular, at $t = 0$, using (3.21) and (3.22):

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

The specific form of $c'(0)$ is unimportant. What matters is that each $c'(0)$ of this form certainly belongs to $\mathrm{T}_x\mathcal{M}$ (3.20). Since the right-most matrix above has rank $d-k$, 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 $\mathcal{M}$, for each $x \in \mathcal{M}$, the set $\mathrm{T}_x\mathcal{M}$ is a linear subspace of $\mathcal{E}$ of some fixed dimension. These subspaces are the linearizations of the smooth set $\mathcal{M}$.

Definition 3.10. *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*) is called the* dimension of $\mathcal{M}$, *denoted by* $\dim \mathcal{M}$.

|| one.

Under Definition 3.6, $\dim \mathcal{M} = \dim \mathcal{E}$ for open submanifolds, and $\dim \mathcal{M} = \dim \mathcal{E} - k$ otherwise.

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 3.11. *The set* $\mathbb{R}^n$ *is a linear manifold of dimension* n *with tangent spaces* $\mathrm{T}_x\mathcal{M} = \mathbb{R}^n$ *for all* $x \in \mathbb{R}^n$.

|| on

Example 3.12. *The sphere* $\mathrm{S}^{n-1} = \{x \in \mathbb{R}^n : x^\top x = 1\}$ *is the zero level set of* $h(x) = x^\top x - 1$*, smooth from* $\mathbb{R}^n$ *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* $x \in \mathrm{S}^{n-1}$*. As a result,* S^{n-1} *is an embedded submanifold of* $\mathbb{R}^n$ *of dimension* $n-1$*. Furthermore, its tangent spaces are given by* $\mathrm{T}_x\mathrm{S}^{n-1} = \ker \mathrm{D}h(x) = \{v \in \mathbb{R}^n : x^\top v = 0\}$.

Example 3.13. *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:*

$$\mathrm{Sym}(2)_1 = \left\{ X = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} : x_1x_3 - x_2^2 = 0 \text{ and } X \neq 0 \right\}.$$

Consider $h\colon U \to \mathbb{R}$ *defined by* $h(X) = x_1x_3 - x_2^2$ *with* $U = \mathrm{Sym}(2)\backslash\{0\}$*: the open subset of* $\mathrm{Sym}(2)$ *obtained by removing the zero matrix. Clearly,* h *is smooth and* $h^{-1}(0) = \mathrm{Sym}(2)_1 \cap U = \mathrm{Sym}(2)_1$*. Furthermore,*

$$\mathrm{D}h(X)[\dot{X}] = \dot{x}_1x_3 + x_1\dot{x}_3 - 2x_2\dot{x}_2 = \begin{bmatrix} x_3 & -2x_2 & x_1 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix}.$$

Here, $\dot{X}$ 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 operator 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 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* $\ker \mathrm{D}h(X)$*:*

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

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 example, it is an exercise to show that Cartesian products of manifolds are manifolds.

Proposition 3.14. *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* $\dim \mathcal{M} + \dim \mathcal{M}'$ *such that*

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

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

In closing this section, we equip embedded submanifolds of $\mathcal{E}$ with the topology induced by $\mathcal{E}$. 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 3.15. *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*.*

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.

Definition 3.16. *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 3.17. *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 3.18. *Give a proof of Proposition 3.14.*

Exercise 3.19. *Give a proof of Proposition 3.17. In particular, deduce that the relative interior of the simplex,*

$$\Delta_+^{n-1} = \{x \in \mathbb{R}^n : x_1 + \cdots + x_n = 1 \text{ and } x_1, \ldots, x_n > 0\}, \tag{3.24}$$

is an embedded submanifold of $\mathbb{R}^n$. This is useful to represent non-vanishing discrete probability distributions. See also Exercise 3.33.

Exercise 3.20. *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 3.21. *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 (3.20)) is too low-dimensional.*

Exercise 3.22. *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 (3.20) is a linear subspace of dimension one in $\mathbb{R}^2$ for all $x \in \mathcal{P}$. This shows that Definition 3.6 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 construct a diffeomorphism F as in the proof of Theorem 3.8; then, derive a contradiction from the fact that $\mathcal{P}$ around the origin does not look like a one-dimensional curve. This exercise touches upon a more fundamental feature of Definition 3.6, namely, the fact that it only admits surfaces which, upon "zooming very close" to a point on the surface, can hardly be distinguished from linear subspaces of the same dimension in the same embedding space. The cross, cusp and double parabola all fail.*

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 3.23. *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 (in the usual sense) 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. We say F is* smooth *if it is smooth at all $x \in \mathcal{M}$.*

The following proposition states that a smooth map admits a smooth extension $\bar{F}$ in a neighborhood of *all* of $\mathcal{M}$. While this is not needed to establish results hereafter, it is convenient to shorten proofs and discussions; so much so that we typically think of it as the definition of a smooth map. See Section 3.10 for a proof sketch.

A neighborhood of $\mathcal{M}$ in $\mathcal{E}$ is an open subset of $\mathcal{E}$ which contains $\mathcal{M}$.

Proposition 3.24. *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 it admits a* smooth extension *$\bar{F}\colon U \to \mathcal{E}'$ in a neighborhood U of $\mathcal{M}$ in $\mathcal{E}$, so that $F(x) = \bar{F}(x)$ for all $x \in \mathcal{M}$, that is, F is the* restriction *of $\bar{F}$ to $\mathcal{M}$: $F = \bar{F}|_{\mathcal{M}}$.*

In particular, a function $f\colon \mathcal{M} \to \mathbb{R}$ is smooth if and only if there exists a smooth extension $\bar{f}\colon U \subseteq \mathcal{E} \to \mathbb{R}$ which coincides with f on $\mathcal{M}$: $f = \bar{f}|_{\mathcal{M}}$.

Definition 3.25. *A* scalar field *on a manifold $\mathcal{M}$ is a function $f\colon \mathcal{M} \to \mathbb{R}$. 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 $\mathfrak{F}(\mathcal{M})$.*

Exercise 3.26. *Give an example of an embedded submanifold $\mathcal{M}$ in a linear space $\mathcal{E}$ and a smooth function $f\colon \mathcal{M} \to \mathbb{R}$ for which there does not exist a smooth extension $\bar{f}\colon \mathcal{E} \to \mathbb{R}$ smooth on all of $\mathcal{E}$. Hint: use Example 3.13, or use Exercise 3.19 and consider a circle with a missing point.*

3.4 *The differential of a smooth map*

Think of $\mathbb{R}^n \to \mathbb{R}^m$.

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

$$\mathrm{D}\bar{F}(x)[v] = \lim_{t \to 0} \frac{\bar{F}(x + tv) - \bar{F}(x)}{t}. \tag{3.25}$$

This tells us how $\bar{F}(x)$ changes when we push x along v. Applying this definition to a smooth map $F\colon \mathcal{M} \to \mathcal{M}'$ 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.

Okay

A natural alternative is to rely on Definition 3.7: for any tangent vector v of $\mathcal{M}$ at x, there exists a smooth curve c on $\mathcal{M}$ passing through x with velocity v. Then, $t \mapsto F(c(t))$ itself defines a smooth (why?) curve on $\mathcal{M}'$ passing through $F(x)$ at $t = 0$, with a certain velocity. By definition, that velocity is a tangent vector of $\mathcal{M}'$ at $F(x)$.

One could also propose to resort to a smooth extension of F: we show in a moment that these two perspectives are equivalent. The smooth curve approach is preferred because it generalizes better.

Definition 3.27. *The differential of $F\colon \mathcal{M} \to \mathcal{M}'$ at x is a linear operator $\mathrm{D}F(x)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_{F(x)}\mathcal{M}'$ defined by:*

$$\mathrm{D}F(x)[v] = \left.\frac{\mathrm{d}}{\mathrm{d}t} F(c(t))\right|_{t=0}, \tag{3.26}$$

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

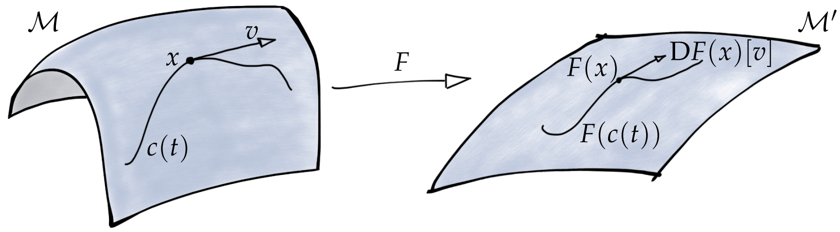


Figure 3.2: $F\colon \mathcal{M} \to \mathcal{M}'$ is a map between two manifolds. Given $x \in \mathcal{M}$ and v tangent at x, we select a smooth curve $c(t)$ passing through x with velocity v. Then, $F(c(t))$ is a curve on $\mathcal{M}'$ passing through $F(x)$ with velocity $\mathrm{D}F(x)[v]$. The latter tells us how $F(x)$ changes when we push x along v.

We must clarify two things: (a) that (3.26) 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. Let $\mathcal{M}$ and $\mathcal{M}'$ be embedded submanifolds of $\mathcal{E}$ and $\mathcal{E}'$. Then, for a smooth map $F\colon \mathcal{M} \to \mathcal{M}'$, there exists a smooth extension $\bar{F}\colon U \subseteq \mathcal{E} \to \mathcal{E}'$ such that $F = \bar{F}|_{\mathcal{M}}$. Using that $F \circ c = \bar{F} \circ c$ and the usual chain rule for $\bar{F} \circ c$ (since it is a composition of functions between open subsets of linear spaces), we find:

$$\begin{aligned} \mathrm{D}F(x)[v] &= \frac{\mathrm{d}}{\mathrm{d}t} F(c(t))\bigg|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \bar{F}(c(t))\bigg|_{t=0} = \mathrm{D}\bar{F}(c(0))[c'(0)] = \mathrm{D}\bar{F}(x)[v]. \end{aligned} \tag{3.27}$$

This can be summarized as follows.

Proposition 3.28. *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 equation (3.26) depends on c only through $c(0)$ and $c'(0)$, as required. One may wonder whether (3.27) depends on the choice of smooth extension $\bar{F}$: it does not. Indeed, if $\tilde{F}$ is another smooth extension of F, then for all smooth curves c with $c(0) = x$ and $c'(0) = v$ we have

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

Example 3.29. *Given a real, symmetric matrix $A \in \mathrm{Sym}(n)$, the* Rayleigh quotient *at a nonzero vector $x \in \mathbb{R}^n$ 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:*

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

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 $\mathbb{R}^n$ by $\bar{f}(x) = x^\top A x$, hence f is smooth according to Definition 3.23. Using this smooth extension, we can also obtain an *expression for its differential. Indeed, for all* $v \in \mathbb{R}^n$,

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

Hence, by (3.27)*, for all* $v \in \mathrm{T}_x\mathrm{S}^{n-1} = \{v \in \mathbb{R}^n : x^\top v = 0\}$,

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

Exercise 3.30. *For smooth maps* $F_1, F_2 \colon \mathcal{M} \to \mathcal{E}'$ *and real numbers* a_1, a_2*, show that* $F \colon x \mapsto a_1F_1(x) + a_2F_2(x)$ *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 3.31. *For smooth maps* $f \in \mathfrak{F}(\mathcal{M})$ *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:*

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

Exercise 3.32. *Let* $F \colon \mathcal{M} \to \mathcal{M}'$ *and* $G \colon \mathcal{M}' \to \mathcal{M}''$ *be smooth, where* $\mathcal{M}, \mathcal{M}', \mathcal{M}''$ *are embedded submanifolds of* $\mathcal{E}, \mathcal{E}', \mathcal{E}''$ *respectively. Show composition preserves smoothness, that is,* $G \circ F \colon x \mapsto G(F(x))$ *is smooth. Furthermore, show that we have a chain rule:*

$$\mathrm{D}(G \circ F)(x)[v] = \mathrm{D}G(F(x))[\mathrm{D}F(x)[v]]. \tag{3.28}$$

Exercise 3.33. *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 3.19. We call* $\mathcal{N}$ *an* embedded submanifold of $\mathcal{M}$*; see also Section 8.14.*

3.5 *Vector fields and the tangent bundle*

For a function $f \colon \mathcal{M} \to \mathbb{R}$, the gradient (still to be defined) associates a tangent vector to each point of $\mathcal{M}$. Such a map V from x to an element of $\mathrm{T}_x\mathcal{M}$ is called a *vector field* on $\mathcal{M}$. 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 3.34. *The* tangent bundle *of a manifold $\mathcal{M}$ is the disjoint union of the tangent spaces of $\mathcal{M}$:*

$$\mathrm{T}\mathcal{M} = \{(x, v) : x \in \mathcal{M} \text{ and } v \in \mathrm{T}_x\mathcal{M}\}. \tag{3.29}$$

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 foot *or* base *of v is x.*

OK

Theorem 3.35. *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 $2 \dim \mathcal{M}$.*

The topology we choose for $\mathrm{T}\mathcal{M}$ is the embedded submanifold topology, as in Definition 3.15. This is different from the so-called disjoint union topology, which we never use.

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.

We restrict the domain U if need be to secure the stronger property $\operatorname{rank} \mathrm{D}h(x) = k$ for all $x \in U$: this is always possible, see Lemma 3.62. 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 $H\colon U \times \mathcal{E} \to \mathbb{R}^{2k}$ 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 $\mathcal{L}(x, v)\colon \mathcal{E} \to \mathcal{E}$ is some linear operator 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. 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 indeed 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 $\dim \mathrm{T}\mathcal{M} = \dim \mathrm{T}_{(x,v)}\mathrm{T}\mathcal{M} = 2 \dim \mathcal{E} - 2k = 2 \dim \mathcal{M}$. (We mention that the tangent spaces to $\mathrm{T}\mathcal{M}$ are *not* of the form $\mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M}$: see the discussion in Section 5.1.) $\square$

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 $\mathrm{T}\mathcal{M}$ is a manifold, we can now use Definition 3.23 to define smooth vector fields as particular smooth maps from $\mathcal{M}$ to $\mathrm{T}\mathcal{M}$.

Definition 3.36. *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 vectors fields is denoted by $\mathfrak{X}(\mathcal{M})$.*

Some authors call smooth vector fields simply vector fields.

$\Gamma(TM)$ in Fredrich Schuller's notes.

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 ambient space.

Proposition 3.37. *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 $U \subseteq \mathcal{E}$ (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 3.24, 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. $\square$

Exercise 3.38. *For $f \in \mathfrak{F}(\mathcal{M})$ and $V, W \in \mathfrak{X}(\mathcal{M})$, show that the vector fields fV and $V + W$ defined by $(fV)(x) = f(x)V(x)$ and $(V + W)(x) = V(x) + W(x)$ are smooth.*

For pointwise scaling, we purposefully write fV and not Vf. Later, we will give a different meaning to the notation Vf.

In closing, we note a useful identification for the tangent bundle of a product manifold $\mathcal{M} \times \mathcal{M}'$, based on a reordering of parameters:

$$\begin{aligned} \mathrm{T}(\mathcal{M} \times \mathcal{M}') &= \{((x, x'), (v, v')) : x \in \mathcal{M}, v \in \mathrm{T}_x\mathcal{M}, x' \in \mathcal{M}', v' \in \mathrm{T}_{x'}\mathcal{M}'\} \\ &\equiv \{((x, v), (x', v')) : x \in \mathcal{M}, v \in \mathrm{T}_x\mathcal{M}, x' \in \mathcal{M}', v' \in \mathrm{T}_{x'}\mathcal{M}'\} \\ &= \mathrm{T}\mathcal{M} \times \mathrm{T}\mathcal{M}'. \end{aligned} \tag{3.30}$$

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

Ohh

of essentially all Riemannian optimization algorithms. We can achieve this by following any smooth curve $c: I \to \mathcal{M}$ such that $c(0) = x$ and $c'(0) = v$, but of course there exist many such curves. A *retraction* picks a particular curve for every possible $(x, v) \in \mathrm{T}\mathcal{M}$. Furthermore, this choice of curve depends smoothly on (x, v), in a sense we make precise below. Let us start with an example.

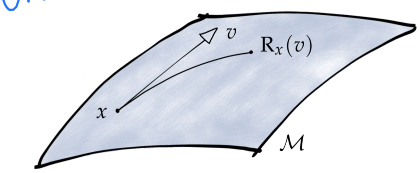


Figure 3.3: Retractions allow us to move away from x along $v \in \mathrm{T}_x\mathcal{M}$ while staying on $\mathcal{M}$.

This is an example of a retraction based on *metric projection*: we study them later in Proposition 5.42.

Example 3.39. *Let x be a point on the sphere S^{n-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 $\mathbb{R}^n$ then to project back to the sphere:*

$$\mathrm{R}_x(v) \triangleq \frac{x + v}{\|x + v\|} = \frac{x + v}{\sqrt{1 + \|v\|^2}}. \tag{3.31}$$

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

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

Evidently, $c(0) = x$. This holds because $\mathrm{R}_x(0) = x$. Furthermore, one can compute $c'(0) = v$, that is: locally around x, up to first order, the retraction curve moves along v. Another way to state this is via the chain rule:

$$v = c'(0) = \mathrm{DR}_x(0)[v],$$

where $\mathrm{DR}_x(0)$ is understood as per Definition 3.27 for $\mathrm{R}_x: \mathrm{T}_x\mathrm{S}^{n-1} \to \mathrm{S}^{n-1}$. In other words: $\mathrm{DR}_x(0): \mathrm{T}_x\mathrm{S}^{n-1} \to \mathrm{T}_x\mathrm{S}^{n-1}$ is the identity map.

Another reasonable choice is to move away from x along the great circle on S^{n-1} traced out by v. Specifically,

$$\mathrm{R}_x(v) \triangleq \cos(\|v\|)x + \frac{\sin(\|v\|)}{\|v\|}v, \tag{3.32}$$

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^{n-1} passing through x at $t = 0$ with velocity $c'(0) = v$.

Such curves are called *geodesics*: we study them in Section 5.8. The retraction is called the *exponential map*: we study those later, in Section 10.2.

More generally, we consider the following important definition.

Definition 3.40. *A* retraction *on $\mathcal{M}$ is a smooth map $\mathrm{R}: \mathrm{T}\mathcal{M} \to \mathcal{M}$ with the following properties. For each $x \in \mathcal{M}$, let $\mathrm{R}_x: \mathrm{T}_x\mathcal{M} \to \mathcal{M}$ be the restriction of* R *at x, so that $\mathrm{R}_x(v) = \mathrm{R}(x, v)$. Then,*

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

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

Smoothness of R is understood in the sense of Definition 3.23 for a map from $\mathrm{T}\mathcal{M}$ to $\mathcal{M}$, that is, R is smooth if and only if there exists a smooth function $\bar{\mathrm{R}}$ from a neighborhood of $\mathrm{T}\mathcal{M}$ in $\mathcal{E} \times \mathcal{E}$ into $\mathcal{E}$ such that $\mathrm{R} = \bar{\mathrm{R}}|_{\mathrm{T}\mathcal{M}}$. In Example 3.39, consider $\bar{\mathrm{R}}(x, v) = (x+v)/\sqrt{1 + \|v\|^2}$: this is a smooth extension of (3.31) to all of $\mathcal{E} \times \mathcal{E}$, confirming smoothness. It is instructive to check that, for linear manifolds, $\mathrm{R}_x(v) = x + v$ is a retraction.

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 the domain. For example, this is the case for the manifold of fixed-rank matrices (Section 7.5).

Exercise 3.41. *Let $\mathcal{M}, \mathcal{M}'$ be equipped with retractions R, R'. Show that $\mathrm{R}'' \colon \mathrm{T}(\mathcal{M} \times \mathcal{M}') \to \mathcal{M} \times \mathcal{M}'$ 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: 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, unless it is clear from context in which case we often omit it.

Definition 3.42. *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 $x \in \mathcal{M}$.*

Notes:

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

Definition 3.44. *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 3.45. *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 3.46. *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. Notice that a Riemannian submanifold is not merely a submanifold with some Riemannian structure: the words single out a precise choice of metric.

Example 3.47. *Endow $\mathbb{R}^n$ with the standard metric $\langle u, v \rangle = u^\top v$ and consider the sphere $\mathrm{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$, embedded in $\mathbb{R}^n$. With the inherited metric $\langle u, v \rangle_x = \langle u, v \rangle = u^\top v$ on each tangent space $\mathrm{T}_x\mathrm{S}^{n-1}$, the sphere becomes a Riemannian submanifold of $\mathbb{R}^n$.*

3.8 Riemannian gradients

Let $\mathcal{M}$ be a Riemannian manifold, that is, a manifold endowed 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 3.48. *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 these identities:*

$$\forall (x, v) \in \mathrm{T}\mathcal{M}, \qquad \mathrm{D}f(x)[v] = \langle v, \mathrm{grad} f(x) \rangle_x, \tag{3.33}$$

where $\mathrm{D}f(x)$ is as in Definition 3.27 and $\langle \cdot, \cdot \rangle_x$ is the Riemannian metric.

It is an exercise to show that $\mathrm{grad} f(x)$ is indeed uniquely determined by (3.33) for each x in $\mathcal{M}$, confirming $\mathrm{grad} f$ is well defined.

Following the definition, to compute the gradient of f, a direct approach is to obtain an expression for $\mathrm{D}f(x)[v]$ and to manipulate it until it takes the form $\langle v, \cdot \rangle_x$: this yields the gradient by identification—more on this in Section 4.7. An indirect approach is through the following proposition, based on retractions.

Proposition 3.49. *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 $x \in \mathcal{M}$,*

See also Exercise 10.67.

$$\mathrm{grad} f(x) = \mathrm{grad}(f \circ \mathrm{R}_x)(0), \tag{3.34}$$

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.

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

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

since $\mathrm{R}_x(0) = x$ and $\mathrm{D}\mathrm{R}_x(0)$ 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. $\square$

For the important special case of Riemannian submanifolds, we can make a more explicit statement. Indeed, $\mathcal{E}$ is then itself equipped with a metric $\langle \cdot, \cdot \rangle$, so that combining (3.27) with (3.33) we find:

$$\begin{aligned} \forall (x, v) \in \mathrm{T}\mathcal{M}, \quad \langle v, \mathrm{grad}f(x)\rangle_x &= \mathrm{D}f(x)[v] \\ &= \mathrm{D}\bar{f}(x)[v] = \left\langle v, \mathrm{grad}\bar{f}(x)\right\rangle, \end{aligned} \tag{3.35}$$

where $\bar{f}$ is a smooth extension of f. Now comes the crucial observation: $\mathrm{T}_x\mathcal{M}$ 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

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

with one component in $\mathrm{T}_x\mathcal{M}$ and another orthogonal to $\mathrm{T}_x\mathcal{M}$, that is, $\mathrm{grad}\bar{f}(x)_\| \in \mathrm{T}_x\mathcal{M}$ 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 (3.35) that, for all (x, v) in $\mathrm{T}\mathcal{M}$,

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

Now we use that $\mathrm{grad}\bar{f}(x)_\|$ is tangent at x and that the Riemannian metric is merely a restriction of the Euclidean metric to the tangent spaces to state:

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

Since the gradient is uniquely defined, it follows that, for Riemannian submanifolds,

$$\mathrm{grad}f(x) = \mathrm{grad}\bar{f}(x)_\|. \tag{3.36}$$

In other words: to compute $\mathrm{grad}f$, first obtain an expression for the (classical) gradient of any smooth extension of f, then orthogonally project to the tangent spaces. Since an expression for $\bar{f}$ is typically available in optimization, this is a practical recipe.

This key formula motivates us to introduce orthogonal projectors.

Definition 3.50. *Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$ equipped with the metric $\langle \cdot, \cdot \rangle$. With*

$$\mathrm{Proj}_x \colon \mathcal{E} \to \mathrm{T}_x\mathcal{M} \subseteq \mathcal{E},$$

we denote the projector from $\mathcal{E}$ to $\mathrm{T}_x\mathcal{M}$, orthogonal with respect to $\langle \cdot, \cdot \rangle$. Being a projector, Proj_x is a linear operator such that $\mathrm{Proj}_x \circ \mathrm{Proj}_x = \mathrm{Proj}_x$. Being orthogonal, $\langle u - \mathrm{Proj}_x(u), v \rangle = 0$ for all $v \in \mathrm{T}_x\mathcal{M}$ and $u \in \mathcal{E}$.

The above discussion can be summarized into the following useful proposition. Note that, for an open submanifold, Proj_x is the identity operator.

Proposition 3.51. *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*

$$\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad} \bar{f}(x)), \tag{3.37}$$

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

Example 3.52. *We continue with the Rayleigh quotient from Example 3.29: $f(x) = x^\top A x$. Equip $\mathbb{R}^n$ with the standard Euclidean metric $\langle u, v \rangle = u^\top v$. Then, using $A = A^\top$, for all $v \in \mathbb{R}^n$,*

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

Hence, by identification with Definition 3.48,

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

To get a notion of gradient for f on S^{n-1}, we need to choose a Riemannian metric for S^{n-1}. One convenient choice is to turn S^{n-1} into a Riemannian submanifold of $\mathbb{R}^n$ by endowing it with the induced Riemannian metric. In that scenario, Proposition 3.51 suggests we should determine the orthogonal projectors of S^{n-1}. Since, for the Euclidean metric,

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

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

$$\mathrm{Proj}_x(u) = u - (x^\top u)x = (I_n - xx^\top)u. \tag{3.38}$$

It follows that the Riemannian gradient of f on S^{n-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 $x \in \mathrm{S}^{n-1}$,

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

In other words: all points where the gradient vanishes are eigenvectors of A. (Conversely: the gradient vanishes at all unit-norm eigenvectors of A.) This fact is crucially important to understand the behavior of optimization algorithms for f on S^{n-1}.

Orthogonal projectors are self-adjoint (Definition 3.3)—this fact will prove useful later on.

Proposition 3.53. *Let Proj_x be the orthogonal projector from $\mathcal{E}$ to a linear subspace of $\mathcal{E}$. Then, Proj_x is self-adjoint, that is,*

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

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

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

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

This concludes the proof. □

Exercise 3.54. *Show that $\mathrm{grad} f(x)$ is uniquely defined by (3.33).*

The following 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 $\mathbb{R}^n$ and it is a Riemannian manifold, it is *not* a Riemannian submanifold of $\mathbb{R}^n$.

Exercise 3.55. *In Exercise 3.19, we considered the relative interior of the simplex,*

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

as an embedded submanifold of $\mathbb{R}^n$. Its tangent spaces are given by

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

Show that $\langle u, v \rangle_x = \sum_{i=1}^n \frac{u_i v_i}{x_i}$ defines a Riemannian metric on $\mathcal{M}$. Then, considering a smooth function $f \colon \mathcal{M} \to \mathbb{R}$ and a smooth extension $\bar{f}$ on a neighborhood of $\mathcal{M}$ in $\mathbb{R}^n$ (equipped with the canonical Euclidean metric), give an expression for $\mathrm{grad} f(x)$ in terms of $\mathrm{grad} \bar{f}(x)$.

This is called the Fisher–Rao metric.

Exercise 3.56. *Let $\mathcal{E}$ be a Euclidean space, and let $\mathcal{L}(\mathcal{E}, \mathcal{E})$ denote the set of linear operators from $\mathcal{E}$ into itself: this is a linear space. If $\mathcal{E}$ is identified with*

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.

$\mathbb{R}^d$, 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

$$\mathrm{Proj}\colon x \mapsto \mathrm{Proj}_x$$

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

Exercise 3.57. *Let $\mathcal{M}, \mathcal{M}'$ be Riemannian manifolds with metrics $\langle\cdot,\cdot\rangle^a$ and $\langle\cdot,\cdot\rangle^b$. The product $\mathcal{M} \times \mathcal{M}'$ is itself a manifold. Make it into a Riemannian manifold with the* product metric*: for all $(u, u'), (v, v')$ in the tangent space $\mathrm{T}_{(x,x')}(\mathcal{M} \times \mathcal{M}')$,*

$$\left\langle (u, u'), (v, v') \right\rangle_{(x,x')} = \langle u, v\rangle_x^a + \left\langle u', v'\right\rangle_{x'}^b.$$

For a smooth function $f\colon \mathcal{M} \times \mathcal{M}' \to \mathbb{R}$, show that

$$\mathrm{grad} f(x, y) = \left(\mathrm{grad} g_y(x), \mathrm{grad} g_x(y)\right),$$

where $g_y(x) = f(x, y)$ is a function on $\mathcal{M}$ (fixed y) and $g_x(y) = f(x, y)$ is a function on $\mathcal{M}'$ (fixed x).

3.9 *Local frames**

This section introduces a technical tool which proves useful in certain proofs: the reader can safely skip it for a first reading. We show that embedded submanifolds admit *local frames* (as defined below) around any point. 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 3.56 which is restricted to Riemannian submanifolds.

Definition 3.58. *Given a point x on a manifold $\mathcal{M}$ of dimension d, a* local frame *around x is a set of smooth vector fields $W_1, \ldots, W_d$ defined on a neighborhood of x in $\mathcal{M}$ such that, for all y in that neighborhood, the vectors $W_1(y), \ldots, W_d(y)$ form a basis for the tangent space $\mathrm{T}_y\mathcal{M}$.*

[Lee12, pp177–179]

Proposition 3.59. *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. If $\mathcal{M}$ is open in $\mathcal{E}$, the statement is clear: pick a basis $e_1, \ldots, e_d$ of $\mathcal{E}$ and corresponding constant vector fields $E_1(y) = e_1, \ldots, E_d(y) = e_d$. The restriction of these vector fields to $\mathcal{M}$ forms a local frame.

Otherwise, let $h\colon U \to \mathbb{R}^k$ be a local defining function for $\mathcal{M}$ around x, with U open in $\mathcal{E}$ and $\dim \mathcal{M} = d - k$. As we did in the proof of Theorem 3.8, identify $\mathcal{E}$ with $\mathbb{R}^d$ by picking a basis $e_1, \ldots, e_d$ of $\mathcal{E}$, ordered so that

We construct a rather special local frame called a *coordinate frame*: compare with Sections 8.3 and 8.8.

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

with $B \in \mathbb{R}^{k\times k}$ invertible. Consider anew the function $F\colon U \to \mathbb{R}^d$:

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

where $y_1, \ldots, y_d$ denote the coordinates of $y \in \mathcal{E}$. Following the same argument based on the inverse function theorem, possibly after restricting the domain U to a smaller neighborhood of x, the function F admits a smooth inverse. Thus, for all $y \in \mathcal{M} \cap U$, we may consider these smooth curves, for $i = 1, \ldots, d-k$:

$$c_i(t) = F^{-1}\left(F(y) + te_i\right),$$

where e_i is here conflated with its coordinate representation: a canonical basis vector in $\mathbb{R}^d$. Since

$$F(c_i(t)) = F(y) + te_i$$

has its last k entries all equal to zero (owing to $y \in \mathcal{M} \cap U$ and $1 \leq i \leq d-k$), it follows that $h(c_i(t)) = 0$ for all t in the domain of c_i (a neighborhood of $t = 0$), so that c_i is a smooth curve on $\mathcal{M}$. Furthermore, $c_i(0) = y$. As a result, we may define vector fields $W_1, \ldots, W_{d-k}$ on $\mathcal{M} \cap U$ as

$$W_i(y) = c_i'(0).$$

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 and the identity $\mathrm{D}F^{-1}(F(y)) = (\mathrm{D}F(y))^{-1}$ to see that

$$W_i(y) = \mathrm{D}F^{-1}(F(y))e_i = (\mathrm{D}F(y))^{-1}e_i. \tag{3.39}$$

Since $\mathrm{D}F(x)$ is invertible, $\mathrm{D}F(y)$ is invertible for y in a neighborhood of x in $\mathcal{E}$. If need be, reduce the domain U so that $\mathrm{D}F(y)$ is invertible for all $y \in \mathcal{M} \cap U$. Then, equation (3.39) shows $W_1(y), \ldots, W_{d-k}(y)$ are linearly independent for all y in $\mathcal{U} = \mathcal{M} \cap U$: a neighborhood of x in $\mathcal{M}$. □

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

Proposition 3.60. *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 an arbitrary point $x \in \mathcal{M}$ and a local frame $W_1, \ldots, W_d$ defined on a neighborhood $\mathcal{U}$ of x in $\mathcal{M}$, where $\dim \mathcal{M} = d$. By the properties of local frames, there exist unique functions $g_1, \ldots, g_d \colon \mathcal{U} \to \mathbb{R}$ such that

$$\mathrm{grad} f(y) = g_1(y)W_1(y) + \cdots + g_d(y)W_d(y)$$

for all $y \in \mathcal{U}$. If we show $g_1, \ldots, g_d$ are smooth, then $\operatorname{grad} f$ is smooth on $\mathcal{U}$; showing this in a neighborhood of each point, it follows that $\operatorname{grad} f$ is a smooth vector field on $\mathcal{M}$. To show that the g_i's are indeed smooth, consider this 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

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

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 (justify this). Hence, The solution depends smoothly on y. $\square$

It is sometimes useful to orthonormalize a local frame: the Gram–Schmidt procedure achieves this while preserving smoothness [Lee12, Prop. 13.6], [Lee18, Prop. 2.8].

3.10 *Notes and references*

The main sources for this chapter are [AMS08], [Lee12], [Lee18] and [O'N83]. The definitions given here in a restricted setting are compatible with the general theory of differential geometry. As it is useful to be able to connect the definitions we use to a broader body of work, this section gives a detailed outline of equivalence to a modern account by Lee [Lee12]. When we say "in the usual sense," we mean according to the standard definitions in that reference. See also Section 8.14.

A (finite dimensional) linear space $\mathcal{E}$ is a smooth manifold in a natural way [Lee12, Example 1.24]. Our definition of embedded submanifold of a linear space is equivalent to the usual definition. Its statement is in fact very close to the one by Darling [Dar94, p56], who introduces it with the same motivation, namely, to avoid charts at first.

Proposition 3.61. *A subset $\mathcal{M}$ of $\mathcal{E}$ is an embedded submanifold of $\mathcal{E}$ in the usual sense [Lee12, pp98–99] if and only if it is an embedded submanifold of $\mathcal{E}$ under Definition 3.6.*

Proof. By [Lee12, Example 1.26, Prop. 5.1], open subsets of $\mathcal{E}$ are embedded submanifolds of $\mathcal{E}$ of codimension 0, and vice versa. Furthermore, by [Lee12, Prop. 5.16], $\mathcal{M}$ is an embedded submanifold of $\mathcal{E}$ of positive codimension k if and only if every $x \in \mathcal{M}$ has a neighborhood U in $\mathcal{E}$ satisfying $\mathcal{M} \cap U = h^{-1}(0)$ for some smooth function $h\colon U \to \mathbb{R}^k$ such that $\operatorname{rank} \mathrm{D}h(x') = k$ for all $x' \in U$. The only difference with Definition 3.6 is that, in that definition, we only require $\mathrm{D}h$ to have rank k at x. This is easily resolved by Lemma 3.62 below. $\square$

Lemma 3.62. *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 another neighborhood $U' \subseteq U$ of x such that $\operatorname{rank} \mathrm{D}h(x') = k$ for all x' in U'.*

[Lee12, Prop. 4.1]

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 this continuous function on U: $g(x') = \det(A(x')A(x')^\top)$. Notice that $U' = U \backslash g^{-1}(0)$. Since $g^{-1}(0)$ is closed (because it is the preimage of the closed set $\{0\}$), it follows that U' is open. Furthermore, by assumption, x is in U'. Hence, it suffices to restrict the domain of h to U'. $\square$

All maps we call smooth under Definition 3.23 are smooth in the usual sense and vice versa.

Proposition 3.63. *Let $\mathcal{M}, \mathcal{M}'$ be embedded submanifolds of linear spaces $\mathcal{E}, \mathcal{E}'$. A map $F\colon \mathcal{M} \to \mathcal{M}'$ is smooth in the sense of Definition 3.23 if and only if it is smooth in the usual sense [Lee12, p34].*

Proof. Assume F is smooth in the usual sense. We must show it admits smooth local extensions. Since $\mathcal{M}$ is embedded in $\mathcal{E}$, by [Lee12, Thm. 5.8, Ex. 1.22], for each $x \in \mathcal{M}$, there exists a *slice chart* (U, φ) for $\mathcal{M}$ around x, that is, U is a neighborhood of x in $\mathcal{E}$ and $\varphi\colon U \to \varphi(U) \subseteq \mathbb{R}^d$ is a diffeomorphism (smooth map with smooth inverse) such that $\varphi(y) = (\hat{\varphi}(y), 0, \ldots, 0)$ for every $y \in \mathcal{M} \cap U$, with $\dim \mathcal{E} = d$ and $\hat{\varphi}\colon U \to \mathbb{R}^{d-k}$, where $\dim \mathcal{M} = d - k$. Let $T\colon \mathbb{R}^d \to \mathbb{R}^d$ be the linear operator that zeros out the last k entries of a vector. Then, $\varphi^{-1} \circ T \circ \varphi$ is a map from U to U which "projects" U to $\mathcal{M} \cap U$. It is smooth as the composition of smooth maps [Lee12, Prop. 2.10]. We may restrict its codomain to $\mathcal{M} \cap U$ and it is still smooth since $\mathcal{M}$ is embedded in $\mathcal{E}$ [Lee12, Cor. 5.30]: this allows us to compose it with F. Further consider $i\colon \mathcal{M}' \to \mathcal{E}'$, defined by $i(z) = z$: this is the *inclusion map*, smooth by definition of embedded submanifolds [Lee12, pp98–99]. Then, consider $\tilde{F}\colon U \to \mathcal{E}'$, defined by $\tilde{F} = i \circ F \circ \varphi^{-1} \circ T \circ \varphi$. This is indeed smooth since it is a composition of smooth maps, and $\tilde{F}(y) = F(y)$ for all $y \in \mathcal{M} \cap U$, hence this is a local smooth extension. Since such an extension exists for arbitrary $x \in \mathcal{M}$, we conclude that F is smooth in the sense of Definition 3.23.

The other way around, assume F is smooth in the sense of Definition 3.23. For an arbitrary $x \in \mathcal{M}$, use that $\mathcal{M}, \mathcal{M}'$ are embedded in $\mathcal{E}, \mathcal{E}'$ to pick slice charts as above: with U a neighborhood of x in $\mathcal{E}$ and V a neighborhood of $F(x)$ in $\mathcal{E}'$, let $\varphi\colon U \to \varphi(U) \subseteq \mathbb{R}^d$ and $\psi\colon V \to \psi(V) \subseteq \mathbb{R}^{d'}$ be diffeomorphisms such that $\varphi(y) = (\hat{\varphi}(y), 0, \ldots, 0)$ for $y \in \mathcal{M} \cap U$ and $\psi(z) = (\hat{\psi}(z), 0, \ldots, 0)$ for $z \in \mathcal{M}' \cap V$. Then, $(\mathcal{M} \cap U, \hat{\varphi})$ and $(\mathcal{M}' \cap V, \hat{\psi})$ are charts of $\mathcal{M}, \mathcal{M}'$, respectively [AMS08, Prop. 3.3.2]. To show that F is smooth in the usual sense, it is sufficient to show that the coordinate representation of F in these charts, $\hat{\psi} \circ F \circ \hat{\varphi}^{-1}$, is smooth (in the usual sense)—if need be, replace U by $U \cap F^{-1}(V)$ to ensure the composition is well defined. To this end, use our assumption that there exists a local smooth extension $\tilde{F}\colon U \to V$ for F—if need be, the neighborhoods U, V of the charts and of the smooth extension can be reduced to their respective intersections. Then, $\hat{\psi} \circ F \circ \hat{\varphi}^{-1} = T_1 \circ \psi \circ \tilde{F} \circ \varphi^{-1} \circ T_2$, where T_2 pads its input with zeros and T_1 trims the trailing entries of its input (all zeros in this case), so as to ensure equality. This shows smoothness as it is a composition of smooth maps. □

Thus, there is no need to distinguish between these two notions of smoothness. Proposition 3.24 states that any smooth map between embedded submanifolds of linear spaces can be smoothly extended to a neighborhood of its domain, and vice versa. We give a proof below, based on the *tubular neighborhood theorem* [Lee12, Thm. 6.24], [Lee18, Thm. 5.25].

Proof of Proposition 3.24. Let $\mathcal{M}, \mathcal{M}'$ be embedded submanifolds of linear spaces $\mathcal{E}, \mathcal{E}'$ and let $F\colon \mathcal{M} \to \mathcal{M}'$ be a smooth map. We need to show it can be smoothly extended to a map $\bar{F}$ from a neighborhood of $\mathcal{M}$ to $\mathcal{E}'$. Since $\mathcal{M}$ is embedded in $\mathcal{E} \equiv \mathbb{R}^d$, it admits a *tubular neighborhood*: there exists a neighborhood U of $\mathcal{M}$ in $\mathcal{E}$ and a smooth map $r\colon U \to \mathcal{M}$ such that $r(x) = x$ for all $x \in \mathcal{M}$ [Lee12, Prop. 6.25]. Since composition preserves smoothness [Lee12, Prop. 2.10], $F \circ r\colon U \to \mathcal{M}'$ is smooth. Finally, since $\mathcal{M}'$ is embedded in $\mathcal{E}'$, the *inclusion map* $i\colon \mathcal{M}' \to \mathcal{E}'$, defined by $i(x) = x$, is smooth—this is part of the definition [Lee12, pp98–99]. Overall, the map $\bar{F} = i \circ F \circ r\colon U \to \mathcal{E}'$ is smooth, and it verifies $\bar{F}(x) = i(F(r(x))) = F(x)$ for all $x \in \mathcal{M}$. This is a smooth extension for F around $\mathcal{M}$. As a side note, observe that $\bar{F}(x) \in \mathcal{M}'$ for all $x \in U$.

r is a *(topological) retraction*: this is different from (but related to) our notion of retraction.

The other way around, given a map $F\colon \mathcal{M} \to \mathcal{M}'$, assume there exists $\bar{F}\colon U \subseteq \mathcal{E} \to \mathcal{E}'$, smooth on a neighborhood U of $\mathcal{M}$ in $\mathcal{E}$, such that $\bar{F}(x) = F(x)$ for all $x \in \mathcal{M}$. Then, by [Lee12, Thm. 5.27], since $\mathcal{M}$ is an embedded submanifold of $\mathcal{E}$,[1] the map $\bar{F}|_{\mathcal{M}}\colon \mathcal{M} \to \mathcal{E}'$ is a smooth map. Furthermore, since the image of $\bar{F}|_{\mathcal{M}}$ is contained in $\mathcal{M}'$, we may restrict the codomain. By [Lee12, Cor. 5.30], since $\mathcal{M}'$ is an embedded submanifold of $\mathcal{E}'$, the map $F = \bar{F}|_{\mathcal{M}}\colon \mathcal{M} \to \mathcal{M}'$ is smooth. □

[1] Formally: $\mathcal{M}$ is an embedded submanifold of $\mathcal{E}$, and it is included in U which is an open submanifold of $\mathcal{E}$, hence $\mathcal{M}$ is an embedded submanifold of U: simply intersect the domain of any local defining map of $\mathcal{M}$ in $\mathcal{E}$ with U and apply [Lee12, Prop. 5.16]. It is to $\bar{F}$ as a smooth map from U to $\mathcal{M}'$ that we apply [Lee12, Thm. 5.27].

The converse follows from Proposition 3.63: a map $F\colon \mathcal{M} \to \mathcal{M}'$ which admits a smooth extension on a neighborhood of $\mathcal{M}$ in $\mathcal{E}$ is smooth since this smooth extension can be used as a local smooth extension around every point. Thus, we could also present the characterization of smoothness offered by Proposition 3.24 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, that the gradient vector field of a smooth function is smooth.

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: see Exercise 3.26.

The proof of Proposition 3.60 generalizes seamlessly to establish a broader result known as the *musical isomorphism*, or the *tangent–cotangent isomorphism* [Lee12, pp341–343][O'N83, Prop. 3.10]: through the Riemannian metric, smooth vector fields on $\mathcal{M}$ are one-to-one with smooth *one-forms* or *covector fields* on $\mathcal{M}$; these maps transform smooth vector fields into smooth scalar fields—$\mathrm{D}f$ is an example.

Definition 3.46 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 to be called a Riemannian submanifold; this is the case of Lee for example [Lee18, p15].

The rank condition in Definition 3.6 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 necessarily have rank drops.

Here is an example: consider $h\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $h(x) = (x_1x_2)^2$. Its zero-set is a cross in $\mathbb{R}^2$. Yet, $\mathrm{D}h(x) = \left[\, 2x_1x_2^2,\ 2x_1^2x_2 \,\right]$ 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: it has to have maximal rank too. (As we mention in Section 8.14, it is also acceptable to require a constant rank on a neighborhood of the zero-set: h fails that test.)

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 $f : \mathcal{M} \to \mathbb{R}$ 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 further manifolds in later chapters.

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 $y \in \mathcal{M}$. There may not exist a global minimizer. Global minimizers may also not be unique. Defining this notion merely requires $\mathcal{M}$ to be a set and f to be a function: their smoothness is irrelevant.

Formally, if no minimizer exists, problem (4.1) should read "$\inf_{x \in \mathcal{M}} f(x)$".

While it is typically our goal to compute a global minimizer, in general, this goal is out of reach. A more realistic 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 $\mathcal{M}$. Just as for global minimizers, it does not rely on smoothness of either $\mathcal{M}$ or f.

The topology of embedded submanifolds is specified in Definition 3.15.

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 smooth 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 $\mathcal{U}$. Then, we write* $\lim_{k\to\infty} x_k = x$ *and we say the sequence* converges *to x.*

2. *A point $x \in \mathcal{M}$ is an* accumulation point *of S if it is a 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.*

If the topology on $\mathcal{M}$ is *Hausdorff*, which is the case for embedded submanifolds of linear spaces and is typically part of the general definition of a manifold (see Section 8.2), then a sequence on $\mathcal{M}$ has at most one limit.

This chapter focuses on Riemannian gradient descent. With just one additional geometry 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 point. Give an example of a sequence that has a single accumulation point yet no limit point. Give an example of a sequence that has two distinct accumulation points. Show that if a sequence converges to x, then all 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 $\mathcal{M}$. Show 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 these 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 extend this concept to manifolds.

Let $c: I \to \mathcal{M}$ 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:

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

Since $g = f \circ c$ is smooth by composition and it 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 $g'(0) = \langle \mathrm{grad} f(x), v \rangle_x$. Overall, we get this Taylor expansion:

See Section 5.9 for a second-order Taylor expansion.

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

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

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

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

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

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. 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.

It is called this way as it quite literally pulls the cost function back to the tangent spaces.

Exercise 4.3. *Given a smooth curve $c \colon [0,1] \to \mathcal{M}$ with $c(0) = x$ and $c(1) = y$, show 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 5.34 for the next order.)

4.2 First-order optimality conditions

In general, checking whether a point x on $\mathcal{M}$ is a local optimizer 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 optimizer. The following proposition states such a condition. It is called the *first-order necessary optimality condition*, because it involves first-order derivatives, namely, the gradient of f. At the very least, we should ensure accumulation points of sequences generated by optimization algorithms satisfy this condition.

Proposition 4.4. *Let $f \colon \mathcal{M} \to \mathbb{R}$ be a smooth function on a Riemannian manifold. If x is a local minimizer of f, then* $\mathrm{grad} f(x) = 0$.

The Riemannian structure is optional to define necessary optimality conditions: we could state Proposition 4.4 as saying $\mathrm{D}f(x) = 0$ instead.

Proof. For contradiction, assume $\mathrm{grad} f(x) \neq 0$. Since $-\mathrm{grad} f(x)$ is a tangent vector to $\mathcal{M}$ at x, there exists a smooth curve $c \colon I \to \mathcal{M}$ such that $c(0) = x$ and $c'(0) = -\mathrm{grad} f(x)$. Applying the Taylor expansion (4.2) to $f \circ c$ we find:

$$f(c(t)) = f(x) - t\|\mathrm{grad} f(x)\|_x^2 + O(t^2).$$

Combining with the assumption $\|\mathrm{grad} f(x)\|_x > 0$, we conclude that there exists $\varepsilon > 0$ such that $f(c(t)) < f(x)$ for all $t \in (0, \varepsilon)$, which contradicts local optimality of x.

More formally, since x is a local minimizer, there exists a neighborhood $\mathcal{U}$ of x in $\mathcal{M}$ such that $f(x) \leq f(y)$ for all $y \in \mathcal{U}$. Since c is continuous, $c^{-1}(\mathcal{U}) = \{t \in I : c(t) \in \mathcal{U}\}$ is open, and it contains 0 since $c(0) = x \in \mathcal{U}$. Hence, $c^{-1}(\mathcal{U}) \cap (0, \varepsilon)$ is non-empty: there exists $t \in (0, \varepsilon)$ (implying $f(c(t)) < f(x)$) such that $c(t)$ is in $\mathcal{U}$ (implying $f(c(t)) \geq f(x)$): a contradiction. □

This result justifies our interest in points where the gradient vanishes. We give them a name.

Definition 4.5. *Given a smooth function f on a Riemannian manifold $\mathcal{M}$, we call $x \in \mathcal{M}$ a* critical point *or a* stationary point *of f if* $\mathrm{grad} f(x) = 0$.

4.3 *Riemannian gradient descent*

The standard gradient descent algorithm in Euclidean space iterates

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

starting with some $x_0 \in \mathcal{E}$. 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 on $\mathcal{M}$, iterate

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

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

Algorithm 4.1: Riemannian gradient descent (RGD)

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 ways. Define

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

Line-search is about minimizing g approximately: sufficiently well 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 $t_{i+1} = \tau t_i$ 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.

A1. *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.

A2. *At each iteration, the algorithm achieves* sufficient decrease, *in that there exists a constant $c > 0$ such that, for all k,*

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

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.

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.

Proposition 4.6. *Let f be a smooth function satisfying A1 on a Riemannian manifold $\mathcal{M}$. Let $x_0, x_1, x_2, \ldots$ be iterates generated by Algorithm 4.1 satisfying A2 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:

$$\begin{aligned} f(x_0) - f_{\mathrm{low}} &\overset{A1}{\geq} f(x_0) - f(x_K) = \sum_{k=0}^{K-1} f(x_k) - f(x_{k+1}) \\ &\overset{A2}{\geq} Kc \min_{k=0,\ldots,K-1} \|\mathrm{grad} f(x_k)\|^2. \end{aligned}$$

To get a limit statement, observe that $f(x_{k+1}) \leq f(x_k)$ for all k by A2. Then,

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

is a bound on a series of nonnegative numbers. This implies the summands converge to zero:

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

which confirms the limit statement. 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. □

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 A2, we need to understand how $f(x_{k+1})$ compares to $f(x_k)$. Let s_k be the tangent vector such that $x_{k+1} = \mathrm{R}_{x_k}(s_k)$. Then, the Taylor expansion (4.4) 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 $f \circ \mathrm{R}$.

A3. *For a given subset S of the tangent bundle $\mathrm{T}\mathcal{M}$, there exists a constant $L > 0$ such that, for all $(x, s) \in S$,*

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

See Corollary 10.48 and Exercises 10.51 and 10.52.

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

Proposition 4.7. *Let f be a smooth function on a Riemannian manifold $\mathcal{M}$. For a retraction R, let $f \circ \mathrm{R}$ satisfy A3 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, A2 holds with

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

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

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

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

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

The coefficient, quadratic in α_k, is symmetric around its maximizer $1/L$, positive between its roots at 0 and $2/L$. By assumption on α_k,

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

which concludes the proof. $\square$

As a particular case, if the constant L is known beforehand, then we get an explicit algorithm and associated guarantee as a corollary of Propositions 4.6 and 4.7.

Corollary 4.8. *Let f be a smooth function satisfying A1 on a Riemannian manifold $\mathcal{M}$. For a retraction R, let $f \circ \mathrm{R}$ satisfy A3 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

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

In particular, for all $\varepsilon > 0$ there exists k in $0, \ldots, K-1$ such that $\|\mathrm{grad} f(x_k)\| \leq \varepsilon$ provided $K \geq 2L(f(x_0) - f_{\mathrm{low}})\frac{1}{\varepsilon^2}$.

How reasonable is A3? 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 A3 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, under that condition, by the fundamental theorem of calculus and Cauchy–Schwarz:

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

With $c(t) = x + ts$, we start with

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

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). 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 lead to A3, which we henceforth call a *Lipschitz-type* assumption.

Exercise 4.9. *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 A3 holds over the whole tangent bundle.*

4.5 Backtracking line-search

The simplest result in the previous section is Corollary 4.8, which assumes a constant step-size of $1/L$. In practice however, the constant L is seldom known. Even when it is known, it may be large due to a particular behavior of $f \circ \mathrm{R}$ in a limited part of the domain, thus seemingly forcing us to take small steps for the whole sequence of iterates, which evidently is not necessary: 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: $\tau, r \in (0,1)$

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 following lemma and corollary show that, under the regularity condition A3, backtracking line-search produces sufficient decrease, with a constant c which depends on various factors. Importantly, the regularity constant L affects the guarantee but need not be known.

Lemma 4.10. *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 A3 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 available).

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. 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:

Of course, it might also terminate earlier with a longer step-size: we consider the worst case.

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

Furthermore, 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. $\square$

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 4.11. *Let f be a smooth function satisfying A1 on a Riemannian manifold $\mathcal{M}$. For a retraction R, let $f \circ \mathrm{R}$ satisfy A3 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 4.10.

Proof. By Lemma 4.10, 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 4.6,

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

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

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

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

As a concluding 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, $-\bar{\alpha}_k \mathrm{grad} f(x_k)$ is invariant under positive scaling of f. For the first iteration, this can be done for example by setting

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

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)

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

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 theorem).

4.6 *Local convergence**

We discussed *global* convergence of RGD, establishing that, under some assumptions, all accumulation points of the sequence of iterates are critical points. We further discussed how fast the gradient norm converges to zero.

Importantly, global convergence does *not* mean convergence to a global optimizer; it merely means convergence regardless of initialization x_0.

These results, however, do not predict convergence to a single point, and the worst-case rate is slow (it is called *sublinear*), whereas in practice we typically do observe convergence to a single point, and the rate of convergence is, eventually, *linear*, as we define now.

Definition 4.12. *In a metric space equipped with a distance* dist, *a sequence* $a_0, a_1, a_2, \ldots$ converges at least linearly *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* $\lim_{k\to\infty} \frac{\epsilon_{k+1}}{\epsilon_k} = \mu$ *for some* $\mu \in (0,1)$. *Then,* $\log_{10}(1/\mu)$ *is the* rate of convergence. *If the above holds with $\mu = 0$, the convergence is* superlinear.

This is also called *R-linear convergence*, as opposed to the more restricted notion of *Q-linear convergence* which forces $\epsilon_k = \mathrm{dist}(a_k, a)$.

This definition applies to sequences of real numbers (for example, $f(x_k)$ or $\|\mathrm{grad} f(x_k)\|$) using the absolute value distance, as well as to sequences of iterates on $\mathcal{M}$ using a notion of distance on the manifold: see Section 10.1.

The observed behavior is elucidated in two important results called the capture theorem [AMS08, 4.4.2] and the local convergence theorem for line-search–based methods [AMS08, Thm. 4.5.6]. Informally, these state that, if $x_\star$ is a critical point and the Hessian at $x_\star$ (defined in Section 5.5) is positive definite, then there exists a neighborhood of $x_\star$ such that, if the sequence of iterates enters that neighborhood, then RGD (under some conditions on the cost function and on the step-sizes) converges to $x_\star$, and the sequence $f(x_0), f(x_1), f(x_2), \ldots$ converges at least linearly to $f(x_\star)$. Using this, one can show $x_0, x_1, x_2, \ldots$ also converges at least linearly (to $x_\star$), with half the convergence rate.

Furthermore, the rate of convergence is better if the condition number of the Hessian, that is, the ratio of its largest to smallest eigenvalue, is smaller.

4.7 Computing gradients

The gradient[1] of a function $f: \mathcal{M} \to \mathbb{R}$ on a Riemannian manifold is defined in full generality as the unique vector field $\mathrm{grad} f$ on $\mathcal{M}$ such that, for all $x \in \mathcal{M}$ and $v \in \mathrm{T}_x\mathcal{M}$,

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

This suggests a general strategy to obtain an explicit expression for $\mathrm{grad} f$:

1. Determine an expression for $\mathrm{D}f(x)[v]$, 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; then, to use the notion of adjoint[2] of a linear operator to isolate v.

In working out directional derivatives, three rules get most of the work done:

1. The chain rule: as for (3.28), let $F: \mathcal{M} \to \mathcal{M}'$ and $G: \mathcal{M}' \to \mathcal{M}''$ 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:

$$\mathrm{D}H(x)[v] = \mathrm{D}G(F(x))[\mathrm{D}F(x)[v]]. \tag{4.14}$$

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:

$$\mathrm{D}H(x)[v] = \mathrm{D}F(x)[v]G(x) + F(x)\mathrm{D}G(x)[v]. \tag{4.15}$$

This rule holds for any type of product. For example, with the entrywise product $H(x) = F(x) \odot G(x)$, we have

$$\mathrm{D}H(x)[v] = \mathrm{D}F(x)[v] \odot G(x) + F(x) \odot \mathrm{D}G(x)[v]. \tag{4.16}$$

Likewise, with the Kronecker product $H(x) = F(x) \otimes G(x)$,

$$\mathrm{D}H(x)[v] = \mathrm{D}F(x)[v] \otimes G(x) + F(x) \otimes \mathrm{D}G(x)[v]. \tag{4.17}$$

[1] See Definitions 3.48 and 8.52.

[2] See Definition 3.3.

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:

$$\mathrm{D}h(x)[v] = \langle \mathrm{D}F(x)[v], G(x) \rangle + \langle F(x), \mathrm{D}G(x)[v] \rangle . \tag{4.18}$$

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 5.4 and 5.26.

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 3.51 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, here we focus on the Euclidean case.

Let $f \colon \mathcal{E} \to \mathbb{R}$ be defined by $f(x) = \langle F(x), G(x) \rangle$, where F and G map $\mathcal{E}$ to $\mathcal{E}'$: two possibly distinct Euclidean spaces of real matrices (that is, matrices in $\mathcal{E}$ and $\mathcal{E}'$ may have different sizes). We equip both of them with the standard inner product $\langle U, V \rangle = \mathrm{Tr}(U^\top V)$ for real matrices (3.15), bearing in mind that, formally, these are different inner products since they are defined on different spaces. To stress this distinction, we write $\langle \cdot, \cdot \rangle_{\mathcal{E}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{E}'}$, so that $f(x) = \langle F(x), G(x) \rangle_{\mathcal{E}'}$. From (4.18), we know

E.g., $f(X) = \mathrm{Tr}(F(X)) = \langle F(X), I \rangle$ or $f(X) = \|F(X)\|^2 = \langle F(X), F(X) \rangle$.
This discussion extends directly to complex matrices with inner product (3.18).

$$\mathrm{D}f(x)[v] = \langle \mathrm{D}F(x)[v], G(x) \rangle_{\mathcal{E}'} + \langle F(x), \mathrm{D}G(x)[v] \rangle_{\mathcal{E}'} .$$

Letting $\mathrm{D}F(x)^*$ denote the adjoint of $\mathrm{D}F(x)$ with respect to the inner products on $\mathcal{E}$ and $\mathcal{E}'$, and similarly for $\mathrm{D}G(x)^*$, we get

$$\mathrm{D}f(x)[v] = \langle v, \mathrm{D}F(x)^*[G(x)] \rangle_{\mathcal{E}} + \langle \mathrm{D}G(x)^*[F(x)], v \rangle_{\mathcal{E}} .$$

Using symmetry and linearity of the inner product and the definition of gradient of f, it follows that

$$\mathrm{grad} f(x) = \mathrm{D}F(x)^*[G(x)] + \mathrm{D}G(x)^*[F(x)]. \tag{4.19}$$

This highlights the importance of computing adjoints of linear operators in obtaining gradients. Formulas (3.16) and (3.19) are particularly helpful in this respect. We illustrate this in examples below.

Example 4.13. *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*

$$\begin{aligned}\mathrm{D}F(X)[U] &= UX^{k-1} + XUX^{k-2} + X^2UX^{k-3} + \cdots + X^{k-2}UX + X^{k-1}U \\ &= \sum_{\ell=1}^{k} X^{\ell-1}UX^{k-\ell}. \end{aligned} \tag{4.20}$$

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* (3.16),

$$\begin{aligned}\langle \mathrm{D}F(X)[U], V\rangle &= \sum_{\ell=1}^{k} \left\langle X^{\ell-1}UX^{k-\ell}, V\right\rangle \\ &= \sum_{\ell=1}^{k} \left\langle U, (X^\top)^{\ell-1}V(X^\top)^{k-\ell}\right\rangle \\ &= \left\langle U, \mathrm{D}F(X^\top)[V]\right\rangle. \end{aligned} \tag{4.21}$$

Similarly, for $F(X) = X^k$ *defined on* $\mathbb{C}^{n\times n}$ *equipped with the usual inner product* (3.18), *the expression for* $\mathrm{D}F(X)$ *is unchanged, and* $\mathrm{D}F(X)^* = \mathrm{D}F(X^*)$.

Example 4.14. *Consider* $F(X) = X^{-1}$ *defined over square matrices (real or complex). This function is characterized by*

$$F(X)X = I$$

whenever X *is invertible. Differentiating at* X *along* U *on both sides, the product rule yields*

$$\mathrm{D}F(X)[U]X + F(X)U = 0.$$

Hence,

$$\mathrm{D}F(X)[U] = -X^{-1}UX^{-1}. \tag{4.22}$$

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 4.15. *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*

$$\mathrm{D}\tilde{g}(X)[U] = \tilde{g}'(X) \odot U. \tag{4.23}$$

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* (3.16), *we have*

$$\langle \mathrm{D}\tilde{g}(X)[U], V\rangle = \left\langle \tilde{g}'(X) \odot U, V\right\rangle = \left\langle U, \tilde{g}'(X) \odot V\right\rangle = \langle U, \mathrm{D}\tilde{g}(X)[V]\rangle.$$

(In general, there does not 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 4.16. *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*

$$G(X) = \sum_{k=0}^{\infty} a_k X^k. \tag{4.24}$$

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`.

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. 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.

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,

$$G\left(\begin{bmatrix} X & U \\ 0 & X \end{bmatrix}\right) = \begin{bmatrix} G(X) & \mathrm{D}G(X)[U] \\ 0 & G(X) \end{bmatrix}. \tag{4.25}$$

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* $\mathrm{DG}(X)[U]$ *involving* Loewner matrices *[Hig08, Cor. 3.12].*

We can gain insight into the adjoint of the directional derivative of a matrix function through (4.24) *and Example 4.13. Indeed,*

$$\begin{aligned}\mathrm{DG}(X)[U] &= \sum_{k=0}^{\infty} a_k \mathrm{D}\Big(X \mapsto X^k\Big)(X)[U] \\ &= \sum_{k=0}^{\infty} a_k \sum_{\ell=1}^{k} X^{\ell-1} U X^{k-\ell}.\end{aligned} \tag{4.26}$$

From this identity, assuming the Taylor expansion coefficients a_k are real, it is straightforward to see that the adjoint with respect to the usual inner product obeys $\mathrm{DG}(X)^* = \mathrm{DG}(X^*)$. *Indeed,*

This holds for matrix exponential, logarithm and square root.

$$\begin{aligned}\langle \mathrm{DG}(X)[U], V\rangle &= \left\langle \sum_{k=0}^{\infty} a_k \sum_{\ell=1}^{k} X^{\ell-1} U X^{k-\ell}, V \right\rangle \\ &= \sum_{k=0}^{\infty} a_k \sum_{\ell=1}^{k} \left\langle X^{\ell-1} U X^{k-\ell}, V \right\rangle \\ &= \sum_{k=0}^{\infty} a_k \sum_{\ell=1}^{k} \left\langle U, (X^*)^{\ell-1} V (X^*)^{k-\ell} \right\rangle = \langle U, \mathrm{DG}(X^*)[V]\rangle .\end{aligned} \tag{4.27}$$

Of course, $X^* = X^\top$ in the real case.

Example 4.17. *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] and [AMS08, Ex. 8.1.5] among others.*

Example 4.18. *The directional derivative of* $g(X) = \log(\det(X))$ *is given by* $\mathrm{D}g(X)[U] = \mathrm{Tr}(X^{-1}U)$*, provided that if* $\det(X)$ *is real, then it is positive. Indeed, using* $\det(AB) = \det(A)\det(B)$ *and* $\log(ab) = \log(a) + \log(b)$,

$$\begin{aligned}\log(\det(X + tU)) &= \log(\det(X(I_n + tX^{-1}U))) \\ &= \log(\det(X)) + \log(\det(I_n + tX^{-1}U)) \\ &= \log(\det(X)) + t\,\mathrm{Tr}(X^{-1}U) + O(t^2),\end{aligned} \tag{4.28}$$

where we also used $\det(I_n + tA) = 1 + t\,\mathrm{Tr}(A) + O(t^2)$. *This can be seen in various ways; one of them is that if $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of A, then* $\log(\det(I_n + tA)) = \sum_{i=1}^{n} \log(1 + \lambda_i t)$, *then using* $\log(1 + \lambda t) =$

$\lambda t + O(t^2)$. *In particular, if we restrict* X *to be positive definite 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 directional derivative $\mathrm{D}h(X)[U] = -\operatorname{Tr}(X^{-1}U)$.

Example 4.19. *We now work out a gradient as a full example. Consider this silly 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 3.57 for pointers regarding gradients on a product manifold. Define* $Q(X, Y) = A \odot \left[\exp(X^{-1}B)Y\right]$*, so that*

$$f(X, Y) = \frac{1}{2}\|Q(X, Y)\|^2 = \frac{1}{2}\langle Q(X, Y), Q(X, Y)\rangle.$$

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:*

$$\mathrm{D}Q(X, Y)[\dot{X}, \dot{Y}] = A \odot \left[\mathrm{D}\exp(X^{-1}B)[U]Y + \exp(X^{-1}B)\dot{Y}\right],$$

where $\mathrm{D}\exp$ *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*

$$\mathrm{D}f(X, Y)[\dot{X}, \dot{Y}] = \left\langle A \odot \left[\mathrm{D}\exp(W)[-X^{-1}\dot{X}W]Y + \exp(W)\dot{Y}\right], Q(X, Y)\right\rangle.$$

We must now 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 operators: recall Section 3.1. First using the adjoint of entrywise multiplication with respect to the usual inner product as in (3.16), then linearity of the inner product:*

$$\begin{aligned}\mathrm{D}f(X, Y)[\dot{X}, \dot{Y}] = &\left\langle \mathrm{D}\exp(W)[-X^{-1}\dot{X}W]Y, A \odot Q(X, Y)\right\rangle \\ &+ \left\langle \exp(W)\dot{Y}, A \odot Q(X, Y)\right\rangle.\end{aligned}$$

Let $Z = A \odot Q(X, Y)$ *for short; using the adjoint of matrix multiplication for both terms:*

$$\mathrm{D}f(X, Y)[\dot{X}, \dot{Y}] = \left\langle \mathrm{D}\exp(W)[-X^{-1}\dot{X}W], ZY^\top\right\rangle + \left\langle \dot{Y}, \exp(W)^\top Z\right\rangle.$$

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 usual real inner product over complex matrices,* $\mathrm{grad} f$ *takes on the same expression except all transposes are replaced by conjugate-transposes, and* $Z = \overline{A} \odot Q(X,Y)$.

Something to keep in mind. 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].

These principles are at the heart of *automatic differentiation* (AD): algorithms that can automatically compute the derivatives of a function, from code to compute that function. AD can significantly speed up development time. Packages such as PyManopt offer AD for optimization on manifolds [TKW16].

Exercise 4.20. *Prove rules* (4.14), (4.15), (4.16), (4.17) *and* (4.18).

Exercise 4.21. *The (principal) matrix square root function $F(X) = X^{1/2}$ is well defined provided any real eigenvalue of X is 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 4.22. *For a matrix function whose Taylor series has real coefficients, show* $\mathrm{D}F(X)[U^*] = (\mathrm{D}F(X^*)[U])^*$*. Combining with* (4.27)*, this yields another expression for the adjoint of the differential of a matrix function:* $\mathrm{D}F(X)^*[U] = \mathrm{D}F(X^*)[U] = (\mathrm{D}F(X)[U^*])^*$.

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. 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):

In Manopt, these tests are implemented as `checkgradient`.
See Section 6.7 to check Hessians.

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

This says that, for all $x \in \mathcal{M}$ and $v \in \mathrm{T}_x\mathcal{M}$, with all retractions,

$$E(t) \triangleq \left| f(\mathrm{R}_x(tv)) - f(x) - t \langle \mathrm{grad} f(x), v \rangle_x \right| = O(t^2). \tag{4.29}$$

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[3]) when t is small:

[3] If the Taylor remainder happens to be $O(t^k)$ with $k > 2$, we would get a slope of k. This is rare.

$$\log(E(t)) \approx 2\log(t) + \text{constant}.$$

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 6.4.5). 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 4.23. *With some symmetric matrix A and size n, recall the cost function $f(X) = -\frac{1}{2}\mathrm{Tr}(X^\top A X)$ 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 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) -.5*inner(X, A*X);

problem.grad = @(X) St.proj(X, A*X); % Oops: forgot minus sign
checkgradient(problem); % Top panel

problem.grad = @(X) St.proj(X, -A*X); % This is better
checkgradient(problem); % Bottom panel
```

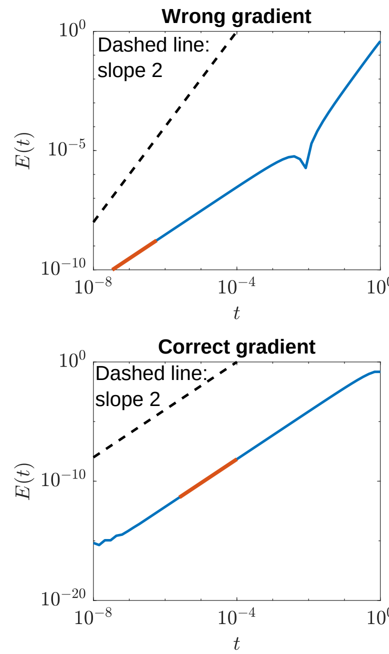


Figure 4.1: Example 4.23 illustrates a numerical procedure to check one's gradient computation code. The dashed line has a slope of two: this serves as a visual reference. The blue curve represents the function $E(t)$ (4.29), in a log–log plot. Part of the blue curve is overlaid with red. The (average) slopes of those red parts are nearly 1 (top) and 2 (bottom), strongly suggesting the top gradient is incorrect, and suggesting the bottom gradient is correct (as is indeed the case).

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 (but now difficult to find) 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 around 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, ZRS16a, ZRS16b, KSM18, SKM19].

There is also recent work focused on nonsmooth cost functions on smooth manifolds, including proximal point methods and subgradient methods [BFM17, CMSZ18], 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.

First-order necessary optimality conditions are described without reference to a Riemannian structure in [BH18], for a broader class of optimization problems where 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], with reference to the Riemannian structure.

After Proposition 4.6, we observed that (under the stated conditions) the statement guarantees all limit points of RGD are critical points, but it does not guarantee convergence of the iterates: there could be more than one limit 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, Polyak)–Łojasiewicz inequality for real analytic functions [Łoj65].

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 problems and to design algorithms. In particular, we found that local minimizers are critical points, that is, their gradient vanishes. Furthermore, we showed that following the negative gradient allows us to compute critical points, under a regularity condition.

A tool of choice in these developments has been a type of first-order Taylor expansion of the cost function along a curve (typically obtained through retraction.) Concretely, with $c\colon I \to \mathcal{M}$ a smooth curve 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

$$g(t) = g(0) + tg'(0) + O(t^2) = f(x) + t\,\langle \mathrm{grad} f(x), v \rangle_x + O(t^2).$$

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

$$g(t) = f(x) + t\,\langle \mathrm{grad} f(x), v \rangle_x + \frac{t^2}{2} g''(0) + O(t^3).$$

In the same way we expressed $g'(0)$ in terms of f through its gradient, our target now is to express $g''(0)$ in terms of properties of f.

Evidently, $g''(0)$ involves second-order derivatives of f, which leads us to ponder: can we define an equivalent of the *Hessian* of a function on a manifold? Let us first review the case of a function defined on $\mathbb{R}^n$ equipped with the canonical inner product $\langle u, v \rangle = u^\top v$.

In a first encounter, the gradient and Hessian of $f\colon \mathbb{R}^n \to \mathbb{R}$ are often defined in terms of partial derivatives as follows:

$$\mathrm{grad} f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}, \quad \mathrm{Hess} f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{bmatrix}.$$

This casts the Hessian as a symmetric matrix of size n. It is more revealing to think of that matrix as a linear operator instead. Specifically, consider the directional derivative of the gradient vector field $\mathrm{grad} f(x)$ along $v \in \mathbb{R}^n$ (this tells us how much the gradient changes if x is perturbed along v, up to first order):

$$\begin{aligned}
\mathrm{D}(\mathrm{grad} f)(x)[v] &= \begin{bmatrix} \mathrm{D}\left(\frac{\partial f}{\partial x_1}\right)(x)[v] \\ \vdots \\ \mathrm{D}\left(\frac{\partial f}{\partial x_n}\right)(x)[v] \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) v_1 + \cdots + \frac{\partial^2 f}{\partial x_n \partial x_1}(x) v_n \\ \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) v_1 + \cdots + \frac{\partial^2 f}{\partial x_n \partial x_n}(x) v_n \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \mathrm{Hess} f(x)[v].
\end{aligned}$$

In other words, the Hessian of f at x is the linear operator which, applied to v, yields the directional derivative of the gradient vector field at x along v. We adopt this perspective as the definition of the Hessian in Euclidean spaces.

Thus, to extend the concept of Hessian to manifolds, we need to identify a good notion of derivative of vector fields on Riemannian manifolds: we call it a *connection* or *covariant derivative*. It 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:

$$\begin{aligned}
f(c(t)) = f(x) + t \langle \mathrm{grad} f(x), v \rangle_x &+ \frac{t^2}{2} \langle \mathrm{Hess} f(x)[v], v \rangle_x \\
&+ \frac{t^2}{2} \left\langle \mathrm{grad} f(x), c''(0) \right\rangle_x + O(t^3),
\end{aligned}$$

where $\mathrm{Hess} f(x)$ is the *Riemannian Hessian* of f 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 essential properties. For example, as a linear operator from $\mathrm{T}_x\mathcal{M}$ to $\mathrm{T}_x\mathcal{M}$, it is self-adjoint with respect to the Riemannian metric $\langle \cdot, \cdot \rangle_x$.

We already mentioned the role of 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*.

For example, one famous optimization algorithm in $\mathbb{R}^n$ is *Newton's method*. For a cost function $f\colon \mathbb{R}^n \to \mathbb{R}$ and a current iterate $x_k \in \mathbb{R}^n$, this method approximates the cost function locally around x_k with a quadratic model in v:

$$f(x_k + v) \approx m(v) = f(x_k) + \langle \mathrm{grad} f(x_k), v\rangle + \frac{1}{2}\langle \mathrm{Hess} f(x_k)[v], v\rangle .$$

Assuming the Hessian of f at x_k is positive definite, m has a unique minimizer, located at its critical point. The gradient of the quadratic model with respect to v is $\mathrm{grad}\, m(v) = \mathrm{grad} f(x_k) + \mathrm{Hess} f(x_k)[v]$, hence the minimizer is attained at the solution of the linear system

$$\mathrm{Hess} f(x_k)[v] = -\mathrm{grad} f(x_k).$$

With this v, the point $x_{k+1} = x_k + v$ minimizes m. This reasoning leads to Newton's method: given an initial guess $x_0 \in \mathbb{R}^n$, iterate

$$x_{k+1} = x_k - (\mathrm{Hess} f(x_k))^{-1}[\mathrm{grad} f(x_k)].$$

Once we have a proper notion of Hessian for functions on a manifold, it is straightforward to generalize this algorithm. For all its qualities, Newton's method is not without critical flaws and limitations: we address these too.

5.1 *The case for a new derivative of vector fields*

Remember the function $f(x) = \frac{1}{2}x^\top Ax$ defined on the sphere S^{n-1} as a Riemannian submanifold of $\mathbb{R}^n$ with the canonical Euclidean metric (A is some symmetric matrix). Its gradient is the following smooth vector field on S^{n-1}:

See Example 3.52.

$$V(x) = \mathrm{grad} f(x) = Ax - (x^\top Ax)x.$$

Since V is a smooth map from $\mathcal{M}$ to $\mathrm{T}\mathcal{M}$ (two manifolds), we already have a notion of differential for it, as put forward by Definition 3.27. Explicitly, with the smooth extension

There is some notational abuse here in that points of $\mathrm{T}\mathcal{M}$ are pairs (x, v): as is common, we omit the first component.

$$\bar{V}(x) = Ax - (x^\top Ax)x$$

defined on all of $\mathbb{R}^n$, apply (3.27) with a tangent vector $u \in \mathrm{T}_x\mathcal{M}$:

$$\begin{aligned} \mathrm{D}V(x)[u] = \mathrm{D}\bar{V}(x)[u] &= Au - (x^\top Ax)u - (u^\top Ax + x^\top Au)x \\ &= \mathrm{Proj}_x(Au) - (x^\top Ax)u - (u^\top Ax)x, \qquad (5.1) \end{aligned}$$

where $\mathrm{Proj}_x(v) = v - (x^\top v)x$ is the orthogonal projector from $\mathbb{R}^n$ to $\mathrm{T}_x\mathcal{M}$. Evidently, $\mathrm{D}V(x)[u]$ is *not* (in general) tangent to $\mathcal{M}$ at x: the first two terms in (5.1) are tangent, but the third one is not if it is

Geodesics are curves with 0-acceleration.

Notion of second derivative hence leads to notion of connections-

nonzero. Thus, if we were to use this notion of differential to define Hessians as derivatives of gradient vector fields, 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 operator to and from $\mathrm{T}_x\mathcal{M}$, 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 notion of derivative for vector fields on Riemannian 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, classical differentiation of vector fields on Euclidean spaces should qualify: let us have a look at some elementary properties of the latter 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*, traditionally denoted by ∇. We do not need a Riemannian metric yet.

Read "nabla" for ∇.

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 these properties for arbitrary $U, V, W \in \mathfrak{X}(\mathcal{M})$, $u, w \in \mathrm{T}_x\mathcal{M}$ and $a, b \in \mathbb{R}$:

This is also called an *affine connection*. Formally, we should write $\nabla_{(x,u)}V$, but the base point x is typically clear from context. The standard definition introduces $(U, V) \mapsto \nabla_U V$ as a map ∇ from $\mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M})$ to $\mathfrak{X}(\mathcal{M})$. Definition 5.1 is equivalent (see Section 5.6) and simplifies a number of discussions.

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* ∇*.*

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, for example as follows:

$$\nabla_u V = \mathrm{Proj}_x(\mathrm{D}\bar{V}(x)[u]), \tag{5.4}$$

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. This 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,

$$\nabla_u V = \mathrm{Proj}_x\left(\bar{\nabla}_u \bar{V}\right). \tag{5.5}$$

If $\mathcal{M}$ is an open submanifold of $\mathcal{E}$, the claim is clear since Proj_x is identity. 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 repeatedly that $\bar{\nabla}$ is a connection and that Proj_x is linear, it is straightforward to verify linearity in the first argument:

$$\begin{aligned}
\nabla_{au+bw} V &= \mathrm{Proj}_x\left(\bar{\nabla}_{au+bw} \bar{V}\right) \\
&= \mathrm{Proj}_x\left(a\bar{\nabla}_u \bar{V} + b\bar{\nabla}_w \bar{V}\right) \\
&= a\nabla_u V + b\nabla_w V,
\end{aligned}$$

then linearity in the second argument:

$$\begin{aligned}
\nabla_u(aV + bW) &= \mathrm{Proj}_x\left(\bar{\nabla}_u(a\bar{V} + b\bar{W})\right) \\
&= \mathrm{Proj}_x\left(a\bar{\nabla}_u \bar{V} + b\bar{\nabla}_u \bar{W}\right) \\
&= a\nabla_u V + b\nabla_u W,
\end{aligned}$$

and the Leibniz rule (with $f \in \mathfrak{F}(\mathcal{M})$ and smooth extension $\bar{f} \in \mathfrak{F}(O)$):

$$\begin{aligned}\nabla_u(fV) &= \mathrm{Proj}_x\big(\bar{\nabla}_u(\bar{f}\bar{V})\big)\\ &= \mathrm{Proj}_x\big(\mathrm{D}\bar{f}(x)[u]\cdot\bar{V}(x) + \bar{f}(x)\bar{\nabla}_u\bar{V}\big)\\ &= \mathrm{D}f(x)[u]\cdot V(x) + f(x)\nabla_u V.\end{aligned}$$

We use that $\bar{f}\bar{V}$ is a smooth extension for fV on O.

Finally, we can see that $\nabla_U V$ is smooth as per Exercise 3.56. ☐

Not only do connections exist, but actually: there exist infinitely many of them on any manifold $\mathcal{M}$. For example, we could 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 in relation to the Riemannian structure on $\mathcal{M}$. As it turns out, connection (5.4) satisfies these additional properties if $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$.

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 operator on each tangent space.

In order to state the desired properties, we need to introduce a few notational definitions.

Definition 5.3. *For $U, V \in \mathfrak{X}(\mathcal{M})$ and $f \in \mathfrak{F}(\mathcal{U})$ with $\mathcal{U}$ open in $\mathcal{M}$, let:*

(a) $Uf \in \mathfrak{F}(\mathcal{U})$ *such that* $(Uf)(x) = \mathrm{D}f(x)[U(x)]$;

Mind the difference between Uf and fU.

(b) $[U, V]\colon \mathfrak{F}(\mathcal{U}) \to \mathfrak{F}(\mathcal{U})$ *such that* $[U, V]f = U(Vf) - V(Uf)$; *and*

Even in linear spaces $[U, V]f$ is nonzero in general. For example, in $\mathbb{R}^2$ consider $U(x) = \begin{bmatrix}1, 0\end{bmatrix}^\top$, $V(x) = \begin{bmatrix}0, x_1x_2\end{bmatrix}^\top$ and $f(x) = x_2$. Then, $[U, V]f = f$.

(c) $\langle U, V\rangle \in \mathfrak{F}(\mathcal{M})$ *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* of such action, $[U, V]$, is called the *Lie bracket*. The last definition illustrates how the Riemannian metric allows us to take the product of two vector fields to form a scalar field. Notice that

$$Uf = \langle \mathrm{grad} f, U\rangle, \tag{5.6}$$

owing to definitions (a) and (c).

Theorem 5.4. *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*. The other property is a type of product rule for differentiation of inner products.

Before we prove this theorem, 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).

Proposition 5.5. *The Riemannian connection on a Euclidean space $\mathcal{E}$ with any Euclidean metric $\langle \cdot, \cdot \rangle$ is $\nabla_u V = \mathrm{D}V(x)[u]$. It is called the* canonical Euclidean connection.

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: $\langle \mathrm{Hess} f[U], V\rangle = \langle \mathrm{Hess} f[V], U\rangle$. Hence,

This is the Clairaut–Schwarz theorem, which you may remember as the fact that the Hessian matrix is symmetric.

$$\begin{aligned}[U, V]f &= U(Vf) - V(Uf) \\ &= \langle \mathrm{grad} f, \nabla_U V - \nabla_V U\rangle \\ &= (\nabla_U V - \nabla_V U)f,\end{aligned}$$

concluding the proof. □

For embedded submanifolds of Euclidean spaces, the connection we defined by projection to tangent spaces is always symmetric. To show this, it is convenient to introduce this notation in analogy with (5.5):

$$\nabla_U V = \mathrm{Proj}\big(\bar{\nabla}_{\bar{U}} \bar{V}\big), \tag{5.7}$$

where $\bar{\nabla}$ is the canonical Euclidean connection on $\mathcal{E}$, 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, that is, $(\nabla_U V)(x) = \mathrm{Proj}_x((\bar{\nabla}_{\bar{U}} \bar{V})(x)) = \mathrm{Proj}_x(\bar{\nabla}_{\bar{U}(x)} \bar{V})$.

Theorem 5.6. *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:

$$\begin{aligned}[U, V]f &= U(Vf) - V(Uf) \\ &= U\big((\bar{V}\bar{f})|_{\mathcal{M}}\big) - V\left((\bar{U}\bar{f})|_{\mathcal{M}}\right) \\ &= \big(\bar{U}(\bar{V}\bar{f})\big)|_{\mathcal{M}} - \big(\bar{V}(\bar{U}\bar{f})\big)|_{\mathcal{M}} \\ &= \big([\bar{U}, \bar{V}]\bar{f}\big)|_{\mathcal{M}} \\ &= \big((\bar{\nabla}_{\bar{U}} \bar{V} - \bar{\nabla}_{\bar{V}} \bar{U})\bar{f}\big)|_{\mathcal{M}} \\ &= (\bar{W}\bar{f})|_{\mathcal{M}},\end{aligned} \tag{5.8}$$

where we defined $\bar{W} = \bar{\nabla}_{\bar{U}} \bar{V} - \bar{\nabla}_{\bar{V}} \bar{U} \in \mathfrak{X}(O)$. Assume for now that $\bar{W}$ is a smooth extension of a vector field W on $\mathcal{M}$. If that is the case, then,

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 on $\mathcal{M}$. Yet, we are about to show that their difference is.

$$W = \bar{W}|_{\mathcal{M}} = \mathrm{Proj}(\bar{W}) = \mathrm{Proj}\big(\bar{\nabla}_{\bar{U}} \bar{V} - \bar{\nabla}_{\bar{V}} \bar{U}\big) = \nabla_U V - \nabla_V U.$$

Furthermore, $(\bar{W}\bar{f})|_{\mathcal{M}} = Wf$, so that continuing from (5.8):

$$[U, V]f = (\bar{W}\bar{f})|_{\mathcal{M}} = Wf = (\nabla_U V - \nabla_V U)f,$$

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 (5.8) to h, we find:

$$0 = [U, V]h = (\bar{W}\bar{h})|_{\mathcal{M}\cap O'}.$$

Evaluate this at x:

$$0 = (\bar{W}\bar{h})(x) = \mathrm{D}\bar{h}(x)[\bar{W}(x)]\,.$$

In other 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. $\square$

If furthermore $\mathcal{M}$ inherits the metric from its embedding Euclidean space, then ∇ as defined above is the Riemannian connection.

Theorem 5.7. *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 5.6, it remains to check compatibility with the metric, that is, property 5 in Theorem 5.4. 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:

$$U\langle V, W\rangle = (\bar{U}\langle \bar{V}, \bar{W}\rangle)\,|_{\mathcal{M}} = \Big(\langle \bar{\nabla}_{\bar{U}}\bar{V}, \bar{W}\rangle + \langle \bar{V}, \bar{\nabla}_{\bar{U}}\bar{W}\rangle\Big)\Big|_{\mathcal{M}}. \tag{5.9}$$

Pick $x \in \mathcal{M}$. Then, $\bar{W}(x) = W(x) = \mathrm{Proj}_x(W(x))$. Furthermore, recall that Proj_x is self-adjoint (Proposition 3.53). Consequently,

For any two vectors $u, v \in \mathcal{E}$,

$$\langle u, \mathrm{Proj}_x(v)\rangle = \langle \mathrm{Proj}_x(u), v\rangle\,.$$

$$\begin{aligned}\langle \bar{\nabla}_{\bar{U}}\bar{V}, \bar{W}\rangle(x) &= \big\langle (\bar{\nabla}_{\bar{U}}\bar{V})(x), \mathrm{Proj}_x(W(x))\big\rangle \\ &= \big\langle \mathrm{Proj}_x\big((\bar{\nabla}_{\bar{U}}\bar{V})(x)\big), W(x)\big\rangle_x \\ &= \langle \nabla_U V, W\rangle(x).\end{aligned}$$

Combining twice with (5.9), 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. $\square$

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 still defines a symmetric connection, but it may not be the Riemannian connection.

We return to Theorem 5.4 and provide the missing proof using the following technical observation: a Lie bracket "is" a smooth vector field.

Proposition 5.8. *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 may identify $[U, V]$ with that smooth vector field.*

Proof. Say $\mathcal{M}$ is embedded in the Euclidean space $\mathcal{E}$. In Theorem 5.6 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 $\mathcal{E}$.) $\square$

Importantly, $[U, V]$ is defined independently of any connection. The proof shows that for any symmetric connection ∇ it so happens that $[U, V]$ is equivalent to $\nabla_U V - \nabla_V U$, independently of the choice of symmetric connection.

Proof of Theorem 5.4. We first prove uniqueness. Assume ∇ is a connection satisfying the two listed properties. As per Proposition 5.8, we can treat Lie brackets as vector fields. With this in mind, consider the so-called *Koszul formula*:

$$\begin{aligned} 2\langle \nabla_U V, W \rangle = U\langle V, W\rangle + V\langle W, U\rangle - W\langle U, V\rangle \\ - \langle U, [V, W]\rangle + \langle V, [W, U]\rangle + \langle W, [U, V]\rangle. \end{aligned} \tag{5.10}$$

The right-hand side is independent of the connection ∇. This being said, because ∇ is symmetric, we know that $[U, V] = \nabla_U V - \nabla_V U$. Using this and compatibility with the metric, it is straightforward to check that (5.10) holds. Moreover,[1] if two vector fields $U_1, U_2 \in \mathfrak{X}(\mathcal{M})$ satisfy $\langle U_1, W\rangle = \langle U_2, W\rangle$ for all $W \in \mathfrak{X}(\mathcal{M})$, then $U_1 = U_2$. This implies that $\nabla_U V$ is uniquely defined by the Koszul formula, that is, the Riemannian connection is unique.

[1] With $d = \dim \mathcal{M}$ and any $x \in \mathcal{M}$, consider vector fields $W_1, \ldots, W_d \in \mathfrak{X}(\mathcal{M})$ such that $W_1(x), \ldots, W_d(x)$ form a basis for $\mathrm{T}_x\mathcal{M}$. These imply $U_1(x) = U_2(x)$. Repeat this argument for each x.

The Koszul formula also shows how to compute the Riemannian connection, thus showing existence. That $\nabla_U V$ as defined by (5.10) is smooth can be reasoned using the musical isomorphism mentioned in

Section 3.10. The other basic properties of connections as well as the two special properties of Riemannian connections can all be checked from (5.10). $\square$

Exercise 5.9. *A* derivation *on $\mathcal{M}$ is a map $\mathcal{D}\colon \mathfrak{F}(\mathcal{M}) \to \mathfrak{F}(\mathcal{M})$ such that, for all $V \in \mathfrak{X}(\mathcal{M}), f, g \in \mathfrak{F}(\mathcal{M})$ 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 in Definition 5.3) is a derivation.

Exercise 5.10. *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 5.8.*

5.5 Riemannian Hessians

The Riemannian Hessian is defined as the covariant derivative of the gradient vector field with respect to the Riemannian connection ∇ (Theorem 5.4). At any point x on the manifold, the Hessian defines a linear operator from the tangent space $\mathrm{T}_x\mathcal{M}$ into itself.

Definition 5.11. *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 a linear operator* $\mathrm{Hess} f(x)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ *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* $\mathrm{Hess} f[U] = \nabla_U \mathrm{grad} f$.

In practical terms, for a Riemannian submanifold of a Euclidean space $\mathcal{E}$, this means that to compute the Riemannian Hessian we can consider a smooth extension of the Riemannian gradient vector field to $\mathcal{E}$, differentiate it in the classical sense, then orthogonally project the result to the tangent space.

Corollary 5.12. *Let $\mathcal{M}$ be a Riemannian submanifold of a Euclidean space. Considering a smooth function $f\colon \mathcal{M} \to \mathbb{R}$, let $\bar{G}$ be a smooth extension of* $\mathrm{grad} f$. *Then, by Definition 5.11 and Theorem 5.7,*

$$\mathrm{Hess} f(x)[u] = \mathrm{Proj}_x(\mathrm{D}\bar{G}(x)[u]).$$

$\bar{G}$ 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 $x \in \mathcal{M}$. See also Corollary 5.15 below.

Chapter 7 derives consequences of this corollary for various manifolds. In practice, the analytic expression for $\mathrm{grad} f$ typically provides a suitable smooth extension $\bar{G}$ automatically. Section 4.7 provides further details on how to compute the directional derivatives of $\bar{G}$. We illustrate the overall procedure in the following example.

Example 5.13. *Consider the cost function $\bar{f}(x) = \frac{1}{2}x^\top Ax$ for some symmetric matrix $A \in \mathbb{R}^{n\times n}$ and its restriction $f = \bar{f}|_{\mathrm{S}^{n-1}}$ to the sphere. We already determined the Euclidean and Riemannian gradients of $\bar{f}$ and f, respectively:*

$$\begin{aligned}
&\mathrm{grad}\bar{f}(x) = Ax, \\
&\mathrm{grad}f(x) = \mathrm{Proj}_x(\mathrm{grad}\bar{f}(x)) = (I_n - xx^\top)Ax = Ax - (x^\top Ax)x.
\end{aligned}$$

To obtain the Riemannian Hessian of f, we further differentiate a smooth extension of $\mathrm{grad}f$ *in $\mathbb{R}^n$, and project the result. A valid smooth extension of* $\mathrm{grad}f$ *is the smooth vector field in $\mathbb{R}^n$ given by:*

$$\bar{G}(x) = Ax - (x^\top Ax)x.$$

Applying the usual product rule (see also Section 4.7), its differential is:

$$\mathrm{D}\bar{G}(x)[u] = Au - (u^\top Ax + x^\top Au)x - (x^\top Ax)u.$$

Project to the tangent space at x to reveal the Hessian:

$$\begin{aligned}
\mathrm{Hess}f(x)[u] = \mathrm{Proj}_x(\mathrm{D}\bar{G}(x)[u]) &= \mathrm{Proj}_x(Au) - (x^\top Ax)u \\
&= Au - (x^\top Au)x - (x^\top Ax)u.
\end{aligned}$$

This operator is formally defined only on $\mathrm{T}_x\mathrm{S}^{n-1}$ (not on all of $\mathbb{R}^n$).

The next proposition illustrates how both special properties of the Riemannian connection together lead to symmetry of the Hessian. By the spectral theorem (Theorem 3.4), this implies $\mathrm{Hess}f(x)$ admits $\dim\mathcal{M}$ real eigenvalues as well as a basis of eigenvectors in $\mathrm{T}_x\mathcal{M}$, orthonormal with respect to $\langle\cdot,\cdot\rangle_x$.

Proposition 5.14. *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 and using compatibility of the Riemannian connection with the Riemannian metric,

$$\begin{aligned}
\langle \mathrm{Hess}f[U], V\rangle &= \langle \nabla_U \mathrm{grad}f, V\rangle \\
&= U\langle \mathrm{grad}f, V\rangle - \langle \mathrm{grad}f, \nabla_U V\rangle \\
&= U(Vf) - (\nabla_U V)f.
\end{aligned}$$

Similarly,

$$\langle U, \mathrm{Hess}f[V]\rangle = V(Uf) - (\nabla_V U)f.$$

Thus, recalling the definition of Lie bracket,

$$\begin{aligned}\langle \mathrm{Hess} f[U], V\rangle - \langle U, \mathrm{Hess} f[V]\rangle &= U(Vf) - V(Uf) - (\nabla_U V)f + (\nabla_V U)f \\ &= [U,V]f - (\nabla_U V - \nabla_V U)f \\ &= 0,\end{aligned}$$

by symmetry of the Riemannian connection. □

In closing this section, let us detail implications of Corollary 5.12: we do not need these going forward, but they are instructive. Consider a smooth function f on a Riemannian submanifold $\mathcal{M}$ of a Euclidean space $\mathcal{E}$. Let $\bar{f}$ be a smooth extension of f to a neighborhood of $\mathcal{M}$ in $\mathcal{E}$. By Proposition 3.51, we know that $\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad}\bar{f}(x))$: call this vector field G. We claim that

$$\bar{G}(x) = \mathrm{Proj}_x(\mathrm{grad}\bar{f}(x))$$

is a meaningful smooth extension of G to a neighborhood of $\mathcal{M}$ in $\mathcal{E}$. Indeed, $\mathrm{grad}\bar{f}(x)$ is defined in such a neighborhood. Moreover, from Exercise 3.56 we remember that $x \mapsto \mathrm{Proj}_x$ is a smooth map from $\mathcal{M}$ to the space $\mathcal{L}(\mathcal{E},\mathcal{E})$ of linear maps from $\mathcal{E}$ to $\mathcal{E}$, that is, there exists a smooth extension of $x \mapsto \mathrm{Proj}_x$ to a neighborhood of $\mathcal{M}$ in $\mathcal{E}$: pick any such extension to define $\bar{G}$. By a standard product rule, we work out the differential of $\bar{G}$ at x along $u \in \mathrm{T}_x\mathcal{M}$ to be

$$\mathrm{D}\bar{G}(x)[u] = \mathcal{P}_u(\mathrm{grad}\bar{f}(x)) + \mathrm{Proj}_x(\mathrm{Hess}\bar{f}(x)[u]),$$

where we introduced this notation for the differentials of the projector:

$$\mathcal{P}_u \triangleq \mathrm{D}(x \mapsto \mathrm{Proj}_x)(x)[u] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Proj}_{c(t)}\right|_{t=0}, \tag{5.11}$$

with any smooth curve $c\colon I \to \mathcal{M}$ such that $c(0) = x$ and $c'(0) = u$—this last identity is merely Definition 3.27. Returning to Corollary 5.12, we have found that

$$\mathrm{Hess} f(x)[u] = \mathrm{Proj}_x(\mathrm{Hess}\bar{f}(x)[u]) + \mathrm{Proj}_x\left(\mathcal{P}_u(\mathrm{grad}\bar{f}(x))\right). \tag{5.12}$$

We can say more if we study $\mathcal{P}_u$. To this end, define $P(t) = \mathrm{Proj}_{c(t)}$ with a curve c as above. 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:

$$\mathcal{P}_u = \mathcal{P}_u \circ \mathrm{Proj}_x + \mathrm{Proj}_x \circ \mathcal{P}_u. \tag{5.13}$$

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:

$$\mathcal{P}_u \circ \mathrm{Proj}_x^\perp = \mathrm{Proj}_x \circ \mathcal{P}_u \qquad \text{and} \qquad \mathcal{P}_u \circ \mathrm{Proj}_x = \mathrm{Proj}_x^\perp \circ \mathcal{P}_u. \tag{5.14}$$

Combining (5.12) and (5.14) warrants the following statement.

Corollary 5.15. *Let $\mathcal{M}$ be a Riemannian submanifold of a Euclidean space $\mathcal{E}$. Considering 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:* [AMT13]

$$\operatorname{Hess} f(x)[u] = \operatorname{Proj}_x(\operatorname{Hess} \bar{f}(x)[u]) + \mathcal{P}_u(\operatorname{Proj}_x^\perp(\operatorname{grad} \bar{f}(x))),$$

where $\mathcal{P}_u$ is the differential of $x \mapsto \operatorname{Proj}_x$ at x along u and $\operatorname{Proj}_x^\perp = \operatorname{Id} - \operatorname{Proj}_x$.

This shows that, for Riemannian submanifolds of Euclidean spaces, the Riemannian Hessian is the projected Euclidean Hessian plus a correction term which depends only on the normal part of the Euclidean gradient. See Section 5.11 for a discussion of $\mathcal{P}_u$ and $\operatorname{Hess} f(x)$ in relation to the *second fundamental form* and the *Weingarten map*.

Exercise 5.16. *Continuing Example 5.13, show that if $\operatorname{grad} f(x) = 0$ and $\operatorname{Hess} f(x)$ is positive semidefinite on $\mathrm{T}_x\mathrm{S}^{n-1}$ (i.e., $\langle u, \operatorname{Hess} f(x)[u]\rangle_x \geq 0$ for all $u \in \mathrm{T}_x\mathrm{S}^{n-1}$), then x is a global minimizer of f, i.e., 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 of a point simply by imposing conditions on derivatives at that point.)* See also Exercise 9.48.

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.

Definition 5.17. *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 satisfies these properties for arbitrary $U, V, W \in \mathfrak{X}(\mathcal{M})$, $f, g \in \mathfrak{F}(\mathcal{M})$ 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 5.17, 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 5.17 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 $U(x) = u$.

This 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 5.17.

Proposition 5.18. *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 $W_1, \ldots, W_d$ on a neighborhood $\mathcal{U}$ of x in $\mathcal{M}$ (Proposition 3.59). Given a vector field $U \in \mathfrak{X}(\mathcal{M})$ with $U(x) = 0$, there exist unique smooth functions $g_1, \ldots, g_d \in \mathfrak{F}(\mathcal{U})$ such that

$$U|_{\mathcal{U}} = g_1 W_1 + \cdots + g_d W_d.$$

Clearly, $U(x) = 0$ implies $g_1(x) = \cdots = g_d(x) = 0$. By a technical lemma given hereafter (Lemma 5.24), it is legitimate to write:

$$\begin{aligned}(\nabla_U V)(x) &= (\nabla_{g_1 W_1 + \cdots + g_d W_d} V)(x) \\ &= g_1(x)(\nabla_{W_1} V)(x) + \cdots + g_d(x)(\nabla_{W_d} V)(x) = 0, \end{aligned} \tag{5.15}$$

which concludes the proof. □

In the proof above, it is not immediately clear why (5.15) 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 a standard exercise in real analysis to build such functions.

Lemma 5.19. *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 $\|y - x\| \in (r_1, r_2)$.* [Lee12, Lem. 2.22]

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 5.20. *Let V_1, V_2 be smooth vectors 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 5.19 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 5.21. *Let U_1, U_2 be smooth vectors 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 5.20. 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*.

Lemma 5.22. *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 $g|_{\mathcal{U}'} = f|_{\mathcal{U}'}$.* [Lee12, Lem. 2.26]

Proof. For $\mathcal{M}$ an embedded submanifold of a Euclidean space $\mathcal{E}$, we know from Proposition 3.17 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 5.20, 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 $g|_{\mathcal{U}'} = f|_{\mathcal{U}'}$. □

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.

Lemma 5.23. *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 $V|_{\mathcal{U}'} = U|_{\mathcal{U}'}$.* [Lee12, Lem. 8.6]

Equipped with the last three lemmas, we can finally state the technical result necessary to support the proof of Proposition 5.18.

Lemma 5.24. *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 $U|_{\mathcal{U}} = g_1 W_1 + \cdots + g_d W_d$ for some $g_1, \ldots, g_d \in \mathfrak{F}(\mathcal{U})$ and $W_1, \ldots, W_d \in \mathfrak{X}(\mathcal{U})$. Then,*

$$(\nabla_U V)(x) = g_1(x)(\nabla_{W_1} V)(x) + \cdots + g_d(x)(\nabla_{W_d} V)(x),$$

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 5.22 and 5.23, we know there exist smooth extensions $\tilde{g}_1, \ldots, \tilde{g}_d \in \mathfrak{F}(\mathcal{M})$ and $\tilde{W}_1, \ldots, \tilde{W}_d \in \mathfrak{X}(\mathcal{M})$ that coincide with $g_1, \ldots, g_d$ and $W_1, \ldots, W_d$ on a neighborhood $\mathcal{U}' \subseteq \mathcal{U}$ of x, so that $\tilde{U} = \tilde{g}_1 \tilde{W}_1 + \cdots + \tilde{g}_d \tilde{W}_d$ is a smooth vector field on $\mathcal{M}$ which agrees with U locally: $U|_{\mathcal{U}'} = \tilde{U}|_{\mathcal{U}'}$. Thus, by Lemma 5.21,

$$\begin{aligned}
(\nabla_U V)(x) &= (\nabla_{\tilde{U}} V)(x) \\
&= (\nabla_{\tilde{g}_1 \tilde{W}_1 + \cdots + \tilde{g}_d \tilde{W}_d} V)(x) \\
&= \tilde{g}_1(x)(\nabla_{\tilde{W}_1} V)(x) + \cdots + \tilde{g}_d(x)(\nabla_{\tilde{W}_d} V)(x) \\
&= g_1(x)(\nabla_{W_1} V)(x) + \cdots + g_d(x)(\nabla_{W_d} V)(x).
\end{aligned}$$

The stated definition of $(\nabla_{W_i} V)(x)$ is independent of the choice of smooth extension owing to Lemma 5.21. □

Anticipating our needs for Section 5.7, we note that Lemmas 5.20, 5.22 and 5.23 also allow us to make sense of the notation

$$(\nabla_u (gW))(x) = \mathrm{D}g(x)[u] \cdot W(x) + g(x) \cdot (\nabla_u W)(x), \tag{5.16}$$

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 5.20 and 5.23.

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 $f : \mathcal{M} \to \mathbb{R}$ evaluated along a smooth curve $c : I \to \mathcal{M}$. We already reasoned that the first derivative of g is

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

To obtain a second order expansion of g, we must differentiate g' with respect to t. The 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, to each value of t in I (the domain of c), these maps each associate 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 other words: we need a derivative for vector fields on c.

Below, instead of $(\mathrm{grad} f) \circ c$ we write simply $\mathrm{grad} f \circ c$, where it is clear that we mean the gradient of f (which is a vector field) composed with c. This is as opposed to taking the gradient of $f \circ c$, which we would write $\mathrm{grad}(f \circ c)$ or, more naturally, $(f \circ c)'$.

Definition 5.25. *Given a smooth curve $c : I \to \mathcal{M}$, the map $Z : I \to \mathrm{T}\mathcal{M}$ is a* smooth vector field on c *if $Z(t)$ is in $\mathrm{T}_{c(t)}\mathcal{M}$ for all $t \in I$, and if it is smooth as a map from I (an open submanifold of $\mathbb{R}$) to $\mathrm{T}\mathcal{M}$. The set of smooth vector fields on c is denoted by $\mathfrak{X}(c)$.*

A connection ∇ on a manifold $\mathcal{M}$ induces a natural notion of derivative of vector fields along a curve, with desirable properties. The proof of the following statement involves local frames, which we discussed in Section 3.9. If the reader skipped that section, it is safe to skip the next proof and to consider Proposition 5.27: this is sufficient to work on Riemannian submanifolds.

Theorem 5.26. *Let $c : I \to \mathcal{M}$ be a smooth curve on a manifold equipped with a connection ∇. There exists a unique operator $\frac{\mathrm{D}}{\mathrm{d}t} : \mathfrak{X}(c) \to \mathfrak{X}(c)$ satisfying these three 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 $t \in I$.*

This operator is called the induced[2] covariant derivative*. Furthermore, if $\mathcal{M}$ is a Riemannian manifold with metric $\langle \cdot, \cdot \rangle$ and ∇ is compatible with the metric (e.g., if it 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 $\langle Y, Z \rangle(t) = \langle Y(t), Z(t) \rangle_{c(t)}$.

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 $Z(t_1) \neq Z(t_2)$ (even though Z is smooth). Then, we would not know how to define $U(x)$: should it be equal to $Z(t_1)$ or $Z(t_2)$? To some extent, this is why we need $\frac{\mathrm{D}}{\mathrm{d}t}$.

[2] Induced by the connection ∇.

Proof. We first prove uniqueness under properties 1–3. Pick an arbitrary $\bar{t} \in I$. There exists a local frame $W_1, \ldots, W_d \in \mathfrak{X}(\mathcal{U})$ defined on a neighborhood $\mathcal{U}$ of $c(\bar{t})$ in $\mathcal{M}$ (see Proposition 3.59). Since c is a fortiori 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 $g_1, \ldots, g_d \colon J \to \mathbb{R}$ such that

$$\forall t \in J, \qquad Z(t) = g_1(t)W_1(c(t)) + \cdots + g_d(t)W_d(c(t)).$$

Using the first two properties of the covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$, we get

$$\forall t \in J, \qquad \frac{\mathrm{D}}{\mathrm{d}t}Z(t) = \sum_{i=1}^{d} g_i'(t)W_i(c(t)) + g_i(t)\frac{\mathrm{D}}{\mathrm{d}t}(W_i \circ c)(t).$$

Now using the third property, we find

$$\forall t \in J, \qquad \frac{\mathrm{D}}{\mathrm{d}t}Z(t) = \sum_{i=1}^{d} g_i'(t)W_i(c(t)) + g_i(t)\nabla_{c'(t)}W_i. \tag{5.17}$$

See the discussion around eq. (5.16) for how to interpret $\nabla_{c'(t)}W_i$, considering W_i is only defined locally around $c(t)$.

This expression is fully defined 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 defined by the connection ∇ and the three stated properties.

To prove existence, simply consider (5.17) 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 (5.17) defines 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 holds. To this end, expand Y in the local frame:

$$\forall t \in J, \qquad Y(t) = f_1(t)W_1(c(t)) + \cdots + f_d(t)W_d(c(t)).$$

Then, using also the expansion of Z, for $t \in J$,

$$\langle Y, Z \rangle = \sum_{i,j=1}^{d} f_i g_j \left\langle W_i \circ c, W_j \circ c \right\rangle.$$

Differentiate this with respect to t:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle Y, Z \rangle = \sum_{i,j=1}^{d} (f_i' g_j + f_i g_j') \left\langle W_i \circ c, W_j \circ c \right\rangle + f_i g_j \frac{\mathrm{d}}{\mathrm{d}t}\left\langle W_i \circ c, W_j \circ c \right\rangle. \tag{5.18}$$

On the other hand, by uniqueness we know that (5.17) is a valid expression for $\frac{\mathrm{D}}{\mathrm{d}t}Z$ so that

$$\left\langle Y, \frac{\mathrm{D}}{\mathrm{d}t}Z \right\rangle = \sum_{i,j=1}^{d} f_i g_j' \left\langle W_i \circ c, W_j \circ c \right\rangle + f_i g_j \left\langle W_i \circ c, \nabla_{c'} W_j \right\rangle.$$

Similarly,

$$\left\langle \frac{\mathrm{D}}{\mathrm{d}t} Y, Z \right\rangle = \sum_{i,j=1}^{d} f_i' g_j \left\langle W_i \circ c, W_j \circ c \right\rangle + f_i g_j \left\langle \nabla_{c'} W_i, W_j \circ c \right\rangle.$$

Summing up these identities and comparing to (5.18), we find that property 4 holds if

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle W_i \circ c, W_j \circ c \rangle = \left\langle \nabla_{c'} W_i, W_j \circ c \right\rangle + \left\langle W_i \circ c, \nabla_{c'} W_j \right\rangle.$$

This is indeed the case owing to compatibility of ∇ with the metric, since

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\left\langle W_i, W_j \right\rangle \circ c \right)(t)$$

is the directional derivative of $\left\langle W_i, W_j \right\rangle$ at $c(t)$ along $c'(t)$. $\square$

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 smooth vector field Z along c on $\mathcal{M}$ is also a smooth 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

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \mathrm{Proj}_{c(t)} \left(\frac{\mathrm{d}}{\mathrm{d}t} Z(t) \right) \tag{5.19}$$

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 5.27. *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 (5.19) is the induced covariant derivative, that is, it satisfies properties 1–3 in Theorem 5.26. 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

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \mathrm{Proj}_{c(t)} \left(\bar{\nabla}_{c'(t)} \bar{U} \right) = \nabla_{c'(t)} U,$$

as desired for property 3.

Property 4 follows as a consequence of Theorem 5.26, but we verify it explicitly as this is an important special case. 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

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle Y, Z\rangle &= \left\langle \frac{\mathrm{d}}{\mathrm{d}t}Y, \mathrm{Proj}_c Z \right\rangle + \left\langle \mathrm{Proj}_c Y, \frac{\mathrm{d}}{\mathrm{d}t}Z \right\rangle \\ &= \left\langle \mathrm{Proj}_c \frac{\mathrm{d}}{\mathrm{d}t}Y, Z \right\rangle + \left\langle Y, \mathrm{Proj}_c \frac{\mathrm{d}}{\mathrm{d}t}Z \right\rangle = \left\langle \frac{\mathrm{D}}{\mathrm{d}t}Y, Z \right\rangle + \left\langle Y, \frac{\mathrm{D}}{\mathrm{d}t}Z \right\rangle.\end{aligned}$$

Throughout, we used that $\langle \cdot, \cdot \rangle$ is the metric both in the embedding space and on $\mathcal{M}$. □

Example 5.28. *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 5.26 to Definition 5.11 for the Riemannian Hessian, we get this expression:*

$$\mathrm{Hess} f(x)[u] = \nabla_u \mathrm{grad} f = \left.\frac{\mathrm{D}}{\mathrm{d}t}\mathrm{grad} f(c(t))\right|_{t=0}, \tag{5.20}$$

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 $\mathcal{M}$.*

Example 5.29. *For a smooth function f on a Riemannian submanifold $\mathcal{M}$ of a Euclidean space $\mathcal{E}$, we can apply (5.19) to (5.20) to find*

$$\begin{aligned}\mathrm{Hess} f(x)[u] &= \mathrm{Proj}_x\left(\lim_{t\to 0} \frac{\mathrm{grad} f(c(t)) - \mathrm{grad} f(c(0))}{t}\right) \\ &= \lim_{t\to 0} \frac{\mathrm{Proj}_x(\mathrm{grad} f(c(t))) - \mathrm{grad} f(x)}{t},\end{aligned} \tag{5.21}$$

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 $\bar{t} > 0$,*

$$\mathrm{Hess} f(x)[u] \approx \frac{\mathrm{Proj}_x(\mathrm{grad} f(\mathrm{R}_x(\bar{t}u))) - \mathrm{grad} f(x)}{\bar{t}}. \tag{5.22}$$

This is called 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 discuss this concept in more generality in Section 10.6.*

Exercise 5.30. *In the proof of Theorem 5.26, show that the operator (5.17) satisfies properties 1–3.*

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 5.31. *Let $c\colon I \to \mathcal{M}$ be a smooth curve. The* acceleration of c *is a smooth vector field on c defined as:*

$$c'' = \frac{\mathrm{D}}{\mathrm{d}t} c',$$

where c' is the velocity of c.

For $\mathcal{M}$ embedded in a linear space $\mathcal{E}$, it is convenient to distinguish notationally between the acceleration of c on the manifold (as defined above) and the classical acceleration of c in its 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 then coincide. When $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$, the induced covariant derivative takes on a convenient form (5.19), 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.

Example 5.32. *Consider the sphere $\mathrm{S}^{n-1} = \{x \in \mathbb{R}^n : x^\top x = 1\}$ equipped with the Riemannian submanifold geometry of $\mathbb{R}^n$ with the canonical metric. For a given $x \in \mathrm{S}^{n-1}$ and $v \in \mathrm{T}_x\mathrm{S}^{n-1}$ (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 velocity and acceleration of c in $\mathbb{R}^n$ are easily derived:

$$\dot{c}(t) = -\|v\|\sin(t\|v\|)x + \cos(t\|v\|)v,$$
$$\ddot{c}(t) = -\|v\|^2\cos(t\|v\|)x - \|v\|\sin(t\|v\|)v = -\|v\|^2 c(t).$$

The velocity $c'(t)$ matches $\dot{c}(t)$. Owing to (5.23), to get the acceleration of c on S^{n-1}, we project:

$$c''(t) = \mathrm{Proj}_{c(t)}\ddot{c}(t) = (I_n - c(t)c(t)^\top)\ddot{c}(t) = 0.$$

Thus, c is a curve with zero acceleration on the sphere (even though its acceleration in $\mathbb{R}^n$ 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.

Definition 5.33. *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 $\mathbb{R}$.*

By default, we mean for $\mathcal{M}$ to be equipped with the Riemannian connection ∇ and the induced covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$.

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.

5.9 A second-order Taylor expansion on curves

M, D

On a Riemannian manifold $\mathcal{M}$, consider a real function $f \in \mathfrak{F}(\mathcal{M})$ and a smooth curve $c\colon I \to \mathcal{M}$ 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 $\mathbb{R}$ 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,

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

so that

$$g'(0) = \left\langle \mathrm{grad} f(x), v \right\rangle_x.$$

Moreover, using in turn properties 4 and 3 of Theorem 5.26 regarding the covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$ induced by the Riemannian connection ∇, followed by Definition 5.11 for the Hessian, we compute:

$$\begin{aligned}
g''(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \mathrm{grad} f(c(t)), c'(t) \right\rangle_{c(t)} \\
\text{(property 4)} \quad &= \left\langle \frac{\mathrm{D}}{\mathrm{d}t}\mathrm{grad} f(c(t)), c'(t) \right\rangle_{c(t)} + \left\langle \mathrm{grad} f(c(t)), \frac{\mathrm{D}}{\mathrm{d}t}c'(t) \right\rangle_{c(t)} \\
\text{(property 3)} \quad &= \left\langle \nabla_{c'(t)}\mathrm{grad} f, c'(t) \right\rangle_{c(t)} + \left\langle \mathrm{grad} f(c(t)), c''(t) \right\rangle_{c(t)} \\
&= \left\langle \mathrm{Hess} f(c(t))[c'(t)], c'(t) \right\rangle_{c(t)} + \left\langle \mathrm{grad} f(c(t)), c''(t) \right\rangle_{c(t)},
\end{aligned}$$

so that

$$g''(0) = \left\langle \mathrm{Hess} f(x)[v], v \right\rangle_x + \left\langle \mathrm{grad} f(x), c''(0) \right\rangle_x. \tag{5.24}$$

These all combine to form:

$$f(c(t)) = f(x) + t \langle \mathrm{grad} f(x), v \rangle_x + \frac{t^2}{2} \langle \mathrm{Hess} f(x)[v], v \rangle_x + \frac{t^2}{2} \left\langle \mathrm{grad} f(x), c''(0) \right\rangle_x + O(t^3). \quad (5.25)$$

Of particular interest for optimization is the Taylor expansion of f along a retraction curve. This is the topic of the next section.

Exercise 5.34. *Given a smooth curve $c\colon [0,1] \to \mathcal{M}$ on a Riemannian manifold $\mathcal{M}$ with $c(0) = x$ and $c(1) = y$, show that there exists $t \in (0,1)$ such that*

$$f(y) = f(x) + \left\langle \mathrm{grad} f(x), c'(0) \right\rangle_x + \frac{1}{2} \left\langle \mathrm{Hess} f(c(t))[c'(t)], c'(t) \right\rangle_{c(t)} + \frac{1}{2} \left\langle \mathrm{grad} f(c(t)), c''(t) \right\rangle_{c(t)}. \quad (5.26)$$

Hint: use the mean value theorem.

This last exercise notably yields the following statement.

Lemma 5.35. *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 all $t \in [0,1]$ (for some $\mu \in \mathbb{R}$). Then,*

$$f(y) \geq f(x) + \langle \mathrm{grad} f(x), v \rangle_x + \frac{\mu}{2} \|v\|_x^2.$$

Proof. This follows from (5.26) because $\|c'(t)\|_{c(t)}$ is constant. $\square$

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, we may deduce 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 $\mathcal{M}$)—see 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.

5.10 *Second-order retractions*

Continuing from the Taylor expansion (5.25) 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 (5.25) yields

$$f(\mathrm{R}_x(tv)) = f(x) + t\,\langle \mathrm{grad} f(x), v\rangle_x + \frac{t^2}{2}\,\langle \mathrm{Hess} f(x)[v], v\rangle_x \\ + \frac{t^2}{2}\,\big\langle \mathrm{grad} f(x), c''(0)\big\rangle_x + O(t^3). \quad (5.27)$$

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 either $\mathrm{grad} f(x) = 0$ or $c''(0) = 0$. The latter happens in particular if c is a geodesic,[3] but notice that we only need the acceleration at $t = 0$ to vanish. This suggests the following definition.

[3] This is the case if R is the exponential map: see Section 10.2.

Definition 5.36. *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. In particular, if $\mathcal{M}$ is a Riemannian submanifold of a Euclidean space, then the map $(x, v) \mapsto \mathrm{R}_x(v)$ which consists in projecting[4] $x + v$ to the manifold (an operation which may only be defined for small v) is a second-order retraction: see Propositions 5.42 and 5.43. This covers some of the retractions we discuss on the sphere, Stiefel, the orthogonal group and the set of fixed-rank matrices in Chapter 7. In the following example, we verify this explicitly for the sphere.

[4] This is called metric projection.

Example 5.37. *Consider this retraction on the sphere S^{n-1}:*

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

This retraction is second order. Indeed, with $c(t) = \mathrm{R}_x(tv)$:

$$\begin{aligned} c(t) &= \frac{x + tv}{\sqrt{1 + t^2\|v\|^2}} = \left(1 - \frac{1}{2}\|v\|^2 t^2 + O(t^4)\right)(x + tv), \\ \dot{c}(t) &= -\|v\|^2 t(x + tv) + \left(1 - \frac{1}{2}\|v\|^2 t^2\right) v + O(t^3), \\ \ddot{c}(t) &= -\|v\|^2(x + tv) - \|v\|^2 tv - \|v\|^2 tv + O(t^2) \\ &= -\|v\|^2(x + 3tv) + O(t^2). \end{aligned}$$

Of course, $c'(0) = \dot{c}(0) = v$. As for acceleration, $\ddot{c}(0) = -\|v\|^2 x$, so that:

$$c''(0) = \mathrm{Proj}_x(\ddot{c}(0)) = 0,$$

as announced.

Using this new definition, we turn (5.27) into the following useful statement regarding the pullback $f \circ \mathrm{R}_x$.

Proposition 5.38. *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*

$$f(\mathrm{R}_x(s)) = f(x) + \frac{1}{2} \langle \mathrm{Hess} f(x)[s], s \rangle_x + O(\|s\|_x^3). \tag{5.28}$$

If R *is a second-order retraction, then for all points $x \in \mathcal{M}$ we have*

$$f(\mathrm{R}_x(s)) = f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{1}{2} \langle \mathrm{Hess} f(x)[s], s \rangle_x + O(\|s\|_x^3). \tag{5.29}$$

Proof. Simply rewrite (5.27) with $s = tv$. □

That the first identity holds for all retractions is useful to study the behavior of optimization algorithms close to critical points.

Proposition 5.38 suggests an alternative way to compute the Riemannian Hessian. Indeed, the direct way is to use the definition as we did in Example 5.13. 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 3.49 stating $\mathrm{grad} f(x) = \mathrm{grad}(f \circ \mathrm{R}_x)(0)$.

Proposition 5.39. *If the retraction is second order or if* $\mathrm{grad} f(x) = 0$*, then*

See also Exercise 10.67.

$$\mathrm{Hess} f(x) = \mathrm{Hess}(f \circ \mathrm{R}_x)(0),$$

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.

Proof. If R is second order, expand $\hat{f}_x(s) = f(\mathrm{R}_x(s))$ using (5.29):

$$\hat{f}_x(s) = f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{1}{2} \langle \mathrm{Hess} f(x)[s], s \rangle_x + O(\|s\|_x^3).$$

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 (5.28). □

5.11 Notes and references

Definition 5.1 for connections is not standard. 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 (see Section 5.6). Leading with Definition 5.1 makes it possible to skip these technicalities at first.

The pointwise dependence 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 in that full 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 5.20) or along any smooth curve passing through x with velocity u (owing to the chain rule property in Theorem 5.26)—see also [Lee18, Prop. 4.26].

Existence and uniqueness of the Riemannian connection is proved in most Riemannian geometry textbooks, including [Lee12, Prop. 4.12 and 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 5.6. 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 5.4, 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 5.8). Our proof is non-standard and restricted to embedded submanifolds. We provide a general proof later in Section 8.8. In the meantime, we get some insight along these lines: Exercise 5.9 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 5.10 claims Lie brackets are derivations, so that Lie brackets are indeed smooth vector fields. (The proof in Section 8.8 follows yet another path.)

We follow the definition of Riemannian Hessian preferred by Absil et al. [AMS08, §5.5]. The definition of second-order retractions and

Proposition 5.39 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].

For $\mathcal{M}$ a Riemannian submanifold of a Euclidean space $\mathcal{E}$, Corollary 5.15 illustrates the relevance of the directional derivatives of the orthogonal projector (5.11),

$$\mathcal{P}_u = \mathrm{D}(x \mapsto \mathrm{Proj}_x)(x)[u].$$

With $\mathrm{Proj}_x^\perp = \mathrm{Id} - \mathrm{Proj}_x$ the orthogonal projector to the normal space $\mathrm{N}_x\mathcal{M} = (\mathrm{T}_x\mathcal{M})^\perp$, we found in (5.14) that

$$\mathcal{P}_u \circ \mathrm{Proj}_x^\perp = \mathrm{Proj}_x \circ \mathcal{P}_u \qquad \text{and} \qquad \mathcal{P}_u \circ \mathrm{Proj}_x = \mathrm{Proj}_x^\perp \circ \mathcal{P}_u.$$

Thus, if $v \in \mathrm{N}_x\mathcal{M}$ is a normal vector at x, then $v = \mathrm{Proj}_x^\perp(v)$ and

$$\mathcal{P}_u(v) = \mathcal{P}_u(\mathrm{Proj}_x^\perp(v)) = \mathrm{Proj}_x(\mathcal{P}_u(v)).$$

The output is necessarily a tangent vector at x. This motivates us to define the following bilinear map called the *Weingarten map*:

$$\mathcal{W}\colon \mathrm{T}_x\mathcal{M} \times \mathrm{N}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}\colon (u, v) \mapsto \mathcal{W}(u, v) = \mathcal{P}_u(v). \tag{5.30}$$

Likewise, if $w \in \mathrm{T}_x\mathcal{M}$ is a tangent vector at x, then $w = \mathrm{Proj}_x(w)$ and

$$\mathcal{P}_u(w) = \mathcal{P}_u(\mathrm{Proj}_x(w)) = \mathrm{Proj}_x^\perp(\mathcal{P}_u(w)).$$

The output is necessarily a normal vector at x, leading us to define this bilinear map called the *second fundamental form*:

Read “two” for II.

$$\mathrm{II}\colon \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M} \to \mathrm{N}_x\mathcal{M}\colon (u, w) \mapsto \mathrm{II}(u, w) = \mathcal{P}_u(w). \tag{5.31}$$

For an arbitrary $z \in \mathcal{E}$, these combined definitions allow us to write

$$\mathcal{P}_u(z) = \mathrm{II}(u, \mathrm{Proj}_x(z)) + \mathcal{W}(u, \mathrm{Proj}_x^\perp(z)). \tag{5.32}$$

With respect to Corollary 5.15 in particular, we get an identity for the Hessian described in [AMT13]:

Compare (5.33) with (7.77) which provides explicit expressions in terms of a local defining function h for $\mathcal{M}$.

$$\mathrm{Hess} f(x)[u] = \mathrm{Proj}_x(\mathrm{Hess}\bar{f}(x)[u]) + \mathcal{W}(u, \mathrm{Proj}_x^\perp(\mathrm{grad}\bar{f}(x))). \tag{5.33}$$

In quadratic form, we also have (with $\langle \cdot, \cdot \rangle$ denoting both the Euclidean inner product and the Riemannian metric at x since $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$)

$$\begin{aligned}
\langle w, \mathrm{Hess} f(x)[u] \rangle &= \langle w, \mathrm{Hess}\bar{f}(x)[u] \rangle + \langle w, \mathcal{W}(u, \mathrm{Proj}_x^\perp(\mathrm{grad}\bar{f}(x))) \rangle \\
&= \langle w, \mathrm{Hess}\bar{f}(x)[u] \rangle + \langle w, \mathcal{P}_u(\mathrm{Proj}_x^\perp(\mathrm{grad}\bar{f}(x))) \rangle \\
&= \langle w, \mathrm{Hess}\bar{f}(x)[u] \rangle + \langle \mathcal{P}_u(w), \mathrm{Proj}_x^\perp(\mathrm{grad}\bar{f}(x)) \rangle \\
&= \langle w, \mathrm{Hess}\bar{f}(x)[u] \rangle + \langle \mathrm{II}(u, w), \mathrm{grad}\bar{f}(x) \rangle, \qquad (5.34)
\end{aligned}$$

where we used that Proj_x, $\mathrm{Proj}_x^\perp$ and $\mathcal{P}_u$ are self-adjoint on $\mathcal{E}$.

While it is not obvious from the definition (5.31), we may surmise from (5.34) that II is symmetric in its inputs. This is indeed the case.

Proposition 5.40. *For all $u, w \in \mathrm{T}_x\mathcal{M}$ we have $\mathrm{II}(u, w) = \mathrm{II}(w, u)$.* [Lee18, Prop. 8.1]

Proof. For an arbitrary tangent vector $w \in \mathrm{T}_x\mathcal{M}$, consider the smooth vector field $W \in \mathfrak{X}(\mathcal{M})$ defined by $W(y) = \mathrm{Proj}_y(w)$. For $u \in \mathrm{T}_x\mathcal{M}$,

$$\mathrm{D}W(x)[u] = \mathrm{D}\Big(y \mapsto \mathrm{Proj}_y(w)\Big)(x)[u] = \mathcal{P}_u(w) = \mathrm{II}(u, w).$$

Moreover, with $\bar{\nabla}$ and ∇ denoting the Riemannian connections on $\mathcal{E}$ and $\mathcal{M}$ respectively, remember from Propositions 5.5 and 3.28 that

$$\bar{\nabla}_u \bar{W} = \mathrm{D}\bar{W}(x)[u] = \mathrm{D}W(x)[u] = \mathrm{II}(u, w). \tag{5.35}$$

(This also implies $\nabla_u W = \mathrm{Proj}_x(\bar{\nabla}_u \bar{W}) = 0$ since $\mathrm{II}(u, w)$ is in $\mathrm{N}_x\mathcal{M}$.) Likewise, define $U \in \mathfrak{X}(\mathcal{M})$ by $U(y) = \mathrm{Proj}_y(u)$ and let $\bar{U}$ be any smooth extension of U. Then, symmetry of $\bar{\nabla}$ implies

$$\mathrm{II}(u, w) - \mathrm{II}(w, u) = \bar{\nabla}_u \bar{W} - \bar{\nabla}_w \bar{U} = (\bar{\nabla}_{\bar{U}} \bar{W} - \bar{\nabla}_{\bar{W}} \bar{U})(x) = [\bar{U}, \bar{W}](x).$$

We learned from the proof of Theorem 5.6 that $[\bar{U}, \bar{W}](x)$ is tangent to $\mathcal{M}$ at x; yet $\mathrm{II}(u, w) - \mathrm{II}(w, u)$ is normal to $\mathcal{M}$ at x. Thus, we conclude that $\mathrm{II}(u, w) - \mathrm{II}(w, u) = 0$. □

Definitions (5.31) and (5.30) 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 above) extend as is to the more general situation of a Riemannian submanifold of a Riemannian manifold, as defined in Section 8.14.

We conclude this discussion with two additional useful facts about II. First, given any $V \in \mathfrak{X}(\mathcal{M})$ and smooth extension $\bar{V}$, the *Gauss formula* states that [Lee18, Thm. 8.2]

$$\bar{\nabla}_u \bar{V} = \mathrm{Proj}_x(\bar{\nabla}_u \bar{V}) + \mathrm{Proj}_x^\perp(\bar{\nabla}_u \bar{V}) = \nabla_u V + \mathrm{II}(u, v), \tag{5.36}$$

where u and $v = V(x)$ are tangent at x. Second, if γ is a geodesic of $\mathcal{M}$ satisfying $\gamma(0) = x$ and $\gamma'(0) = u$, then $\mathrm{II}(u, u)$ is the *extrinsic* acceleration of γ in the embedding space [Lee18, Prop. 8.10]: this informs us regarding the *extrinsic curvature* of $\mathcal{M}$ in its embedding space, and may be useful to interpret (5.34).

The extension lemmas (Lemmas 5.22 and 5.23) 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 5.19) 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].

The characterization of geodesics on Riemannian submanifolds as curves with normal acceleration appears in [Lee18, Cor. 5.2] and also in [O'N83, Cor. 4.10].

We close with a word about retractions of the form

$$\mathrm{R}_x(v) = \operatorname*{arg\,min}_{x' \in \mathcal{M}} \|x' - (x+v)\| \tag{5.37}$$

for $\mathcal{M}$ an embedded submanifold of a Euclidean space $\mathcal{E}$ equipped with the inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$. The distance $\mathrm{dist}_{\mathcal{M}} \colon \mathcal{E} \to \mathbb{R}$ from a point of $\mathcal{E}$ to the set $\mathcal{M}$ is defined as

$$\mathrm{dist}_{\mathcal{M}}(y) = \inf_{x \in \mathcal{M}} \|x - y\|. \tag{5.38}$$

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. In general, A could be neither open nor closed: consider $\mathcal{M} = \{(t, t^2) : t \in \mathbb{R}\}$ for which $A = \mathbb{R}^2 \backslash \{(0, t) : t > 1/2\}$ [DH94, Ex. 6.1]. However, strong properties hold on the interior of A (that is, on the largest subset of A which is open in $\mathcal{E}$).

Theorem 5.41. *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.*

[DH94, Thms 3.8 and 3.13, Cor. 3.14, Thm. 4.1]

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 for all $x \in \mathcal{M}$ we have $\mathrm{D}P_{\mathcal{M}}(x) = \mathrm{Proj}_x$ (the orthogonal projector from $\mathcal{E}$ to $\mathrm{T}_x\mathcal{M}$).*

Now consider this subset of the tangent bundle of $\mathcal{M}$:

$$\mathcal{O} = \{(x, v) \in \mathrm{T}\mathcal{M} : x + v \in \Omega\}. \tag{5.39}$$

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. We use metric projection to define a retraction on $\mathcal{O}$.

Proposition 5.42. *On an embedded submanifold $\mathcal{M}$ of a Euclidean space $\mathcal{E}$ with norm $\|\cdot\|$, this map is a retraction:*

$$\mathrm{R}\colon \mathcal{O} \to \mathcal{M}\colon (x, v) \mapsto \mathrm{R}(x, v) = \mathrm{R}_x(v) = P_{\mathcal{M}}(x + v).$$

Proof. Clearly, $\mathrm{R}_x(0) = P_{\mathcal{M}}(x) = x$ and R is smooth on its domain by composition. Moreover, for all $(x, u) \in \mathrm{T}\mathcal{M}$ it holds that

$$\mathrm{DR}_x(0)[u] = \mathrm{D}P_{\mathcal{M}}(x)[u] = \mathrm{Proj}_x(u) = u,$$

confirming that $\mathrm{DR}_x(0)$ is the identity on $\mathrm{T}_x\mathcal{M}$. □

Absil and Malick show that this retraction is part of a large family of second-order retractions (recall Definition 5.36) [AM12, Ex. 23]. For the case at hand, Razvan-Octavian Radu shared the short proof below.

Proposition 5.43. *If $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$, the retraction in Proposition 5.42 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 5.41, we know that $x + tv - c(t)$ is orthogonal to $\mathrm{T}_{c(t)}\mathcal{M}$ for all t. (We can 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 3.56). 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 necessarily true that if $v \in \mathcal{O}_x$ then $tv \in \mathcal{O}_x$ for all $t \in [0,1]$. Indeed, consider this example [DH94, Ex. 6.4]: $\mathcal{M} = \{(t, \cos(t)) : t \in \mathbb{R}\} \subset \mathbb{R}^2$, for which A is all of $\mathbb{R}^2$ except for certain half lines: $A = \mathbb{R}^2 \setminus \cup_{k\in\mathbb{Z}} \{(k\pi, t) : (-1)^{k+1}t > 0\}$. Now consider $x = (0,1) \in \mathcal{M}$. The domain $\mathcal{O}_x \subset \mathrm{T}_x\mathcal{M} = \{(t,0) : t \in \mathbb{R}\}$ is the set of pairs $(t,0)$ such that $(t,1)$ is in the interior of A. Thus, $\mathcal{O}_x$ is the set of pairs $(t,0)$ such that t/π is not an odd integer. This set is not star-shaped (but it can be reduced to a star-shaped domain by restriction to $\{(t,0) : -\pi < t < \pi\}$). Another example is metric projection to the set of matrices of fixed rank $\mathbb{R}_r^{m\times n}$ (7.47): if $X \in \mathbb{R}_r^{m\times n}$, then $V = -X \in \mathrm{T}_X\mathbb{R}_r^{m\times n}$; consider the line $t \mapsto X + tV$: projection of $X + tV$ to $\mathbb{R}_r^{m\times n}$ is well defined for all t except $t = 1$.

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 $\mathcal{E}$ [DH94, Thm. 5.3]. 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 whenever $\mathcal{M}$ is the boundary of a nonempty, 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 [DH94, Thm. 5.1]. Examples are the metric projections onto the sphere (7.7) and (more generally) onto the Stiefel manifold (7.22): the sphere is the boundary of the unit Euclidean ball in $\mathbb{R}^n$, and the Stiefel manifold is the boundary of the unit operator-norm ball in $\mathbb{R}^{n\times p}$ $(p < n)$. This extends to $\mathrm{SO}(n)$ and (with some care regarding its two components) to $\mathrm{O}(n)$ (Section 7.4).

From Theorem 5.41 it is also 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 5.43 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) [BV19].

6
Second-order optimization algorithms

First-order derivatives of the cost function $f \colon \mathcal{M} \to \mathbb{R}$, through the gradient, provide exploitable information for optimization algorithms. In the same way, second-order derivatives, through the Hessian, play an important role for the development of richer algorithms.

In this chapter, we discuss a Riemannian version of the famous Newton method: a pillar of both optimization and numerical analysis. As each iteration of this algorithm involves solving a linear system of equations on a tangent space of $\mathcal{M}$, we take a moment to discuss the (linear) conjugate gradients method. For all its qualities, Newton's method unfortunately has many flaws, not the least of which is that its behavior critically depends on initialization, that is, on the choice of $x_0 \in \mathcal{M}$. From the classical literature on optimization in linear spaces, we know that these flaws can be corrected by adding safeguards, yielding the trust-region method. We discuss and analyze a Riemannian version of this algorithm: arguably the most robust algorithm for smooth optimization on manifolds to date. For Riemannian trust regions, too, each iteration requires solving a subproblem. We detail how a variation of the conjugate gradients method, known as truncated conjugate gradients, solves the subproblem in an elegant and robust way.

6.1 Second-order optimality conditions

Before we move on to discuss second-order optimization algorithms, we secure a second-order necessary optimality condition. It involves spectral properties of the Hessian (see Definition 3.5).

Proposition 6.1. *Consider a smooth function $f \colon \mathcal{M} \to \mathbb{R}$ defined on a Riemannian manifold $\mathcal{M}$. If x is a local (or global) minimizer of f, then* $\mathrm{grad} f(x) = 0$ *and* $\mathrm{Hess} f(x) \succeq 0$.

Proof. We established $\mathrm{grad} f(x) = 0$ in Proposition 4.4. For contradiction, assume $\mathrm{Hess} f(x)$ is not positive semidefinite. Then, there exists a tangent vector $v \in \mathrm{T}_x\mathcal{M}$ such that $\langle \mathrm{Hess} f(x)[v], v \rangle_x = -2a < 0$, for some positive a. Let $c\colon I \to \mathcal{M}$ be a smooth curve passing through x with velocity v at $t = 0$. Then, using the Taylor expansion (5.25) and the fact that $\mathrm{grad} f(x) = 0$, we can write

$$\begin{aligned} f(c(t)) &= f(x) + \frac{t^2}{2} \langle \mathrm{Hess} f(x)[v], v \rangle_x + O(t^3) \\ &= f(x) - at^2 + O(t^3). \end{aligned}$$

Hence, there exists $\bar{t}$ such that

$$f(c(t)) < f(x) \quad \text{for all } t \text{ in } (0, \bar{t}],$$

contradicting the fact that x is a local minimizer. □

Proposition 6.2. *Consider a smooth function $f\colon \mathcal{M} \to \mathbb{R}$ defined on a Riemannian manifold $\mathcal{M}$. If $\mathrm{grad} f(x) = 0$ and $\mathrm{Hess} f(x) \succ 0$, then x is a strict local (or global) minimizer of f.*

Proof sketch. Recall equation (5.26): at the critical point x, for any smooth curve $c\colon [0,1] \to \mathcal{M}$ such that $c(0) = x$ and $c(1) = y$, there exists $t \in (0,1)$ such that

$$f(y) = f(x) + \frac{1}{2} \left\langle \mathrm{Hess} f(c(t))[c'(t)], c'(t) \right\rangle_{c(t)} + \frac{1}{2} \left\langle \mathrm{grad} f(c(t)), c''(t) \right\rangle_{c(t)}.$$

This proof sketch notably uses the existence of geodesics between any two sufficiently close points, which we further discuss in Section 10.2. A lighter way to prove this statement is to use charts, which we introduce for general manifolds in Chapter 8.

We claim without proof that there exists a neighborhood $\mathcal{U}$ around x on $\mathcal{M}$ such that, for all $y \in \mathcal{U}$, the curve c can be chosen to map $[0,1]$ into $\mathcal{U}$ and to be a geodesic, that is, $c''(t) = 0$ and (consequently) $\|c'(t)\|_{c(t)} = \|c'(0)\|_x$. Thus, for all $y \in \mathcal{U}$,

$$f(y) \geq f(x) + \frac{\lambda}{2} \|c'(0)\|_x^2, \qquad \text{with} \qquad \lambda = \inf_{z \in \mathcal{U}} \lambda_{\min}(\mathrm{Hess} f(z)).$$

Since $\lambda_{\min}(\mathrm{Hess} f(x)) > 0$, by continuity of the smallest eigenvalue of the Hessian, the neighborhood $\mathcal{U}$ can be reduced to ensure $\lambda > 0$. For $y \neq x$ we have $\|c'(0)\|_x > 0$, hence $f(y) > f(x)$ as desired. □

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 cost function along iterates is dictated by 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 general strategy to design algorithms is choose models $m_{x_k}\colon \mathrm{T}_{x_k}\mathcal{M} \to \mathbb{R}$ which suitably approximate $\hat{f}_{x_k}$, and to pick s_k as an (approximate) minimizer of m_{x_k}. This is beneficial if the model is substantially simpler than the pullback while being close to it. Based on our work building Taylor expansions (Proposition 5.38), we know that for second-order retractions or close to critical points it holds that

$$\hat{f}_x(s) \approx m_x(s) = f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{1}{2} \langle \mathrm{Hess} f(x)[s], s \rangle_x .$$

A minimizer of m_x, if one exists, must be a critical point of m_x. Thus, we compute

$$\mathrm{grad}\, m_x(s) = \mathrm{grad} f(x) + \mathrm{Hess} f(x)[s],$$

and we note that s is a critical point of m_x if and only if

$$\mathrm{Hess} f(x)[s] = -\mathrm{grad} f(x).$$

Since $\mathrm{Hess} f(x)$ is self-adjoint,

$$\begin{aligned} \mathrm{D}m_x(s)[u] &= \langle \mathrm{grad} f(x), u \rangle_x \\ &\quad + \langle \mathrm{Hess} f(x)[s], u \rangle_x ; \end{aligned}$$

Also, $\mathrm{D}m_x(s)[u] = \langle \mathrm{grad}\, m_x(s), u \rangle_x$ by definition. Conclude by identification.

This defines a *linear* system of equations (called the *Newton equations*) in 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. Of course, the critical point of the quadratic function m_x is its minimizer if and only if $\mathrm{Hess} f(x)$ is positive definite. In all other cases, the Newton step does *not* correspond to a minimizer of m_x, which may lead the sequence of iterates 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 notion of local convergence rate introduced in Section 4.6. In contrast to the typical behavior of gradient descent, Newton's method may converge superlinearly. Specifically, it may converge *quadratically* (or even faster), as we now define.

Definition 6.3. *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 $\mu \geq 0$.*

Under this definition, pick K such that

$$\epsilon_k \leq \frac{1}{2(\mu + 0.1)} \quad \text{and} \quad \frac{\epsilon_{k+1}}{\epsilon_k^2} \leq \mu + 0.1$$

for all $k \geq K$. Then,

$$\epsilon_{k+1} \leq \epsilon_k^2(\mu + 0.1) \leq \frac{\epsilon_k}{2}.$$

Thus, the sequence also converges at least linearly (Definition 4.12).

In Section 10.1, we discuss how a Riemannian manifold can be endowed with a natural notion of distance, called the *Riemannian distance*.

Up to technical points, $\mathrm{dist}(x, y)$ is the length of the shortest curve on $\mathcal{M}$ joining x and y. This allows us to make sense of convergence rates for sequences of iterates. Specifically, we have the following result for Riemannian Newton's method.

Theorem 6.4. *Given a smooth function $f \colon \mathcal{M} \to \mathbb{R}$ on a Riemannian manifold $\mathcal{M}$, let $x_\star$ be such that $\mathrm{grad} f(x_\star) = 0$ and $\mathrm{Hess} f(x_\star)$ is invertible.*[1] *There exists a neighborhood $\mathcal{U}$ of $x_\star$ on $\mathcal{M}$ such that, for any $x_0 \in \mathcal{U}$, 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 $x_\star$.*

[AMS08, Thm. 6.3.2]

[1] Such a point may not exist, and it may not be unique.

The retraction does *not* need to be second-order. Essentially, this is due to Proposition 5.38 and the fact that $x_\star$ is a critical point.

From an optimization perspective, Theorem 6.4 is only beneficial in the case $\mathrm{Hess} f(x_\star) \succ 0$. Indeed, by Proposition 6.1, 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 $\mathcal{U}$ may be arbitrarily small.

For this reason, 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. This will prove useful for the safeguarded algorithm as well.

6.3 Computing Newton steps: conjugate gradients

For a smooth function f on a Riemannian manifold $\mathcal{M}$, the Newton step at $x \in \mathcal{M}$ is the tangent vector $s \in \mathrm{T}_x\mathcal{M}$ such that

$$\mathrm{Hess} f(x)[s] = -\mathrm{grad} f(x), \tag{6.1}$$

assuming the linear operator $\mathrm{Hess} f(x)$ is invertible. In practice, we do not have access to $\mathrm{Hess} f(x)$ as a matrix. Rather, we can access the Hessian as an operator, computing $\mathrm{Hess} f(x)[s]$ for any given $s \in \mathrm{T}_x\mathcal{M}$. Solvers for linear systems which only require access to this operator are called *matrix free*.

Upon choosing a basis for $\mathrm{T}_x\mathcal{M}$, one could form a matrix which represents $\mathrm{Hess} f(x)$ in that basis, but this is computationally expensive as it requires to apply $\mathrm{Hess} f(x)$ to the $\dim \mathcal{M}$ basis vectors, just to form that matrix.

We readily know that $\mathrm{Hess} f(x)$ is self-adjoint with respect to $\langle \cdot, \cdot \rangle_x$. Furthermore, the Newton step is only meaningful for optimization if $\mathrm{Hess} f(x)$ is positive definite: we here assume it is. The most famous matrix-free solver for self-adjoint, positive definite linear systems is the *conjugate gradients method* (CG).

To simplify notation, let H denote $\mathrm{Hess} f(x)$ and let b denote the right hand side, $-\mathrm{grad} f(x)$. The solution s satisfies $Hs = b$. Iteratively, CG produces better and better approximations of s as a sequence $v_0, v_1, \ldots$ in $\mathrm{T}_x\mathcal{M}$. To make sense of this statement, we must specify a norm. We could of course work with the norm $\|u\|_x = \sqrt{\langle u, u\rangle_x}$, induced by the Riemannian metric at x. Instead, and this is crucial, CG tracks the approximation error in the H-norm, defined as:

$$\|u\|_H = \sqrt{\langle u, Hu\rangle_x}. \tag{6.2}$$

This is a norm on $\mathrm{T}_x\mathcal{M}$ since H is positive definite. The approximation error for a vector $v \in \mathrm{T}_x\mathcal{M}$ then obeys:

$$\begin{aligned} \|v - s\|_H^2 &= \langle v - s, H(v - s)\rangle_x \\ &= \langle v, Hv\rangle_x - \langle s, Hv\rangle_x - \langle v, Hs\rangle_x + \langle s, Hs\rangle_x \\ &= \langle v, Hv\rangle_x - 2\langle v, b\rangle_x + \langle s, Hs\rangle_x, \end{aligned}$$

where we used that H is self-adjoint and $Hs = b$. Defining

$$g(v) = \frac{1}{2}\langle v, Hv\rangle_x - \langle v, b\rangle_x, \tag{6.3}$$

we find that

$$\|v - s\|_H^2 = 2g(v) + \langle s, Hs\rangle_x. \tag{6.4}$$

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. This is useful specifically because $\|v - s\|_H$ cannot be computed (since s is not yet known), whereas g is computable.

We now show that, if one has access to nonzero tangent vectors $p_0, \ldots, p_{n-1}$ (for some integer $n \geq 1$) such that

$$\forall i \neq j, \qquad \langle p_i, Hp_j\rangle_x = 0,$$

then minimizing g over the space spanned by $p_0, \ldots, p_{n-1}$ is straightforward. Such directions are called *H-conjugate*: they are orthogonal with respect to the inner product induced by H.

First, observe that H-conjugate directions are necessarily linearly independent. Indeed, otherwise, there would exist a linear combination

$$\alpha_0 p_0 + \cdots + \alpha_{n-1} p_{n-1} = 0$$

with at least one nonzero coefficient α_i. This cannot be, as otherwise

$$0 = \left\langle \sum_i \alpha_i p_i, H \sum_j \alpha_j p_j \right\rangle_x = \sum_{ij} \alpha_i \alpha_j \langle p_i, Hp_j\rangle_x = \sum_i \alpha_i^2 \langle p_i, Hp_i\rangle_x \neq 0,$$

using that $\langle p_i, Hp_i \rangle_x > 0$ by positive definiteness of H.

Second, using that any vector v in the span of the H-conjugate directions can be expressed uniquely as

$$v = y_0 p_0 + \cdots + y_{n-1} p_{n-1}$$

with some coefficients $y_0, \ldots, y_{n-1} \in \mathbb{R}$, we can express $g(v)$ as a function of these coefficients:

$$\begin{aligned} g(v) &= \frac{1}{2} \left\langle \sum_i y_i p_i, H \sum_j y_j p_j \right\rangle_x - \left\langle \sum_i y_i p_i, b \right\rangle_x \\ &= \sum_i \left[\frac{1}{2} y_i^2 \langle p_i, Hp_i \rangle_x - y_i \langle p_i, b \rangle_x \right]. \end{aligned}$$

The simplifications brought about by H-conjugacy are sizable. Indeed, at a minimizer, the partial derivatives of g with respect to each coefficient y_i,

$$\frac{\partial g}{\partial y_i} = y_i \langle p_i, Hp_i \rangle_x - \langle p_i, b \rangle_x,$$

must vanish. This translates into simple, decoupled conditions:

$$\forall i \in \{0, \ldots, n-1\}, \qquad y_i = \frac{\langle p_i, b \rangle_x}{\langle p_i, Hp_i \rangle_x}.$$

This reveals the (unique) minimizer in closed form.

A few comments are in order:

1. Given H-conjugate directions $p_0, \ldots, p_{n-1}$, the minimizer of $g(v)$ over the span of $p_0, \ldots, p_{n-1}$ is

   $$v_n = \sum_{i=0}^{n-1} \frac{\langle p_i, b \rangle_x}{\langle p_i, Hp_i \rangle_x} p_i.$$

   Due to (6.4), this is *also* the minimizer of $\|v - s\|_H$ over the same subspace.

2. Given v_n, if we produce an additional H-conjugate direction p_n, then it is easy to compute the minimizer of $g(v)$ (or $\|v - s\|_H$) over the (larger) subspace spanned by $p_0, \ldots, p_n$:

   $$v_{n+1} = v_n + \frac{\langle p_n, b \rangle_x}{\langle p_n, Hp_n \rangle_x} p_n.$$

   Of course, v_{n+1} is at least as good an approximation of s as v_n, since it is obtained through minimization of $\|v - s\|_H$ over a larger subspace.

3. Since H-conjugate directions are linearly independent, if we can produce $\dim \mathcal{M}$ of them, then they span the whole tangent space $\mathrm{T}_x\mathcal{M}$. Since s is the global minimizer of $\|v - s\|_H$ over $\mathrm{T}_x\mathcal{M}$, it follows that $v_{\dim \mathcal{M}} = s$. In other words: after $\dim \mathcal{M}$ updates, the solution would be found.

Evidently, the crux of the effort is to devise an efficient procedure to generate H-conjugate directions. The CG algorithm does this while simultaneously generating the approximate solutions $v_1, v_2, \ldots$: see Algorithm 6.2. The fact that the directions p_n produced by this algorithm are H-conjugate, and the fact that the vectors v_n produced by this algorithm are equivalent to those described above, follows from a standard proof by induction: see for example [TB97, Thm. 38.1]. Notice that CG requires a single evaluation of the operator H (in our setting, a single evaluation of $\mathrm{Hess} f(x)$) per iteration, and we know CG terminates after at most $\dim \mathcal{M}$ iterations since this is the maximum number of H-conjugate directions.

Algorithm 6.2: Conjugate gradients on a tangent space

Input: positive definite operator H on $\mathrm{T}_x\mathcal{M}$ and $b \in \mathrm{T}_x\mathcal{M}$

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 $Hs = b$

$\beta_n = \frac{\|r_n\|_x^2}{\|r_{n-1}\|_x^2}$

$p_n = r_n + \beta_n p_{n-1}$

To further understand how good the intermediate solutions $v_1, v_2, \ldots$ are, it is instructive to determine in which subspace they live. From Algorithm 6.2, we find that v_1 is a multiple of p_0, while v_2 is a linear combination of v_1 and p_1, hence it is a linear combination of p_0 and p_1. Likewise, v_3 is a linear combination of v_2 and p_2, hence it is a linear combination of p_0, p_1 and p_2. Iterating this reasoning, we find that, for each n,

$$v_n \in \mathrm{span}(p_0, \ldots, p_{n-1}).$$

We write $\mathrm{span}(u_1, \ldots, u_m)$ to denote the subspace spanned by vectors $u_1, \ldots, u_m$. Likewise, $\mathrm{span}(X)$ denotes the subspace spanned by the columns of a matrix X.

Similar scrutiny applied to the H-conjugate directions shows

$$\operatorname{span}(p_0, \ldots, p_{n-1}) \subseteq \operatorname{span}(r_0, \ldots, r_{n-1}).$$

Once again, this time reasoning with the *residuals* $r_0, \ldots, r_{n-1}$, we find

It is an exercise to show $r_n = b - Hv_n$; this is why it is called the *residual*.

$$\operatorname{span}(r_0, \ldots, r_{n-1}) \subseteq \operatorname{span}\left(b, Hb, H^2 b, \ldots, H^{n-1} b\right).$$

All these subspaces are, in fact, equal, as otherwise the algorithm would terminate. It follows that v_n is the minimizer of $\|v - s\|_H$ in the subspace

$$\operatorname{span}\left(b, Hb, H^2 b, \ldots, H^{n-1} b\right),$$

called the *Krylov subspace*.

From this last statement, we deduce that v_n can be expanded as

$$v_n = a_0 b + a_1 Hb + a_2 H^2 b + \cdots + a_{n-1} H^{n-1} b,$$

for some coefficients $a_0, \ldots, a_{n-1} \in \mathbb{R}$. Furthermore, we are guaranteed that these coefficients are such that $\|v_n - s\|_H$ is minimized. Stated somewhat differently,

$$v_n = \left(a_0 I + a_1 H + a_2 H^2 + \cdots + a_{n-1} H^{n-1}\right) b = p_{n-1}(H) b,$$

for some polynomial of degree at most $n - 1$,

$$p_{n-1}(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_{n-1} z^{n-1},$$

such that $\|p_{n-1}(H)b - s\|_H$ is minimized. Let $\mathcal{P}_{n-1}$ denote the set of polynomials of degree at most $n - 1$. Since $b = Hs$, the error vector obeys

$$v_n - s = p_{n-1}(H)Hs - s = q_n(H)s,$$

where $q_n(z) = p_{n-1}(z)z - 1$ is a polynomial of degree at most n such that $q_n(0) = 1$. Let $\mathcal{Q}_n$ denote the set of such polynomials. Since any polynomial $q \in \mathcal{Q}_n$ can be written in the form $q(z) = p(z)z - 1$ for some $p \in \mathcal{P}_n$, it follows that the CG method guarantees

$$\|v_n - s\|_H = \min_{p \in \mathcal{P}_{n-1}} \|(p(H)H - I)s\|_H = \min_{q \in \mathcal{Q}_n} \|q(H)s\|_H. \tag{6.5}$$

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 operator 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. Furthermore, let $\lambda_1, \ldots, \lambda_d$

be associated eigenvalues. The unknown vector s can be expanded in the eigenbasis 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 a polynomial $q(H)$ to s 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,

$$\frac{\|q(H)s\|_H^2}{\|s\|_H^2} = \frac{\langle q(H)s, Hq(H)s \rangle_x}{\langle s, Hs \rangle_x} = \frac{\sum_{i=1}^{d} q(\lambda_i)^2 \lambda_i \langle u_i, s \rangle_x^2}{\sum_{i=1}^{d} \lambda_i \langle u_i, s \rangle_x^2} \leq \max_{i=1,\ldots,d} q(\lambda_i)^2,$$

where the inequality is due to positivity of the eigenvalues. Combining with (6.5), it follows that

$$\|v_n - s\|_H \leq \|s\|_H \cdot \min_{q \in \mathcal{Q}_n} \max_{i=1,\ldots,d} |q(\lambda_i)|. \tag{6.6}$$

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 any of the eigenvalues of H.

Based on these considerations, it follows easily that CG terminates in n iterations if H has only n distinct eigenvalues. Indeed, it suffices to consider a polynomial q of degree n 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

$$\|v_n - s\|_H \leq \|s\|_H \cdot 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^n, \tag{6.7}$$

Applying Definition 4.12, this shows at least linear convergence, with rate $\log_{10}\left(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\right)$. See also Exercise 6.5.

so that the error decreases exponentially fast as CG iterates. This is done 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.

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 preempt this (specifically, because numerically the vectors p_i are not exactly H-conjugate). However, the progressive improvement of the

iterates v_n as predicted by (6.7) is borne out empirically, and the role of the condition number κ is indeed critical. In practice, CG is terminated after a maximum number of iterations, or when a certain relative tolerance is met. For example, we may replace the stopping criterion $r_n = 0$ with $\|r_n\|_x \leq \varepsilon_{\text{tolerance}}\|b\|_x$.

Recall $r_n = b - Hv_n$.

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 $\langle p_{n-1}, Hp_{n-1}\rangle_x$ 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 Hp_{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 6.5. *An alternative to CG is to run gradient descent on $g(v)$ (6.3) 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 $\kappa = \frac{\lambda_{\max}(H)}{\lambda_{\min}(H)}$. In contrast, CG guarantees $\|v_n - s\|_H \leq 2e^{-n/\sqrt{\kappa}}\|s\|_H$.*

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 operator 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 5.38, 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. Furthermore, 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

$$m_k(s) = f(x_k) + \langle \mathrm{grad} f(x_k), s \rangle_{x_k} + \frac{1}{2} \langle H_k(s), s \rangle_{x_k} \tag{6.8}$$

for some operator $H_k \colon \mathrm{T}_{x_k}\mathcal{M} \to \mathrm{T}_{x_k}\mathcal{M}$ 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*:

$$\min_{s \in \mathrm{T}_{x_k}\mathcal{M}} m_k(s) \quad \text{subject to} \quad \|s\|_{x_k} \leq \Delta_k, \tag{6.9}$$

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).

Running RTR, we expect to generate a point $x = x_k$ such that

$$\|\mathrm{grad} f(x)\|_x \leq \varepsilon_g \qquad \text{and} \qquad \mathrm{Hess} f(x) \succeq -\varepsilon_H \,\mathrm{Id}, \tag{6.13}$$

where Id is the identity operator 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, we must first 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.

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 (6.9). In choosing H_k, we aim to strike a balance

Algorithm 6.3: Riemannian trust-region method (RTR)

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 (6.8).

Approximately solve the subproblem (6.9), yielding s_k.

The tentative next iterate is $x_k^+ = \mathrm{R}_{x_k}(s_k)$.

Compute the ratio of actual to model improvement:

$$\rho_k = \frac{f(x_k) - f(x_k^+)}{m_k(0) - m_k(s_k)}. \tag{6.10}$$

Accept or reject the tentative next iterate:

$$x_{k+1} = \begin{cases} x_k^+ & \text{if } \rho_k > \rho' \text{ (accept)}, \\ x_k & \text{otherwise (reject)}. \end{cases} \tag{6.11}$$

Update the trust-region radius:

$$\Delta_{k+1} = \begin{cases} \frac{1}{4}\Delta_k & \text{if } \rho_k < \frac{1}{4}, \\ \min(2\Delta_k, \bar{\Delta}) & \text{if } \rho_k > \frac{3}{4} \text{ and } \|s_k\|_{x_k} = \Delta_k, \\ \Delta_k & \text{otherwise.} \end{cases} \tag{6.12}$$

between model accuracy, computational efficiency, and convenience. With the goal (6.13) 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.

A4. *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:*

$$\forall s \in \mathrm{T}_{x_k}\mathcal{M}, \alpha \geq 0, \qquad H_k(\alpha s) = \alpha H_k(s); \qquad \textit{and} \tag{6.14}$$

2. *H_k is uniformly bounded, that is, there exists $c_0 \geq 0$, independent of k, such that*

$$\forall s \in \mathrm{T}_{x_k}\mathcal{M}, \qquad \left| \langle s, H_k(s) \rangle_{x_k} \right| \leq c_0 \|s\|_{x_k}^2. \tag{6.15}$$

See Corollary 10.41.

An extreme case consists in selecting $H_k = L \cdot \mathrm{Id}$ for some $L > 0$. This is convenient, computationally inexpensive, and allows to solve the subproblem (6.9) 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.

A5. *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 exists $c_1 \geq 0$, independent of k, such that*

$$\|\mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0) - H_k\| \leq \frac{c_1 \Delta_k}{3}, \tag{6.16}$$

where $\|\cdot\|$ denotes the operator norm.

The *operator norm* of a self-adjoint operator 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 5.38 and 5.39 (and Exercise 10.67 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:

$$H_k = \operatorname{Hess}(f \circ \mathrm{R}_{x_k})(0) = \operatorname{Hess} f(x_k). \tag{6.17}$$

Then, A5 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 (6.9) approximately, producing a step s_k. Numerous efficient algorithms have been proposed over the last few decades: we discuss one that is particularly well suited to the Riemannian case in detail 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.

A6. *There exists $c_2 > 0$ such that, for all k with $\|\operatorname{grad} f(x_k)\|_{x_k} > \varepsilon_g$, the step s_k satisfies*

$$m_k(0) - m_k(s_k) \geq c_2 \min\left(\Delta_k, \frac{\varepsilon_g}{c_0}\right) \varepsilon_g, \tag{6.18}$$

where c_0 is the constant in A4.

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.

Exercise 6.6. *The Cauchy point is of the form $s_k^C = -t\operatorname{grad} f(x_k)$ for some $t \geq 0$. Writing $g_k = \operatorname{grad} f(x_k)$ for convenience, under A4, show that*

$$t = \begin{cases} \min\left(\frac{\|g_k\|_{x_k}^2}{\langle g_k, H_k(g_k)\rangle_{x_k}}, \frac{\Delta_k}{\|g_k\|_{x_k}}\right) & \textit{if } \langle g_k, H_k(g_k)\rangle_{x_k} > 0, \\ \frac{\Delta_k}{\|g_k\|_{x_k}} & \textit{otherwise.} \end{cases}$$

Furthermore, show that setting $s_k = s_k^C$ satisfies A6 with $c_2 = \frac{1}{2}$.

Once the gradient is small, if $\varepsilon_H < \infty$, it becomes necessary to focus on second-order optimality conditions.

A7. *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*

$$m_k(0) - m_k(s_k) \geq c_3 \Delta_k^2 \varepsilon_H. \tag{6.19}$$

(H_k has real eigenvalues owing to A5.)

This condition too can be satisfied with explicit, finite procedures, by computing *eigensteps*: moving up to the boundary along a direction which certifies that the smallest eigenvalue of H_k is strictly smaller than $-\varepsilon_H$.

Exercise 6.7. *Under A5, if $\lambda_{\min}(H_k) < -\varepsilon_H$, verify that there exists a tangent vector $u \in \mathrm{T}_{x_k}\mathcal{M}$ with*

$$\|u\|_{x_k} = 1, \quad \langle u, \mathrm{grad} f(x_k)\rangle_{x_k} \leq 0, \quad \text{and} \quad \langle u, H_k(u)\rangle_{x_k} < -\varepsilon_H.$$

Show that $s_k = \Delta_k u$ (called an eigenstep) satisfies A7 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 A7.

Corollary 6.8. *If H_k is linear and self-adjoint for every iteration k, then setting s_k to be a global minimizer of the subproblem (6.9) at every iteration satisfies both A6 and A7 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 $c_2 = c_3 = \frac{\alpha}{2}$.*

6.4.3 *Regularity conditions*

As we did when analyzing the Riemannian gradient method, we require that the cost function be lower-bounded.

A8. *There exists $f_{\mathrm{low}} \in \mathbb{R}$ such that $f(x) \geq f_{\mathrm{low}}$ for all $x \in \mathcal{M}$.*

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.

A9. *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 $(x, s) \in S_g$,*

See Corollary 10.48 and Exercises 10.51 and 10.52.

$$f(\mathrm{R}_x(s)) \leq f(x) + \langle \mathrm{grad} f(x), s\rangle_x + \frac{L_g}{2}\|s\|_x^2.$$

In addition to these, we now also include a second-order Lipschitz-type condition. When $\mathcal{M}$ is a linear space and $\mathrm{R}_x(s) = x + s$, this one holds in particular if $\mathrm{Hess} f(x)$ is Lipschitz continuous with constant L_H. The set S_H is specified later on; it is empty if $\varepsilon_H = \infty$.

A10. *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 $(x, s) \in S_H$,*

See Corollary 10.50 and Exercise 10.79.

$$f(\mathrm{R}_x(s)) \leq f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{1}{2} \langle s, \mathrm{Hess}(f \circ \mathrm{R}_x)(0)[s] \rangle_x + \frac{L_H}{6} \|s\|_x^3.$$

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). Furthermore, the root points of elements in S_g and S_H are only 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.

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

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:

$$\|\mathrm{grad} f(x_k)\|_{x_k} \leq \varepsilon_g \qquad \text{and} \qquad \lambda_{\min}(H_k) \geq -\varepsilon_H. \tag{6.20}$$

We stress that ε_H may be set to infinity if only first-order optimality conditions are targeted. We separate the theorem statement in two scenarios accordingly. See the discussion around eq. (6.17) 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 proof is based on three supporting lemmas which we state and prove below. In a nutshell, they show:

1. The trust-region radius cannot become arbitrarily small, essentially because regularity of the cost function ensures the model m_k is sufficiently accurate for small steps, which ultimately 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), but the total amount of cost decrease is bounded by $f(x_0) - f_{\mathrm{low}}$, so that there cannot be arbitrarily many such 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 now state the main theorem.

Theorem 6.9. *Let $S = \{(x_0, s_0), (x_1, s_1), \ldots\}$ be the pairs of iterates and tentative steps generated by RTR under A4, A5, A6, A7 and A8. Furthermore, assume A9 and A10 hold with constants L_g and L_H on the sets*

$$S_g = \left\{(x_k, s_k) \in S : \|\mathrm{grad} f(x_k)\|_{x_k} > \varepsilon_g\right\}, \text{ and}$$
$$S_H = \left\{(x_k, s_k) \in S : \|\mathrm{grad} f(x_k)\|_{x_k} \le \varepsilon_g \text{ and } \lambda_{\min}(H_k) < -\varepsilon_H\right\}.$$

Define

$$\lambda_g = \frac{1}{4}\min\left(\frac{1}{c_0}, \frac{c_2}{L_g + c_0}\right) \qquad \text{and} \qquad \lambda_H = \frac{3}{4}\frac{c_3}{L_H + c_1}. \tag{6.21}$$

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*

$$t \le \frac{3}{2}\frac{f(x_0) - f_{\mathrm{low}}}{\rho' c_2 \lambda_g}\frac{1}{\varepsilon_g^2} + \frac{1}{2}\log_2\left(\frac{\Delta_0}{\lambda_g \varepsilon_g}\right) = O\left(\frac{1}{\varepsilon_g^2}\right). \tag{6.22}$$

(In this scenario, A5, A7 and A10 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*

$$t' \le \frac{3}{2}\frac{f(x_0) - f_{\mathrm{low}}}{\rho' c_3 \lambda^2}\frac{1}{\varepsilon^2 \varepsilon_H} + \frac{1}{2}\log_2\left(\frac{\Delta_0}{\lambda \varepsilon}\right) = O\left(\frac{1}{\varepsilon^2 \varepsilon_H}\right), \tag{6.23}$$

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, $f(x_{t'}) \le f(x_t) \le f(x_0)$.

This first lemma lower-bounds the trust-region radius.

Lemma 6.10. *Under the assumptions of Theorem 6.9, if $x_0, \ldots, x_n$ are iterates generated by RTR such that none of them satisfy the termination conditions (6.20), then*

$$\Delta_k \ge \min(\Delta_0, \lambda_g \varepsilon_g, \lambda_H \varepsilon_H) \tag{6.24}$$

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. (6.12)), this guarantees Δ_k cannot decrease further. By definition of ρ_k (6.10), 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 A9 and A4:

$$
\begin{aligned}
&f(\mathrm{R}_{x_k}(s_k)) - m_k(s_k) \\
&\qquad = f(\mathrm{R}_{x_k}(s_k)) - f(x_k) - \langle \mathrm{grad} f(x_k), s_k \rangle_{x_k} - \frac{1}{2} \langle H_k(s_k), s_k \rangle_{x_k} \\
&\qquad \leq \frac{L_g + c_0}{2} \|s_k\|_{x_k}^2.
\end{aligned}
$$

Furthermore, the denominator is lower-bounded by A6:

$$
m_k(0) - m_k(s_k) \geq c_2 \min\left(\Delta_k, \frac{\varepsilon_g}{c_0}\right) \varepsilon_g.
$$

Hence, using $\|s_k\|_{x_k} \leq \Delta_k$,

$$
1 - \rho_k \leq \frac{1}{2} \frac{L_g + c_0}{c_2 \varepsilon_g} \frac{\Delta_k^2}{\min\left(\Delta_k, \frac{\varepsilon_g}{c_0}\right)}.
$$

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 (6.21), 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 A10, A5 and $\|s_k\|_{x_k} \leq \Delta_k$:

$$
\begin{aligned}
&f(\mathrm{R}_{x_k}(s_k)) - m_k(s_k) \\
&\qquad = f(\mathrm{R}_{x_k}(s_k)) - f(x_k) - \langle \mathrm{grad} f(x_k), s_k \rangle_{x_k} - \frac{1}{2} \left\langle \mathrm{Hess}(f \circ \mathrm{R}_{x_k})(0)[s_k], s_k \right\rangle_{x_k} \\
&\qquad\qquad + \frac{1}{2} \left\langle (\mathrm{Hess}(f \circ \mathrm{R}_{x_k}(0) - H_k)[s_k], s_k \right\rangle_{x_k} \\
&\qquad \leq \frac{L_H + c_1}{6} \Delta_k^3,
\end{aligned}
$$

and the denominator is lower-bounded by A7:

$$
m_k(0) - m_k(s_k) \geq c_3 \Delta_k^2 \varepsilon_H. \tag{6.25}
$$

Combining, we get

$$
1 - \rho_k \leq \frac{L_H + c_1}{6 c_3 \varepsilon_H} \Delta_k.
$$

Again, considering (6.21), 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 thus 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. $\square$

This second lemma upper-bounds the total number of successful (that is, accepted) steps before termination conditions are met.

Lemma 6.11. *Under the assumptions of Theorem 6.9, if $x_0, \ldots, x_n$ are iterates generated by RTR such that none of them satisfy the termination conditions (6.20), define the set of* successful steps *among those as*

$$S_n = \{k \in \{0, \ldots, n\} : \rho_k > \rho'\},$$

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 6.9, the number of successful steps is bounded as*

$$|S_n| \leq \frac{f(x_0) - f_{\text{low}}}{\rho' c_2} \frac{1}{\lambda_g \varepsilon_g^2}. \tag{6.26}$$

Similarly, in the second scenario we have

$$|S_n| \leq \frac{f(x_0) - f_{\text{low}}}{\rho' c_3} \frac{1}{\min(\lambda_g \varepsilon_g, \lambda_H \varepsilon_H)^2 \varepsilon_H}. \tag{6.27}$$

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 (6.10) combined with A6 and A7 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 6.10 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,

$$f(x_k) - f(x_{k+1}) \geq \rho' \min\left(c_2 \lambda_g \varepsilon_g^2, c_2 \lambda_H \varepsilon_g \varepsilon_H, c_3 \lambda_g^2 \varepsilon_g^2 \varepsilon_H, c_3 \lambda_H^2 \varepsilon_H^3\right). \tag{6.28}$$

In the first scenario, $\varepsilon_H = \infty$ and this 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 A1 (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 (6.28) 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

$$f(x_0) - f_{\text{low}} \geq |S_n|\rho' c_3 \min(\lambda_g\varepsilon_g, \lambda_H\varepsilon_H)^2\varepsilon_H.$$

Solve for $|S_n|$ to conclude. $\square$

This third and last lemma lower-bounds the number of successful steps before termination as a fraction of the total number of iterations before termination.

Lemma 6.12. *Under the assumptions of Theorem 6.9, if $x_0, \ldots, x_n$ are iterates generated by RTR such that none of them satisfy the termination conditions (6.20), using the notation S_n and U_n of Lemma 6.11, it holds that*

$$|S_n| \geq \frac{2}{3}(n+1) - \frac{1}{3}\max\left(0, \log_2\left(\frac{\Delta_0}{\lambda_g\varepsilon_g}\right), \log_2\left(\frac{\Delta_0}{\lambda_H\varepsilon_H}\right)\right). \quad (6.29)$$

Proof. The proof rests on the lower bound for Δ_k from Lemma 6.10. 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 6.10 gives

$$\Delta_n \geq \min\left(\Delta_0, \lambda_g\varepsilon_g, \lambda_H\varepsilon_H\right).$$

Combine, divide by Δ_0 and take the log in base 2:

$$|S_n| - 2|U_n| \geq \min\left(0, \log_2\left(\frac{\lambda_g\varepsilon_g}{\Delta_0}\right), \log_2\left(\frac{\lambda_H\varepsilon_H}{\Delta_0}\right)\right).$$

Use $|S_n| + |U_n| = n + 1$ to conclude. $\square$

With these lemmas available, the main theorem follows easily.

Proof of Theorem 6.9. For each scenario, Lemmas 6.11 and 6.12 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

$$n \leq \frac{3}{2}\frac{f(x_0) - f_{\text{low}}}{\rho' c_2}\frac{1}{\lambda_g\varepsilon_g^2} + \frac{1}{2}\log_2\left(\frac{\Delta_0}{\lambda_g\varepsilon_g}\right) - 1.$$

Thus, by contraposition, after a number of iterations larger than the right hand side, an iterate with sufficiently small gradient must have been produced. The same argument applies in the second scenario. $\square$

6.4.5 *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 compact; 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 (6.10) 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 standard [IEE08]. 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,[2] in computing the numerator for ρ_k we truly compute

[2] By Sterbenz's theorem, this is true if f_1, f_2 are within a factor 2 of each other.

$$f_1 - f_2 = f(x_k) - f(x_k^+) + \varepsilon_1 f(x_k) - \varepsilon_2 f(x_k^+).$$

Unfortunately, the relative error can be catastrophic. Indeed,

$$\frac{|(f_1 - f_2) - (f(x_k) - f(x_k^+))|}{|f(x_k) - f(x_k^+)|} \leq \varepsilon_{\mathrm{M}} \frac{|f(x_k)| + |f(x_k^+)|}{|f(x_k) - f(x_k^+)|}.$$

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,

$$m_k(0) - m_k(s_k) = -\langle s_k, \operatorname{grad} f(x_k)\rangle_{x_k} - \frac{1}{2}\langle s_k, H_k(s_k)\rangle_{x_k}. \tag{6.30}$$

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.

A standard fix [CGT00, §17.4.2] to these numerical issues is to regularize the computation of ρ_k, as

$$\rho_k = \frac{f(x_k) - f(x_k^+) + \delta_k}{-\langle s_k, \operatorname{grad} f(x_k)\rangle_{x_k} - \frac{1}{2}\langle s_k, H_k(s_k)\rangle_{x_k} + \delta_k}, \tag{6.31}$$

with

$$\delta_k = \max(1, |f(x_k)|)\varepsilon_{\mathrm{M}}\rho_{\mathrm{reg}}. \tag{6.32}$$

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 (6.31) toward 1, which leads to step acceptance as expected. This is, of course, merely a heuristic to (try to) address an inescapable limitation of inexact arithmetic.

6. Care should be put in implementations to minimize the number of calls to the operator H_k. For example, in the subproblem solver described in Section 6.5 below, exactly one call to this operator is needed per iteration, and furthermore the vector $H_k(s_k)$ is a by-product of that algorithm (when H_k is linear), so that one can avoid an explicit computation of it to evaluate the denominator of ρ_k.

6.5 *The trust-region subproblem: truncated CG*

The trust-region subproblem (6.9) consists in approximately solving a problem of the form

$$\min_{s \in \mathrm{T}_x\mathcal{M}} m(s) = \frac{1}{2}\langle s, Hs\rangle_x - \langle b, s\rangle_x \text{ subject to } \|s\|_x \leq \Delta, \tag{6.33}$$

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 $H = H_k$, $b = -\operatorname{grad} f(x_k)$ and $\Delta = \Delta_k$.

We consider the important particular case where H is a linear, self-adjoint operator (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 equations (6.3) and (6.33).

The general idea of the *truncated CG* method (tCG), Algorithm 6.4, is to run CG on the cost function (6.33) while

1. Keeping an eye out for signs that H is not 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.

If the inner products $\langle p_i, Hp_i \rangle_x$ are positive for $i = 1, \ldots, n-2$, then, since $p_1, \ldots, p_{n-2}$ are linearly independent, they form a basis for a subspace of $\mathrm{T}_x\mathcal{M}$, and H is positive definite on that subspace. Thus, up to that point, all the properties of CG hold. If, however, upon computing 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

$$\begin{aligned} &\|v_{n-1} + tp_{n-1}\|_x^2 - \Delta^2 \\ &\qquad = \|p_{n-1}\|_x^2 t^2 + 2t \langle v_{n-1}, p_{n-1} \rangle_x + \|v_{n-1}\|_x^2 - \Delta^2. \end{aligned} \tag{6.34}$$

The product of these roots is negative since $\|v_{n-1}\|_x < \Delta$, 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 respects the trust-region constraint. This corresponds exactly to the positive root of the quadratic in eq. (6.34). 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 curvature is 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 allows tCG to terminate if

$$\|r_n\|_x \leq \|\mathrm{grad} f(x_k)\|_{x_k} \cdot \min(\|\mathrm{grad} f(x_k)\|_{x_k}^{\theta}, \kappa). \tag{6.35}$$

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 [ABG07, Thm. 4.13]. Intuitively, superlinear convergence occurs because, close to a critical point x with positive definite Hessian, 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 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.

We note in passing that tCG can be *preconditioned*, just like regular CG [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 6.13. *Show that v_1 is the Cauchy point. Since iterates monotonically improve $m(v_n)$, this implies tCG guarantees A6 with $c_2 = \frac{1}{2}$.*

Exercise 6.14. *Show that if $\mathrm{grad} f(x) = 0$, then tCG returns immediately, even if H_k has negative eigenvalues. This shows that tCG, as set up here, does* not *guarantee A7. See [CGT00, §7.5.4] for a fix based on Lanczos iterations.*

Exercise 6.15. *In Algorithm 6.4, show that Hs can be computed as indicated.*

6.6 *Local convergence*

Under suitable assumptions, once iterates of RTR are close enough to a critical point x where $\mathrm{Hess} f(x)$ is positive definite, the algorithm converges superlinearly provided the subproblem is solved with sufficient accuracy; tCG provides the necessary assurances. See [ABG07] or [AMS08, §7].

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 $\|s\|_x \leq \Delta$, and Hs as a by-product

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)

  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 pstve root of a quadratic, $\|v_{n-1} + tp_{n-1}\|_x^2 = \Delta^2$.)

    output $s = v_n$ and $Hs = b - r_{n-1} + tHp_{n-1}$

  Else

    $v_n = v_{n-1}^+$

  $r_n = r_{n-1} - \alpha_n Hp_{n-1}$

  If $\|r_n\|_x \leq \|r_0\|_x \min(\|r_0\|_x^\theta, \kappa)$

    output $s = v_n$ and $Hs = b - r_n$

  $\beta_n = \frac{\|r_n\|_x^2}{\|r_{n-1}\|_x^2}$

  $p_n = r_n + \beta_n p_{n-1}$

6.7 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 Manopt, these tests are 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 operator.

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 (5.27): if R is a second-order retraction, or if x is a critical point, then

$$f(\mathrm{R}_x(tv)) = f(x) + t \langle \mathrm{grad} f(x), v \rangle_x + \frac{t^2}{2} \langle \mathrm{Hess} f(x)[v], v \rangle_x + O(t^3). \tag{6.36}$$

This says that, under the stated conditions,

$$E(t) \triangleq \left| f(\mathrm{R}_x(tv)) - f(x) - t \langle \mathrm{grad} f(x), v \rangle_x - \frac{t^2}{2} \langle \mathrm{Hess} f(x)[v], v \rangle_x \right| = O(t^3).$$

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 $f(x)$, $\langle \mathrm{grad} f(x), v\rangle_x$ and $\langle \mathrm{Hess} f(x)[v], v\rangle_x$;
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.

6.8 *Notes and references*

First- and second-order necessary optimality conditions are further studied in [YZS14, BH18], 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]. The reference material for the discussion of conjugate gradients in Section 6.3 is [TB97, Lect. 38].

Trust-region methods in Euclidean space are discussed in great detail by Conn et al.[3]; see also [NW06] for a shorter treatment. Absil et al.[4] 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], with certain parts appearing verbatim in that reference (in particular, the proofs of Lemmas 6.11 and 6.12). 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.

[3] 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

[4] 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

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 [YP06]. Cartis et al. give a thorough treatment including complexity bounds [CGT11b, CGT11a]. Qi proposed a first extension of ARC to Riemannian manifolds [Qi11]. An iteration complexity analysis akin to the one we give here for RTR appears in [ZZ18, ABBC20]. 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, and we derive the relevant geometric tools that are useful to optimize over them.

We lead with first-order tools as introduced in Chapter 3, and we follow up with second-order tools as described in Chapter 5. This whole chapter is meant to be consulted periodically for illustration, while reading those chapters. Pointers to implementations in the toolbox Manopt are also included. Similar implementations are or will become available in PyManopt and Manopt.jl as well.

manopt.org [BMAS14]
pymanopt.org [TKW16]
manoptjl.org [Ber19]

Remember from Chapter 3 that products of embedded submanifolds are embedded submanifolds, and that the geometric toolbox of a product is easily obtained from the geometric toolboxes of its parts. In particular, given a manifold $\mathcal{M}$, it is straightforward to work with tuples of k elements on $\mathcal{M}$, that is, with $\mathcal{M}^k = \mathcal{M} \times \cdots \times \mathcal{M}$.

In Manopt, see `productmanifold` and `powermanifold`.

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 $x \in \mathcal{E}$,

In Manopt, see `euclideanfactory` and `euclideancomplexfactory`.

$$\mathrm{T}_x\mathcal{E} = \mathcal{E}. \tag{7.1}$$

An obvious (and reasonable) choice of retraction is

$$\mathrm{R}_x(v) = x + v, \tag{7.2}$$

though Definition 3.40 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 course the identity map:

$$\mathrm{Proj}_x(u) = u. \tag{7.3}$$

Smoothness of a function $f\colon \mathcal{E} \to \mathbb{R}$ 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). If we make $\mathcal{M}$ into a 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 3.51 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 $\mathcal{E}$,

In Manopt, see `euclideansubspacefactory`, and special cases `symmetricfactory`, `skewsymmetricfactory`, ...

$$\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad}\bar{f}(x)). \tag{7.4}$$

For example, with the usual inner product on $\mathcal{E} = \mathbb{R}^{n\times n}$ (3.15), with $\mathcal{M} = \mathrm{Sym}(n)$ as a Riemannian submanifold, $\mathrm{Proj}_X(Z) = \frac{Z+Z^\top}{2}$ so that the gradient of a function on $\mathrm{Sym}(n)$ is simply the symmetric part of its classical gradient on all of $\mathbb{R}^{n\times n}$.

Second-order tools

Covariant derivatives (∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$) on $\mathcal{E}$ 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 5.36). In fact, it is the exponential map (Section 10.2).

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 $\mathbb{R}^n$ 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

In Manopt, see `spherefactory` and `spherecomplexfactory`.

$$\mathrm{S}^{d-1} = \{x \in \mathcal{E} : \|x\| = 1\}. \tag{7.5}$$

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

$$\mathrm{T}_x \mathrm{S}^{d-1} = \{v \in \mathcal{E} : \langle x, v \rangle = 0\}, \tag{7.6}$$

and $\dim \mathrm{S}^{d-1} = \dim \mathcal{E} - 1 = d - 1$. One possible retraction is

$$\mathrm{R}_x(v) = \frac{x+v}{\|x+v\|} = \frac{x+v}{\sqrt{1+\|v\|^2}}. \tag{7.7}$$

The orthogonal projector to the tangent space at x is

$$\mathrm{Proj}_x \colon \mathcal{E} \to \mathrm{T}_x\mathrm{S}^{d-1} \colon u \mapsto \mathrm{Proj}_x(u) = u - \langle x, u \rangle x. \tag{7.8}$$

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 $\bar{f} \colon U \to \mathbb{R}$ in a neighborhood U of S^{d-1} in $\mathcal{E}$, the gradient of f is given by Proposition 3.51 as

In Manopt, formulas such as (7.9) which convert the Euclidean gradient of a smooth extension into a Riemannian gradient are available for each manifold as `egrad2rgrad`.

$$\mathrm{grad} f(x) = \mathrm{Proj}_x(\mathrm{grad} \bar{f}(x)) = \mathrm{grad} \bar{f}(x) - \left\langle x, \mathrm{grad} \bar{f}(x) \right\rangle x. \tag{7.9}$$

In particular, x is a critical point of f if and only if $\mathrm{grad}\bar{f}(x)$ is parallel to x.

A product of k spheres is called an *oblique manifold*. For example, the product of k spheres in $\mathbb{R}^n$ is denoted by $\mathrm{OB}(n, k) = (\mathrm{S}^{n-1})^k$. Its elements are typically represented using matrices in $\mathbb{R}^{n \times k}$ (or $\mathbb{R}^{k \times n}$) 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: this is nothing but an alternative way of representing S^1.

In Manopt, see: `obliquefactory`, `obliquecomplexfactory`, `complexcirclefactory`.

Second-order tools

Since S^{n-1} is a Riemannian submanifold of the Euclidean space $\mathcal{E}$, covariant derivatives (∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$) on S^{n-1} coincide with the usual vector field derivatives (of smooth extensions) in $\mathcal{E}$, followed by orthogonal projection to tangent spaces (Theorem 5.7, Proposition 5.27).

We can use this to obtain a formula for the Riemannian Hessian of $f \colon \mathrm{S}^{n-1} \to \mathbb{R}$, with smooth extension $\bar{f} \colon U \to \mathbb{R}$ defined on a neighborhood U of S^{n-1} in $\mathcal{E}$. Following Example 5.13, 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^{n-1} in $\mathcal{E}$. Then,

$$
\begin{aligned}
\operatorname{Hess} f(x)[v] &= \nabla_v \operatorname{grad} f \\
&= \operatorname{Proj}_x(\mathrm{D}\bar{G}(x)[v]) \\
&= \operatorname{Proj}_x\Big(\operatorname{Hess}\bar{f}(x)[v] - \left[\langle v, \operatorname{grad}\bar{f}(x)\rangle + \langle x, \operatorname{Hess}\bar{f}(x)[v]\rangle\right] x \\
&\qquad\qquad - \langle x, \operatorname{grad}\bar{f}(x)\rangle\, v\Big) \\
&= \operatorname{Proj}_x(\operatorname{Hess}\bar{f}(x)[v]) - \langle x, \operatorname{grad}\bar{f}(x)\rangle\, v. \qquad (7.10)
\end{aligned}
$$

In Manopt, formulas such as (7.10) which convert the Euclidean gradient and Hessian of a smooth extension into a Riemannian Hessian are available for each manifold as `ehess2rhess`.

The retraction (7.7) is a second-order retraction (see Definition 5.36, Example 5.37 and Proposition 5.43). Geodesics on S^{n-1} are given in Example 5.32.

7.3 *The Stiefel manifold: orthonormal matrices*

For $p \leq n$, let $\mathbb{R}^{n\times p}$ be endowed with the standard inner product $\langle U, V\rangle = \operatorname{Tr}(U^\top V)$. The (compact) *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:

$$\operatorname{St}(n,p) = \left\{X \in \mathbb{R}^{n\times p} : X^\top X = I_p\right\}, \qquad (7.11)$$

See `stiefelfactory` in Manopt; for complex matrices: `stiefelcomplexfactory`.

The *non-compact* Stiefel manifold refers to the open subset of matrices of rank p in $\mathbb{R}^{n\times p}$. By default, we mean compact.

Some authors use the notation $\operatorname{St}(p,n)$ for the same set—we prefer the present notation as it is reminiscent of the size of the matrices.

where I_p is the identity matrix of size p. In particular, $\operatorname{St}(n,1)$ is the unit sphere in $\mathbb{R}^n$. We call matrices in $\operatorname{St}(n,p)$ *orthonormal matrices* and we reserve the word *orthogonal matrix* for square orthonormal matrices.

Consider the following function:

$$h\colon \mathbb{R}^{n\times p} \to \operatorname{Sym}(p)\colon X \mapsto h(X) = X^\top X - I_p, \qquad (7.12)$$

where $\operatorname{Sym}(p)$ 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 claim that h is a defining function for $\operatorname{St}(n,p)$. Indeed, h is smooth and $h^{-1}(0) = \operatorname{St}(n,p)$: it remains to check that the differential of h has rank k for all $X \in \operatorname{St}(n,p)$. To this end, consider $\mathrm{D}h(X)\colon \mathbb{R}^{n\times p} \to \operatorname{Sym}(p)$:

$$
\begin{aligned}
\mathrm{D}h(X)[V] &= \lim_{t\to 0} \frac{h(X+tV) - h(X)}{t} \\
&= \lim_{t\to 0} \frac{(X+tV)^\top(X+tV) - X^\top X}{t} \\
&= X^\top V + V^\top X. \qquad (7.13)
\end{aligned}
$$

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 $\operatorname{Sym}(p)$ has dimension k, we must show that the image of $\mathrm{D}h(X)$ is all of $\operatorname{Sym}(p)$, that is,

$\mathrm{D}h(X)$ is *surjective*. To this end, 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

$$\dim \mathrm{St}(n,p) = \dim \mathbb{R}^{n\times p} - \dim \mathrm{Sym}(p) = np - \frac{p(p+1)}{2}. \tag{7.14}$$

The tangent spaces are subspaces of $\mathbb{R}^{n\times p}$:

$$\mathrm{T}_X\mathrm{St}(n,p) = \ker \mathrm{D}h(X) = \left\{V \in \mathbb{R}^{n\times p} : X^\top V + V^\top X = 0\right\}. \tag{7.15}$$

It is sometimes convenient to parameterize tangent vectors in explicit form. First, complete 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:

This matrix $X_\perp$ is never created explicitly in practice: it is merely a useful mathematical construct.

$$X^\top X = I_p, \qquad X_\perp^\top X_\perp = I_{n-p}, \qquad \text{and} \qquad X^\top X_\perp = 0. \tag{7.16}$$

Since $[\, X \; X_\perp \,]$ is, in particular, invertible, any matrix $V \in \mathbb{R}^{n\times p}$ can be written as

$$V = \begin{bmatrix} X & X_\perp \end{bmatrix} \begin{bmatrix} \Omega \\ B \end{bmatrix} = X\Omega + X_\perp B, \tag{7.17}$$

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,

$$\mathrm{T}_X\mathrm{St}(n,p) = \left\{X\Omega + X_\perp B : \Omega \in \mathrm{Skew}(p), B \in \mathbb{R}^{(n-p)\times p}\right\}, \tag{7.18}$$

where we used the decomposition (7.17) with respect to an arbitrary choice of $X_\perp \in \mathbb{R}^{n\times(n-p)}$ satisfying (7.16), and

$$\mathrm{Skew}(p) = \{\Omega \in \mathbb{R}^{p\times p} : \Omega^\top = -\Omega\} \tag{7.19}$$

is the set of skew-symmetric matrices of size p.

One popular retraction for $\mathrm{St}(n,p)$ is the *Q-factor retraction*:

$$\mathrm{R}_X(V) = Q, \tag{7.20}$$

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`.

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 $V \in \mathrm{T}_X\mathrm{St}(n, p)$,

$$(X+V)^\top(X+V) = I_p + V^\top V \tag{7.21}$$

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] (see also the erratum for that reference), from which it is straightforward to verify that $\mathrm{DR}_X(0)$ is the identity map.

Another popular retraction for $\mathrm{St}(n, p)$ is the *polar retraction*:

$$\begin{aligned} \mathrm{R}_X(V) &= (X+V)\left((X+V)^\top(X+V)\right)^{-1/2} \\ &= (X+V)(I_p + V^\top V)^{-1/2}, \end{aligned} \tag{7.22}$$

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 $X + V$. Clearly, $\mathrm{R}_X(0) = X$ and R is smooth. It is an exercise to check that $\mathrm{DR}_X(0)$ is the identity map. In fact, the polar retraction is a metric projection retraction: see Proposition 5.42.

If $M = X + V = U\Sigma W^\top$ is a thin singular value decomposition, the polar factor of M is $UW^\top$.

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:

As an alternative to this derivation of the normal space, one can 'guess' the final result (7.23), verify that this subspace is orthogonal to the tangent space, and count dimensions to conclude.

$$\begin{aligned} \mathrm{N}_X\mathrm{St}(n, p) &= (\mathrm{T}_X\mathrm{St}(n, p))^\perp \\ &= \left\{U \in \mathbb{R}^{n\times p} : \langle U, V\rangle = 0 \text{ for all } V \in \mathrm{T}_X\mathrm{St}(n, p)\right\} \\ &= \big\{U \in \mathbb{R}^{n\times p} : \langle U, X\Omega + X_\perp B\rangle = 0 \\ &\qquad\qquad \text{for all } \Omega \in \mathrm{Skew}(p), B \in \mathbb{R}^{(n-p)\times p}\big\}. \end{aligned}$$

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:

$$\begin{aligned}
\mathrm{N}_X\mathrm{St}(n,p) &= \big\{U \in \mathbb{R}^{n\times p} : \langle XA + X_\perp C, X\Omega + X_\perp B\rangle = 0 \\
&\qquad\qquad\qquad \text{for all } \Omega \in \mathrm{Skew}(p), B \in \mathbb{R}^{(n-p)\times p}\big\} \\
&= \big\{U \in \mathbb{R}^{n\times p} : \langle A, \Omega\rangle = 0 \text{ and } \langle C, B\rangle = 0 \\
&\qquad\qquad\qquad \text{for all } \Omega \in \mathrm{Skew}(p), B \in \mathbb{R}^{(n-p)\times p}\big\} \\
&= \{XA : A \in \mathrm{Sym}(p)\}, \qquad (7.23)
\end{aligned}$$

where we used that the orthogonal complement of $\mathrm{Skew}(p)$ in $\mathbb{R}^{p\times p}$ is $\mathrm{Sym}(p)$. Thus, orthogonal projectors obey:

$$U - \mathrm{Proj}_X(U) = XA,$$

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,

$$\begin{aligned}
\mathrm{Proj}_X(U) &= U - X\frac{X^\top U + U^\top X}{2} && (7.24) \\
&= (I - XX^\top)U + X\frac{X^\top U - U^\top X}{2}. && (7.25)
\end{aligned}$$

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 3.51:

$$\mathrm{grad} f(X) = \mathrm{Proj}_X(\mathrm{grad}\bar f(X)) = \mathrm{grad}\bar f(X) - X\,\mathrm{sym}(X^\top \mathrm{grad}\bar f(X)), \qquad (7.26)$$

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

Since $\mathrm{St}(n,p)$ is 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 5.7, Proposition 5.27).

We can 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,

$$\begin{aligned}
\mathrm{Hess}f(X)[V] &= \nabla_V \mathrm{grad}f \\
&= \mathrm{Proj}_X(\mathrm{D}\bar{G}(X)[V]) \\
&= \mathrm{Proj}_X\Big(\mathrm{Hess}\bar{f}(X)[V] - V\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X)) - XS\Big) \\
&= \mathrm{Proj}_X(\mathrm{Hess}\bar{f}(X)[V]) - V\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X)), \qquad (7.27)
\end{aligned}$$

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 (7.22) is a second-order retraction (Definition 5.36, Proposition 5.43), but the Q-factor retraction (7.20) is not. Geodesics on $\mathrm{St}(n,p)$ are given in [AMS08, eq. (5.26)].

Exercise 7.1. *Show that the polar retraction* R *on* $\mathcal{M} = \mathrm{St}(n,p)$ (7.22) *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 $\mathbb{R}^{n\times n}$:

In Manopt, see `rotationsfactory` and `unitaryfactory`.

$$\mathrm{O}(n) = \{X \in \mathbb{R}^{n\times n} : X^\top X = XX^\top = I_n\}. \qquad (7.28)$$

It is a group equipped with matrix multiplication as its group operation. Being a special case of the Stiefel manifold, $\mathrm{O}(n)$ also is 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*. It has dimension

$$\dim \mathrm{O}(n) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}, \qquad (7.29)$$

and tangent spaces given by

$$\mathrm{T}_X\mathrm{O}(n) = \left\{X\Omega \in \mathbb{R}^{n\times n} : \Omega \in \mathrm{Skew}(n)\right\} = X\mathrm{Skew}(n). \qquad (7.30)$$

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 tangent vectors to 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

$$\mathrm{Proj}_X(U) = X\frac{X^\top U - U^\top X}{2} = X\,\mathrm{skew}(X^\top U), \tag{7.31}$$

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 3.51:

$$\mathrm{grad} f(X) = X\,\mathrm{skew}(X^\top \mathrm{grad} \bar{f}(X)). \tag{7.32}$$

Of course, this is equivalent to the corresponding formula (7.26) 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:

$$1 = \det(I_n) = \det(XX^\top) = \det(X)^2.$$

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 $\mathrm{O}(n)$.

Optimization algorithms we discuss do not jump between connected components of a manifold. Hence, when initializing, it is important to start in the appropriate component. Geometrically, orthogonal matrices 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*:

$$\mathrm{SO}(n) = \{X \in \mathrm{O}(n) : \det(X) = +1\}. \tag{7.33}$$

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 this open subset of $\mathbb{R}^{n\times n}$: $\{X \in \mathbb{R}^{n\times n} : \det(X) > 0\}$.

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 (7.32) and

$$\dim \mathrm{SO}(n) = \frac{n(n-1)}{2}, \tag{7.34}$$

$$\mathrm{T}_X\mathrm{SO}(n) = X\mathrm{Skew}(n), \tag{7.35}$$

$$\mathrm{Proj}_X(U) = X\,\mathrm{skew}(X^\top U). \tag{7.36}$$

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

Since $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ are 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 5.7, Proposition 5.27).

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 (7.27) we get:

$$\mathrm{Hess} f(X)[V] = \mathrm{Proj}_X\Big(\mathrm{Hess}\bar{f}(X)[V] - V\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X))\Big). \tag{7.37}$$

Writing $V = X\Omega$ for some $\Omega \in \mathrm{Skew}(n)$, this also reads

$$\mathrm{Hess} f(X)[X\Omega] = X\,\mathrm{skew}\Big(X^\top \mathrm{Hess}\bar{f}(X)[V] - \Omega\,\mathrm{sym}(X^\top \mathrm{grad}\bar{f}(X))\Big),$$

making the skew-symmetric representation of $\mathrm{Hess} f(X)[X\Omega]$ more apparent.

For both $\mathrm{O}(n)$ and $\mathrm{SO}(n)$, the polar retraction (7.22) is a second-order retraction (Definition 5.36, Proposition 5.43), but the Q-factor retraction (7.20) is not. It is an exercise to show that

$$c(t) = X\exp(t\Omega) \tag{7.38}$$

is a geodesic on $\mathrm{O}(n)$ (or $\mathrm{SO}(n)$) such that $c(0) = X$ and $c'(0) = X\Omega$.[1]

[1] This happens because the Riemannian metric is bi-invariant, so that the Lie exponential map and the Riemannian exponential map coincide. It is known that the Lie exponential map is given by the matrix exponential exp.

Exercise 7.2. *Show that* $c(t)$ *as defined by* (7.38) *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* (4.25) *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 7.3. *Work out a geometric toolbox for the* unitary group

$$\mathrm{U}(n) = \{X \in \mathbb{C}^{n\times n} : X^*X = I_n\} \tag{7.39}$$

as a Riemannian submanifold of $\mathbb{C}^{n\times n}$ *with the usual inner product* (3.18).

7.5 *Fixed-rank matrices*

The set of real matrices of size $m \times n$ and rank r,

$$\mathbb{R}_r^{m\times n} = \{X \in \mathbb{R}^{m\times n} : \operatorname{rank}(X) = r\}, \tag{7.40}$$

In Manopt, see: `fixedrankembeddedfactory`.

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 $\mathbb{R}^{m\times n}$.

One popular technique to optimize over matrices of rank up 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).

For an arbitrary $X \in \mathbb{R}_r^{m\times n}$, we now build a local defining function. 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

We cannot use $h(X) = \operatorname{rank}(X) - r$ as a defining function because it is not continuous, let alone smooth.

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$$

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 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 U, and 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

$$h\colon U \to \mathbb{R}^{(m-r)\times(n-r)} : Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \mapsto h(Y) = Y_{22} - Y_{21}Y_{11}^{-1}Y_{12},$$

with the same block-matrix structure as before. By our discussion above, $h^{-1}(0) = \mathbb{R}_r^{m\times n} \cap U$. Furthermore, h is smooth in 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 this identity for the differential of the matrix inverse (recall Example 4.14):

$$\mathrm{D}\Big(M \mapsto M^{-1}\Big)(M)[H] = -M^{-1}HM^{-1}. \tag{7.41}$$

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 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

$$\begin{aligned}\dim \mathbb{R}_r^{m\times n} &= \dim \mathbb{R}^{m\times n} - \dim \mathbb{R}^{(m-r)\times(n-r)} \\ &= mn - (m-r)(n-r) \\ &= r(m+n-r).\end{aligned} \tag{7.42}$$

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. Hence, large matrices of small rank can be encoded with a small number of numbers. To exploit this key feature in numerical algorithms, we must *represent* X appropriately in memory. One convenient choice is as a thin singular value decomposition:

$$X = U\Sigma V^\top, \qquad U \in \mathrm{St}(m,r), \qquad \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}, \tag{7.43}$$
$$V \in \mathrm{St}(n,r),$$

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 (7.43).

Since we know that each tangent space has dimension as in (7.42), 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 (3.20), 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 (7.18), this means 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 all of the following velocities are in the tangent space of $\mathbb{R}_r^{m\times n}$ at X:

$$\begin{aligned} c'(0) &= (U\Omega + U_\perp B)\Sigma V^\top + UAV^\top + U\Sigma(V\Omega' + V_\perp C)^\top \\ &= U(\underbrace{\Omega\Sigma + A - \Sigma\Omega'}_{M})V^\top + \underbrace{U_\perp B\Sigma}_{U_p} V^\top + U(\underbrace{V_\perp C\Sigma^\top}_{V_p})^\top. \end{aligned} \tag{7.44}$$

Since Σ is invertible, we find that any matrix of the form

$$UMV^\top + U_pV^\top + UV_p^\top$$

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

$$r^2 + mr + nr - 2r^2 = r(m+n-r),$$

which coincides with the dimension of $\mathrm{T}_X\mathbb{R}_r^{m\times n}$ by (7.42). Thus, we have found the whole tangent space:

$$\begin{aligned}\mathrm{T}_X\mathbb{R}_r^{m\times n} = \big\{UMV^\top + U_pV^\top + UV_p^\top :\ & \\ & M \in \mathbb{R}^{r\times r}, U_p \in \mathbb{R}^{m\times r}, V_p \in \mathbb{R}^{n\times r}, \text{ and} \\ & U^\top U_p = 0, V^\top V_p = 0\big\}.\end{aligned} \tag{7.45}$$

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:

$$\mathrm{T}_X\mathbb{R}_r^{m\times n} = \left\{ \begin{bmatrix} U & U_\perp \end{bmatrix} \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} \begin{bmatrix} V & V_\perp \end{bmatrix}^\top : A, B, C \text{ are arbitrary} \right\}. \tag{7.46}$$

This reveals the dimension of the tangent space even more explicitly.

To build a retraction for $\mathbb{R}_r^{m\times n}$, we use this common approach: 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, the proposal is:

$$\mathrm{R}_X(H) = \operatorname*{arg\,min}_{Y \in \mathbb{R}_r^{m\times n}} \|X + H - Y\|^2. \tag{7.47}$$

Notice that the standard retraction for the sphere and the polar retraction for the Stiefel manifold are also of that nature. This is called *metric projection*: see the discussion around (5.37) for more.

It is well known that solutions to this optimization problem (when they exist) are given by the singular value decomposition of $X + H$, truncated at rank r. With X and H represented as above, this can be done 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 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,

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.

$$X + H = Q_U \underbrace{R_U \begin{bmatrix} \Sigma + M & I_r \\ I_r & 0 \end{bmatrix} R_V^\top}_{\approx \tilde{U}\tilde{\Sigma}\tilde{V}^\top} Q_V^\top.$$

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 $X + H$:

$$\mathrm{R}_X(H) = (Q_U \tilde{U})\tilde{\Sigma}(Q_V \tilde{V})^\top. \tag{7.48}$$

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 $\begin{bmatrix} \Sigma+M & I_r \\ I_r & 0 \end{bmatrix}$ with $\begin{bmatrix} \Sigma+tM & tI_r \\ tI_r & 0 \end{bmatrix}$. 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 follows from general properties of retractions obtained by projection to the manifold from the ambient space: see Proposition 5.42. See also [AO15] for details on this and several other retractions on $\mathbb{R}_r^{m\times n}$.

Important caveat!

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 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 (7.46), it is clear that the normal space at $X = U\Sigma V^\top$ is given by:

$$\mathrm{N}_X\mathbb{R}_r^{m\times n} = \left\{ U_\perp W V_\perp^\top : W \in \mathbb{R}^{(m-r)\times(n-r)} \right\}. \tag{7.49}$$

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 (7.45),

$$\mathrm{Proj}_X(Z) = UMV^\top + U_p V^\top + UV_p^\top \tag{7.50}$$

for some M, U_p, V_p with $U^\top U_p = V^\top V_p = 0$. Combined, these state

$$Z = UMV^\top + U_p V^\top + UV_p^\top + U_\perp W V_\perp^\top.$$

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:

$$P_U Z P_V = UMV^\top, \quad P_U^\perp Z P_V = U_p V^\top, \quad \text{and} \quad P_U Z P_V^\perp = U V_p^\top.$$

Hence,

$$\begin{aligned} \mathrm{Proj}_X(Z) &= P_U Z P_V + P_U^\perp Z P_V + P_U Z P_V^\perp \\ &= U(U^\top Z V)V^\top + (I_m - UU^\top) Z V V^\top + UU^\top Z (I_n - VV^\top). \end{aligned} \tag{7.51}$$

In the notation of (7.50), this is a tangent vector at X represented by

$$M = U^\top Z V, \quad U_p = ZV - UM, \quad \text{and} \quad V_p = Z^\top U - VM^\top. \tag{7.52}$$

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 3.51 as

$$\mathrm{grad} f(X) = \mathrm{Proj}_X(\mathrm{grad} \bar{f}(X)),$$

to be computed using (7.50) and (7.52). 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.

Second-order tools

Since $\mathbb{R}_r^{m\times n}$ is 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 5.7, Proposition 5.27).

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 (7.51)), 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 $H = UMV^\top + U_p V^\top + UV_p^\top$ 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 (7.44), 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:

$$P_{U(t)} = U(t)U(t)^\top, \text{ and}$$

$$\dot{P}_U \triangleq \left.\frac{\mathrm{d}}{\mathrm{d}t} P_{U(t)}\right|_{t=0} = U(0)U'(0)^\top + U'(0)U(0)^\top = U\Sigma^{-1}U_p^\top + U_p\Sigma^{-1}U^\top.$$

Likewise, define $\dot{P}_V$ through

$$P_{V(t)} = V(t)V(t)^\top, \quad \text{and} \quad \dot{P}_V \triangleq \left.\frac{\mathrm{d}}{\mathrm{d}t} P_{V(t)}\right|_{t=0} = V\Sigma^{-1}V_p^\top + V_p\Sigma^{-1}V^\top.$$

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 (7.50) to claim

$$\mathrm{Hess} f(X)[H] = \mathrm{Proj}_X(\mathrm{D}\bar{G}(X)[H]) = U\tilde{M}V^\top + \tilde{U}_p V^\top + U\tilde{V}_p^\top, \quad (7.53)$$

for matrices $\tilde{M}, \tilde{U}_p, \tilde{V}_p$ given as in (7.52). Explicitly,

$$\begin{aligned}\tilde{M} &= U^\top \mathrm{D}\bar{G}(X)[H]V = U^\top \dot{Z} V, \\ \tilde{U}_p &= \mathrm{D}\bar{G}(X)[H]V - U\tilde{M} = P_U^\perp\left(\dot{Z}V + ZV_p\Sigma^{-1}\right), \\ \tilde{V}_p &= (\mathrm{D}\bar{G}(X)[H])^\top U - V\tilde{M}^\top = P_V^\perp\left(\dot{Z}^\top U + Z^\top U_p\Sigma^{-1}\right), \qquad (7.54)\end{aligned}$$

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.

where $Z = \operatorname{grad}\bar{f}(X)$, $\dot{Z} = \operatorname{Hess}\bar{f}(X)[H]$, $X = U\Sigma V^\top$ and H is represented by the triplet (M, U_p, V_p). We may reorganize this as:

$$\begin{aligned} \operatorname{Hess} f(X)[H] &= \operatorname{Proj}_X(\operatorname{Hess}\bar{f}(X)[H]) \\ &+ \left[P_U^\perp \operatorname{grad}\bar{f}(X) V_p \Sigma^{-1}\right] V^\top + U \left[P_V^\perp (\operatorname{grad}\bar{f}(X))^\top U_p \Sigma^{-1}\right]^\top. \end{aligned} \tag{7.55}$$

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 5.15. 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 (7.48) is second order as a consequence of Proposition 5.43.

Exercise 7.4. *Taking inspiration from the discussion of retraction (7.48), 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 (7.43) 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 (7.45), 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`.

7.6 The hyperboloid model

In Manopt, see: `hyperbolicfactory`

Consider the bilinear map $\langle\cdot,\cdot\rangle_M$ on $\mathbb{R}^{n+1}$ defined by

$$\langle u, v\rangle_M = -u_0 v_0 + u_1 v_1 + \cdots + u_n v_n = u^\top J v \tag{7.56}$$

with $J = \operatorname{diag}(-1, 1, \ldots, 1)$. 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 this subset of $\mathbb{R}^{n+1}$ (sometimes denoted by H^n):

$$\begin{aligned} \mathcal{M} &= \{x \in \mathbb{R}^{n+1} : \langle x, x\rangle_M = -1\} \\ &= \{x \in \mathbb{R}^{n+1} : x_0^2 = 1 + x_1^2 + \cdots + x_n^2\}. \end{aligned} \tag{7.57}$$

Its defining function is $h(x) = \langle x, x\rangle_M + 1$, with 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

$$\mathrm{T}_x\mathcal{M} = \{u \in \mathbb{R}^{n+1} : \langle x, u\rangle_M = 0\}. \tag{7.58}$$

This manifold has two connected components, corresponding to x_0 being positive and negative. (Indeed, any continuous curve on $\mathcal{M}$ connecting a point with positive x_0 to a point with negative x_0 would have to go through a point with $x_0 = 0$, which is impossible.) With $n = 2$, these two components are hyperboloids 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 $(x, u) \in \mathrm{T}\mathcal{M}$,

$$\begin{aligned}
\langle u, u \rangle_M &= u_1^2 + \cdots + u_n^2 - u_0^2 \\
&= u_1^2 + \cdots + u_n^2 - \frac{1}{x_0^2}(x_1 u_1 + \cdots + x_n u_n)^2 \\
&\geq u_1^2 + \cdots + u_n^2 - \frac{1}{x_0^2}(x_1^2 + \cdots + x_n^2)(u_1^2 + \cdots + u_n^2) \\
&= (u_1^2 + \cdots + u_n^2)\left(1 - \frac{x_0^2 - 1}{x_0^2}\right) \\
&= \frac{1}{x_0^2}(u_1^2 + \cdots + u_n^2) \\
&\geq 0.
\end{aligned}$$

(On the first line we use $\langle x, u \rangle_M = 0$ 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 essential geometric trait of $\mathcal{M}$ (with $n \geq 2$) is that its sectional curvatures are constant, equal to -1. All such manifolds 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*. See [Lee18, p62] for a comparison.

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

$$\mathrm{N}_x\mathcal{M} = \left\{ v \in \mathbb{R}^{n+1} : \langle u, v \rangle_M = 0 \text{ for all } u \in \mathrm{T}_x\mathcal{M} \right\} = \mathrm{span}(x). \tag{7.59}$$

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:

$$\mathrm{Proj}_x(z) = z + \langle x, z \rangle_M \cdot x. \tag{7.60}$$

With this tool in hand, we can construct a useful formula to compute gradients of functions on $\mathcal{M}$.

Proposition 7.5. *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:*

$$\mathrm{grad} f(x) = \mathrm{Proj}_x\big(J \mathrm{grad} \bar{f}(x)\big), \tag{7.61}$$

where $J = \mathrm{diag}(-1, 1, \ldots, 1)$ and Proj_x is defined by (7.60).

Note that $J\mathrm{grad}\bar{f}(x)$ 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.

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

$$\begin{aligned}
\mathrm{D}f(x)[u] &= \mathrm{D}\bar{f}(x)[u] \\
&= \big\langle \mathrm{grad}\bar{f}(x), u \big\rangle \\
&= \big\langle J\mathrm{grad}\bar{f}(x), u \big\rangle_M \\
&= \big\langle J\mathrm{grad}\bar{f}(x), \mathrm{Proj}_x(u) \big\rangle_M \\
&= \big\langle \mathrm{Proj}_x\big(J\mathrm{grad}\bar{f}(x)\big), u \big\rangle_M .
\end{aligned}$$

(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 $\langle\cdot,\cdot\rangle_M$, being an orthogonal projector.) The claim follows by uniqueness. ☐

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

$$\nabla_u V = \mathrm{Proj}_x(\mathrm{D}\bar{V}(x)[u]), \tag{7.62}$$

Compare this with Theorem 5.7 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.

where $\bar{V}$ is any smooth extension of V to a neighborhood of $\mathcal{M}$ in $\mathbb{R}^{n+1}$ and $\mathrm{D}\bar{V}(x)[u] = \lim_{t\to 0} \frac{1}{t}(\bar{V}(x+tu) - \bar{V}(x))$ is the usual directional derivative. It is an exercise to check that ∇ is the Riemannian connection for $\mathcal{M}$.

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

Compare this with Proposition 5.27.

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \mathrm{Proj}_{c(t)}\left(\frac{\mathrm{d}}{\mathrm{d}t} Z(t)\right) \tag{7.63}$$

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 $Z(t) \in \mathrm{T}_{c(t)}\mathcal{M} \subset \mathbb{R}^{n+1}$.

It is an exercise to check that, for arbitrary $(x, u) \in \mathrm{T}\mathcal{M}$,

Compare this with the geodesics on the sphere, Example 5.32.

$$\begin{aligned}
c(t) = \mathrm{Exp}_x(tu) &= \cosh(\|tu\|_M)x + \frac{\sinh(\|tu\|_M)}{\|tu\|_M} tu \\
&= \cosh(t\|u\|_M)x + \frac{\sinh(t\|u\|_M)}{\|u\|_M} u
\end{aligned} \tag{7.64}$$

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).

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 7.6. *(Continued from Proposition 7.5.) The Hessian of f is related to that of $\bar{f}$ as follows:*

$$\mathrm{Hess} f(x)[u] = \mathrm{Proj}_x\big(J\mathrm{Hess}\bar{f}(x)[u]\big) + \big\langle x, J\mathrm{grad}\bar{f}(x)\big\rangle_M \cdot u, \tag{7.65}$$

where $J = \mathrm{diag}(-1, 1, \ldots, 1)$ and Proj_x is defined by (7.60).

Proof. Consider this 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

$$\begin{aligned}
\mathrm{Hess} f(x)[u] &= \nabla_u \mathrm{grad} f \\
&= \mathrm{Proj}_x(\mathrm{D}\bar{G}(x)[u]) \\
&= \mathrm{Proj}_x\big(J\mathrm{Hess}\bar{f}(x)[u] + qx + \big\langle J\mathrm{grad}\bar{f}(x), x\big\rangle_M \cdot u\big) \\
&= \mathrm{Proj}_x\big(J\mathrm{Hess}\bar{f}(x)[u]\big) + \big\langle J\mathrm{grad}\bar{f}(x), x\big\rangle_M \cdot u,
\end{aligned}$$

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 7.7. *Check that $\langle \cdot, \cdot \rangle_M$ indeed defines a Riemannian metric on $\mathcal{M}$. Verify that ∇ (7.62) is the Riemannian connection for $\mathcal{M}$, that $\frac{\mathrm{D}}{\mathrm{d}t}$ (7.63) is the covariant derivative induced by ∇ and that $c(t)$ (7.64) 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 $\|\cdot\|$. The set

$$\mathcal{M} = \{x \in \mathcal{E} : h(x) = 0\} \tag{7.66}$$

is an embedded submanifold of $\mathcal{E}$ of dimension $\dim \mathcal{E} - k$ if $\mathrm{D}h(x)$ has full rank k for all $x \in \mathcal{M}$: we make this assumption here. In contrast with Definition 3.6, we require the whole manifold to be defined with a single defining function h. 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 $h_i \colon \mathcal{E} \to \mathbb{R}$, note the linear operator

$$\mathrm{D}h(x)[v] = \left(\langle \mathrm{grad} h_1(x), v \rangle, \ldots, \langle \mathrm{grad} h_k(x), v \rangle \right)^\top \tag{7.67}$$

and its adjoint

$$\mathrm{D}h(x)^*[\alpha] = \sum_{i=1}^{k} \alpha_i \mathrm{grad} h_i(x). \tag{7.68}$$

The tangent spaces are given by

$$\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x) = \{v \in \mathcal{E} : \langle \mathrm{grad} h_i(x), v \rangle = 0 \text{ for all } i\}. \tag{7.69}$$

The fact that $\mathrm{D}h(x)$ has full rank k means the gradients of the constraints at x are linearly independent; they form a basis for the normal space at x:

$$\mathrm{N}_x\mathcal{M} = (\ker \mathrm{D}h(x))^\perp = \mathrm{span}\big(\mathrm{grad} h_1(x), \ldots, \mathrm{grad} h_k(x)\big). \tag{7.70}$$

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

$$v = \mathrm{Proj}_x(v) + \mathrm{D}h(x)^*[\alpha]. \tag{7.71}$$

This decomposes v into its tangent and normal parts at x. Explicitly, α is the unique solution to the following least-squares problem:

$$\alpha = \operatorname*{arg\,min}_{\alpha \in \mathbb{R}^k} \|v - \mathrm{D}h(x)^*[\alpha]\|^2 = \left(\mathrm{D}h(x)^*\right)^\dagger [v],$$

where the dagger denotes Moore–Penrose pseudo-inversion, so that

$$\mathrm{Proj}_x(v) = v - \mathrm{D}h(x)^* \left[\left(\mathrm{D}h(x)^*\right)^\dagger [v] \right]. \tag{7.72}$$

A natural retraction for $\mathcal{M}$ is metric projection: it relies on the Euclidean metric to define:

$$\mathrm{R}_x(v) = \operatorname*{arg\,min}_{y \in \mathcal{E}} \|x + v - y\| \text{ subject to } h(y) = 0. \tag{7.73}$$

This is well defined for small enough v (it is even second order), but $\mathrm{R}_x(v)$ may not be uniquely defined for larger v—see Propositions 5.42 and 5.43. Moreover, it is difficult to compute in general.

Endow $\mathcal{M}$ with the Riemannian submanifold structure. Given a smooth function $\bar{f} \colon \mathcal{E} \to \mathbb{R}$ and its restriction $f = \bar{f}|_{\mathcal{M}}$, the Euclidean gradient decomposes using (7.71) as

$$\mathrm{grad} \bar{f}(x) = \mathrm{grad} f(x) + \sum_{i=1}^{k} \lambda_i(x) \mathrm{grad} h_i(x) \tag{7.74}$$

with $\lambda(x) = (\mathrm{D}h(x)^*)^\dagger [\mathrm{grad}\bar{f}(x)]$ and

$$\mathrm{grad}f(x) = \mathrm{Proj}_x(\mathrm{grad}\bar{f}(x)). \tag{7.75}$$

Second-order tools

Since $\mathcal{M}$ is 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 5.7, Proposition 5.27). We use this to determine the Riemannian Hessian of $f = \bar{f}|_{\mathcal{M}}$.

Notice that

$$\lambda(x) = (\mathrm{D}h(x)^*)^\dagger[\mathrm{grad}\bar{f}(x)] \tag{7.76}$$

is a smooth function on the open subset of $\mathcal{E}$ consisting in all points x where $\mathrm{D}h(x)$ has full rank k. Thus, we can differentiate $\mathrm{grad}f(x)$ (7.74) (understood as a smooth extension of itself in $\mathcal{E}$) 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

$$\mathrm{Hess}f(x)[v] = \mathrm{Proj}_x\left(\mathrm{Hess}\bar{f}(x)[v] - \sum_{i=1}^k \lambda_i(x)\mathrm{Hess}h_i(x)[v]\right). \tag{7.77}$$

This can be summarized with the pleasantly symmetric identities

See also Section 7.8 for an alternative derivation.

$$\mathrm{grad}f(x) = \mathrm{grad}\bar{f}(x) - \sum_{i=1}^k \lambda_i(x)\mathrm{grad}h_i(x), \tag{7.78}$$

$$\mathrm{Hess}f(x) = \mathrm{Proj}_x \circ \left(\mathrm{Hess}\bar{f}(x) - \sum_{i=1}^k \lambda_i(x)\mathrm{Hess}h_i(x)\right) \circ \mathrm{Proj}_x, \tag{7.79}$$

with $\lambda(x)$ as defined above, and with the understanding that the linear operator on the right-hand side of (7.79) 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 Corollary 5.15 and eq. (5.33).

Exercise 7.8. *Consider the equality constrained optimization problem*

$$\min_{x\in\mathcal{E}} \bar{f}(x) \qquad \textit{subject to} \qquad h(x) = 0, \tag{7.80}$$

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* (7.80) *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* (7.80).

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. Furthermore, 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 Riemannian conditions we established in Propositions 4.4 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 the derivations in [AMS08].

The construction of geometric tools for optimization on $\mathbb{R}_r^{m \times n}$ as a Riemannian submanifold of $\mathbb{R}^{m \times n}$ follows work by Vandereycken.[2] Similarly, one can derive tools for optimization over fixed-rank tensors in both tensor train (TT) and Tucker format [UV13, KSV14, HS18].

[2] B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013

Here are a few other manifolds of interest for applications:

- The Stiefel manifold with the canonical metric [EAS98];
- The Grassmann manifold $\mathrm{Gr}(n, p)$ (see Chapter 9) as an embedded submanifold of $\mathbb{R}^{n \times n}$, where each subspace is identified with an orthogonal projector of rank p [AMT13, SI14], [BH15, Def. 2.3, §4.2];
- Matrices with positive entries (see Section 11.4) and positive definite matrices (see Section 11.5);

- Positive semidefinite matrices with a fixed rank [JBAS10, MA20];
- 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 $\mathbb{R}^2$ or higher dimensions (see [FCPJ04] and many references therein).

There are a few modern applications of optimization on hyperbolic space in machine learning, notably for hierarchical embeddings [NK17, JMM19, KMU⁺19].

About the Riemannian Hessian on a manifold defined by $h(x) = 0$ as given in (7.79): here is an alternative derivation. It is longer, but may generalize more naturally to higher-order Riemannian derivatives. Let $c \colon I \to \mathcal{E}$ be a smooth curve on $\mathcal{M}$ with $c(0) = x$ and $c'(0) = v$ for some arbitrary v in $\mathrm{T}_x\mathcal{M}$. Compute derivatives of $\bar{f} \circ c \colon I \to \mathbb{R}$ in the Euclidean way:

$$
\begin{aligned}
(\bar{f} \circ c)'(t) &= \left\langle \mathrm{grad}\bar{f}(c(t)), c'(t) \right\rangle, \\
(\bar{f} \circ c)''(t) &= \left\langle \mathrm{Hess}\bar{f}(c(t))[c'(t)], c'(t) \right\rangle + \left\langle \mathrm{grad}\bar{f}(c(t)), \frac{\mathrm{d}}{\mathrm{d}t} c'(t) \right\rangle.
\end{aligned}
$$

Alternatively, we can use the fact that $\bar{f} \circ c$ is equivalent to $f \circ c$ to get that the above are equal to:

$$
\begin{aligned}
(f \circ c)'(t) &= \left\langle \mathrm{grad} f(c(t)), c'(t) \right\rangle_{c(t)}, \\
(f \circ c)''(t) &= \left\langle \mathrm{Hess} f(c(t))[c'(t)], c'(t) \right\rangle_{c(t)} + \left\langle \mathrm{grad} f(c(t)), \frac{\mathrm{D}}{\mathrm{d}t} c'(t) \right\rangle_{c(t)}.
\end{aligned}
$$

The identity $(\bar{f} \circ c)''(t) = (f \circ c)''(t)$ at $t = 0$ reads:

$$
\begin{aligned}
\langle \mathrm{Hess} f(x)[v], v \rangle_x = &\left\langle \mathrm{Hess}\bar{f}(x)[v], v \right\rangle \\
&+ \left\langle \mathrm{grad}\bar{f}(x), \frac{\mathrm{d}}{\mathrm{d}t} c'(0) \right\rangle - \left\langle \mathrm{grad} f(x), \frac{\mathrm{D}}{\mathrm{d}t} c'(0) \right\rangle_x.
\end{aligned}
$$

We now use that $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$ in several ways. First, there is no need to distinguish between inner products in $\mathcal{E}$ and in $\mathrm{T}_x\mathcal{M}$; second, we know that $\frac{\mathrm{D}}{\mathrm{d}t} c'(0) = \mathrm{Proj}_x \frac{\mathrm{d}}{\mathrm{d}t} c'(0)$, and that $\mathrm{grad} f(x)$ is tangent at x, so that we may replace $\frac{\mathrm{D}}{\mathrm{d}t} c'(0)$ with $\frac{\mathrm{d}}{\mathrm{d}t} c'(0)$.

Further calling upon (7.74) to handle $\mathrm{grad}\bar{f}(x) - \mathrm{grad}f(x)$ yields

$$\langle \mathrm{Hess}f(x)[v], v\rangle_x = \left\langle \mathrm{Hess}\bar{f}(x)[v], v\right\rangle + \sum_{i=1}^{k} \lambda_i(x) \left\langle \mathrm{grad}h_i(x), \frac{\mathrm{d}}{\mathrm{d}t}c'(0)\right\rangle.$$

Our last step aims to eliminate the presence of $c(t)$. To this end, notice that $c(t)$ is on $\mathcal{M}$ for all t, so that $h(c(t)) = 0$ for all t. Consequently, the first and second derivatives of $h_i \circ c$ vanish identically for each i:

$$(h_i \circ c)'(t) = \left\langle \mathrm{grad}h_i(c(t)), c'(t)\right\rangle,$$
$$(h_i \circ c)''(t) = \left\langle \mathrm{Hess}h_i(c(t))[c'(t)], c'(t)\right\rangle + \left\langle \mathrm{grad}h_i(c(t)), \frac{\mathrm{d}}{\mathrm{d}t}c'(t)\right\rangle.$$

In particular, for $t = 0$ we deduce that

$$\left\langle \mathrm{grad}h_i(x), \frac{\mathrm{d}}{\mathrm{d}t}c'(0)\right\rangle = -\left\langle \mathrm{Hess}h_i(x)[v], v\right\rangle.$$

Combining with our work above yields an explicit expression for the Riemannian Hessian of f at x,

$$\langle \mathrm{Hess}f(x)[v], v\rangle_x = \left\langle \mathrm{Hess}\bar{f}(x)[v], v\right\rangle - \sum_{i=1}^{k} \lambda_i(x) \left\langle \mathrm{Hess}h_i(x)[v], v\right\rangle,$$

valid for all $v \in \mathrm{T}_x\mathcal{M}$. This indeed matches (7.79). This derivation parallels the one given in [Lue72, §III].

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, new manifolds enter the picture. In particular, we touch upon the Grassmann manifold, consisting 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 this broader 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 *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.

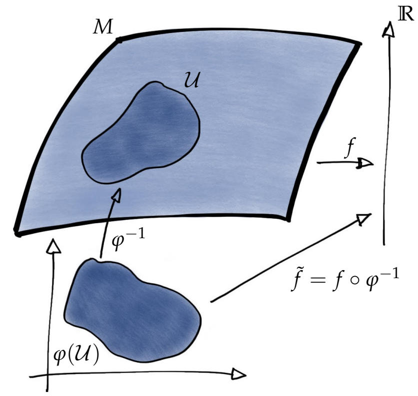
Definition 8.1. *A d-dimensional* chart *on a set M is a pair $(\mathcal{U}, \varphi)$ consisting of a subset $\mathcal{U}$ of M (called the* domain*) and a map $\varphi\colon \mathcal{U} \to \mathbb{R}^d$ such that:*

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.

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.



$$\begin{tikzcd} M \arrow[r, "f"] & \mathbb{R} \\ \varphi(\mathcal{U}) \subseteq \mathbb{R}^d \arrow[u, "\varphi^{-1}"] \arrow[ur, "\tilde{f}"'] & \end{tikzcd}$$

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 $f \colon M \to \mathbb{R}$ on a more general class of sets M. Let $(\mathcal{U}, \varphi)$ be a d-dimensional chart around $x \in M$. Then,

$$\tilde{f} = f \circ \varphi^{-1} \colon \varphi(\mathcal{U}) \to \mathbb{R}$$

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.

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*).*

We could also allow non-overlapping charts to have different dimensions, but this serves little purpose in optimization.

Compatible Charts.

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*.

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* $x \in \mathcal{U}$.

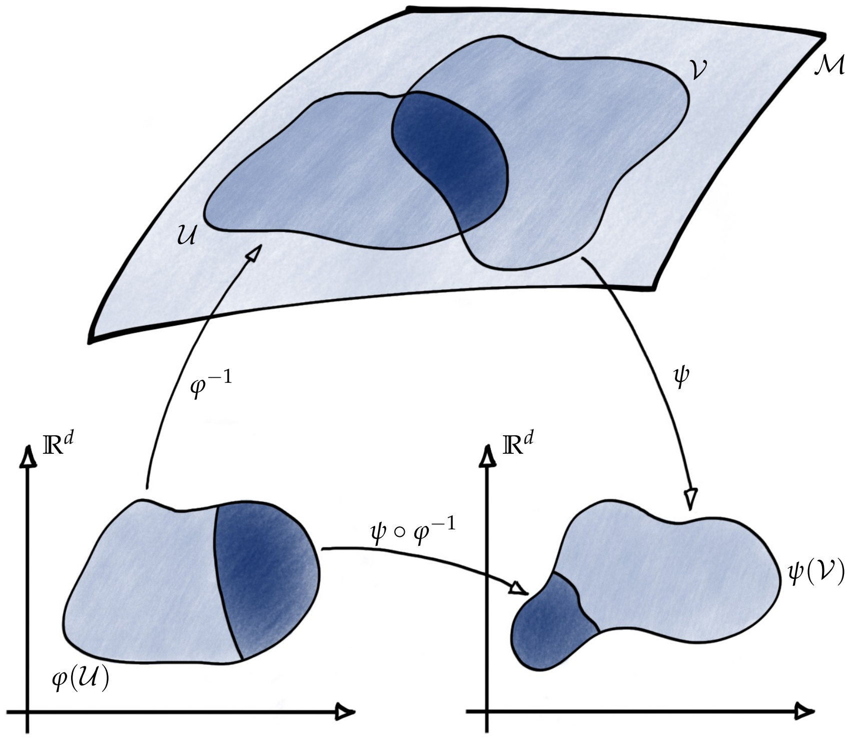


Figure 8.1: Overlapping charts $(\mathcal{U}, \varphi)$ and $(\mathcal{V}, \psi)$ on a manifold of dimension d. The darker area on the manifold corresponds to the intersection $\mathcal{U} \cap \mathcal{V}$ of the chart domains. In the coordinate spaces (bottom), the darker areas correspond to the open 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.

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*. 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 the notations* $\mathcal{M}$ *and* M.

We can now define smoothness of maps between manifolds*.

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 $\varphi(x)$ *(in the usual sense), 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.

$$\begin{tikzcd} \mathcal{U} \subseteq \mathcal{M} \arrow[r, "F"] & \mathcal{V} \subseteq \mathcal{M}' \arrow[d, "\psi"] \\ \varphi(\mathcal{U}) \subseteq \mathbb{R}^d \arrow[u, "\varphi^{-1}"] \arrow[r, "\tilde{F}"] & \psi(\mathcal{V}) \subseteq \mathbb{R}^{d'} \end{tikzcd}$$

Ok

It is an exercise to verify that this definition is independent of the choice of charts, and that composition preserves smoothness.

Okay!

Example 8.6. *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 8.7. *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 8.8. *Let $(\mathcal{U}, \varphi)$ be a chart of $\mathcal{M}$. The local parameterization $\varphi^{-1} \colon \varphi(\mathcal{U}) \to \mathcal{M}$ 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 8.9. *Consider the unit circle, $S^1 = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$. One possible atlas is made of four charts, each defined on a half circle—dubbed North, East, South and West—as follows:*

$$
\begin{aligned}
\mathcal{U}_N &= \{x \in S^1 : x_2 > 0\}, & \varphi_N(x) &= x_1, \\
\mathcal{U}_E &= \{x \in S^1 : x_1 > 0\}, & \varphi_E(x) &= x_2, \\
\mathcal{U}_S &= \{x \in S^1 : x_2 < 0\}, & \varphi_S(x) &= x_1, \\
\mathcal{U}_W &= \{x \in S^1 : x_1 < 0\}, & \varphi_W(x) &= x_2.
\end{aligned}
$$

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. $\mathcal{U}_N \cap \mathcal{U}_E = \{x \in S^1 : x_1 > 0 \text{ and } x_2 > 0\}$;
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, $(S^1, \mathcal{A}^+)$ is a manifold.*

Earlier, using Definition 3.6, 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 8.10. *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*

To be revisited.

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. $\mathcal{U}_i \cap \mathcal{U}_j = \{\pi(x) : x \in \mathbb{R}^n, x_i \neq 0 \text{ and } x_j \neq 0\}$*;*
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,*

$$\begin{aligned}(\varphi_j \circ \varphi_i^{-1})&(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n) \\ &= \left(\frac{z_1}{z_j}, \ldots, \frac{z_{i-1}}{z_j}, \frac{1}{z_j}, \frac{z_{i+1}}{z_j}, \ldots, \frac{z_{j-1}}{z_j}, \frac{z_{j+1}}{z_j}, \ldots, \frac{z_n}{z_j}\right)\end{aligned}$$

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 8.13.

Exercise 8.11. *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 itself an atlas of M.*

Exercise 8.12. *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.*

Exercise 8.13. *For the set $M = \mathbb{R}$, consider these two 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 $f\colon \mathbb{R} \to \mathbb{R}$ which is* not *smooth in the usual sense (equivalently, as a function from $\mathcal{M}$ to $\mathbb{R}$) but 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.

Topology ≡ convergence + local optimal.

We start with a few reminders. After discussing two desirable properties of topologies, we restrict the definition of manifold to those whose atlas topology is favorable.

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 8.14. *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 pair $(M, \mathcal{T})$ consisting of a set with a topology is a *topological space*. Given two topological spaces $(M, \mathcal{T})$, $(M', \mathcal{T}')$ and a map $F: M \to M'$, we say that F is *continuous* if and only 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 8.15. *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 $\mathcal{B}$.

Definition 8.16. *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.*

It is important to consider the *maximal* atlas here, as otherwise we may miss some open sets.

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.

Proposition 8.17. *In the atlas topology, any chart $\varphi: \mathcal{U} \to \varphi(\mathcal{U})$ is continuous and its inverse is also continuous (i.e., it is a* homeomorphism.*)*

[BC70, Prop. 2.4.3]

A welcome consequence of the latter proposition is that, with the atlas topologies on manifolds* $\mathcal{M}$ and $\mathcal{M}'$, any function $F: \mathcal{M} \to \mathcal{M}'$ 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 8.18. *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 8.19. *A topology is* second-countable *if there is a countable basis for its topology.*

Remarkably, a manifold* whose atlas topology is Hausdorff and second-countable admits a Riemannian metric [Lee12, Prop. 13.3]. The standard construction relies on *partitions of unity*, which exist (essentially) if and only if the topology is as prescribed—see [BC70, §3.4].

At last, we can give a proper definition of manifolds.

Definition 8.20. *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.

Proposition 8.21. *Let $\mathcal{A}$ be an atlas for the set M. Assume both:* [Lee12, Lem. 1.35]

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. We use it in Section 8.3. The "only if" direction of the statement is a direct consequence of Proposition 8.17.

Proposition 8.22. *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 are homeomorphisms with respect to $\mathcal{T}$.* [BC70, Prop. 3.1.1]

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 8.23. *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 8.24. *In all examples from Section 8.1, we have constructed atlases with a finite number of charts. Hence, by Proposition 8.21, 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.*

Exercise 8.25. *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 8.21. You may need to add a few charts to the atlases.*

Exercise 8.26. *Check that Definition 8.23 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 8.27. *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

In Example 8.24, we confirmed that linear spaces and their open subsets are manifolds in the sense of Definition 8.20. They are also embedded submanifolds in the sense of Definition 3.6, which we gave early on, in a restricted setting. In this section, we further show that all sets we have thus far called embedded submanifolds are indeed manifolds once equipped with a certain atlas, that the topology we have been using on these sets is the corresponding atlas topology, and that our early notion of smooth maps between embedded submanifolds of linear spaces coincides with the more general notion of smooth maps between manifolds.

Proposition 8.28. *A subset $\mathcal{M}$ of a linear space $\mathcal{E}$ which is an embedded submanifold as per Definition 3.6 admits an atlas which makes it a manifold in the sense of Definition 8.20. With this smooth structure, the atlas topology coincides with the subspace topology as given in Definition 3.15. Furthermore, given $\mathcal{M}$ and $\mathcal{M}'$ both embedded submanifolds in linear spaces, a map $F \colon \mathcal{M} \to \mathcal{M}'$ is smooth in the sense of Definition 3.23 if and only if it is smooth in the sense of Definition 8.5 with the atlases constructed here.*

Proof. The claim is in three parts.

Part 1. We show that $\mathcal{M}$ admits an atlas which makes it a manifold*. Chart construction starts along the same lines as in the proof of Theorem 3.8, where we used the inverse function theorem (IFT), Theorem 3.9. With $d = \dim \mathcal{E}$, fix a basis for $\mathcal{E}$ in order to identify it with $\mathbb{R}^d$. For any $x \in \mathcal{M}$, let $h \colon U \to \mathbb{R}^k$ be a local defining function, that is, U is a neighborhood of x in $\mathcal{E}$, h is smooth, $\mathrm{D}h(x) \in \mathbb{R}^{k \times d}$ has full rank k, and for all $y \in U$ it holds that $h(y) = 0 \iff y \in \mathcal{M}$. Pick k indices among $1, \ldots, d$ such that the corresponding columns of $\mathrm{D}h(x)$ are linearly independent (this is possible since $\mathrm{D}h(x)$ has rank k). Let $\pi \colon \mathbb{R}^d \to \mathbb{R}^{d-k}$ be the operator that extracts the other entries of a vector in $\mathbb{R}^d$. For example, if the last k columns of $\mathrm{D}h(x)$ are independent, we could define $\pi(y)$ as the vector with entries $(y_1, \ldots, y_{d-k})$. We can now define the following map:

$$F \colon U \to \mathbb{R}^d \colon y \mapsto F(y) = \begin{bmatrix} \pi(y) \\ h(y) \end{bmatrix}.$$

The differential of F at x is invertible. Indeed,

$$\mathrm{D}F(x)[u] = \begin{bmatrix} \mathrm{D}\pi(x)[u] \\ \mathrm{D}h(x)[u] \end{bmatrix} = \begin{bmatrix} \pi(u) \\ \mathrm{D}h(x)[u] \end{bmatrix}.$$

This is zero if and only if $\pi(u) = 0$ (indicating u has at most k nonzero entries, corresponding to linearly independent columns of $\mathrm{D}h(x)$) and $\mathrm{D}h(x)[u] = 0$, which implies $u = 0$, confirming $\mathrm{D}F(x)$ is invertible. By the IFT, we may replace U with a possibly smaller neighborhood of x in $\mathcal{E}$ such that $F \colon U \to F(U)$ is a diffeomorphism, that is, F and its inverse are smooth. With these ingredients, we are in a position to propose a tentative chart for $\mathcal{M}$ around x:

$$\mathcal{U} = \mathcal{M} \cap U, \quad \text{and} \quad \varphi \colon \mathcal{U} \to \varphi(\mathcal{U}) \colon y \mapsto \varphi(y) = \pi(y) = \mathrm{trim}(F(y)), \tag{8.1}$$

where $\mathrm{trim} \colon \mathbb{R}^d \to \mathbb{R}^{d-k}$ extracts the first $d-k$ components of a vector. This map is invertible since h is identically zero on $\mathcal{U}$. Indeed,

$$\varphi^{-1}(z) = F^{-1}(\mathrm{zpad}(z)), \tag{8.2}$$

where $\mathrm{zpad} \colon \mathbb{R}^{d-k} \to \mathbb{R}^d$ pads a vector with k zeros at the end, so that $\mathrm{trim} \circ \mathrm{zpad}$ is identity on $\mathbb{R}^{d-k}$, and $\mathrm{zpad} \circ \mathrm{trim}$ is identity on $F(\mathcal{U})$. It remains to verify that $\varphi(\mathcal{U})$ is open in $\mathbb{R}^{d-k}$. To this end, notice that

$$\begin{aligned}
\mathcal{U} &= \mathcal{M} \cap U \\
&= \{y \in U : h(y) = 0\}, \\
&= \{y \in U : \text{ the last } k \text{ entries of } F(y) \text{ are zero}\} \\
&= F^{-1}\left(F(U) \cap \left(\mathbb{R}^{d-k} \times \{0\}\right)\right),
\end{aligned}$$

where 0 here denotes the zero vector in $\mathbb{R}^k$. Thus,

$$\varphi(\mathcal{U}) = \mathrm{trim}\left(F(U) \cap \left(\mathbb{R}^{d-k} \times \{0\}\right)\right),$$

which is open in $\mathbb{R}^{d-k}$ since $F(U)$ is open in $\mathbb{R}^d$ and using properties of the standard topology on real spaces. This confirms that $(\mathcal{U}, \varphi)$ is a $(d-k)$-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 $F \colon U \to F(U) \subseteq \mathbb{R}^{d-k}$ and $G \colon V \to G(V) \subseteq \mathbb{R}^{d-k}$. Then, the change of coordinates map,

$$\psi \circ \varphi^{-1} = \mathrm{trim} \circ G \circ F^{-1} \circ \mathrm{zpad},$$

is smooth by composition of smooth maps, from $\varphi(\mathcal{U} \cap \mathcal{V})$ to $\psi(\mathcal{U} \cap \mathcal{V})$. These domains are open since

$$\begin{aligned}
\varphi(\mathcal{U} \cap \mathcal{V}) &= \mathrm{trim}\left(F(U \cap V) \cap \left(\mathbb{R}^{d-k} \times \{0\}\right)\right), \text{ and} \\
\psi(\mathcal{U} \cap \mathcal{V}) &= \mathrm{trim}\left(G(U \cap V) \cap \left(\mathbb{R}^{d-k} \times \{0\}\right)\right).
\end{aligned}$$

Furthermore, the change of coordinates map is clearly smoothly invertible, which finishes the construction of our atlas, turning $\mathcal{M}$ into a manifold*.

Part 2. That the atlas and the subspace topologies coincide follows from Proposition 8.22. 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 $\mathbb{R}^{d-k}$ 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 $\mathrm{trim} \circ F$ from U to $\mathcal{U}$. That φ^{-1} is continuous is also clear since it is equal to the continuous map $F^{-1} \circ \mathrm{zpad}$, 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.

Part 3. The equivalence of the two definitions of smoothness for maps between manifolds is given in Proposition 3.63. □

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.

We present one standard definition of tangent vectors on manifolds. Another standard definition is through the notion of *derivation*. These two perspectives are equivalent.

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:

$$c_1 \sim c_2 \qquad \Longleftrightarrow \qquad (\varphi \circ c_1)'(0) = (\varphi \circ c_2)'(0). \tag{8.3}$$

These derivatives are well defined as $\varphi \circ c_i$ are smooth functions (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.

Do check later!

Definition 8.29 (tangent vector, tangent space)**.** *The equivalence class of a curve $c \in C_x$ is the set of curves that are equivalent to c as per (8.3):*

$$[c] = \{\tilde{c} \in C_x : c \sim \tilde{c}\}.$$

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, that is, the set of all equivalence classes:*

$$\boxed{\mathrm{T}_x\mathcal{M} = C_x/\sim\, = \{[c] : c \in C_x\}.}$$

Given a chart $(\mathcal{U}, \varphi)$ around x, the map

$$\theta_x^\varphi \colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}^d \colon [c] \mapsto \theta_x^\varphi([c]) = (\varphi \circ c)'(0) \tag{8.4}$$

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 $\mathbb{R}^d$:

ok

$$a \cdot [c_1] + b \cdot [c_2] \triangleq (\theta_x^\varphi)^{-1}\big(a \cdot \theta_x^\varphi([c_1]) + b \cdot \theta_x^\varphi([c_2])\big). \tag{8.5}$$

This structure, again, is independent of the choice of chart. Thus, the tangent space is a linear space in its own right.

Theorem 8.30. *Tangent spaces are linear spaces of dimension* $\dim \mathcal{M}$, *with the linear structure given through* (8.5).

Done ok

When $\mathcal{M}$ is an embedded submanifold of a linear space, the two definitions of tangent spaces we have seen are essentially equivalent, 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 8.31. *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 3.7 and 8.29 are, essentially, equivalent.*

Proof. To see this, reconsider the proof of Theorem 3.8, specifically eq. (3.23), together with the charts in (8.1). □

Exercise 8.32. *Show that the equivalence relation* (8.3) *is independent of the choice of chart $(\mathcal{U}, \varphi)$ around x. Show that θ_x^φ* (8.4) *is bijective. Show that the linear structure on $\mathrm{T}_x\mathcal{M}$ defined by* (8.5) *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 $F: \mathcal{M} \to \mathcal{M}'$ 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 (8.3) is specifically crafted so that this map between tangent spaces does not depend on the choice of c in v. This yields a proper notion of differential for maps between manifolds. It is essentially equivalent to Definition 3.26 for the case of embedded submanifolds of linear spaces.

Definition 8.33. *Given manifolds $\mathcal{M}$ and $\mathcal{M}'$, the differential of a smooth map $F: \mathcal{M} \to \mathcal{M}'$ at x is a linear operator $\mathrm{D}F(x): \mathrm{T}_x\mathcal{M} \to \mathrm{T}_{F(x)}\mathcal{M}'$ defined by:*

$$\mathrm{D}F(x)[v] = [t \mapsto F(c(t))], \tag{8.6}$$

where c is a smooth curve on $\mathcal{M}$ passing through x at $t = 0$ such that $v = [c]$.

On the right of (8.6), brackets indicate taking an equivalence class of curves. On the left, brackets around v are meant to distinguish between x (the point at which we differentiate) and v (the direction along which we differentiate).

Definition 8.34. *Let $\mathfrak{F}(\mathcal{M})$ denote the set of smooth scalar fields on $\mathcal{M}$, that is, the set of smooth functions $f: \mathcal{M} \to \mathbb{R}$. Then, identifying[1] the tangent spaces of $\mathbb{R}$ with $\mathbb{R}$ itself, a particular case of Definition 8.33 for $f \in \mathfrak{F}(\mathcal{M})$ defines the differential $\mathrm{D}f(x): \mathrm{T}_x\mathcal{M} \to \mathbb{R}$ as:*

$$\boxed{\mathrm{D}f(x)[v] = (f \circ c)'(0),} \tag{8.7}$$

where $v = [c]$. More generally, for a smooth map $F: \mathcal{M} \to \mathcal{N}$ where $\mathcal{N}$ is an embedded submanifold of a linear space $\mathcal{E}$, identifying the tangent spaces of $\mathcal{N}$ to subspaces of $\mathcal{E}$ and with $v = [c]$ a tangent vector at $x \in \mathcal{M}$, we write

$$\boxed{\mathrm{D}F(x)[v] = (F \circ c)'(0),} \tag{8.8}$$

where $F \circ c$ is seen as a map into $\mathcal{E}$.

[1] Identification is through Theorem 8.31.

In my notation $C^\infty(M)$ - set of smooth scalar fields $C^\infty(M) := \{f: \{.$

Exercise 8.35. *Verify that equation (8.6) 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 operator with respect to the linear structure (8.5) on tangent spaces.*

Exercise 8.36. *For smooth maps $F_1, F_2: \mathcal{M} \to \mathcal{E}$ (with $\mathcal{E}$ a linear space) and real numbers a_1, a_2, show that $F: x \mapsto a_1F_1(x) + a_2F_2(x)$ 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 $G: \mathcal{M} \to \mathcal{E}$, show that the product map $fG: x \mapsto f(x)G(x)$ 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 $F: \mathcal{M} \to \mathcal{M}'$ and $G: \mathcal{M}' \to \mathcal{M}''$ be smooth. In Exercise 8.12, we verified the composition $G \circ F$ is smooth. Now, show that

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

that is, we have a chain rule.

8.6 *Tangent bundles and vector fields*

Identically to Definition 3.34, we define the *tangent bundle* as the disjoint union of all tangent spaces, now provided by Definition 8.29.

Definition 8.37. *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 the notations (x, v) and v, when the context is clear.

Definition 8.38. *The projection $\pi: \mathrm{T}\mathcal{M} \to \mathcal{M}$ extracts the root 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 3.35), tangent bundles of manifolds are manifolds in a natural way. (Smoothness of π is understood through Definition 8.5.)

Theorem 8.39. *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 $\pi: \mathrm{T}\mathcal{M} \to \mathcal{M}$ 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 $\tilde{\varphi}: \tilde{\mathcal{U}} \to \tilde{\varphi}(\tilde{\mathcal{U}}) \subseteq \mathbb{R}^{2d}$ as

$$\tilde{\varphi}(x, v) = (\varphi(x), \theta_x^\varphi(v)), \tag{8.9}$$

where θ_x^φ is defined by (8.4). See [Lee12, Prop. 3.18] for details. $\square$

The smooth structure on tangent bundles makes differentials of smooth maps into smooth maps themselves.

Proposition 8.40. *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.* [Lee12, Prop. 3.21]

Proof. Write $\mathrm{D}F$ in coordinates using charts from Theorem 8.39, then use Proposition 8.46 below. $\square$

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 8.41. *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 8.42. *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 $\mathcal{U}$.* [Lee12, Prop. 8.1]

Proof. Using Definition 8.5 about smooth maps and the charts of $\mathrm{T}\mathcal{M}$ defined by (8.9), we conclude that V is smooth if and only if, for every chart $(\mathcal{U}, \varphi)$ of $\mathcal{M}$,

$$\tilde{V} = \tilde{\varphi} \circ V \circ \varphi^{-1} \colon \varphi(\mathcal{U}) \subseteq \mathbb{R}^d \to \mathbb{R}^{2d}$$

is smooth, where $\tilde{\varphi}(x, v) = (\varphi(x), \theta_x^\varphi(v))$. For $z = \varphi(x)$,

$$\tilde{V}(z) = (z, \theta_x^\varphi(V(x))),$$

so that $\tilde{V}$ is smooth if and only if $x \mapsto \theta_x^\varphi(V(x))$ is smooth on $\mathcal{U}$. $\square$

Let V be a vector field on $\mathcal{M}$. As we did in Definition 5.3, 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 $Vf \colon \mathcal{U} \to \mathbb{R}$ determined by

$$(Vf)(x) = \mathrm{D}f(x)[V(x)]. \tag{8.10}$$

Based on the latter, we mention the following characterization of smooth vector fields which is sometimes useful. The proof in the direction we need is an exercise in Section 8.8.

Proposition 8.43. *A vector field V on a manifold $\mathcal{M}$ is smooth if and only if Vf is smooth for all $f \in \mathfrak{F}(\mathcal{M})$.* [Lee12, Prop. 8.14]

Exercise 8.44. *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

Now equipped with broader notions of smooth maps, tangent bundles and differentials, the definition of retraction, Definition 3.40, generalizes verbatim. We re-state it here for convenience. Since $\mathrm{T}_x\mathcal{M}$ is a linear space, it has a natural smooth structure as well: we use it below to give meaning to smoothness of R_x. Furthermore, with this structure, the tangent spaces of $\mathrm{T}_x\mathcal{M}$ are identified with $\mathrm{T}_x\mathcal{M}$ itself (Theorem 8.31), so that $\mathrm{T}_0(\mathrm{T}_x\mathcal{M})$—the domain of $\mathrm{DR}_x(0)$ below—is identified with $\mathrm{T}_x\mathcal{M}$.

Definition 8.45. *A* retraction *on* $\mathcal{M}$ *is a smooth map* $\mathrm{R}\colon \mathrm{T}\mathcal{M} \to \mathcal{M}$ *with the following properties. For each* $x \in \mathcal{M}$*, let* $\mathrm{R}_x\colon \mathrm{T}_x\mathcal{M} \to \mathcal{M}$ *be the restriction of* R *at* x*, so that* $\mathrm{R}_x(v) = \mathrm{R}(x, v)$*. Then,*

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*.)*

8.8 Coordinate vector fields as local frames

Let $(\mathcal{U}, \varphi)$ be a chart on a d-dimensional manifold $\mathcal{M}$. Consider the following vector fields on $\mathcal{U}$, called *coordinate vector fields*:

Here and in many places, we use that $\mathcal{U}$ itself is a manifold; specifically, an open submanifold of $\mathcal{M}$: see Definition 8.23.

$$W_i(x) = \left[t \mapsto \varphi^{-1}(\varphi(x) + te_i)\right], \qquad i = 1, \ldots, d, \tag{8.11}$$

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 (8.4), they correspond to the constant coordinate vector fields of $\mathbb{R}^d$:

$$\theta_x^\varphi(W_i(x)) = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(\varphi^{-1}(\varphi(x) + te_i))\bigg|_{t=0} = e_i. \tag{8.12}$$

As a corollary, we obtain a generalization of Proposition 3.59: local frames exist around any point on a manifold (see Definition 3.58).

Proposition 8.46. *Coordinate vector fields (8.11) 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 Proposition 8.42 and (8.12). Now consider the linear structure on $\mathrm{T}_x\mathcal{M}$ defined by (8.5): $W_1(x), \ldots, W_d(x)$ are linearly independent if and only if they are so after being pushed through θ_x^φ, which is clearly the case owing to (8.12). $\square$

Corollary 8.47. *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* $V|_\mathcal{U} = g_1W_1 + \cdots + g_dW_d$*. These functions are smooth if and only if* $V|_\mathcal{U}$ *is smooth.*

We use that V is smooth on $\mathcal{M}$ if and only if it is smooth when restricted to any chart domain $\mathcal{U}$ equipped with the open submanifold geometry.

Proof. That functions $g_i\colon \mathcal{U} \to \mathbb{R}$ such that $V|_\mathcal{U} = \sum_i g_iW_i$ exist and are unique follows from linear independence of $W_1(x), \ldots, W_d(x)$. The smoothness equivalence follows from Proposition 8.42 and

$$\theta_x^\varphi(V(x)) = \sum_{i=1}^d g_i(x)\theta_x^\varphi(W_i(x)) = (g_1(x), \ldots, g_d(x))^\top, \tag{8.13}$$

where we used $\theta_x^\varphi(W_i(x)) = e_i$ by (8.12). $\square$

Exercise 8.48. *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 8.43).*

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 8.49. *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 8.50. *A metric $\langle \cdot, \cdot \rangle_x$ on $\mathcal{M}$ is a* Riemannian metric *if it varies smoothly with x, in the sense that if V, W are two smooth vector fields on $\mathcal{M}$ then the function $x \mapsto \langle V(x), W(x) \rangle_x$ is smooth from $\mathcal{M}$ to $\mathbb{R}$.*

Definition 8.51. *A manifold with a Riemannian metric is a* Riemannian manifold.

Definition 8.52. *Let $f : \mathcal{M} \to \mathbb{R}$ 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 these identities:*

$$\forall (x, v) \in \mathrm{T}\mathcal{M}, \qquad \mathrm{D}f(x)[v] = \langle v, \mathrm{grad} f(x) \rangle_x, \tag{8.14}$$

where $\mathrm{D}f(x)$ is as in Definition 8.34 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 3.60 extends as is, using local frames provided by Proposition 8.46 for example.

Proposition 8.53. *For $f \in \mathfrak{F}(\mathcal{M})$, the gradient $\mathrm{grad} f$ is smooth.* [Lee12, p343]

Proposition 3.49 also holds true in the general case, with the same proof. We restate the claim here.

Proposition 8.54. *Let $f : \mathcal{M} \to \mathbb{R}$ be a smooth function on a Riemannian manifold $\mathcal{M}$ equipped with a retraction R. Then, for all $x \in \mathcal{M}$,* See also Exercise 10.67.

$$\mathrm{grad} f(x) = \mathrm{grad}(f \circ \mathrm{R}_x)(0), \tag{8.15}$$

where $f \circ \mathrm{R}_x : \mathrm{T}_x\mathcal{M} \to \mathbb{R}$ 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.

8.10 Lie brackets as vector fields

Recall Definition 5.3 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:

$$[U, V]f = U(Vf) - V(Uf). \tag{8.16}$$

We now extend Proposition 5.8 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 8.55. *Lie brackets of coordinate vector fields* (8.11) *vanish identically, that is,*

$$[W_i, W_j]f = 0$$

for all $1 \le i, j \le d$ *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 Definition 8.34:

$$\begin{aligned}
(W_i f)(x) &= \mathrm{D}f(x)[W_i(x)] \\
&= \frac{\mathrm{d}}{\mathrm{d}t} f\left(\varphi^{-1}\left(\varphi(x) + te_i\right)\right)\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \tilde{f}(\varphi(x) + te_i)\Big|_{t=0} \\
&= \mathrm{D}\tilde{f}(\varphi(x))[e_i] \\
&= \langle \mathrm{grad}\tilde{f}(\varphi(x)), e_i \rangle,
\end{aligned} \tag{8.17}$$

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. $\square$

Proposition 8.56. *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 (8.11). By Corollary 8.47, 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 5.9),

$$U(Vf) = \sum_{i,j} g_i W_i(h_j W_j f) = \sum_{i,j} g_i (W_i h_j)(W_j f) + g_i h_j W_i(W_j f).$$

With similar considerations for $V(Uf)$,

$$V(Uf) = \sum_{i,j} h_j W_j(g_i W_i f) = \sum_{i,j} h_j (W_j g_i)(W_i f) + h_j g_i W_j(W_i f),$$

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 8.55, it follows that, on the domain $\mathcal{U}$, there is a unique smooth vector field, namely,

$$\sum_{i,j} g_i (W_i h_j) W_j - h_j (W_j g_i) W_i, \tag{8.18}$$

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 $\mathcal{M}$, which we identify with $[U, V]$. $\square$

8.11 *Riemannian connections and Hessians*

The notion of connection applies in the general case. For convenience we repeat Definition 5.17 here. (Definition 5.1 also extends as is.)

Definition 8.57. *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 satisfies the following three properties for arbitrary $U, V, W \in \mathfrak{X}(\mathcal{M})$, $f, g \in \mathfrak{F}(\mathcal{M})$ 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 5.4 regarding the existence and uniqueness of a Riemannian connection extends without difficulty. We use Proposition 8.56 (stating Lie brackets are vector fields) to state the symmetry condition in a more standard way.

Theorem 8.58. *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 5.18 in the embedded case, connections are pointwise operators in U. The proof for 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 8.46; second, we must reaffirm the technical Lemma 5.24 which allows us to make sense of ∇ when applied to locally defined smooth vector fields (such as coordinate vector fields for example).

Proposition 8.59. *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 5.11 for Riemannian Hessians to general manifolds.

Definition 8.60. *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 a linear operator $\mathrm{Hess}f(x) \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ 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 5.14, extends verbatim.

Proposition 8.61. *The Riemannian Hessian is self-adjoint with respect to the Riemannian metric. That is, for all $x \in \mathcal{M}$ and $u, v \in \mathrm{T}_x\mathcal{M}$,*

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

8.12 Covariant derivatives, velocity and geodesics

Recall Definition 5.25: 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 if it 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 5.26, 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 8.46) and we reaffirm the notation (5.16) justified in the embedded case.

Theorem 8.62. *Let $c\colon I \to \mathcal{M}$ be a smooth curve on a manifold equipped with a connection ∇. There exists a unique operator $\frac{\mathrm{D}}{\mathrm{d}t}\colon \mathfrak{X}(c) \to \mathfrak{X}(c)$ satisfying these three 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 $t \in I$.*

This operator is called the induced covariant derivative*. Furthermore, if $\mathcal{M}$ is a Riemannian manifold with metric $\langle \cdot, \cdot \rangle$ and ∇ is compatible with the metric (e.g., if it 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 $\langle Y, Z\rangle(t) = \langle Y(t), Z(t)\rangle_{c(t)}$.

The general definition of tangent vectors, Definition 8.29, makes it straightforward to define *velocity* along a curve. Then, using the induced covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$, we may define acceleration along a curve similarly to Definition 5.31, and geodesics as in Definition 5.33.

Definition 8.63. *Let $c\colon I \to \mathcal{M}$ be a smooth curve. The* velocity *of c is a smooth vector field on c defined for all t in I 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 (8.3). The acceleration *of c is a smooth vector field on c defined as the covariant derivative of the velocity:*

$$c'' = \frac{\mathrm{D}}{\mathrm{d}t}c'.$$

A geodesic *is a smooth curve $c\colon I \to \mathcal{M}$ such that $c''(t) = 0$ for all $t \in I$.*

8.13 Taylor expansions and second-order retractions

Using the general tools constructed thus far, the reasoning that lead to second-order Taylor expansions for embedded submanifolds and which culminated in eq. (5.25) 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 $c'(0) = v$,

$$f(c(t)) = f(x) + t \langle \operatorname{grad} f(x), v \rangle_x + \frac{t^2}{2} \langle \operatorname{Hess} f(x)[v], v \rangle_x + \frac{t^2}{2} \langle \operatorname{grad} f(x), c''(0) \rangle_x + O(t^3). \quad (8.19)$$

Definition 5.36 extends as is to the general case.

Definition 8.64. *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 formulate Proposition 5.38 in the general case, with the same proof.

Proposition 8.65. *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 $\operatorname{grad} f(x) = 0$), then*

$$f(\mathrm{R}_x(s)) = f(x) + \frac{1}{2} \langle \operatorname{Hess} f(x)[s], s \rangle_x + O(\|s\|_x^3). \quad (8.20)$$

If R is a second-order retraction, then for any point $x \in \mathcal{M}$ we have

$$f(\mathrm{R}_x(s)) = f(x) + \langle \operatorname{grad} f(x), s \rangle_x + \frac{1}{2} \langle \operatorname{Hess} f(x)[s], s \rangle_x + O(\|s\|_x^3). \quad (8.21)$$

Finally, Proposition 5.39 and its proof also extend as is.

Proposition 8.66. *If the retraction is second order or if $\operatorname{grad} f(x) = 0$, then*

See also Exercise 10.67.

$$\operatorname{Hess} f(x) = \operatorname{Hess}(f \circ \mathrm{R}_x)(0),$$

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 this chapter, we considered a general definition of manifolds, and we showed that embedded submanifolds of linear spaces are manifolds. Now, we discuss the more general concept of embedded submanifold of a manifold. 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 $\mathcal{M}$, which may or may not interact nicely with the smooth structure of $\overline{\mathcal{M}}$. 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, and it satisfies an additional condition spelled out below, $\mathcal{M}$ is called a *submanifold* of $\overline{\mathcal{M}}$.

Definition 8.67. *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 that $\mathcal{M}$ is an* (immersed) submanifold *of $\overline{\mathcal{M}}$.*

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.

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 $\overline{\mathcal{M}}$. 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 8.68. *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*.*

Some authors call them *regular* submanifolds [BC70], or even just submanifolds [O'N83].

Theorem 8.69. *A subset $\mathcal{M}$ of a manifold $\overline{\mathcal{M}}$ admits at most one smooth structure that makes $\mathcal{M}$ an embedded submanifold of $\overline{\mathcal{M}}$.*

[Lee12, Thm. 5.31]

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 smooth structure.

The next result gives a complete characterization of embedded submanifolds. Compare it to Definition 3.6 where $\overline{\mathcal{M}}$ is a linear space $\mathcal{E}$.

Proposition 8.70. *A subset $\mathcal{M}$ of a manifold $\overline{\mathcal{M}}$ is an* embedded submanifold *of $\overline{\mathcal{M}}$ if and only if either of the following holds:*

[Lee12, Prop. 5.16] and Lemma 3.62.

1. *$\mathcal{M}$ is an open subset of $\overline{\mathcal{M}}$. Then, we also call $\mathcal{M}$ an* open submanifold *as in Definition 8.23 and* $\dim \mathcal{M} = \dim \overline{\mathcal{M}}$.

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*

$$h^{-1}(0) = \mathcal{M} \cap \overline{\mathcal{U}} \qquad \textit{and} \qquad \operatorname{rank} \mathrm{D}h(x) = k.$$

Then, $\dim \mathcal{M} = \dim \overline{\mathcal{M}} - k$ *and h is called a* local defining function*.*

The tangent spaces of $\mathcal{M}$ are subspaces of those of $\overline{\mathcal{M}}$:

$$\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x) \subseteq \mathrm{T}_x\overline{\mathcal{M}}, \tag{8.22}$$

where h is any local defining function for $\mathcal{M}$ around x.

In Proposition 8.70, 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 this corollary.

Corollary 8.71. *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 $\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x)$.*

[Lee12, Cor. 5.14]
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 in $\mathbb{R}^d$ 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 this corollary 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 8.72. *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 $\mathrm{T}_x\mathcal{M} = \ker \mathrm{D}h(x)$.*

[Lee12, Thm. 5.12]

As we discovered in Chapters 3 and 5, geometric tools for submanifolds embedded in linear spaces are related to their counterparts on that linear space in a straightforward way. This is true more generally for submanifolds embedded in manifolds, and the proofs we have considered extend to the general case with little friction. We now summarize these results stated in the general case.

Let $\mathcal{M}$ be an embedded submanifold of $\overline{\mathcal{M}}$. Smooth maps to and from $\mathcal{M}$ are related to smooth maps to and from $\overline{\mathcal{M}}$ as follows, where $\mathcal{N}$ is an arbitrary manifold. In a nutshell: if $\bar{F}$ is a smooth map on $\overline{\mathcal{M}}$, then $\bar{F}|_{\mathcal{M}}$ is also smooth (we may restrict the domain without losing smoothness); likewise, if F is a smooth map to $\overline{\mathcal{M}}$ whose image happens to be included in $\mathcal{M}$, then F is also smooth as a map to $\mathcal{M}$ (we may restrict (or extend) the codomain without losing smoothness).

See Section 3.10 for a related discussion.

Proposition 8.73. *If the map $\bar{F}\colon \overline{\mathcal{M}} \to \mathcal{N}$ is smooth (at $x \in \mathcal{M}$), then the restriction $F = \bar{F}|_{\mathcal{M}}$ is smooth (at x). The other way around, if $F\colon \mathcal{M} \to \mathcal{N}$ is smooth at x, there exists a neighborhood $\overline{\mathcal{U}}$ of x in $\overline{\mathcal{M}}$ and a smooth map $\bar{F}\colon \overline{\mathcal{U}} \to \mathcal{N}$ such that $F(y) = \bar{F}(y)$ for all $y \in \overline{\mathcal{U}} \cap \mathcal{M}$. If F is smooth, then there exists a neighborhood $\overline{\mathcal{U}}$ of $\mathcal{M}$ in $\overline{\mathcal{M}}$ and a smooth map $\bar{F}\colon \overline{\mathcal{U}} \to \mathcal{N}$ such that $F = \bar{F}|_{\mathcal{M}}$. Such maps $\bar{F}$ are called* smooth extensions *of F.*

Proof. The first part holds because $\bar{F}|_{\mathcal{M}} = \bar{F} \circ i$ is smooth by composition, where $i\colon \mathcal{M} \to \overline{\mathcal{M}}$ is the inclusion map, smooth by definition of embedded submanifolds: see [Lee12, Thm. 5.27]. For the second part, see the proof of Proposition 3.24 in Section 3.10. □

Proposition 8.74. *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).*

[Lee12, Cor. 5.30]

From (8.22) we know $\mathrm{T}_x\mathcal{M}$ is a linear subspace of $\mathrm{T}_x\overline{\mathcal{M}}$. If we equip the submanifold $\mathcal{M}$ with a Riemannian metric simply by restricting the Riemannian metric $\langle\cdot,\cdot\rangle_x$ of $\overline{\mathcal{M}}$ to the tangent spaces of $\mathcal{M}$, then $\mathcal{M}$ is a *Riemannian submanifold* of $\overline{\mathcal{M}}$, and we recover many useful identities. The linear map Proj_x which projects vectors from $\mathrm{T}_x\overline{\mathcal{M}}$ to $\mathrm{T}_x\mathcal{M}$ orthogonally with respect to $\langle\cdot,\cdot\rangle_x$ features abundantly.

We restrict this notion to embedded submanifolds, but some authors also allow immersed submanifolds to be Riemannian submanifolds [Lee18, p15].

In the remainder of this section, $\mathcal{M}$ is a Riemannian submanifold of $\overline{\mathcal{M}}$.

For example, 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 $x \in \mathcal{M}$,

$$\mathrm{grad} f(x) = \mathrm{Proj}_x\big(\mathrm{grad} \bar{f}(x)\big). \tag{8.23}$$

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 $\mathcal{M}$):

$$\nabla_U V = \mathrm{Proj}(\bar{\nabla}_{\bar{U}} \bar{V}). \tag{8.24}$$

$\bar{U}, \bar{V}$ 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 $u \in \mathrm{T}_x\mathcal{M}$:

$$\nabla_u V = \mathrm{Proj}_x(\bar{\nabla}_u \bar{V}). \tag{8.25}$$

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 $u \in \mathrm{T}_x\mathcal{M} \subseteq \mathrm{T}_x\overline{\mathcal{M}}$,

$$\mathrm{Hess} f(x)[u] = \nabla_u \mathrm{grad} f = \mathrm{Proj}_x\big(\bar{\nabla}_u \bar{G}\big). \tag{8.26}$$

The Weingarten map $\mathcal{W}$ (5.30) and second fundamental form II (5.31) are also defined in the general context of Riemannian submanifolds of Riemannian manifolds, so that expressions (5.33) and (5.34) for the Hessian still hold.

A similarly simple expression is valid for covariant derivatives of vector fields along curves, in analogy to (5.19):

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \mathrm{Proj}_{c(t)}\left(\frac{\bar{\mathrm{D}}}{\mathrm{d}t} Z(t)\right), \tag{8.27}$$

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):

$$c''(t) = \mathrm{Proj}_{c(t)}(\ddot{c}(t)). \tag{8.28}$$

Section 5.11 provides pointers regarding normal parts of $\bar{\nabla}_{\bar{U}}\bar{V}$ and $\ddot{c}$.

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 the note on pp21–22). 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 the basics of topology. The relevance of the topological conditions imposed in Section 8.2 for optimization is spelled out in [AMS08, §3.1.2].

Embedded submanifolds are called *regular submanifolds* by Brickell and Clark, and simply *submanifolds* by O'Neill. Furthermore, we mean Riemannian submanifolds to be embedded, whereas Lee allows them to be immersed as well, pointing out when it is necessary for them to be embedded.

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].

9
Quotient manifolds

The Grassmannian $\mathrm{Gr}(n,p)$ is the set of linear subpsaces 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

$$\max_{[X] \in \mathrm{St}(n,p)/\sim} f([X]), \qquad\qquad f([X]) = \bar{f}(X). \tag{9.4}$$

Crucially, f 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 $[X_0], [X_1], \ldots$, also actually generates 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, $\mathrm{Proj}_X^{\mathrm{St}}$ (7.25), the gradient is given by

$$\frac{1}{2}\mathrm{grad}\bar{f}(X) = \mathrm{Proj}_X^{\mathrm{St}}(AX) \tag{9.6}$$
$$= (I_n - XX^\top)AX + X\frac{X^\top AX - X^\top AX}{2} = (I_n - XX^\top)AX.$$

(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

$$\forall X \in \mathrm{St}(n,p), Q \in \mathrm{O}(p), \qquad \mathrm{grad}\bar{f}(XQ) = \mathrm{grad}\bar{f}(X)\cdot Q.$$

This statement relies on the fact: $V \in \mathrm{T}_X\mathrm{St}(n,p) \implies VQ \in \mathrm{T}_{XQ}\mathrm{St}(n,p)$.

Hence, if the retraction has the property 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 (7.22) can be written as

$$\mathrm{R}_X^{\mathrm{pol}}(V) = (X+V)(I_p + V^\top V)^{-1/2},$$

so that

$$\mathrm{R}_{XQ}^{\mathrm{pol}}(VQ) = (X+V)Q\cdot\left(Q^\top\left[I_p + V^\top V\right]Q\right)^{-1/2} = \mathrm{R}_X^{\mathrm{pol}}(V)\cdot Q. \tag{9.9}$$

Also, the QR retraction (7.20) 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,

$\mathrm{R}^{\mathrm{QR}}_{XQ}(VQ)$ 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?

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 quotients 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 mentioned for gradient descent (Section 4.6), Newton's method (Theorem 6.4) and trust regions (Section 6.6): 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 9.38, 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), but it does provide the stronger theoretical perspective, aptly explaining why we do get fast local convergence.

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 (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 $x \in \overline{\mathcal{M}}$.*

Owing to the rank condition, π is called a *submersion*. By construction, it is also surjective.

As an exercise, one can show that the projective space RP^{n-1} with smooth structure as in Example 8.10 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.

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.*

[BC70, Prop. 6.3.1], [Lee12, Thm. 4.31]

This statement should be compared to Theorem 8.69 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 8.71 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 8.71 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 9.7. 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 (7.30) *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:*

$$0 = c'(0) = \mathrm{D}\pi(\bar{c}(0))[\bar{c}'(0)] = \mathrm{D}\pi(X)[X\Omega].$$

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 what is called the *rank theorem* in differential geometry.

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,*

[Lee12, Thm. 4.12] [BC70, Prop. 6.1.1]

$$\tilde{\pi} = \varphi \circ \pi \circ \bar{\varphi}^{-1} \colon \bar{\varphi}(\overline{\mathcal{U}}) \subseteq \mathbb{R}^{n+k} \to \varphi(\mathcal{U}) \subseteq \mathbb{R}^n, \tag{9.13}$$

is simply the function $\tilde{\pi}(z_1, \ldots, z_{n+k}) = (z_1, \ldots, z_n)$.

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 $\overline{\mathcal{U}}$.

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.

Exercise 9.6. *Show that the projective space RP^{n-1} with the smooth structure of Example 8.10 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 9.7. *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 this with the introduction, where $\overline{\mathcal{M}} = \mathrm{St}(n, p)$.)*

Exercise 9.8. *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 $\mathcal{M}$. Deduce that, for any function $f \colon \mathcal{M} \to \mathbb{R}$, a point $x \in \overline{\mathcal{M}}$ is a local minimizer for $\bar{f} = f \circ \pi$ if and only if $\pi(x)$ is a local minimizer for f.*

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 9.16). 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 9.19).

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 9.9. *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 8.27).*

Example 9.10. *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 9.11. *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 9.12. *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. (A similar definition holds 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 9.13. *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 9.14. *A group action θ is* free *if only the identity fixes any x. Explicitly: a left action is free if $\theta(g, x) = x \implies g = e$ for all $x \in \overline{\mathcal{M}}$.*

Definition 9.15. *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, any compact subset of $\overline{\mathcal{M}} \times \overline{\mathcal{M}}$ maps to a compact subset of $\mathcal{G} \times \overline{\mathcal{M}}$ through ϑ^{-1}. The definition is similar for right actions.

Recall that a set S is *compact* if, given any collection of open sets whose union contains S (that is, an *open cover*), one can select a finite number of open sets from that collection whose union still contains S (called a *finite subcover*).

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. If not, see Section 9.16 for some pointers, specifically Proposition 9.58.

Proposition 9.16. *Every smooth action by a compact Lie group is proper.*

[Lee12, Cor. 21.6]

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]. If the action is not free, then different orbits could have different dimensions, which is impossible for a quotient manifold.

We can now state the main theorem of this section. This is our tool of choice to identify quotient manifolds.

Theorem 9.17. *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 $\dim \mathcal{G}$.*

[Lee12, Thm. 21.10]

Example 9.18. *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 9.16, 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 9.17 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 9.7 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 9.19. *In each item below, a Lie group $\mathcal{G}$ (recall Example 9.10) 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.*

1. $\overline{\mathcal{M}} = \mathrm{SO}(n)^k$, $\mathcal{G} = \mathrm{SO}(n)$, $\theta((R_1,\ldots,R_k),Q) = (R_1Q,\ldots,R_kQ)$.
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. $\overline{\mathcal{M}} = \mathbb{R}_d^{d\times n}$, $\mathcal{G} = \mathrm{SE}(d)$, $\theta((R,t),X) = RX + t\mathbf{1}^\top$.

Recall that $\mathbb{R}_k^{m\times n}$ is the set of matrices of size $m \times n$ and rank k.

Describe the equivalence classes and the significance of the quotient manifolds.

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 9.20. *Given a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with canonical projection π and any manifold $\mathcal{N}$, a map $F\colon \mathcal{M} \to \mathcal{N}$ is smooth if and only if $\bar{F} = F \circ \pi\colon \overline{\mathcal{M}} \to \mathcal{N}$ is smooth.*

[Lee12, Thm. 4.29]

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 (we omit a proof).

$$\begin{tikzcd} \overline{\mathcal{M}} \arrow[d, "\pi"'] \arrow[dr, "\bar{F} = F \circ \pi"] & \\ \mathcal{M} \arrow[r, "F"'] & \mathcal{N} \end{tikzcd}$$

Proposition 9.21. *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 $S \colon \mathcal{U} \to \overline{\mathcal{M}}$ (called a local section) such that $\pi \circ S$ is the identity map on $\mathcal{U}$ and $S([x]) = x$.*

[Lee12, Thm. 4.26] [BC70, Prop. 6.1.4]

Proof of Theorem 9.20. 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 9.21 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. □

We note another result, this one about maps *into* quotient manifolds.

Proposition 9.22. *Let $\bar{F} \colon \mathcal{N} \to \overline{\mathcal{M}}$ 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 $F = \pi \circ \bar{F} \colon \mathcal{N} \to \mathcal{M}$ is smooth. The other way around, if $F \colon \mathcal{N} \to \mathcal{M}$ 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 $\bar{F} \colon F^{-1}(\mathcal{U}) \to \overline{\mathcal{M}}$.*

$$\begin{tikzcd} \mathcal{N} \arrow[r, "\bar{F}"] \arrow[dr, "F"'] & \overline{\mathcal{M}} \arrow[d, "\pi"] \\ & \mathcal{M} \end{tikzcd}$$

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 9.21: 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 $[x] \in \mathcal{M}$ itself is an equivalence class, and a tangent vector $\xi \in \mathrm{T}_{[x]}\mathcal{M}$ is an equivalence class of smooth curves on $\mathcal{M}$ passing through $[x]$. 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 $\mathrm{T}_x\overline{\mathcal{M}}$, which naturally suggests that we also consider the orthogonal complement of V_x.

Technically, the Riemannian metric on $\overline{\mathcal{M}}$ 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 $\mathrm{T}_x\overline{\mathcal{M}}$. However, this leaves a lot of freedom that we do not need. We opt for this more directive definition of horizontal space, while noting that other authors use the terminology in a broader sense [AMS08, §3.5.8].

Definition 9.23. *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:*

$$\mathrm{H}_x = (\mathrm{V}_x)^\perp = \left\{u \in \mathrm{T}_x\overline{\mathcal{M}} : \langle u, v\rangle_x = 0 \text{ for all } v \in \mathrm{V}_x\right\}.$$

Then, $\mathrm{T}_x\overline{\mathcal{M}} = \mathrm{V}_x \oplus \mathrm{H}_x$ is a direct sum of linear spaces.

We can now make the following claim: 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. As such, 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 9.24. *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*

$$u = (\mathrm{D}\pi(x)|_{\mathrm{H}_x})^{-1}[\xi] = \mathrm{lift}_x(\xi). \tag{9.15}$$

These 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 9.25. *In Example 9.18 we determined the tangent spaces of fibers of* $\mathcal{M} = \mathrm{St}(n,p)/\mathrm{O}(p)$*, so that*

$$\mathrm{V}_X = \{X\Omega : \Omega \in \mathrm{Skew}(p)\}.$$

With the usual Riemannian metric on $\mathrm{St}(n,p)$*, namely,* $\langle U, V\rangle_X = \mathrm{Tr}(U^\top V)$*, it follows that*

$$\mathrm{H}_X = \{U \in \mathrm{T}_X\mathrm{St}(n,p) : \langle U, X\Omega\rangle_X = 0 \textit{ for all } \Omega \in \mathrm{Skew}(p)\}.$$

In this definition, we conclude that $X^\top U$ *must be symmetric. Yet, from (7.15) 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 this other smooth curve on $\mathrm{St}(n, p)$*: $\tilde{c}(t) = \bar{c}(t)Q$. Clearly, $\tilde{c}(0) = XQ = Y$ and $\tilde{c}'(0) = \bar{c}'(0)Q = U_X Q$. Since by construction $\pi \circ \bar{c}$ and $\pi \circ \tilde{c}$ are the* same *curve on $\mathcal{M}$, we may conclude that*

$$\xi = (\pi \circ \tilde{c})'(0) = \mathrm{D}\pi(\tilde{c}(0))[\tilde{c}'(0)] = \mathrm{D}\pi(Y)[U_X Q].$$

Crucially, $U_X Q$ is a horizontal vector at Y. Uniqueness of horizontal lifts then tells us that $U_Y = U_X Q$ is the horizontal lift of ξ at Y, that is,

$$\mathrm{lift}_{XQ}(\xi) = \mathrm{lift}_X(\xi) \cdot Q. \tag{9.19}$$

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*, 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 horizontal vector field $\bar{V}$ 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.

Theorem 9.26. *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:*

$$V \circ \pi = \mathrm{D}\pi \circ \bar{V}. \tag{9.22}$$

Furthermore, V is smooth on $\mathcal{M}$ if and only if $\bar{V}$ is smooth on $\overline{\mathcal{M}}$.

$$\begin{tikzcd} \overline{\mathcal{M}} \arrow[r, "\bar{V}"] \arrow[d, "\pi"'] & \mathrm{T}\overline{\mathcal{M}} \arrow[d, "\mathrm{D}\pi"] \\ \mathcal{M} \arrow[r, "V"] & \mathrm{T}\mathcal{M} \end{tikzcd}$$

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 (8.11).

Proposition 9.27. *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 $\mathbb{R}^n$):*

$$\mathrm{D}\pi(x)[\bar{W}_i(x)] = \begin{cases} (\mathrm{D}\varphi(\pi(x)))^{-1}[e_i] & \text{if } i \in \{1, \ldots n\}, \\ 0 & \text{if } i \in \{n+1, \ldots, n+k\}. \end{cases}$$

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 (8.11):

$$\bar{W}_i(\bar{\varphi}^{-1}(z)) = \mathrm{D}\bar{\varphi}^{-1}(z)[\bar{e}_i], \tag{9.23}$$

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

$$\begin{aligned} (\dot{z}_1, \ldots, \dot{z}_n) &= \mathrm{D}\tilde{\pi}(z)[\dot{z}] \\ &= \mathrm{D}\varphi(\pi(\bar{\varphi}^{-1}(z))) \left[\mathrm{D}\pi(\bar{\varphi}^{-1}(z)) \left[\mathrm{D}\bar{\varphi}^{-1}(z)[\dot{z}]\right]\right]. \end{aligned}$$

Introducing the notation $x = \bar{\varphi}^{-1}(z)$, the expression simplifies to:

$$(\dot{z}_1, \ldots, \dot{z}_n) = \mathrm{D}\varphi(\pi(x)) \left[\mathrm{D}\pi(x) \left[\mathrm{D}\bar{\varphi}^{-1}(z)[\dot{z}]\right]\right]. \tag{9.24}$$

In particular, for $\dot{z} = \bar{e}_i$ we recognize the coordinate vector fields as in (9.23) so that

$$\mathrm{D}\varphi(\pi(x)) \left[\mathrm{D}\pi(x) \left[\bar{W}_i(x)\right]\right] = \begin{cases} e_i & \text{if } i \in \{1, \ldots, n\}, \\ 0 & \text{if } i \in \{n+1, \ldots, n+k\}. \end{cases}$$

To conclude, note that $\mathrm{D}\varphi(\pi(x))$ is invertible since φ is a chart. □

We can now give a proof for our main result characterizing smoothness of vector fields on quotient manifolds.

Proof of Theorem 9.26. Equation (9.22) follows from the definition of $\bar{V}$ (9.21) and from the properties of horizontal lifts (9.16). Using Theorem 9.20 then equation (9.22), we find these 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 8.40, 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 9.27 and defined on some neighborhood $\overline{\mathcal{U}}$ of x'. By Corollary 8.47, 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 9.27, 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 8.39), 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, recall that, in defining horizontal spaces, we invoked a Riemannian metric $\langle \cdot, \cdot \rangle_x$ on $\overline{\mathcal{M}}$. We use it here too, noting once again that this structure is not strictly necessary. 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 $\bar{V}$ is smooth at an arbitrary x', hence $\bar{V}$ is smooth. □

In light of the above result, actions of smooth vector fields on smooth functions on the quotient manifold are easily understood in the total space.

Proposition 9.28. *For a quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with canonical projection π, consider $V \in \mathfrak{X}(\mathcal{M})$ and $f \in \mathfrak{F}(\mathcal{M})$ 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 (8.10), 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 9.26 is that we can construct local frames (Definition 3.58) 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 9.32 below.

Proposition 9.29. *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 $\pi(\overline{\mathcal{U}})$.

That the local frame is orthonormal means that, for all x in the domain,

$$\langle \hat{W}_i(x), \hat{W}_j(x) \rangle_x = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

See also Section 9.7 for a comment regarding orthonormality of $W_1, \ldots, W_n$.

Proof. By Proposition 8.46, the coordinate vector fields $\bar{W}_1, \ldots, \bar{W}_{n+k}$ provided by Proposition 9.27 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. □

We can use this last proposition to show that the lift operator is smooth. To make sense of this statement, we resort to local sections (Proposition 9.21).

Proposition 9.30. *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 9.29, 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 8.39), 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 9.31. *Work out the details for the proof of Proposition 9.29.*

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 $\mathcal{M}$ 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 9.32. *If the retraction $\overline{\mathrm{R}}$ on the total space $\overline{\mathcal{M}}$ satisfies* [AMS08, Prop. 4.1.3]

$$x \sim y \implies \overline{\mathrm{R}}_x(\mathrm{lift}_x(\xi)) \sim \overline{\mathrm{R}}_y(\mathrm{lift}_y(\xi)) \tag{9.28}$$

for all $([x], \xi) \in \mathrm{T}\mathcal{M}$, then (9.27) defines a retraction R on $\mathcal{M} = \overline{\mathcal{M}}/\sim$.

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 $\zeta \in \mathrm{T}_{[x]}\mathcal{M}$,

$$\mathrm{DR}_{[x]}(0)[\zeta] = \mathrm{D}\pi(x)\left[\mathrm{D}\overline{\mathrm{R}}_x(0)[\mathrm{Dlift}_x(0)[\zeta]]\right] = \zeta,$$

where we used $\mathrm{Dlift}_x(0) = \mathrm{lift}_x$ since it is a linear operator, $\mathrm{D}\overline{\mathrm{R}}_x(0)$ 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 operator on $\mathrm{T}_{[x]}\mathcal{M}$.

To verify smoothness, invoke Proposition 9.30 to select a local section $S\colon \mathcal{U} \to \overline{\mathcal{M}}$; then, $\mathrm{R}_{[x]}(\xi) = \pi\Big(\overline{\mathrm{R}}_{S([x])}\Big(\mathrm{lift}_{S([x])}(\xi)\Big)\Big)$ 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 $\mathrm{T}\mathcal{M}$. □

Example 9.33. *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 9.32. 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 $\mathrm{St}(n, p)/\mathrm{O}(p)$:

$$\mathrm{R}^{\mathrm{QR}}_{[X]}(\xi) = [\mathrm{qfactor}(X + \mathrm{lift}_X(\xi))]\,, \text{ and} \tag{9.29}$$

$$\mathrm{R}^{\mathrm{pol}}_{[X]}(\xi) = [\mathrm{pfactor}(X + \mathrm{lift}_X(\xi))]\,, \tag{9.30}$$

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 (7.20) and (7.22).*

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 $\langle \xi, \zeta \rangle_{[x]}$ may 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

$$x \sim y \implies \langle \mathrm{lift}_x(\xi), \mathrm{lift}_x(\zeta) \rangle_x = \left\langle \mathrm{lift}_y(\xi), \mathrm{lift}_y(\zeta) \right\rangle_y. \tag{9.32}$$

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 8.50 that a metric is Riemannian if for every two 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 $\overline{\mathcal{M}}$:

$$\bar{f}(x) = \langle \bar{U}(x), \bar{V}(x) \rangle_x = \langle \mathrm{lift}_x(U([x])), \mathrm{lift}_x(V([x])) \rangle_x = f([x]).$$

The function $\bar{f}$ is smooth since $\bar{U}$ and $\bar{V}$ are smooth by Theorem 9.26. Furthermore, $\bar{f} = f \circ \pi$, which shows f is smooth by Theorem 9.20. This discussion yields the following result, which doubles as a definition of *Riemannian quotient manifold*. Notice also how, by construction, lifting commutes with taking inner products, that is,

The term "Riemannian quotient manifold" is coined in [AMS08, p53].

$$\forall U, V \in \mathfrak{X}(\mathcal{M}), \qquad \langle U, V \rangle \circ \pi = \langle \mathrm{lift}(U), \mathrm{lift}(V) \rangle. \tag{9.33}$$

Theorem 9.34. *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 9.29 are orthonormal.

Example 9.35. *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 $Q \in \mathrm{O}(p)$,

$$\begin{aligned} \left\langle \mathrm{lift}_{XQ}(\xi), \mathrm{lift}_{XQ}(\zeta) \right\rangle_{XQ} &= \langle UQ, VQ \rangle_{XQ} = \mathrm{Tr}((UQ)^\top (VQ)) \\ &= \mathrm{Tr}(U^\top V) = \langle U, V \rangle_X = \langle \mathrm{lift}_X(\xi), \mathrm{lift}_X(\zeta) \rangle_X. \end{aligned}$$

This confirms that the tentative metric is invariant under the choice of lifting point: condition (9.32) is fulfilled, thus Theorem 9.34 tells us this is a Riemannian metric on the quotient manifold, turning it into a Riemannian quotient manifold.

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 8.52. Being a smooth vector field on the quotient manifold, the gradient is an abstract object. Prompted by the discussions of the last 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 9.34, this can be done rather easily, as we now show.

Consider $f : \mathcal{M} \to \mathbb{R}$ 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:

$$\forall([x], \xi) \in \mathrm{T}\mathcal{M}, \qquad \mathrm{D}f([x])[\xi] = \langle \operatorname{grad} f([x]), \xi \rangle_{[x]}.$$

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:

$$\forall(x, u) \in \mathrm{T}\overline{\mathcal{M}}, \qquad \mathrm{D}\bar{f}(x)[u] = \left\langle \operatorname{grad} \bar{f}(x), u \right\rangle_x.$$

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:

$$\begin{aligned} \forall u \in \mathrm{T}_x\overline{\mathcal{M}}, \qquad \left\langle \operatorname{grad} \bar{f}(x), u \right\rangle_x &= \mathrm{D}\bar{f}(x)[u] \\ &= \mathrm{D}f(\pi(x))[\mathrm{D}\pi(x)[u]] \\ &= \langle \operatorname{grad} f([x]), \mathrm{D}\pi(x)[u] \rangle_{[x]}. \end{aligned}$$

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), we get

$$\begin{aligned} \forall u \in \mathrm{H}_x, \qquad \left\langle \operatorname{grad} \bar{f}(x), u \right\rangle_x &= \langle \operatorname{grad} f([x]), \mathrm{D}\pi(x)[u] \rangle_{[x]} \\ &= \langle \operatorname{lift}_x(\operatorname{grad} f([x])), u \rangle_x. \end{aligned}$$

This tells us that the horizontal part of $\operatorname{grad} \bar{f}(x)$ is equal to the lift of $\operatorname{grad} f([x])$ at x. What about the vertical part? That one is necessarily zero, owing to the fact that $\bar{f}$ is constant along fibers. Indeed,

$$\forall v \in \mathrm{V}_x, \quad \left\langle \operatorname{grad} \bar{f}(x), v \right\rangle_x = \mathrm{D}\bar{f}(x)[v] = \mathrm{D}f([x])[\mathrm{D}\pi(x)[v]] = 0,$$

since $\mathrm{D}\pi(x)[v] = 0$ for $v \in \mathrm{V}_x$. This leads to a simple conclusion.

Proposition 9.36. *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*

$$\operatorname{lift}_x(\operatorname{grad} f([x])) = \operatorname{grad} \bar{f}(x), \tag{9.34}$$

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 need only compute the gradient of the lifted function, $\bar{f} = f \circ \pi$. In other words: taking gradients commutes with lifting. Compare with (8.23) for Riemannian submanifolds.

Example 9.37. *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 9.36 to relate it to the gradient of f on the Riemannian quotient manifold, we find:

$$\mathrm{lift}_X(\mathrm{grad} f([X])) = \mathrm{grad}\bar{f}(X) = 2(I_n - XX^\top)AX. \tag{9.35}$$

Notice how the gradient of $\bar{f}$ is necessarily horizontal: comparing with the explicit description of the horizontal and vertical spaces given in Example 9.25, 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

$$[x_{k+1}] = \mathrm{R}_{[x_k]}(-\alpha_k \mathrm{grad} f([x_k])) \tag{9.36}$$

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.

Having 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 9.36 conveniently tells us that

$$\mathrm{lift}_{x_k}(\mathrm{grad} f([x_k])) = \mathrm{grad}\bar{f}(x_k), \tag{9.37}$$

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 (9.36) we deduce that

$$\begin{aligned} [x_{k+1}] &= \mathrm{R}_{[x_k]}(-\alpha_k \mathrm{grad} f([x_k])) \\ &= \left[\overline{\mathrm{R}}_{x_k}(-\alpha_k \, \mathrm{lift}_{x_k}(\mathrm{grad} f([x_k])))\right] \\ &= \left[\overline{\mathrm{R}}_{x_k}(-\alpha_k \mathrm{grad} \bar{f}(x_k))\right]. \end{aligned} \tag{9.38}$$

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

$$x_{k+1} = \overline{\mathrm{R}}_{x_k}(-\alpha_k \mathrm{grad} \bar{f}(x_k)) \tag{9.39}$$

on $\overline{\mathcal{M}}$ is *equivalent* to running the abstract iteration (9.36) on $\mathcal{M}$. Interestingly, iteration (9.39) 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 (9.39) to be a proper instantiation of (9.36), we must make sure that the step-size α_k is chosen as a function of $[x_k]$ but not as a function of 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:

$$\forall t, \qquad \bar{f}(\overline{\mathrm{R}}_x(-t\mathrm{grad}\bar{f}(x))) = f\Big(\mathrm{R}_{[x]}(-t\mathrm{grad} f([x]))\Big).$$

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 9.38. *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 these identities as in (5.20):

$$\mathrm{Hess}\bar{f}(x)[v] = \bar{\nabla}_v \mathrm{grad}\bar{f} = \frac{\bar{\mathrm{D}}}{\mathrm{d}t}\mathrm{grad}\bar{f}(\bar{c}(t))\bigg|_{t=0}.$$

By Proposition 9.36, 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

$$\operatorname{grad}\bar{f}(\bar{c}(t)) = \operatorname{lift}_{\bar{c}(t)}(\operatorname{grad}f([\bar{c}(t)])) = \operatorname{lift}_{\bar{c}(t)}(\operatorname{grad}f([x])) = 0.$$

Thus, we conclude that $\operatorname{Hess}\bar{f}(x)[v] = 0$. □

Thus, the standard theory does not predict fast local convergence for RGD on $\bar{f}$.

The good news is: the trivial eigenvalues associated to vertical directions do not appear in the spectrum of the Hessian of f on the quotient manifold (see Exercise 9.43). 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 on the quotient manifold.

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 8.58 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 this formula:

$$\operatorname{lift}(\nabla_U V) = \operatorname{Proj}^{\mathrm{H}}\left(\bar{\nabla}_{\bar{U}}\bar{V}\right), \tag{9.40}$$

where $\operatorname{lift}\colon \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\overline{\mathcal{M}})$ returns the horizontal lift of a vector field (9.21), $\bar{U} = \operatorname{lift}(U)$, $\bar{V} = \operatorname{lift}(V)$, and $\operatorname{Proj}^{\mathrm{H}}\colon \mathfrak{X}(\overline{\mathcal{M}}) \to \mathfrak{X}(\overline{\mathcal{M}})$ orthogonally projects each tangent vector to the horizontal space at its base.

The proof of this statement is based on the Koszul formula (5.10), which we first encountered in the proof of Theorem 5.4. Recall that this formula completely characterizes the Riemannian connection in terms of the Riemannian metric and Lie brackets: for all $U, V, W \in \mathfrak{X}(\mathcal{M})$,

$$\begin{aligned} 2\left\langle \nabla_U V, W\right\rangle = U\left\langle V, W\right\rangle + V\left\langle W, U\right\rangle - W\left\langle U, V\right\rangle \\ - \left\langle U, [V,W]\right\rangle + \left\langle V, [W,U]\right\rangle + \left\langle W, [U,V]\right\rangle. \end{aligned}$$

To make progress, we must first understand how Lie brackets on the quotient manifold are related to Lie brackets of horizontal lifts.

Proposition 9.39. *For any two smooth vector fields $U, V \in \mathfrak{X}(\mathcal{M})$ and their horizontal lifts $\bar{U}, \bar{V} \in \mathfrak{X}(\overline{\mathcal{M}})$,*

$$\mathrm{lift}([U, V]) = \mathrm{Proj}^{\mathrm{H}}([\bar{U}, \bar{V}]). \tag{9.41}$$

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 $\bar{f} = f \circ \pi$:

$$(Vf) \circ \pi = \bar{V}\bar{f}, \qquad \langle U, V \rangle \circ \pi = \langle \bar{U}, \bar{V} \rangle. \tag{9.42}$$

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. □

With $\bar{U}, \bar{V}, \bar{W}$ denoting the horizontal lifts of U, V and W, using (9.41) and (9.42) several times we find that:

$$\begin{aligned}
&(U \langle V, W \rangle) \circ \pi = \bar{U}(\langle V, W \rangle \circ \pi) = \bar{U} \left\langle \bar{V}, \bar{W} \right\rangle, \text{ and} \\
&\langle U, [V, W] \rangle \circ \pi = \langle \bar{U}, \mathrm{lift}([V, W]) \rangle = \left\langle \bar{U}, [\bar{V}, \bar{W}] \right\rangle,
\end{aligned}$$

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 shows 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 (9.40) holds, as announced. This discussion warrants the following theorem.

Theorem 9.40. *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* [O'N83, Lem. 7.45]

$$\mathrm{lift}(\nabla_U V) = \mathrm{Proj}^{\mathrm{H}}\left(\bar{\nabla}_{\bar{U}} \bar{V}\right)$$

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 (8.24) for Riemannian submanifolds.

Exercise 9.41. *Show that* [GHL04, Prop. 3.35], [dC92, Ex. 8.9]

$$\bar{\nabla}_{\bar{U}} \bar{V} = \mathrm{lift}(\nabla_U V) + \frac{1}{2}\mathrm{Proj}^{\mathrm{V}}([\bar{U}, \bar{V}]),$$

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*

Given a smooth function f on a Riemannian quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$, the Hessian of f is defined for $([x], \xi) \in \mathrm{T}\mathcal{M}$ by

$$\mathrm{Hess} f([x])[\xi] = \nabla_\xi \mathrm{grad} f. \tag{9.43}$$

For any vector field $V \in \mathfrak{X}(\mathcal{M})$, Theorem 9.40 tells us that

$$\mathrm{lift}_x(\nabla_\xi V) = \mathrm{Proj}_x^{\mathrm{H}}(\bar{\nabla}_u \bar{V}), \tag{9.44}$$

where $u = \mathrm{lift}_x(\xi)$ and $\bar{V} = \mathrm{lift}(V)$. Recall from Proposition 9.36 that

$$\mathrm{lift}(\mathrm{grad} f) = \mathrm{grad} \bar{f}, \tag{9.45}$$

with $\bar{f} = f \circ \pi$. Combining, we find that

$$\mathrm{lift}_x(\mathrm{Hess} f([x])[\xi]) = \mathrm{Proj}_x^{\mathrm{H}}\left(\bar{\nabla}_u \mathrm{grad} \bar{f}\right). \tag{9.46}$$

Finally, since $\bar{\nabla}_u \mathrm{grad} \bar{f} = \mathrm{Hess} \bar{f}(x)[u]$, we get the following result.

Proposition 9.42. *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*

$$\mathrm{lift}_x(\mathrm{Hess} f([x])[\xi]) = \mathrm{Proj}_x^{\mathrm{H}}\left(\mathrm{Hess} \bar{f}(x)[u]\right), \tag{9.47}$$

for all $x \in \overline{\mathcal{M}}$ and $\xi \in \mathrm{T}_{[x]}\mathcal{M}$, with $u = \mathrm{lift}_x(\xi)$.

See Example 9.46 for an illustration. We conclude with an exercise linking first- and second-order critical points on the total space and on the quotient space.

Exercise 9.43. *Let $f : \mathcal{M} \to \mathbb{R}$ be smooth on a Riemannian quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ with canonical projection π, and let $\bar{f} = f \circ \pi$. From Proposition 9.36, it is clear 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.*

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 9.38.) Deduce that x is a second-order critical point for $\bar{f}$ if and only if $[x]$ is second-order critical for f.

9.12 A word about Riemannian Newton's method

We describe 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

$$[x_{k+1}] = \mathrm{R}_{[x_k]}(\xi_k), \tag{9.48}$$

where $\xi_k \in \mathrm{T}_{[x_k]}\mathcal{M}$ is the solution of the linear equation

$$\mathrm{Hess}f([x_k])[\xi_k] = -\mathrm{grad}f([x_k]), \tag{9.49}$$

which we assume to be unique. In the spirit of Section 9.9, we now discuss how to run (9.48) in practice.

Let us assume that $\mathcal{M}$ is a Riemannian quotient manifold with canonical projection π. We lift both sides of (9.49) to the horizontal space at x_k. With $s_k = \mathrm{lift}_{x_k}(\xi_k)$, Propositions 9.36 and 9.42 tell us that the linear equation is equivalent to

$$\mathrm{Proj}^{\mathrm{H}}_{x_k}\left(\mathrm{Hess}\bar{f}(x_k)[s_k]\right) = -\mathrm{grad}\bar{f}(x_k), \tag{9.50}$$

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 operators 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

$$\mathrm{Proj}^{\mathrm{H}}_{x_k}\left(\mathrm{Hess}\bar{f}(x_k)\left[\mathrm{Proj}^{\mathrm{H}}_{x_k}(s_k)\right]\right) = -\mathrm{grad}\bar{f}(x_k). \tag{9.51}$$

This is a linear system with symmetric operator

$$\mathrm{Proj}^{\mathrm{H}}_{x_k} \circ \mathrm{Hess}\bar{f}(x_k) \circ \mathrm{Proj}^{\mathrm{H}}_{x_k}. \tag{9.52}$$

By construction, if (9.49) has a unique solution ξ_k, then (9.51) has a unique horizontal solution $s_k = \mathrm{lift}_{x_k}(\xi_k)$. (If we solve (9.51) in the whole tangent space $\mathrm{T}_{x_k}\overline{\mathcal{M}}$, then all solutions are of the form $s_k + v$ with $v \in \mathrm{V}_{x_k}$, 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 (9.48) we see that

$$[x_{k+1}] = \mathrm{R}_{[x_k]}(\xi_k) = \left[\overline{\mathrm{R}}_{x_k}(\mathrm{lift}_{x_k}(\xi_k))\right] = \left[\overline{\mathrm{R}}_{x_k}(s_k)\right]. \tag{9.53}$$

In summary, to run Newton's method on f in practice, we may iterate

$$x_{k+1} = \overline{\mathrm{R}}_{x_k}(s_k) \tag{9.54}$$

with $s_k \in \mathrm{H}_{x_k}$ the horizontal solution of (9.51) (unique if and only if the solution of (9.49) 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 (9.54) but with s_k a solution of this linear system over $\mathrm{T}_{x_k}\overline{\mathcal{M}}$ (if one exists):

$$\mathrm{Hess}\bar{f}(x_k)[s_k] = -\mathrm{grad}\bar{f}(x_k). \tag{9.55}$$

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 9.38 tells us that $\mathrm{Hess}\bar{f}(x_k)$ converges to a singular operator. Thus, we must expect difficulties in solving the linear system on the total space. (However, see Exercise 9.45 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 9.44. *In light of eq. (9.51), consider the linear system $Hs = b$ with*

$$H = \mathrm{Proj}_x^{\mathrm{H}} \circ \mathrm{Hess}\bar{f}(x) \circ \mathrm{Proj}_x^{\mathrm{H}} \qquad \text{and} \qquad b = -\mathrm{grad}\bar{f}(x)$$

defined at some point $x \in \overline{\mathcal{M}}$. Show that the eigenvalues of H (a self-adjoint operator 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 9.38, 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 behave when applied to the system $Hs = b$, especially minding horizontality. How many iterations need to be run at most?

CG is Algorithm 6.2.

Exercise 9.45. *Continuing from Exercise 9.44, 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. Consider the Newton system for $\bar{f}$ at a point x near $\tilde{x}$ on the total space, ignoring the quotient structure:*

$$\mathrm{Hess}\bar{f}(x)[s] = -\mathrm{grad}\bar{f}(x), \qquad\qquad s \in \mathrm{T}_x\overline{\mathcal{M}}. \tag{9.56}$$

From Lemma 9.38, 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 (9.56) exactly with a standard algorithm such as Gaussian elimination can lead to catastrophic failure when x is close to $\tilde{x}$.

However, it is much more common to (try to) solve (9.56) with CG. A typical observation then would be that roughly $\dim\mathcal{M}$ iterations of CG on (9.56) *are fairly well behaved. The next iteration would break CG. Nevertheless, since the iterates of CG are monotonically improving approximate solutions of* (9.56)*, it would be reasonable to return the best solution reached so far.*[1] *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* 9.44*).*

Explain this observation. (Hint: use local frames as in Proposition 9.29*, use Lemma* 9.38 *at* $\tilde{x}$*, show that* $\left\langle v, \mathrm{Hess}\bar{f}(x)[v]\right\rangle_x = 0$ *for all* $x \in \overline{\mathcal{M}}$ *and* $v \in \mathrm{V}_x$*, and track the Krylov space implicitly generated by CG.)*

[1] Matlab's implementation of CG, `pcg`, does so.

Exercise 9.45 loosely explains 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 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.

9.13 *Total space embedded in a linear space*

For all quotient manifolds described in Exercise 9.19, 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 $\overline{\mathcal{M}}$. 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 (3.37) that

$$\mathrm{grad}\bar{f}(x) = \mathrm{Proj}_x\left(\mathrm{grad}\bar{\bar{f}}(x)\right), \tag{9.57}$$

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 9.36 tells us that

$$\mathrm{lift}_x(\mathrm{grad}f([x])) = \mathrm{Proj}_x\left(\mathrm{grad}\bar{\bar{f}}(x)\right). \tag{9.58}$$

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:

$$\mathrm{lift}_x(\mathrm{grad} f([x])) = \mathrm{Proj}_x^{\mathrm{H}}\left(\mathrm{grad}\bar{\bar{f}}(x)\right), \tag{9.59}$$

where $\mathrm{Proj}_x^{\mathrm{H}}$ is the orthogonal projector from $\mathcal{E}$ to 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 $\mathrm{H}_x \subseteq \mathrm{T}_x\overline{\mathcal{M}}$, and apply to (9.58).

Similarly, we can express the Hessian of f in terms of $\bar{\bar{f}}$. Indeed, Proposition 9.42 states

$$\mathrm{lift}_x(\mathrm{Hess} f([x])[\xi]) = \mathrm{Proj}_x^{\mathrm{H}}(\mathrm{Hess}\bar{f}(x)[u]) \tag{9.60}$$

with $u = \mathrm{lift}_x(\xi)$. Moreover, we know from connections on Riemannian submanifolds (5.4) that

$$\mathrm{Hess}\bar{f}(x)[u] = \bar{\nabla}_u \mathrm{grad}\bar{f}(x) = \mathrm{Proj}_x\left(\mathrm{D}\bar{\bar{G}}(x)[u]\right), \tag{9.61}$$

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 $\mathrm{Proj}_x\left(\mathrm{grad}\bar{\bar{f}}(x)\right)$ or $\mathrm{Proj}_x^{\mathrm{H}}\left(\mathrm{grad}\bar{\bar{f}}(x)\right)$ as they are both equal to $\mathrm{grad}\bar{f}(x)$ but one may lead to more convenient intermediate expressions than the other. Then,

$$\mathrm{lift}_x(\mathrm{Hess} f([x])[\xi]) = \mathrm{Proj}_x^{\mathrm{H}}\left(\mathrm{D}\bar{\bar{G}}(x)[u]\right). \tag{9.62}$$

These formulas are best illustrated through an example.

Example 9.46. *Consider $\mathrm{Gr}(n,p) = \mathrm{St}(n,p)/\mathrm{O}(p)$ 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 $\bar{\bar{f}}(X) = \frac{1}{2}\mathrm{Tr}(X^\top AX)$ 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:*

$$\mathrm{grad}\bar{\bar{f}}(X) = AX \qquad \textit{and} \qquad \mathrm{Hess}\bar{\bar{f}}(X)[U] = AU.$$

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

$$\mathrm{Proj}_X^{\mathrm{H}}(Z) = (I_n - XX^\top)Z = Z - X(X^\top Z). \tag{9.63}$$

Choosing to work with (9.59) (rather than (9.58)) because $\mathrm{Proj}_X^{\mathrm{H}}$ is somewhat simpler than Proj_X (the projector to tangent spaces of the Stiefel manifold), we deduce that the lifted gradient of f is:

$$\begin{aligned}\mathrm{lift}_X(\mathrm{grad} f([X])) &= (I_n - XX^\top)\mathrm{grad}\bar{\bar{f}}(X)\\ &= AX - X(X^\top AX).\end{aligned} \tag{9.64}$$

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 (9.62). For any $U = \mathrm{lift}_X(\xi)$, since U is horizontal, we get after some simplifications:

$$\begin{aligned} \mathrm{lift}_X\big(\mathrm{Hess} f([X])[\xi]\big) &= (I_n - XX^\top)\mathrm{Hess}\bar{\bar{f}}(X)[U] - UX^\top\mathrm{grad}\bar{\bar{f}}(X) \\ &= AU - X(X^\top AU) - U(X^\top AX). \end{aligned} \tag{9.65}$$

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 $\langle U, V\rangle = \mathrm{Tr}(U^\top V)$.

Notice how intermediate formulas in (9.64) and (9.65) provide convenient expressions directly in terms of the Euclidean gradient and Hessian of $\bar{\bar{f}}$.

Example 9.47. *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 (9.58) 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: $\mathrm{D}\bar{\bar{G}}(X)[U] = \mathrm{Proj}\Big(\mathrm{Hess}\bar{\bar{f}}(X)[U]\Big)$. Then, we conclude via (9.62) that*

$$\mathrm{lift}_X\big(\mathrm{Hess} f([X])[\xi]\big) = \mathrm{Proj}_X^{\mathrm{H}}\Big(\mathrm{Hess}\bar{\bar{f}}(X)[U]\Big),$$

where $U = \mathrm{lift}_X(\xi)$ and we used $\mathrm{Proj}_X^{\mathrm{H}} \circ \mathrm{Proj} = \mathrm{Proj}_X^{\mathrm{H}}$.

Exercise 9.48. *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 9.46, establish the following facts about $\min_{[X]\in \mathrm{Gr}(n,p)} \frac{1}{2}\mathrm{Tr}(X^\top AX)$:*

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. The other claims are standard.

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.*

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:

$$c'(t) = \mathrm{D}\pi(\bar{c}(t))[\bar{c}'(t)]. \tag{9.66}$$

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

$$\mathrm{lift}_{\bar{c}(t)}(c'(t)) = \mathrm{Proj}^{\mathrm{H}}_{\bar{c}(t)}(\bar{c}'(t)). \tag{9.67}$$

The latter simplifies in a way that proves particularly useful for our purpose if the velocity of $\bar{c}$ is everywhere horizontal.

Definition 9.49. *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 9.50. *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*

$$\frac{\mathrm{D}}{\mathrm{d}t} Z = \mathrm{D}\pi(\bar{c})\left[\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\bar{Z}\right], \tag{9.68}$$

where $\bar{Z} = \mathrm{lift}(Z)$ is the horizontal lift of Z to the curve $\bar{c}$.

See Section 9.16 for a comment regarding the relevance of the horizontality assumption.

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 8.62.

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. Explicitly, we must show that, for all $U \in \mathfrak{X}(\mathcal{M})$, the proposed formula 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 proposed formula yields

$$\mathrm{D}\pi(\bar{c})\left[\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,\mathrm{lift}\,(U \circ c)\right] = \mathrm{D}\pi(\bar{c})\left[\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,(\bar{U} \circ \bar{c})\right] = \mathrm{D}\pi(\bar{c})\left[\bar{\nabla}_{\bar{c}'}\bar{U}\right], \quad (9.69)$$

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 ∇ (9.44). More specifically, using that vertical vectors are in the kernel of $\mathrm{D}\pi(\bar{c})$ and $\bar{c}' = \mathrm{lift}(c')$ as in (9.67) owing to horizontality,

$$\mathrm{D}\pi(\bar{c})\left[\bar{\nabla}_{\bar{c}'}\bar{U}\right] = \mathrm{D}\pi(\bar{c})\left[\mathrm{Proj}^{\mathrm{H}}(\bar{\nabla}_{\bar{c}'}\bar{U})\right] = \nabla_{c'} U.$$

This concludes the proof. $\square$

Under the same assumptions, the formula in Theorem 9.50 can be stated equivalently as:

$$\mathrm{lift}\left(\frac{\mathrm{D}}{\mathrm{d}t}Z\right) = \mathrm{Proj}^{\mathrm{H}}\left(\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\bar{Z}\right), \quad (9.70)$$

which is more informative regarding numerical representation.

Exercise 9.51. *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 9.50, 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

$$c''(t) = \frac{\mathrm{D}}{\mathrm{d}t}c'(t) = \mathrm{D}\pi(\bar{c}(t))\left[\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\bar{c}'(t)\right] = \mathrm{D}\pi(\bar{c}(t))\left[\bar{c}''(t)\right], \quad (9.71)$$

which we can also write as

$$\operatorname{lift}_{\bar{c}(t)}(c''(t)) = \operatorname{Proj}^{\mathrm{H}}_{\bar{c}(t)}\left(\bar{c}''(t)\right). \tag{9.72}$$

In particular, if $\bar{c}$ is horizontal, we have $\|c''(t)\|_{c(t)} = \|\operatorname{Proj}^{\mathrm{H}}_{\bar{c}(t)}\bar{c}''(t)\|_{\bar{c}(t)}$: acceleration can only decrease in going to the quotient.

Recall that geodesics are curves with zero acceleration. A direct consequence of (9.71) is that horizontal geodesics on the total space descend to geodesics on the quotient manifold.

Corollary 9.52. *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 geodesic on $\overline{\mathcal{M}}$ and it is a horizontal curve, then $c = \pi \circ \bar{c}$ is a geodesic on $\mathcal{M}$.*

[O'N83, Lem. 7.46]

In this last corollary, it is in fact 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 9.52 also holds. The proof (omitted) relies on standard results from ordinary differential equations.

Proposition 9.53. *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.*

See Section 9.16 for a discussion.

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 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 9.54. *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 9.17 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 $\langle U, V\rangle_X = \operatorname{Tr}(U^\top V)$ on $\mathbb{R}^{d\times n}_d$, we can further turn $\mathcal{M}$ into a Riemannian quotient manifold.*

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 9.52 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:

$$\|QX - Y\|^2 = \|X\|^2 + \|Y\|^2 - 2\,\mathrm{Tr}(QXY^\top).$$

It can be shown that this distance is minimized with respect to $Q \in \mathrm{O}(d)$ if $QXY^\top$ is symmetric and positive semidefinite. Specifically, if $XY^\top = U\Sigma V^\top$ is an SVD decomposition, then the minimum is attained by $Q = VU^\top$ (the polar factor of $XY^\top$). 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,

$$\bar{c}(t)\bar{c}(t)^\top = (1-t)^2 XX^\top + t(1-t)(XY^\top + YX^\top) + t^2 YY^\top$$

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 $\mathbb{R}^{d\times n}_d$ 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.

Massart and Absil discuss the geometry of $\mathcal{M}$ in detail [MA20].

Another direct consequence of (9.71) is that second-order retractions (Definition 8.64) on $\overline{\mathcal{M}}$ which satisfy condition (9.28) and whose curves are horizontal yield second-order retractions on the quotient manifold.

Corollary 9.55. *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 $u \in \mathrm{H}_x$, 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 (9.71) and evaluate at $t = 0$. (This also makes it clear that $\bar{c}$ needs only be horizontal in a neighborhood of $t = 0$.) $\square$

Example 9.56. *Recall the polar retraction (7.22) on* $\mathrm{St}(n, p)$*:*

$$\mathrm{R}_X(U) = (X + U)\left(I_p + U^\top U\right)^{-1/2}. \tag{9.73}$$

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

$$\bar{c}'(t) = U\left(I_p + t^2 U^\top U\right)^{-1/2} + (X + tU)\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(I_p + t^2 U^\top U\right)^{-1/2}\right).$$

This curve is horizontal if, for all t*, the matrix*

$$\begin{aligned}\bar{c}(t)^\top \bar{c}'(t) = {} & \left(I_p + t^2 U^\top U\right)^{-1/2}(tU^\top U)\left(I_p + t^2 U^\top U\right)^{-1/2} \\ & + \left(I_p + t^2 U^\top U\right)^{+1/2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(I_p + t^2 U^\top U\right)^{-1/2}\right)\end{aligned}$$

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* $\mathrm{Gr}(n, p)$*, it is second order.*

Exercise 9.57. *Consider a Riemannian quotient manifold* $\mathcal{M} = \overline{\mathcal{M}}/\sim$ *with canonical projection* π*. Let* $\bar{c}$ *be a smooth curve on* $\overline{\mathcal{M}}$ *such that* $\bar{c}(0) = x$ *and* $\bar{c}'(0) = u$*, and let* $\bar{f} = f \circ \pi$ *be a smooth function on* $\overline{\mathcal{M}}$*. Define* $c = \pi \circ \bar{c}$ *and* $\xi = \mathrm{D}\pi(x)[u]$*. Obtain a Taylor expansion for* $f \circ c$ *in two ways:*

1. *Apply the general formula (8.19) to* $f \circ c$*; and*
2. *Apply (8.19) to* $\bar{f} \circ \bar{c}$*.*

Of course, the two formulas are equal since $f \circ c = \bar{f} \circ \bar{c}$*. Assuming* $\bar{c}$ *is a horizontal curve, use Propositions 9.36 and 9.42 and eq. (9.71) to verify that the formulas also match term-by-term.*

9.16 *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].

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 the 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 9.17, which states free, proper and smooth group actions yield quotient manifolds). For a partial converse to the latter, see [Lee12, Pb. 21-5].

When the group acting on the total space is not compact, it may be delicate to determine whether the action is proper. The following characterization may help in this regard.

Proposition 9.58. *Let $\mathcal{G}$ be a Lie group acting smoothly on a manifold $\overline{\mathcal{M}}$. The following are equivalent:* [Lee12, Prop. 21.5]

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 9.26 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 9.26 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 9.50, one may wonder what goes wrong if $\bar{c}$ is not horizontal. In that case, to proceed from (9.69) we separate $\bar{c}'(t)$ into its horizontal and vertical parts, and we use linearity:

$$\begin{aligned} \mathrm{D}\pi(\bar{c})\left[\bar{\nabla}_{\bar{c}'}\bar{U}\right] &= \mathrm{D}\pi(\bar{c})\left[\bar{\nabla}_{\mathrm{Proj}^{\mathrm{H}}(\bar{c}')}\bar{U}\right] + \mathrm{D}\pi(\bar{c})\left[\bar{\nabla}_{\mathrm{Proj}^{\mathrm{V}}(\bar{c}')}\bar{U}\right] \\ &= \nabla_{c'}U + \mathrm{D}\pi(\bar{c})\left[\bar{\nabla}_{\mathrm{Proj}^{\mathrm{V}}(\bar{c}')}\bar{U}\right]. \end{aligned}$$

For the horizontal part, the same argument as in the proof of Theorem 9.50 based on (9.44) and (9.67) 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),

$$\begin{aligned} \bar{\nabla}_{X\Omega}\bar{U} &= \mathrm{Proj}_X\Big(AX\Omega - X\Omega(X^\top AX) - X(\Omega^\top XAX + X^\top AX\Omega)\Big) \\ &= (I_n - XX^\top)AX\Omega, \end{aligned}$$

where Proj_X is the orthogonal projector to $\mathrm{T}_X\mathrm{St}(n, p)$ (7.25). 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 9.53: this concerns the possibility of horizontally lifting curves $c\colon I \to \mathcal{M}$ from the quotient manifold to 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 this ODE where the 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 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 9.53 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[2] 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 $(-2, 2)$.

[2] math.stackexchange.com/questions/3524475

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.)

Proposition 9.59. *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 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 $\overline{\mathcal{M}}$ is not connected, apply the claim to each complete connected component.

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 9.17 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 9.60. *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.

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 smooth 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. There, 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.

Exponential Map:
Retractions whose curves are geodesics.

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, Chs 1 and 7] for an introduction.

10.1 Distance, geodesics and completeness

A *distance* on a set $\mathcal{M}$ is a function dist: $\mathcal{M} \times \mathcal{M} \to \mathbb{R}$ 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 piecewise smooth curve on $\mathcal{M}$ joining x and y, or the infimum over the lengths of such curves.

Two points of $\mathcal{M}$ belong to the same connected component if and only if there exists a continuous curve in $\mathcal{M}$ joining them.

Let $\mathcal{M}$ be a Riemannian manifold. Given a piecewise smooth curve $c \colon [0, 1] \to \mathcal{M}$, we define the *length* of c as the integral of its *speed* $\|c'(t)\|_{c(t)}$:

$$L(c) = \int_0^1 \|c'(t)\|_{c(t)} \mathrm{d}t. \tag{10.1}$$

Piecewise smooth means for some choice of $0 = a_0 < a_1 < \cdots < a_k < a_{k+1} = 1$ the restrictions $c|_{(a_i, a_{i+1})}$ are smooth for $i = 0, \ldots, k$, and c is continuous. Integration is summed on the open intervals.

(Here and below, it is clear how definitions extend to curves on an arbitrary domain $[a, b]$.) 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 smooth curves from $[0, 1]$ to $\mathcal{M}$ such that $c(0) = x$ and $c(1) = y$. We have the following important result.

Theorem 10.1. *If $\mathcal{M}$ is connected (meaning there exists a continuous curve connecting every two 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.*

[Lee18, Thm. 2.55]

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. (Owing to second-countability of the atlas topology (see

Section 8.2), a manifold has finitely many, or countably infinitely many connected components.)

If the infimum in (10.2) is attained for some curve c (which is not always the case), we call c a *minimizing curve*. Remarkably, up to parameterization, these are geodesics (Definition 5.33). 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.

Think of $\mathcal{M} = \mathbb{R}^2 \backslash \{0\}$ as a Riemannian submanifold of $\mathbb{R}^2$, and consider connecting x and $-x$.

Theorem 10.2. *Every minimizing curve admits a constant-speed parameterization such that it is a geodesic, called a* minimizing geodesic.

[Lee18, Thm. 6.4]

This theorem admits a partial converse. Note that 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 10.3. *Every geodesic $c\colon [0,1] \to \mathcal{M}$ is* locally minimizing, *that is, every $t \in [0,1]$ has a neighborhood $I \subseteq [0,1]$ such that, if $a, b \in I$ satisfy $a < b$, then the restriction $c|_{[a,b]}$ is a minimizing curve.*

[Lee18, Thm. 6.15]

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 10.4. *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 10.5. *A Riemannian manifold is* geodesically complete *if every geodesic can be extended to a geodesic defined on the whole real line.*

The following is an important classical result that justifies omitting to specify whether we mean metric or geodesic completeness.

Theorem 10.6 (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.*

[Lee18, Thm. 6.19, Pb. 6-14]
[O'N83, Thm. 5.21]
The last part is the *Heine–Borel property*.

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).

Theorem 10.7. *If $\mathcal{M}$ is complete, then any two points x, y in the same connected component are connected by a minimizing geodesic $c\colon [0,1] \to \mathcal{M}$ such that $c(0) = x$, $c(1) = y$ and $\operatorname{dist}(x, y) = L(c)$.*

[Lee18, Cor. 6.21] [O'N83, Prop. 5.22]

Example 10.8. *Compact Riemannian manifolds are complete.*

Example 10.9. *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 $\operatorname{dist}(x, y)$ is equal to the Euclidean distance $\|x - y\|$.*

Exercise 10.10. *Show that the length $L(c)$ is independent of the parameterization of the curve c, that is, $L(c \circ h) = L(c)$ for any surjective, monotone, piecewise smooth $h\colon [0,1] \to [0,1]$.*

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 10.11. *Consider the following subset of the tangent bundle:*

$$\mathcal{O} = \{(x, v) \in \mathrm{T}\mathcal{M} : \gamma_v \text{ is defined on an interval containing } [0,1]\}.$$

The exponential map $\operatorname{Exp}\colon \mathcal{O} \to \mathcal{M}$ *is defined by*

$$\operatorname{Exp}(x, v) = \operatorname{Exp}_x(v) = \gamma_v(1).$$

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, $\operatorname{Exp}_x(v) = x + v$. By Definition 10.5, 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 [Lee18, Lem. 5.18]. This allows us to express the exponential map as

$$\operatorname{Exp}_x(tv) = \gamma_v(t), \tag{10.3}$$

which is often more convenient. In particular, the domain of Exp_x is star-shaped around the origin in $\mathrm{T}_x\mathcal{M}$. Conveniently, Exp is smooth.

Star-shaped around the origin means: $v \in \mathcal{O}_x \implies tv \in \mathcal{O}_x$ for all $t \in [0,1]$.

Proposition 10.12. *The exponential map is smooth on its domain $\mathcal{O}$, which is open in $\mathrm{T}\mathcal{M}$.* [Lee18, Prop. 5.19]

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 10.13. *The exponential map is a second-order retraction, with a possibly restricted domain $\mathcal{O} \subseteq \mathrm{T}\mathcal{M}$.* [AMS08, §5.5]

Proof. Proposition 10.12 claims smoothness of Exp. By definition, for any $(x,v) \in \mathrm{T}\mathcal{M}$, $\mathrm{Exp}_x(0) = \gamma_v(0) = x$, and furthermore

$$\mathrm{DExp}_x(0)[v] = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Exp}_x(tv)\bigg|_{t=0} = \gamma_v'(0) = v,$$

so that $\mathrm{DExp}_x(0)$ is the identity map on $\mathrm{T}_x\mathcal{M}$ for all $x \in \mathcal{M}$. Finally, it is clear that this retraction is second order (Definition 5.36) since $\gamma_v''(t)$ is zero for all t, hence in particular for $t = 0$. $\square$

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 [Lee12, Thm. 4.5], 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.

Definition 10.14. *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* [Lee18, p165]

$$B(x,r) = \{v \in \mathrm{T}_x\mathcal{M} : \|v\|_x < r\}.$$

By the inverse function theorem, $\mathrm{inj}(x) > 0$.

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 $F(U)$.

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, $\mathrm{Exp}_x \colon U \to \mathcal{U}$ is a diffeomorphism, with well-defined, smooth inverse $\mathrm{Exp}_x^{-1} \colon \mathcal{U} \to U$. 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 $\|u\| \leq \|v\|$, then u would be included in U, which would contradict invertibility. This suggests the following definition.

$\mathcal{U}$ is a kind of *normal neighborhood* of x. [Lee18, p131]

Definition 10.15. *For $x \in \mathcal{M}$, let* Log_x *denote the* logarithmic map at x,

$$\mathrm{Log}_x(y) = \operatorname*{arg\,min}_{v \in \mathrm{T}_x\mathcal{M}} \|v\|_x \text{ subject to } \mathrm{Exp}_x(v) = y, \tag{10.5}$$

with domain such that this is uniquely defined.

This definition does not exclude sufficiently many points from the domain of Log_x to ensure smoothness: compare with Corollary 10.19 below. See also the discussion of *cut locus* and, in particular, of *conjugate points* in [Lee18, §10].

In particular, with domains $U = B(x, \mathrm{inj}(x))$ and $\mathcal{U} = \mathrm{Exp}_x(U)$ as above, $\mathrm{Log}_x \colon \mathcal{U} \to U$ is a diffeomorphism with inverse $\mathrm{Exp}_x \colon U \to \mathcal{U}$, and it is an exercise to show that

$$\|\mathrm{Log}_x(y)\|_x = \mathrm{dist}(x, y), \tag{10.6}$$

where dist is the Riemannian distance on $\mathcal{M}$. For example, in a Euclidean space, $\mathrm{Log}_x(y) = y - x$ for all x, y. It is important to specify the domain of Exp_x^{-1}, as there may be some arbitrary choices to make (and not all of them yield $\mathrm{Log}_x = \mathrm{Exp}_x^{-1}$). This is illustrated in the following example.

Example 10.16. *On the sphere S^{n-1}, the exponential map is (Example 5.32)*

$$v \mapsto \mathrm{Exp}_x(v) = \cos(\|v\|)x + \frac{\sin(\|v\|)}{\|v\|}v.$$

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 $\arccos \colon [-1, 1] \to [0, \pi]$ is the principal inverse of cos. *Overall, this yields the following expression:*

$$y \mapsto \mathrm{Exp}_x^{-1}(y) = \arccos(x^\top y)\frac{u}{\|u\|}, \tag{10.7}$$

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 *(counter-clockwise) 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.*

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 any retraction, and specialize to the exponential map later on.

Proposition 10.17. *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:*

$$E\colon \mathcal{O} \to \mathcal{M} \times \mathcal{M}\colon (x, v) \mapsto E(x, v) = (x, \mathrm{R}_x(v)). \tag{10.8}$$

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* $E\colon \mathcal{T} \to \mathcal{V}$ *is a diffeomorphism.*

Proof. First, to see that $E\colon \mathcal{T} \to \mathcal{V}$ 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 $E\colon \mathcal{T} \to \mathcal{V}$ 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[1] at each point of $\mathcal{T}$). 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:

$$\mathrm{D}E(x, v) \simeq \begin{bmatrix} I & 0 \\ * & \mathrm{DR}_x(v) \end{bmatrix}.$$

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. □

[1] [Lee12, Thm. 4.5]

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 $\mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M}$.

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 $\Delta\colon \mathcal{M} \to (0, \infty]$:

$$\mathcal{T} = \left\{(x, v) \in \mathrm{T}\mathcal{M} : \|v\|_x < \Delta(x)\right\}. \tag{10.9}$$

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 10.17 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.

Proposition 10.18. *On a Riemannian manifold $\mathcal{M}$, the injectivity radius function* $\mathrm{inj}\colon \mathcal{M} \to (0, \infty]$ *is continuous.*

This does *not* require $\mathcal{M}$ to be connected or complete: see the chapter notes in Section 10.8 for a proof.

Corollary 10.19. *The map $(x, v) \mapsto (x, \mathrm{Exp}_x(v))$ is a diffeomorphism from*

$$\mathcal{T} = \left\{(x, v) \in \mathrm{T}\mathcal{M} : \|v\|_x < \mathrm{inj}(x)\right\}$$

to

$$\mathcal{V} = \left\{(x, y) \in \mathcal{M} \times \mathcal{M} : \mathrm{dist}(x, y) < \mathrm{inj}(x)\right\}.$$

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 10.16 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 10.20. *On a Riemannian manifold $\mathcal{M}$, consider the following open subsets of the tangent bundle* $\mathrm{T}\mathcal{M}$*:*

$$V_\delta(x) = \left\{(x', v') \in \mathrm{T}\mathcal{M} : \mathrm{dist}(x, x') < \delta \text{ and } \|v'\|_{x'} < \delta\right\}.$$

If $\mathcal{M}$ is not connected, add the explicit requirement that x, x' belong to the same connected component, even if $\delta = \infty$.

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 the function $\Delta\colon \mathcal{M} \to (0, \infty]$ by:

$$\Delta(x) = \sup\left\{\delta > 0 : V_\delta(x) \subseteq \mathcal{O} \text{ and } E \text{ is a diffeomorphism on } V_\delta(x)\right\},$$

where E is as defined in (10.8). 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 10.17).

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 10.1.

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

If $\Delta(x) = \infty$ for some x, then $\mathcal{O} = \mathrm{T}\mathcal{M}$ and E is a diffeomorphism on $\mathrm{T}\mathcal{M}$, so that $\Delta(x) = \infty$ for all x in the same connected component: this is also compatible with the claim.

$$V_\delta(x') \subset V_{\Delta(x)}(x).$$

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 facts 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 10.17. □

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 10.20.

Corollary 10.21. *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* (10.9) *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.

Definition 10.22. *The injectivity radius* $\mathrm{inj}(\mathcal{M})$ *of a Riemannian manifold* $\mathcal{M}$ *is the infimum of* $\mathrm{inj}(x)$ *over* $x \in \mathcal{M}$.

The set of manifolds with positive injectivity radius is strictly included in the set of complete manifolds: Exercise 10.25.

Proposition 10.23. *For a compact Riemannian manifold,* $\mathrm{inj}(\mathcal{M}) \in (0, \infty)$.

[Lee18, Lem. 6.16]

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 10.24. *Consider a smooth curve $c \colon I \to \mathcal{M}$ on a Riemannian manifold $\mathcal{M}$. Show that its speed $\|c'(t)\|_{c(t)}$ is constant if and only if $c''(t)$ is orthogonal to $c'(t)$ at all times. In particular, if c is the unique minimizing geodesic from $c(0) = x$ to $c(1) = y$, deduce that $c'(0) = \mathrm{Log}_x(y)$ so that*

$$\|\mathrm{Log}_x(y)\|_x = \|c'(0)\|_x = \mathrm{dist}(x, y).$$

Likewise, deduce that if $\|v\|_x < \mathrm{inj}(x)$ then $\mathrm{dist}(x, \mathrm{Exp}_x(v)) = \|v\|_x$.

Exercise 10.25. *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 10.26. *Let $\mathcal{K}$ be any subset of a Riemannian manifold $\mathcal{M}$ and let $r \colon \mathcal{K} \to \mathbb{R}^+$ be continuous on $\mathbb{R}^+ = \{t \geq 0\}$ (in subspace topologies). Show that $\mathcal{K}$ is compact in $\mathcal{M}$ if and only if*

$$\mathcal{T} = \{(x, s) \in \mathrm{T}\mathcal{M} : x \in \mathcal{K} \text{ and } \|s\|_x \leq r(x)\}$$

is compact in $\mathrm{T}\mathcal{M}$.

10.3 Parallel transport

Consider a Riemannian manifold $\mathcal{M}$ and a tangent vector $u \in \mathrm{T}_x\mathcal{M}$. 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 Riemannian structure is not necessary for all claims here to hold, but it is convenient and fits our general treatment well.

The differential geometric tool of choice for this task is called *parallel transport*. 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 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*, that is, using the concept of covariant derivative (Theorem 5.26), to require

$$Z \in \mathfrak{X}(c), \qquad Z(0) = u, \qquad \text{and} \qquad \frac{\mathrm{D}}{\mathrm{d}t} Z = 0. \tag{10.10}$$

Using standard tools from ordinary differential equations, one can show that such a vector field exists and is unique.

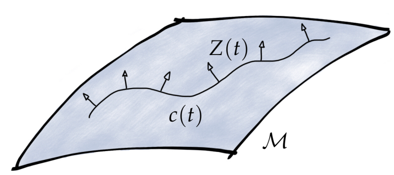


Figure 10.1: Parallel transports 'move' tangent vectors from one tangent space to another, along a specified curve. They can be used to compare or combine tangent vectors at different points by transporting them to a common tangent space.

Definition 10.27. *A vector field $Z \in \mathfrak{X}(c)$ such that $\frac{\mathrm{D}}{\mathrm{d}t}Z = 0$ is* parallel.

Theorem 10.28. *On a Riemannian manifold $\mathcal{M}$ equipped with its Riemannian connection and associated 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 $Z(t_0) = u$.*

[Lee18, Thm. 4.32]

This justifies the following definition.

Definition 10.29. *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,*

Also called *parallel translation*.

$$\mathrm{PT}^c_{t_1 \leftarrow t_0} \colon \mathrm{T}_{c(t_0)}\mathcal{M} \to \mathrm{T}_{c(t_1)}\mathcal{M},$$

is 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.

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:

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.

> *Parallel transport from x to y depends on the choice of curve connecting x and y.*

This is, in fact, a crucial feature of Riemannian geometry, intimately related to the notion of curvature [Lee18, Ch. 7]. Often times, when the curve is not specified, one implicitly means to move along the minimizing geodesic connecting x and y, assuming it exists and is unique.

Proposition 10.30. *The parallel transport operator $\mathrm{PT}^c_{t_1 \leftarrow t_0}$ is linear. It is also an isometry, that is,*

$$\forall u, v \in \mathrm{T}_{c(t_0)}\mathcal{M}, \qquad \langle u, v \rangle_{c(t_0)} = \left\langle \mathrm{PT}^c_{t_1 \leftarrow t_0}(u), \mathrm{PT}^c_{t_1 \leftarrow t_0}(v) \right\rangle_{c(t_1)}.$$

Furthermore, $\left(\mathrm{PT}^c_{t_1 \leftarrow t_0}\right)^* = \left(\mathrm{PT}^c_{t_1 \leftarrow t_0}\right)^{-1} = \mathrm{PT}^c_{t_0 \leftarrow t_1}$, where the star denotes the adjoint. In particular, $\mathrm{PT}^c_{t \leftarrow t}$ is the identity.

Proof. For linearity, consider $u, v \in \mathrm{T}_{c(t_0)}\mathcal{M}$ and $a, b \in \mathbb{R}$, arbitrary. By Theorem 10.28, 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. 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. Isometry implies that the adjoint and the inverse coincide. That the inverse corresponds to parallel transport along the same curve with swapped times t_0, t_1 follows from uniqueness again. □

By Definition 10.29, it is clear that $t_1 \mapsto \mathrm{PT}^c_{t_1 \leftarrow t_0}(u) = Z(t_1)$ is smooth (it is a parallel vector field on c). We can also argue smoothness in t_0, as follows: given $Y \in \mathfrak{X}(c)$, the map $t_0 \mapsto \mathrm{PT}^c_{t_1 \leftarrow t_0}(Y(t_0))$ is smooth. One convenient way of making this point is via *parallel frames*. 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

$$E_i(t) = \mathrm{PT}^c_{t \leftarrow \bar{t}}(e_i), \qquad \text{for } i = 1, \ldots, d. \tag{10.11}$$

Since parallel transport is an isometry, for all t, the vectors $E_i(t)$ form a basis for the tangent space at $c(t)$ (moreover, if the e_i are orthonormal, then so are the $E_i(t)$). There exist unique functions $\alpha_i(t)$ such that

$$Y(t) = \sum_{i=1}^{d} \alpha_i(t) E_i(t). \tag{10.12}$$

These functions are smooth since Y is smooth. Owing to linearity,

$$\mathrm{PT}^c_{t_1 \leftarrow t_0}(Y(t_0)) = \sum_{i=1}^{d} \alpha_i(t_0) E_i(t_1),$$

which is clearly smooth in both t_0 and t_1.

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 10.31. *Given a vector field $Z \in \mathfrak{X}(c)$ defined around t_0,*

$$\frac{\mathrm{D}}{\mathrm{d}t} Z(t) = \mathrm{PT}^c_{t \leftarrow t_0}\left(\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{PT}^c_{t_0 \leftarrow t} Z(t) \right) = \lim_{\delta \to 0} \frac{\mathrm{PT}^c_{t \leftarrow t+\delta} Z(t+\delta) - Z(t)}{\delta}.$$

Proof. Transport any basis $e_1, \ldots, e_d$ of $\mathrm{T}_{c(t_0)}\mathcal{M}$ along c to form $E_i(t) = \mathrm{PT}^c_{t \leftarrow t_0} e_i$. 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 5.26) and $\frac{\mathrm{D}}{\mathrm{d}t} E_i = 0$,

$$\begin{aligned}
\frac{\mathrm{D}}{\mathrm{d}t} Z(t) &= \sum_{i=1}^{d} \alpha_i'(t) E_i(t) \\
&= \mathrm{PT}^c_{t \leftarrow t_0} \sum_{i=1}^{d} \alpha_i'(t) e_i \\
&= \mathrm{PT}^c_{t \leftarrow t_0} \frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{d} \alpha_i(t) e_i = \mathrm{PT}^c_{t \leftarrow t_0} \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{PT}^c_{t_0 \leftarrow t} Z(t),
\end{aligned}$$

as announced. The important point is that $t \mapsto \mathrm{PT}^c_{t_0 \leftarrow t} Z(t)$ is a map into the fixed tangent space $\mathrm{T}_{c(t_0)}\mathcal{M}$, hence why we can take a classical derivative $\frac{\mathrm{d}}{\mathrm{d}t}$. ☐

Exercise 10.32. *On the sphere S^{n-1} as a Riemannian submanifold of $\mathbb{R}^n$ with the usual metric, 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 5.32 for the exponential.)

Exercise 10.33. *Show that if Y is a parallel vector field along a smooth curve c then $Y(t_1) = \mathrm{PT}^c_{t_1 \leftarrow t_0}(Y(t_0))$ for all t_0, t_1 in the domain of c.*

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, the main difficulty is to handle manifolds that are not complete.

In this section, we consider maps between smooth manifolds where the map itself may not be smooth: we define this in analogy with Definition 8.5. Explicitly, we say a map $F\colon \mathcal{M} \to \mathcal{M}'$ is k times (continuously) differentiable at a point $x \in \mathcal{M}$ if $\tilde{F} = \psi \circ F \circ \varphi^{-1}$ is k times (continuously) differentiable at $\varphi(x)$ (in the usual sense), where φ and ψ are charts around x and $F(x)$, respectively. We say F is k times (continuously) differentiable if it has that property at every point in $\mathcal{M}$. If $\mathcal{M}$ is Riemannian, $f\colon \mathcal{M} \to \mathbb{R}$ is (continuously) differentiable iff it has a (continuous) Riemannian gradient, and f is twice (continuously) differentiable iff it has a (continuous) Riemannian Hessian.

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

$$\forall x, y \in A, \qquad \mathrm{dist}_B(F(x), F(y)) \leq L\, \mathrm{dist}_A(x, y),$$

where dist_A, dist_B denote the distances on A and B. In particular:

Definition 10.34. *A function $f\colon \mathcal{M} \to \mathbb{R}$ on a connected manifold $\mathcal{M}$ is L-*Lipschitz continuous *if*

We implicitly assume that manifolds in this section are Riemannian.

$$\forall x, y \in \mathcal{M}, \qquad |f(x) - f(y)| \leq L\, \mathrm{dist}(x, y), \tag{10.13}$$

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.*

This definition is equivalent to the similar condition we state below. This next one is more convenient to study iterates of an optimization algorithm presented as $x_{k+1} = \mathrm{Exp}_{x_k}(s_k)$ for some tangent s_k.

Proposition 10.35. *A function $f\colon \mathcal{M} \to \mathbb{R}$ is L-Lipschitz continuous if and only if*

In this section, we usually omit subscripts for inner products and norms.

$$\forall (x, s) \in \mathcal{O}, \qquad |f(\mathrm{Exp}_x(s)) - f(x)| \leq L\|s\|, \tag{10.14}$$

where $\mathcal{O} \subseteq \mathrm{T}\mathcal{M}$ is the domain of the exponential map (Definition 10.11).

To prove this, we first introduce a lemma which states any continuous curve c can be interpolated by a 'broken geodesic' γ. Of course, if c is piecewise smooth it has a length and we have $L(\gamma) \leq L(c)$.

Lemma 10.36. *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 all i.*

These times define a piecewise smooth 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 $\sum_{i=0}^{n-1} \|s_i\| = L(\gamma)$.

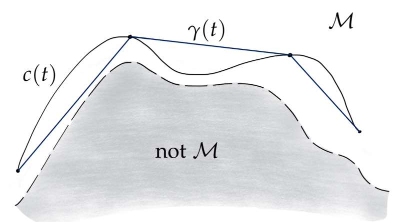


Figure 10.2: In Lemma 10.36, the curve γ is made of a finite number of minimizing geodesic segments, with endpoints on c.

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 `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 $\|s_i\| = L(\gamma|_{[t_i,t_{i+1}]})$. ☐

Proof of Proposition 10.35. 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 10.7, 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 smooth curve $c\colon [0,1] \to \mathcal{M}$ with length $L(c) \leq \mathrm{dist}(x,y) + \varepsilon$ such that $c(0) = x$ and $c(1) = y$. Construct a broken geodesic γ as provided by Lemma 10.36: 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 $\sum_{i=0}^{n-1} \|s_i\| = L(\gamma) \leq L(c)$. Then,

$$|f(x) - f(y)| \leq \sum_{i=0}^{n-1} |f(\gamma(t_i)) - f(\gamma(t_{i+1}))| \overset{(10.14)}{\leq} L \cdot \sum_{i=0}^{n-1} \|s_i\| \leq L \cdot L(c).$$

This reasoning holds for all $\varepsilon > 0$, hence condition (10.13) follows. ☐

If f has a continuous gradient, we express Lipschitz continuity as follows.

Proposition 10.37. *If $f\colon \mathcal{M} \to \mathbb{R}$ has a continuous gradient, then f is L-Lipschitz continuous if and only if*

$$\forall x \in \mathcal{M}, \qquad \|\mathrm{grad} f(x)\| \leq L. \tag{10.15}$$

Proof. For any $(x,s) \in \mathcal{O}$, consider $c(t) = \mathrm{Exp}_x(ts)$ for $t \in [0,1]$. Then,

$$f(c(1)) - f(c(0)) = \int_0^1 (f \circ c)'(t)\mathrm{d}t = \int_0^1 \left\langle \mathrm{grad} f(c(t)), c'(t) \right\rangle \mathrm{d}t.$$

Thus, if the gradient norm is bounded by L at all points along c,

$$|f(\mathrm{Exp}_x(s)) - f(x)| \leq L \int_0^1 \|c'(t)\|\mathrm{d}t = L \cdot L(c) = L\|s\|.$$

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,

$$\begin{aligned}
\|\mathrm{grad} f(x)\| &= \max_{s \in \mathrm{T}_x\mathcal{M}, \|s\|=1} \left\langle \mathrm{grad} f(x), s \right\rangle \\
&= \max_{s \in \mathrm{T}_x\mathcal{M}, \|s\|=1} \mathrm{D}f(x)[s] \\
&= \max_{s \in \mathrm{T}_x\mathcal{M}, \|s\|=1} \lim_{t \to 0} \frac{f(\mathrm{Exp}_x(ts)) - f(x)}{t} \leq L,
\end{aligned}$$

since $f(\mathrm{Exp}_x(ts)) - f(x) \leq L\|ts\| = L|t|$. $\square$

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.[2] 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 $f\colon \mathcal{E} \to \mathbb{R}$, where it is natural to measure $\mathrm{grad} f(x) - \mathrm{grad} f(y)$ in the Euclidean metric, disregarding the root points. With this in mind, it is natural to rely on the notion of 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.

[2] A canonical choice would follow from the *Sasaki metric* [GHL04, §2.B.6].

The following definition is fairly common. Notice how the restriction by the injectivity radius allows us to resolve any ambiguity in terms of choice of path along which to transport.

Definition 10.38. *A vector field V on a connected manifold $\mathcal{M}$ is L-*Lipschitz continuous *if, for all $x, y \in \mathcal{M}$ with* $\mathrm{dist}(x,y) < \mathrm{inj}(x)$,

$$\left\|\mathrm{PT}^{\gamma}_{0 \leftarrow 1} V(y) - V(x)\right\| \leq L\,\mathrm{dist}(x,y), \tag{10.16}$$

For our purpose, this notion is particularly relevant with $V = \mathrm{grad} f$, in which case we would say f has an L-Lipschitz continuous gradient.

where $\gamma\colon [0,1] \to \mathcal{M}$ *is the unique minimizing geodesic connecting* x *to* y*. If* $\mathcal{M}$ *is disconnected, 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 10.39. *A vector field* V *on a manifold* $\mathcal{M}$ *is* L*-Lipschitz continuous if and only if*

$$\forall (x,s) \in \mathcal{O}, \qquad \|P_s^{-1} V(\mathrm{Exp}_x(s)) - V(x)\| \le L\|s\|, \tag{10.17}$$

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 10.36, 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

$$\begin{aligned}
\|\mathrm{PT}^\gamma_{t_0\leftarrow t_n} V(\gamma(t_n)) - V(x)\| &= \|\mathrm{PT}^\gamma_{t_{n-1}\leftarrow t_n} V(\gamma(t_n)) - \mathrm{PT}^\gamma_{t_{n-1}\leftarrow t_0} V(x)\| \\
&\le \|\mathrm{PT}^\gamma_{t_{n-1}\leftarrow t_n} V(\gamma(t_n)) - V(\gamma(t_{n-1}))\| \\
&\quad + \|\mathrm{PT}^\gamma_{t_{n-1}\leftarrow t_0} V(x) - V(\gamma(t_{n-1}))\| \\
&\le L\,\mathrm{dist}(\gamma(t_{n-1}), \gamma(t_n)) \\
&\quad + \|\mathrm{PT}^\gamma_{t_0\leftarrow t_{n-1}} V(\gamma(t_{n-1})) - V(x)\|,
\end{aligned}$$

where in the last step we used (10.16). Repeat this argument on the right-most term $n-1$ times to conclude

$$\|\mathrm{PT}^\gamma_{0\leftarrow 1} V(\mathrm{Exp}_x(s)) - V(x)\| \le L \sum_{i=0}^{n-1} \mathrm{dist}(\gamma(t_i), \gamma(t_{i+1})) = L \cdot L(\gamma),$$

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. $\square$

If the vector field V is continuously differentiable, it is easy to say more about 'long range' vector transports, comparing V at points x and y which may not necessarily be connected by a geodesic.

Proposition 10.40. *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 x to 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,

$$\begin{aligned}
\mathrm{PT}^c_{0\leftarrow 1} V(c(1)) - V(c(0)) &= \sum_{i=1}^d (v_i(1) - v_i(0))e_i \\
&= \sum_{i=1}^d \left(\int_0^1 v_i'(t)\mathrm{d}t\right) e_i = \int_0^1 \mathrm{PT}^c_{0\leftarrow t}\left(\nabla_{c'(t)} V\right) \mathrm{d}t.
\end{aligned}$$

Consequently, using that parallel transports are isometric,

$$\|\mathrm{PT}^c_{0\leftarrow 1} V(y) - V(x)\| \leq \int_0^1 \|\nabla_{c'(t)} V\|\mathrm{d}t \overset{(10.18)}{\leq} L \int_0^1 \|c'(t)\|\mathrm{d}t = L \cdot L(c).$$

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 10.31,

$$\nabla_s V = \frac{\mathrm{D}}{\mathrm{d}t} V(c(t))\bigg|_{t=0} = \lim_{t\to 0} \frac{\mathrm{PT}^c_{0\leftarrow t} V(c(t)) - V(c(0))}{t}.$$

By (10.17), the norm of the numerator is bounded by $|t|L$, which concludes the proof. ☐

Corollary 10.41. *If $f\colon \mathcal{M} \to \mathbb{R}$ 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 10.42. *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 $\operatorname{dist}(x, y) < \operatorname{inj}(x)$, it holds that $\|\mathrm{PT}^{\gamma}_{0\leftarrow 1} V(y) - V(x)\| \leq L \operatorname{dist}(x, y)$ 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:

4. *For any smooth $c\colon [0, 1] \to \mathcal{M}$, $\|\mathrm{PT}^{c}_{0\leftarrow 1} V(c(1)) - V(c(0))\| \leq L \cdot L(c)$.*
5. *For all $(x, s) \in \mathrm{T}\mathcal{M}$, $\|\nabla_s V\| \leq L\|s\|$.*

Particularizing to $V = \operatorname{grad} f$ the above provides a good understanding of functions f with Lipschitz continuous gradients.

Going one degree higher, we now define and discuss functions with a Lipschitz continuous Hessian. The Hessian $\operatorname{Hess} f$ is a *tensor field of order two* on $\mathcal{M}$: for each x, $\operatorname{Hess} f(x)$ is a linear operator from $\mathrm{T}_x\mathcal{M}$ to itself. The following definition applies.

Section 10.7 gives a general take on tensor fields, adopting the viewpoint that $\operatorname{Hess} f(x)$ is a bilinear map from $(\mathrm{T}_x\mathcal{M})^2$ to $\mathbb{R}$, through the Riemannian metric: $\operatorname{Hess} f(x)(u, v) = \langle \operatorname{Hess} f(x)[u], v\rangle$.

Definition 10.43. *Let H be a tensor field of order two on a connected manifold $\mathcal{M}$. We say H is L-*Lipschitz continuous *if for all $x, y \in \mathcal{M}$ such that $\operatorname{dist}(x, y) < \operatorname{inj}(x)$ we have*

$$\|\mathrm{PT}^{\gamma}_{0\leftarrow 1} \circ H(y) \circ \mathrm{PT}^{\gamma}_{1\leftarrow 0} - H(x)\| \leq L \operatorname{dist}(x, y), \tag{10.20}$$

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.

For each x, $H(x)\colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M}$ is linear. We are particularly interested in the special case where $H(x) = \operatorname{Hess} f(x)$.

$\|H(x)\| = \max_{s \in \mathrm{T}_x\mathcal{M}, \|s\|=1} \|H(x)[s]\|$

The proof of the following proposition is left as an exercise.

Proposition 10.44. *A tensor field H of order two on a manifold $\mathcal{M}$ 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 $t = 1$.

We make the following statement now, deferring a formal definition of what the covariant derivative of a tensor field is to Section 10.7.

Proposition 10.45. *If H is a continuously differentiable tensor field of order two on a manifold $\mathcal{M}$, then it is L-Lipschitz continuous if and only if*

$$\forall (x, s) \in \mathrm{T}\mathcal{M}, \qquad \|\nabla_s H\| \leq L\|s\|, \tag{10.22}$$

where ∇ is the Riemannian connection. In that case, for any smooth curve $c\colon [0,1] \to \mathcal{M}$ connecting x to y, it holds that

$$\|\mathrm{PT}^c_{0\leftarrow 1} \circ H(y) \circ \mathrm{PT}^c_{1\leftarrow 0} - H(x)\| \leq L \cdot L(c). \tag{10.23}$$

Proof. The proof is similar to that of Proposition 10.40, using the properties of ∇H afforded by the general definitions we have omitted. For part of the proof, a starting point is to construct an orthonormal frame $E_1, \ldots, E_d$ along a curve c as usual, and to use this frame to expand $H(c(t)) = \sum_{i,j=1}^d h_{ij}(t) E_i(t) \otimes E_j(t)$, where $\otimes$ denotes a tensor (or outer) product: for $u, v \in \mathrm{T}_x\mathcal{M}$, $u \otimes v$ is a linear operator from $\mathrm{T}_x\mathcal{M}$ to itself defined by $(u \otimes v)[s] = \langle v, s\rangle\, u$. ☐

Corollary 10.46. *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*

$$\forall (x,s) \in \mathrm{T}\mathcal{M}, \qquad\qquad \|\nabla_s \mathrm{Hess} f\| \leq L\|s\|,$$

where $\nabla_s \mathrm{Hess} f$ is a self-adjoint linear operator on $\mathrm{T}_x\mathcal{M}$.

A summary of the same kind as Corollary 10.42 holds for tensor fields of order two as well. As one would expect, upon considering the general definitions of tensor fields and their covariant derivatives, all of the above generalizes to higher order tensor fields: see Section 10.7.

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 repeatedly: 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

$$\gamma'(t) = P_{ts}\gamma'(0) = P_{ts}s. \tag{10.25}$$

We use this frequently.

Proposition 10.47. *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:

$$\left| f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle \right| \leq \frac{L}{2}\|s\|^2.$$

Proof. Consider the real function $f \circ \gamma$:

$$
\begin{aligned}
f(\gamma(1)) &= f(\gamma(0)) + \int_0^1 (f \circ \gamma)'(t) \mathrm{d}t \\
&= f(x) + \int_0^1 \left\langle \mathrm{grad} f(\gamma(t)), \gamma'(t) \right\rangle \mathrm{d}t \\
&= f(x) + \int_0^1 \left\langle P_{ts}^{-1} \mathrm{grad} f(\gamma(t)), s \right\rangle \mathrm{d}t,
\end{aligned}
$$

where on the last line we used $\gamma'(t) = P_{ts}s$ 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

$$
\left| f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle \right| \leq \int_0^1 tL\|s\|^2 \mathrm{d}t = \frac{L}{2}\|s\|^2,
$$

as announced. □

Corollary 10.48. *If $f \colon \mathcal{M} \to \mathbb{R}$ has L-Lipschitz continuous gradient, then*

$$
\left| f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle \right| \leq \frac{L}{2}\|s\|^2
$$

for all (x, s) in the domain of the exponential map.

In particular, regularity assumptions A3 and A9 hold for the exponential retraction over its whole domain, provided $\mathrm{grad} f$ is L-Lipschitz (Definition 10.38). See also Exercise 10.52.

Proposition 10.49. *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:

$$
\begin{aligned}
\left| f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle - \frac{1}{2}\langle s, \mathrm{Hess} f(x)[s]\rangle \right| &\leq \frac{L}{6}\|s\|^3, \\
\left\| P_s^{-1} \mathrm{grad} f(\mathrm{Exp}_x(s)) - \mathrm{grad} f(x) - \mathrm{Hess} f(x)[s] \right\| &\leq \frac{L}{2}\|s\|^2.
\end{aligned}
$$

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 $\gamma(t)$: $E_1(t), \ldots, E_d(t)$ 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 $\gamma'(t) = P_{ts}s$, 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,

$$\begin{aligned} P_{ts}^{-1}\mathrm{grad}f(\gamma(t)) = G(t) &= G(0) + \int_0^t G'(\tau)\mathrm{d}\tau \\ &= \mathrm{grad}f(x) + \int_0^t \left(P_{\tau s}^{-1} \circ \mathrm{Hess}f(\gamma(\tau)) \circ P_{\tau s}\right)[s]\mathrm{d}\tau. \end{aligned}$$

Moving $\mathrm{grad}f(x)$ to the left-hand side and subtracting $\mathrm{Hess}f(x)[ts]$ on both sides, we find

$$\begin{aligned} &P_{ts}^{-1}\mathrm{grad}f(\gamma(t)) - \mathrm{grad}f(x) - \mathrm{Hess}f(x)[ts] \\ &\qquad = \int_0^t \left(P_{\tau s}^{-1} \circ \mathrm{Hess}f(\gamma(\tau)) \circ P_{\tau s} - \mathrm{Hess}f(x)\right)[s]\mathrm{d}\tau. \end{aligned}$$

Using the main assumption on $\mathrm{Hess}f$ along γ, it follows that

$$\begin{aligned} &\left\|P_{ts}^{-1}\mathrm{grad}f(\gamma(t)) - \mathrm{grad}f(x) - \mathrm{Hess}f(x)[ts]\right\| \\ &\qquad\qquad \le \int_0^t \tau L\|s\|^2\mathrm{d}\tau = \frac{L}{2}\|ts\|^2. \end{aligned} \tag{10.28}$$

For $t = 1$, this is the announced inequality.

Step 3: Taylor expansion of the function value. With the same start as in the proof of Proposition 10.47 but subtracting the additional 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. □

Corollary 10.50. *If $f\colon \mathcal{M} \to \mathbb{R}$ has L-Lipschitz continuous Hessian, then*

$$\left| f(\mathrm{Exp}_x(s)) - f(x) - \langle s, \mathrm{grad} f(x)\rangle - \frac{1}{2}\langle s, \mathrm{Hess} f(x)[s]\rangle \right| \leq \frac{L}{6}\|s\|^3,$$

$$\text{and} \quad \left\| P_s^{-1}\mathrm{grad} f(\mathrm{Exp}_x(s)) - \mathrm{grad} f(x) - \mathrm{Hess} f(x)[s] \right\| \leq \frac{L}{2}\|s\|^2,$$

for all (x, s) in the domain of the exponential map.

In particular, regularity assumption A10 holds for the exponential retraction over its whole domain, provided Hess f is L-Lipschitz continuous (Definition 10.43). See also Exercise 10.79.

Exercise 10.51. *Let $f\colon \mathcal{M} \to \mathbb{R}$ be differentiable on a manifold $\mathcal{M}$ equipped with a retraction* R. *Assume the pullback $\hat{f}_x = f \circ \mathrm{R}_x$ satisfies*

$$\|\mathrm{grad}\hat{f}_x(ts) - \mathrm{grad}\hat{f}_x(0)\| \leq L\|ts\|$$

for some $s \in \mathrm{T}_x\mathcal{M}$ and all $t \in [0, 1]$. Deduce that

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

Note the implications for the regularity assumptions A3 and A9. It is straightforward to extend this in relation to A10.

Exercise 10.52. *Let $f\colon \mathcal{M} \to \mathbb{R}$ be twice continuously differentiable on a manifold $\mathcal{M}$ equipped with a retraction* R. *Assume*

$$f(\mathrm{R}_x(s)) \leq f(x) + \langle s, \mathrm{grad} f(x)\rangle + \frac{L}{2}\|s\|^2 \tag{10.29}$$

for all (x, s) in a neighborhood of the 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 hold only if* grad f *is L-Lipschitz continuous. With* $\mathrm{R} = \mathrm{Exp}$ *in particular, verify that these three claims are equivalent:*

1. *Inequalities (10.29) hold in a neighborhood of the zero section in* $\mathrm{T}\mathcal{M}$*;*
2. grad f *is L-Lipschitz continuous;*
3. *Inequalities (10.29) hold over the whole domain of* Exp.

Here too, note the implications for the regularity assumptions A3 and A9. See also Exercise 10.79.

Exercise 10.53. *Give a proof of Proposition 10.44, for example by adapting that of Proposition 10.39.*

10.5 *Transporters*

The strong properties of parallel transports 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*. 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

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.

$$\mathrm{T}_{y \leftarrow x} \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_y\mathcal{M},$$

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 10.60.

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. 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.

Here, the two manifolds are the same.

Proposition 10.54. *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 $\mathcal{N}$,*

$$\mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{N}) = \{(x, y, \mathcal{L}) \mid x \in \mathcal{M}, y \in \mathcal{N} \text{ and } \mathcal{L} \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_y\mathcal{N} \text{ is linear}\},$$

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 8.46; 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 10.55. *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* [HGA15, §4.3]

$$\mathrm{T}\colon \mathcal{V} \to \mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{M})\colon (x, y) \mapsto \mathrm{T}_{y \leftarrow x}$$

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 10.54. 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 10.56. *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 8.46. 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 $\mathcal{U}_{\bar{x}} \times \mathcal{U}_{\bar{y}} \subseteq \mathcal{V}$. 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 $\mathcal{L}(\mathrm{T}\mathcal{M}, \mathrm{T}\mathcal{M})$. □

$\mathcal{V}$ 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.

With similar developments, we also get the following result once we equip the manifold with a Riemannian metric.

Proposition 10.57. *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 10.56 (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. One way of choosing such families of curves is via a retraction. Conveniently, the differentials of a retraction also provide a transporter: this is a good alternative when parallel transports are out of reach.

Proposition 10.58. *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 10.21. 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 $v = \mathrm{R}_x^{-1}(y)$.*

   Here we assume $\mathcal{M}$ is Riemannian and refer to the associated parallel transport.

2. *The differentials of the retraction define a transporter on $\mathcal{V}$ via $\mathrm{T}_{y\leftarrow x} = \mathrm{DR}_x(v)$, where $v = \mathrm{R}_x^{-1}(y)$.*

   Intuitively, by perturbing v in $\mathrm{T}_x\mathcal{M}$, we perturb $\mathrm{R}_x(v)$ away from y, thus producing a tangent vector in $\mathrm{T}_y\mathcal{M}$.

Proof. 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 10.59. *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*:*

$$\mathrm{T}_{y\leftarrow x}\circ\mathcal{A}\circ\mathrm{T}_{x\leftarrow y},\qquad (\mathrm{T}_{x\leftarrow y})^*\circ\mathcal{A}\circ\mathrm{T}_{x\leftarrow y},\qquad (\mathrm{T}_{x\leftarrow y})^{-1}\circ\mathcal{A}\circ\mathrm{T}_{x\leftarrow y}.$$

For the first two, we need (x,y) *to be in the domain of* T*; for the last one, we also need* $\mathrm{T}_{x\leftarrow y}$ *to be invertible. If* $\mathcal{A}$ *is self-adjoint, then so is the second operator. If the transporter is obtained through parallel transport as in Proposition 10.58 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 10.30.*

For manifolds embedded in Euclidean spaces, one particularly convenient transporter is given by orthogonal projectors to the tangent spaces. In contrast to Proposition 10.58, it does not involve retractions.

Proposition 10.60. *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 3.56. □

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 10.61. *Consider the Grassmann manifold* $\mathrm{Gr}(n,p) = \mathrm{St}(n,p)/\sim$ *(recall Example 9.18). Equip the total space with the polar retraction (7.22),* $\overline{\mathrm{R}}_X(V) = \mathrm{pfactor}(X+V)$*, 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 $Y = \bar{\mathrm{R}}_X(V)$ 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

$$\begin{aligned} E \colon \mathrm{TGr}(n,p) &\to \mathrm{Gr}(n,p) \times \mathrm{Gr}(n,p) \\ : ([X], \xi) &\mapsto E([X], \xi) = ([X], [\bar{\mathrm{R}}_X(V)]), \end{aligned} \tag{10.31}$$

where $V = \mathrm{lift}_X(\xi)$. In Exercise 10.65, 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

$$\begin{aligned} \mathrm{lift}_Y\Big(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\Big) &= \mathrm{Proj}^{\mathrm{H}}_Y\left(\bar{\mathrm{T}}_{Y\leftarrow X}(\mathrm{lift}_X(\xi))\right) \\ &= \mathrm{Proj}^{\mathrm{H}}_Y\left(\mathrm{lift}_X(\xi)\right), \end{aligned} \tag{10.32}$$

where we used that $\mathrm{Proj}^{\mathrm{H}}_Y \circ \mathrm{Proj}^{\mathrm{St}}_Y = \mathrm{Proj}^{\mathrm{H}}_Y$. We must check that $\mathrm{T}_{[Y]\leftarrow[X]}$ (a) is well defined, and (b) 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 9.25 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 $\bar{\mathrm{R}}_X(V) = Y$, we know that $VQ \in \mathrm{H}_{XQ}$ is such that $\bar{\mathrm{R}}_{XQ}(VQ) = YQ$ (this is specific to the polar retraction by (9.9)), and this is the only horizontal vector at XQ that maps to $[Y]$. Since $\mathrm{Proj}^{\mathrm{H}}_Y = \mathrm{Proj}^{\mathrm{H}}_{YQ} = I_n - YY^\top$, we find that

$$\begin{aligned} \mathrm{lift}_{YQ}\Big(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\Big) &= \mathrm{Proj}^{\mathrm{H}}_{YQ}\left(\mathrm{lift}_{XQ}(\xi)\right) \\ &= (I_n - YY^\top)\,\mathrm{lift}_X(\xi)Q \\ &= \mathrm{lift}_Y\Big(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\Big)\,Q. \end{aligned}$$

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 10.65, 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 $Y = \bar{\mathrm{R}}_X(V)$). 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 10.65 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 $Q = \tilde{U}\tilde{V}^\top$ and $\tilde{U}\tilde{\Sigma}\tilde{V}^\top = X^\top Y$;*
2. *By (10.32), $\mathrm{lift}_{YQ}\left(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\right) = \mathrm{Proj}^{\mathrm{H}}_{YQ}(\mathrm{lift}_X(\xi)) = U - Y(Y^\top U)$;*
3. *Finally, $\mathrm{lift}_Y\left(\mathrm{T}_{[Y]\leftarrow[X]}(\xi)\right) = (U - Y(Y^\top U))Q^\top$.*

Often times, 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 this section, we briefly connect the notion of transporter to that of vector transport.

Definition 10.62. *A* vector transport *on a manifold $\mathcal{M}$ is a smooth map* [AMS08, Def. 8.1.1]

$$(x, u, v) \mapsto \mathrm{VT}_{(x,u)}(v)$$

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 all $(x, u) \in \mathrm{T}\mathcal{M}$:

1. *There exists a retraction R on $\mathcal{M}$ such that $\mathrm{VT}_{(x,u)}$ is a linear map from the tangent space at x to the tangent space at $\mathrm{R}_x(u)$; and*
2. *$\mathrm{VT}_{(x,0)}$ is the identity on $\mathrm{T}_x\mathcal{M}$.*

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 10.21).

Exercise 10.63. *Give a proof of Proposition 10.54.*

Exercise 10.64. *Give a proof of Proposition 10.57.*

Exercise 10.65. *With notation as in Example 10.61, 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*

Compare with Exercise 7.1.

$$V = Y(X^\top Y)^{-1} - X, \tag{10.33}$$

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.

Exercise 10.66. *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 (Exercise 7.2). If we represent tangent vectors $U = X\Omega$ simply as their skew-symmetric part Ω, then this transporter requires no computations.*

Exercise 10.67. *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 $\hat{f} = f \circ \mathrm{R}_x\colon \mathrm{T}_x\mathcal{M} \to \mathbb{R}$.*

[ABBC20, §6]

Prove the following identities: 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 through polarization by

$$\langle \dot{s}, W_s(\dot{s})\rangle_x = \left\langle \mathrm{grad} f(\mathrm{R}_x(s)), c''(0)\right\rangle_{\mathrm{R}_x(s)}, \tag{10.37}$$

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 8.54 and 8.66 (as well as their embedded counter-parts, Propositions 3.49 and 5.39).

For any $u, v \in \mathrm{T}_x\mathcal{M}$, we can use (10.37) to compute $\langle u, W_s(v)\rangle_x$ as it is equal to

$$\frac{1}{4}\Big(\langle u+v, W_s(u+v)\rangle_x - \langle u-v, W_s(u-v)\rangle_x \Big).$$

This is one of several polarization identities. It fully determines W_s because for any orthonormal basis $e_1, \ldots, e_d$ of $\mathrm{T}_x\mathcal{M}$ the symmetric matrix with entries $\langle e_i, W_s(e_j)\rangle_x$ represents W_s in that basis.

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, $\mathrm{Hess}f(x)[u]$. As obtaining an explicit expression for the Hessian may be tedious, it is natural to explore avenues to approximate it numerically. To this end, we consider so-called *finite difference approximations*.

See also Section 4.7 for a word regarding automatic differentiation.

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 10.31, 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

We could also use higher-order finite difference approximations.

$$\mathrm{Hess}f(x)[u] \approx \frac{\mathrm{PT}^c_{0\leftarrow \bar{t}}(\mathrm{grad}f(c(\bar{t}))) - \mathrm{grad}f(x)}{\bar{t}} \tag{10.40}$$

for some well-chosen $\bar{t} > 0$: small enough to be close to the limit (see Corollary 10.50 to quantify this error), large enough to avoid numerical issues.

In light of Section 10.5, we may ask: is it possible to replace the parallel transport in (10.40) with a transporter? We already verified this for a special case in Example 5.29, 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 this 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, 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. To justify (10.41), consider this generalization of Proposition 10.31.

Proposition 10.68. *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 10.56, 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 10.69. *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

$$\operatorname{Hess} f(x)[u] = \lim_{t \to 0} \frac{\mathrm{T}_{x \leftarrow c(t)}(\operatorname{grad} f(c(t))) - \operatorname{grad} f(x)}{t} + \sum_{i=1}^{d} \langle \operatorname{grad} f(x), v_i \rangle_x \frac{\mathrm{D}}{\mathrm{d}t} V_i(0). \tag{10.44}$$

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 5.29 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 10.60), 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 $\mathrm{T}_x\mathcal{M}$ 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 functions $\mathfrak{F}(\mathcal{M})$ as follows:

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

This map if $\mathfrak{F}(\mathcal{M})$-linear in its argument:

$$U(fV + gW) = fU(V) + gU(W)$$

for any $V, W \in \mathfrak{X}(\mathcal{M})$ 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 both of its arguments. These two are examples of *tensor fields* of order one and two, respectively.

Definition 10.70. *A* smooth tensor field *T of order k on a Riemannian 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 tensor field.

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 counter-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.

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 $W_1, \ldots, W_d$ around x (see 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

Using 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].

$$U_i|_{\mathcal{U}} = f_{i,1}W_1 + \cdots + f_{i,d}W_d = \sum_{j=1}^{d} f_{i,j}W_j,$$

where the $f_{i,j}$ are smooth functions on $\mathcal{U}$. Then, working only on the domain $\mathcal{U}$ and using the linearity properties of tensor fields, we find

$$T(U_1, \ldots, U_k) = \sum_{j_1=1}^{d} \cdots \sum_{j_k=1}^{d} f_{1,j_1} \cdots f_{k,j_k} T(W_{j_1}, \ldots, W_{j_k}).$$

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 summary, it is legitimate to think of $T(x)$ as a k-linear map on $\mathrm{T}_x\mathcal{M}$:

$$T(x) \colon \mathrm{T}_x\mathcal{M} \times \cdots \times \mathrm{T}_x\mathcal{M} \to \mathbb{R}.$$

This offers a useful perspective on tensor fields of order k: they associate to each point of $\mathcal{M}$ a k-linear function on the tangent space at that point. Explicitly:

$$T(x)(u_1, \ldots, u_k) = T(U_1, \ldots, U_k)(x),$$

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 $U(x)$ refers to a tangent vector at x, while if we think of U as a tensor field, then $U(x)$ denotes the linear function $v \mapsto \langle U(x), v \rangle_x$ 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 $(\mathrm{T}_x\mathcal{M})^2$ 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 10.54, we can define a *tensor bundle* of order k over $\mathcal{M}$ as:

$$\mathrm{T}^k\mathrm{T}\mathcal{M} = \left\{ (x, L) : x \in \mathcal{M} \text{ and } L \in \mathrm{T}^k\mathrm{T}_x\mathcal{M} \right\}, \text{ where} \tag{10.46}$$

$$\mathrm{T}^k\mathrm{T}_x\mathcal{M} = \left\{ k\text{-linear functions from } (\mathrm{T}_x\mathcal{M})^k \text{ to } \mathbb{R} \right\}.$$

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 with

$$\big\{ (x, L) : x \in \mathcal{M} \text{ and } L \colon \mathrm{T}_x\mathcal{M} \to \mathrm{T}_x\mathcal{M} \text{ is linear} \big\}.$$

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. *Smooth* tensor fields as we defined them above are exactly the smooth sections as defined here [Lee12, Prop. 12.19, Lem. 12.24].

This is commonly taken as the definition of a tensor field.

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.

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 10.71. *Given a smooth tensor field T of order k on a Riemannian manifold $\mathcal{M}$ equipped with its Riemannian connection ∇, the* total covariant derivative ∇T *of T is a smooth tensor field of order $k+1$ on $\mathcal{M}$, defined by*

$$\begin{aligned}\nabla T(U_1, \ldots, U_k, V) &= VT(U_1, \ldots, U_k)\\ &\quad - T(\nabla_V U_1, U_2, \ldots, U_k) - \cdots - T(U_1, \ldots, U_{k-1}, \nabla_V U_k)\end{aligned}$$

for any $U_1, \ldots, U_k, V \in \mathfrak{X}(\mathcal{M})$. We also let $\nabla_V T$ be a smooth tensor field of order k, 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 10.72. *Let $\mathcal{M}$ be a Riemannian manifold with Riemannian connection ∇. A smooth function $f\colon \mathcal{M} \to \mathbb{R}$ is a tensor field of order zero. Differentiating f as a tensor field using Definition 10.71, 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. Since $\mathrm{grad} f$ *is itself thought of* *as a tensor field of order one, we can differentiate it as well, creating a tensor field of order two:*

$$\begin{aligned}\nabla \mathrm{grad} f(U,V) &= V\mathrm{grad} f(U) - \mathrm{grad} f(\nabla_V U) \\ &= V\langle \mathrm{grad} f, U\rangle - \langle \mathrm{grad} f, \nabla_V U\rangle \\ &= \langle \nabla_V \mathrm{grad} f, U\rangle \\ &= \langle \mathrm{Hess} f(V), U\rangle \\ &= \mathrm{Hess} f(U,V).\end{aligned}$$

In other words, $\nabla \mathrm{grad} f$ *is nothing but the Riemannian Hessian* $\mathrm{Hess} f$*, identified with a (symmetric) tensor field of order two. Going one step further, we get a glimpse of the covariant derivative of the Riemannian Hessian:*

$$\begin{aligned}\nabla \mathrm{Hess} f(U,V,W) &= W\mathrm{Hess} f(U,V) \\ &\qquad - \mathrm{Hess} f(\nabla_W U, V) - \mathrm{Hess} f(U, \nabla_W V) \\ &= W\langle \mathrm{Hess} f(U), V\rangle \\ &\qquad - \langle \mathrm{Hess} f(V), \nabla_W U\rangle - \langle \mathrm{Hess} f(U), \nabla_W V\rangle.\end{aligned}$$

This is a smooth tensor field of order three, with the other useful notion being that $\nabla_W \mathrm{Hess} f$ *is a smooth, symmetric tensor field of order two defined by*

$$\nabla_W \mathrm{Hess} f(U,V) = \nabla \mathrm{Hess} f(U,V,W).$$

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

$$(\nabla_v T)(u_1, \ldots, u_k) = \big((\nabla T)(x)\big)(u_1, \ldots u_k, v). \tag{10.47}$$

Thus, $\nabla_v T$ is a k-linear map on $\mathrm{T}_x\mathcal{M}$; if T is symmetric, so is $\nabla_v T$.

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 5.26.

Definition 10.73. *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 $\mathbb{R}$.

Theorem 10.74. *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* $a, b \in \mathbb{R}$*:*

[Lee18, Thm. 4.24]; for an analog of the product rule with tensor fields, see also [Lee18, Prop. 5.15].

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:* $\left(\frac{\mathrm{D}}{\mathrm{d}t}(T \circ c)\right)(t) = \nabla_{c'(t)}T$ *for all* $t \in I$.

This operator is called the induced covariant derivative*.*

In the statement above, we understand $\nabla_{c'(t)}T$ through (10.47).

As a result of Definition 10.71, we also have this chain rule: given $Z \in \mathfrak{X}^k(c)$ and $U_1, \ldots, U_k \in \mathfrak{X}(c)$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(Z(U_1, \ldots, U_k)) &= \frac{\mathrm{d}}{\mathrm{d}t}(Z(t)(U_1(t), \ldots, U_k(t))) \\
&= \left(\frac{\mathrm{D}}{\mathrm{d}t}Z(t)\right)(U_1(t), \ldots, U_k(t)) \\
&\quad + Z(t)\left(\frac{\mathrm{D}}{\mathrm{d}t}U_1(t), U_2(t), \ldots, U_k(t)\right) + \cdots \\
&\quad + Z(t)\left(U_1(t), \ldots, U_{k-1}(t), \frac{\mathrm{D}}{\mathrm{d}t}U_k(t)\right) \\
&= \left(\frac{\mathrm{D}}{\mathrm{d}t}Z\right)(U_1, \ldots, U_k) \\
&\quad + Z\left(\frac{\mathrm{D}}{\mathrm{d}t}U_1, U_2, \ldots, U_k\right) + \cdots \\
&\quad + Z\left(U_1, \ldots, U_{k-1}, \frac{\mathrm{D}}{\mathrm{d}t}U_k\right).
\end{aligned} \tag{10.48}$$

Example 10.75. *If $c \colon I \to \mathcal{M}$ is a geodesic on the Riemannian manifold $\mathcal{M}$ with Riemannian connection ∇, given a smooth function $f \colon \mathcal{M} \to \mathbb{R}$ we know that*

$$(f \circ c)' = \langle \mathrm{grad} f \circ c, c' \rangle \quad \textit{and} \quad (f \circ c)'' = \langle (\mathrm{Hess} f \circ c)[c'], c' \rangle,$$

where we benefit from the fact that $c''(t) = 0$ for all t. Going one step further and identifying the Hessian with a tensor field of order two as before, so that $\langle (\mathrm{Hess} f \circ c)[c'], c' \rangle = (\mathrm{Hess} f \circ c)(c', c')$, we find

$$\begin{aligned}
(f \circ c)''' &= \frac{\mathrm{d}}{\mathrm{d}t}((\mathrm{Hess} f \circ c)(c', c')) \\
&= \left(\frac{\mathrm{D}}{\mathrm{d}t}(\mathrm{Hess} f \circ c)\right)(c', c') \\
&= (\nabla_{c'}\mathrm{Hess} f)(c', c') \\
&= (\nabla \mathrm{Hess} f \circ c)(c', c', c').
\end{aligned}$$

Remark 10.76. *In our discussion of tensor fields, we have used the Riemannian metric to identify tangent vectors $u \in \mathrm{T}_x\mathcal{M}$ with linear functions on $\mathrm{T}_x\mathcal{M}$, as $v \mapsto \langle u, v \rangle_x$. This simplifies the story and is a standard approach on Riemannian manifolds. It is useful to know, however, that tensor fields (and* *many of the tools discussed here) can also be constructed without referencing a Riemannian metric. In that case, one has to distinguish between vectors and* covectors*, leading to the notions of* covariant, contravariant and mixed tensors*. 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 10.77. *Show that ∇T as provided by Definition 10.71 is indeed a tensor field of order $k+1$.*

Exercise 10.78. *For a Riemannian manifold $\mathcal{M}$ with Riemannian connection ∇, show that ∇G is identically zero, where G is the smooth tensor field of order two defined by the Riemannian metric, that is: $G(U,V) = \langle U, V\rangle$.*

See also [Lee18, Prop. 5.5] for further equivalent characterizations of compatibility between the Riemannian metric and connection. We say G is *parallel*.

Exercise 10.79. *Let $f\colon \mathcal{M} \to \mathbb{R}$ be three times continuously differentiable on a Riemannian manifold $\mathcal{M}$ equipped with a retraction R. Assume*

Note the implications for the regularity assumption A10. See also Exercise 10.52.

$$f(\mathrm{R}_x(s)) \leq f(x) + \langle s, \mathrm{grad} f(x)\rangle + \frac{1}{2}\langle s, \mathrm{Hess} f(x)[s]\rangle + \frac{L}{6}\|s\|^3 \quad (10.49)$$

for all (x,s) in a neighborhood of the zero section in the tangent bundle. 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 $c'''(0) = 0$. In particular, the exponential map is a third-order retraction. Show that $\|\nabla_s \mathrm{Hess} f\| \leq L\|s\|$ for all $(x,s) \in \mathrm{T}\mathcal{M}$. Apply Corollary 10.46 to conclude that $\mathrm{Hess} f$ is L-Lipschitz continuous.*

See also Exercise 10.80.

Further apply Corollary 10.50 to conclude that, with $R = \mathrm{Exp}$, these three claims are equivalent:

1. *Inequalities (10.49) hold in a neighborhood of the zero section in $\mathrm{T}\mathcal{M}$;*
2. $\mathrm{Hess} f$ *is L-Lipschitz continuous;*
3. *Inequalities (10.49) hold over the whole domain of* Exp.

Exercise 10.80. *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 (Propositions 5.42 and 5.43). Show that $c'''(0) = -2\mathcal{W}\left(v, \frac{\mathrm{d}}{\mathrm{d}t}c'(0)\right)$, where $\mathcal{W}$ is the Weingarten map (5.30). In general, this is nonzero. Thus, we do not expect the metric projection retraction to be third order in general. Hint: use (5.19) 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 5.43.*

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 Bacák [Bac14] for an introduction to *geodesic metric spaces*, and applications in convex analysis and optimization in Hadamard spaces.

Propositions 10.17 and 10.20 (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 complete manifolds. It is easy to find a published proof that the function $\mathrm{inj}\colon \mathcal{M} \to \mathbb{R}$ is continuous if $\mathcal{M}$ is (connected[3] 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[4] 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 10.81 and 10.82 below. Both proofs rely on continuity in the complete case. They are rearranged to highlight commonalities.

[3] 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.

[4] mathoverflow.net/questions/335032

Our main motivation to study continuity of inj is to reach Corollary 10.19, stating 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].

Parallel transporting a tangent vector u at x to all the points in a normal neighborhood of x along geodesics contained in that neighborhood results in a smooth vector field. This is a well-known fact; details of the argument appear notably in [LB19, 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 10.39. This may be their first occurrence in an optimization context. There too, the motivation is to derive inequalities such as the ones in Proposition 10.47. 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 10.44. 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 10.38, one could also endow the tangent bundle 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: see [dOF18] 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 10.69 handles finite differences of the Riemannian Hessian using an arbitrary transporter (Definition 10.55). An analog result for vector transports (Definition 10.62) appears in [AMS08, Lem. 8.2.2].

In Exercise 10.67, we contemplate the role of the initial acceleration of curves of the form $c(t) = \mathrm{R}_x(s + t\dot{s})$. If R is the exponential map, then c is a geodesic when $s = 0$ but (in general) not otherwise, so that $c''(0)$ is nonzero in general. This tangent vector and its norm are tightly related to curvature of the manifold.

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, but the reader should mind this when combining statement in that section with statements found in other sources.

Lemma 10.81. *The injectivity radius function* $\mathrm{inj}\colon \mathcal{M} \to (0, \infty]$ *is lower-semicontinuous.*

If $\mathrm{inj}(x) = \infty$ at some point x, then in particular Exp_x is defined on all of $\mathrm{T}_x\mathcal{M}$. Thus, the connected component of x is complete, 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.) Hence, we do not need to worry about infinite values.

Proof by Stephen McKeown. For contradiction, assume inj is not lower-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

$$\lim_{k\to\infty} x_k = x \qquad \text{and} \qquad \forall k,\ \mathrm{inj}(x_k) \le r < R = \mathrm{inj}(x).$$

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[5] 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 $d = \dim \mathcal{M}$. 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

$$B(y,\rho) \subseteq B(x,r'') \qquad \Longrightarrow \qquad \begin{cases} \hat{\mathrm{inj}}(\hat{y}) = \mathrm{inj}(y) & \text{if } \mathrm{inj}(y) < \rho, \\ \hat{\mathrm{inj}}(\hat{y}) \ge \rho & \text{otherwise,} \end{cases}$$

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 $\hat{\mathrm{inj}}(\hat{x}) \ge r''$, and
2. There exists k_0 large enough such that, for all $k \ge k_0$, $\mathrm{dist}(x_k, x) < \varepsilon$, so that $B(x_k, r') \subset B(x, r'')$. Moreover, $\mathrm{inj}(x_k) \le r < r'$, so that $\hat{\mathrm{inj}}(\hat{x}_k) = \mathrm{inj}(x_k) \le r$ for all $k \ge k_0$.

Together with the fact that $\hat{\mathrm{inj}}$ is continuous, these yield:

$$r < r'' \le \hat{\mathrm{inj}}(\hat{x}) = \lim_{k\to\infty} \hat{\mathrm{inj}}(\hat{x}_k) \le r,$$

a contradiction. $\square$

[5] φ is constructed from the inverse of Exp_x followed by a linear isometry from $\mathrm{T}_x\mathcal{M}$ to $\mathbb{R}^d$.

Lemma 10.82. *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

$$\lim_{k\to\infty} x_k = x \qquad \text{and} \qquad \forall k,\ \mathrm{inj}(x_k) \ge R > r = \mathrm{inj}(x).$$

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 $\widehat{\mathrm{inj}} \colon \hat{\mathcal{M}} \to (0, \infty]$ 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

$$B(y,\rho) \subseteq B(x_{k_0}, r'') \qquad \Longrightarrow \qquad \begin{cases} \widehat{\mathrm{inj}}(\hat{y}) = \mathrm{inj}(y) & \text{if } \mathrm{inj}(y) < \rho, \\ \widehat{\mathrm{inj}}(\hat{y}) \geq \rho & \text{otherwise,} \end{cases}$$

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 $\widehat{\mathrm{inj}}(\hat{x}) = \mathrm{inj}(x) = r$, and
2. For all $k \geq k_0$, $B(x_k, r') \subset B(x_{k_0}, r'')$ and $\mathrm{inj}(x_k) \geq R > r'$, so that $\widehat{\mathrm{inj}}(\hat{x}_k) \geq r'$.

Together with the fact that $\widehat{\mathrm{inj}}$ is continuous, these yield:

$$r = \mathrm{inj}(x) = \widehat{\mathrm{inj}}(\hat{x}) = \lim_{k\to\infty} \widehat{\mathrm{inj}}(\hat{x}_k) \geq r' > r,$$

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 Riemannian distances, geodesics and completeness, covered in Section 10.1.

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 in the 70s and 80s. Excellent reference books on this topic include one by Udrişte [Udr94] and another by Rapcsák [Rap97]. There are some variations in how geodesic convexity is defined by different authors, in part because notions of convexity useful to differential geometry are somewhat different from those that are useful to optimization; we mostly follow [Rap97, §6].

Applications 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] and more. Zhang and Sra analyze a collection of algorithms specifically designed for optimization of geodesically convex functions [ZS16], providing worst-case iteration complexity results. See also Section 11.6 for recent research questions.

11.1 Convex sets and functions

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, $f \colon S \to \mathbb{R}$ is a *convex function* if S is convex and for all $x, y \in S$ we have:

With some care, we may allow f to take on infinite values [Roc70, HUL01].

$$\forall t \in [0,1], \qquad f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

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:

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) - \frac{t(1-t)\mu}{2}\|x-y\|^2. \quad (11.1)$$

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 that 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.

Perspective needed to generalize convexity.

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.*
2. *If g is differentiable on I:*
   (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 $x \in \operatorname{int} I$.*

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 $x \in I$.

$\operatorname{int} I$ denotes the interior of I.

Proof. For the first point, see [HUL01, p15]. Points 2 and 3 are standard facts: we include proofs for convenience. Note that it is allowed for I not to be open.

2.(a) Assume the inequalities hold. Then, for all $x, y \in I$ and for all $t \in [0,1]$, define $z = (1-t)x + ty$. These two 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

$$g(x + t(y - x)) = g((1 - t)x + ty) \\ \leq (1 - t)g(x) + tg(y) = g(x) + t(g(y) - g(x)).$$

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:

$$g(y) \geq g(x) + (y-x)g'(x) \quad \text{and} \quad g(x) \geq g(y) + (x-y)g'(y).$$

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 this 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. □

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 the notes and references at the end of this chapter for further details.

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 $c\colon [0,1] \to \mathcal{M}$ such that $c(0) = x$, $c(1) = y$ and $c(t)$ is in S for all $t \in [0,1]$.*

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, 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 $c(t) = (1-t)x + ty$.

Check this out.

By Theorem 10.7, any 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. More interestingly, any hemisphere of S^{n-1}, open or closed, is a geodesically convex subset of S^{n-1}. 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 this later in Sections 11.4 and 11.5, notably to handle $\mathbb{R}^n_+$ and $\mathrm{Sym}(n)^+$: entrywise positive vectors, and positive definite matrices.

In a geodesically convex set, any two points are joined by at least one geodesic segment contained in S. Restricting a function f on S to these segments yields real functions on a real interval. If all of these restrictions 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 the set S is geodesically convex and $f \circ c: [0, 1] \to \mathbb{R}$ is (strictly) convex for all geodesic segments $c: [0, 1] \to \mathcal{M}$ whose image is in S (with $c(0) \neq c(1)$).*

This is defined with respect to the Riemannian manifold $\mathcal{M}$ for which $S \subseteq \mathcal{M}$ is geodesically convex. Here too, if $\mathcal{M}$ is a Euclidean space, we recover the standard definition.

In other words, for S a geodesically convex set, $f: S \to \mathbb{R}$ 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,

We say a curve 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 $t \in [0, 1]$.

$$\forall t \in [0, 1], \qquad f(c(t)) \leq (1 - t) f(x) + t f(y).$$

If furthermore whenever $x \neq y$ we have

$$\forall t \in (0, 1), \qquad f(c(t)) < (1 - t) f(x) + t f(y),$$

then 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.*

In analogy with (11.1), we also extend the notion of strong convexity.

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 all geodesic segments $c: [0, 1] \to \mathcal{M}$ whose image is in S we have*

[Udr94, p187]

$$f(c(t)) \leq (1 - t) f(c(0)) + t f(c(1)) - \frac{t(1 - t)\mu}{2} L(c)^2,$$

where $L(c) = \|c'(0)\|_{c(0)}$ is the length of the geodesic segment. The latter 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.

As for standard convexity in Euclidean space, 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. □

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. When a maximizer occurs 'strictly inside' the domain, the situation is much less interesting. We formalize this below, with a definition and a proposition.

Definition 11.8. *Let S be a geodesically convex set on a Riemannian manifold $\mathcal{M}$. The* relative interior *of S, denoted by* relint S, *is the set of points $x \in S$ with the following property: for all $y \in S$, all geodesics $c: [0, 1] \to \mathcal{M}$ 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.*

This definition is an extension from the classical case [Roc70, Thm. 6.4]. Picture a two-dimensional triangle or disk in $\mathbb{R}^3$.

Proposition 11.9. *Let $f: S \to \mathbb{R}$ be geodesically convex. If f attains its maximum at a point x in the relative interior of S, then f is constant on S.*

Akin to [Roc70, Thm. 32.1].

Proof. Pick an arbitrary point $y \in S$. Our goal is to show $f(y) = f(x)$. Consider any geodesic $c: [0, 1] \to \mathcal{M}$ 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 $f(x) \geq f(y)$. Combining, $f(x) = f(y)$ as announced. □

A connected, complete Riemannian manifold is a geodesically convex set, and its relative interior is the whole manifold itself. 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 continuous and 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.

We close this section with two useful statements, without proofs.

Proposition 11.11. *If S_1 is a geodesically convex set in a Riemannian manifold $\mathcal{M}_1$, and similarly for S_2 in $\mathcal{M}_2$, then $S_1 \times S_2$ is geodesically convex in the Riemannian product manifold $\mathcal{M}_1 \times \mathcal{M}_2$.*

[Rap97, Cor. 6.1.2]

Proposition 11.12. *If $f\colon S \to \mathbb{R}$ is geodesically convex, then f is continuous on the interior of S.*

[Rap97, Thm. 6.1.8]

The interior of S is the set of points in S that have a neighborhood (with respect to the topology of $\mathcal{M}$) included in S.

Exercise 11.13. *Show that the sublevel sets (strict and non strict) of a geodesically convex function are geodesically convex sets. (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].) In particular, the set of global minimizers is geodesically convex.*

Exercise 11.14. *Let $i \in I$ index an arbitrary collection of geodesically convex functions $f_i\colon S \to \mathbb{R}$ and scalars $\alpha_i \in \mathbb{R}$. Define the sublevel sets*

[Rap97, Lem. 6.1.1]

$$S_i = \{x \in S : f_i(x) \leq \alpha_i\}.$$

Show that the intersection $S' = \cap_{i\in I} S_i$ is geodesically convex.

Here is a take-away from Exercise 11.14: if $f, f_1, \ldots, f_m$ are geodesically convex functions on a geodesically convex set S, then we call

$$\min_{x\in S} f(x) \qquad \text{subject to} \qquad f_i(x) \leq \alpha_i \text{ for } i = 1, \ldots, m \tag{11.3}$$

a *geodesically convex optimization problem*. The set $S \cap S_1 \cap \cdots \cap S_m$ of points which satisfy the constraints is geodesically convex, hence any local minimizer of f restricted to that set is a global minimizer. If $g_j\colon S \to \mathbb{R}$ for $j = 1, \ldots, p$ are geodesically linear (Definition 11.4), then we can also allow equality constraints of the form $g_j(x) = \beta_j$. Indeed, this constraint is equivalent to the two constraints $g_j(x) \leq \beta_j$ and $-g_j(x) \leq -\beta_j$: both allowed since both g_j and $-g_j$ are geodesically convex.

Exercise 11.15. *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 11.16. *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.*

11.3 *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.

On a technical note, recall from Section 10.2 that a geodesic segment $c \colon [0, 1] \to \mathcal{M}$ admits a unique extension to a maximally large open interval containing $[0, 1]$: this is how we make sense of $c'(0)$ and $c'(1)$.

See for example [HUL01, Thm. B.4.1.1, p110] for the same statement in the Euclidean case.

Theorem 11.17. *Let S be a geodesically convex set on a Riemannian manifold $\mathcal{M}$ and let f be a real function, differentiable[1] in a neighborhood of S. Then, $f \colon S \to \mathbb{R}$ is geodesically convex if and only if for all geodesic segments $c \colon [0, 1] \to \mathcal{M}$ contained in S we have (defining $x = c(0)$):*

$$\forall t \in [0, 1], \qquad f(c(t)) \geq f(x) + t \left\langle \mathrm{grad} f(x), c'(0) \right\rangle_x. \qquad (11.4)$$

Moreover, f is geodesically μ-strongly convex for some $\mu > 0$ if and only if

$$\forall t \in [0, 1], \quad f(c(t)) \geq f(x) + t \left\langle \mathrm{grad} f(x), c'(0) \right\rangle_x + t^2 \frac{\mu}{2} L(c)^2. \quad (11.5)$$

Finally, f is geodesically strictly convex if and only if, whenever $c'(0) \neq 0$,

$$\forall t \in (0, 1], \qquad f(c(t)) > f(x) + t \left\langle \mathrm{grad} f(x), c'(0) \right\rangle_x. \qquad (11.6)$$

[1] In Section 10.4, we defined what it means for a function on a manifold to be differentiable (as opposed to smooth).

Proof. By definition, f 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 is geodesically convex then the inequalities (11.4) hold. The other way around, if the inequalities (11.4) hold, then (by reparameterization of c) we conclude that $f \circ c$ is convex for all c as prescribed, hence f 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 (11.6) holds. Conclude similarly to the first part. $\square$

In particular, if $f: \mathcal{M} \to \mathbb{R}$ is differentiable, and if it is geodesically convex when restricted to a subset S, 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

$$\forall t \in [0, 1], \qquad f(\mathrm{Exp}_x(tv)) \geq f(x) + t \langle \mathrm{grad} f(x), v \rangle_x, \tag{11.7}$$

and the inequality is strict for $t \neq 0$ if f is geodesically strictly convex. Furthermore, if f is geodesically μ-strongly convex, then

$$\forall t \in [0, 1], \quad f(\mathrm{Exp}_x(tv)) \geq f(x) + t \langle \mathrm{grad} f(x), v \rangle_x + t^2 \frac{\mu}{2} \|v\|_x^2. \tag{11.8}$$

These convenient inequalities should be compared with the corresponding inequality we have (in particular) if the gradient of f is L-Lipschitz continuous (Proposition 10.47):

$$\forall t \in [0, 1], \qquad f(\mathrm{Exp}_x(tv)) \leq f(x) + t \langle \mathrm{grad} f(x), v \rangle_x + t^2 \frac{L}{2} \|v\|_x^2.$$

When both types of inequalities hold, we can obtain strong guarantees for optimization algorithms [ZS16].

The following corollary is of particular importance to optimization. Note that we need the geodesically convex domain to be open. This is because it is possible for a global minimizer to have nonzero gradient, provided it lies on the boundary of S.

Corollary 11.18. *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 11.17 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 4.4 to conclude that if x is a global minimizer, then $\mathrm{grad} f(x) = 0$. $\square$

We now turn to a characterization of convexity based on second-order derivatives, also with the requirement that the domain be open.

Theorem 11.19. *Let $f: S \to \mathbb{R}$ be twice differentiable on an open geodesically convex set S. The function f is*

See for example [HUL01, Thm. B.4.3.1, p115] for most of the same statement in the Euclidean case.

1. *Geodesically convex if and only if* $\operatorname{Hess} f(x) \succeq 0$;

2. *Geodesically μ-strongly convex if and only if* $\operatorname{Hess} f(x) \succeq \mu \operatorname{Id}$;

3. *Geodesically strictly convex if (but not only if)* $\operatorname{Hess} f(x) \succ 0$,

all understood to hold for all $x \in S$.

Proof. Similarly to the proof of Theorem 11.17, we start with the fact that f is geodesically convex if and only if $f \circ c \colon [0,1] \to \mathbb{R}$ is convex for all geodesic segments $c \colon [0,1] \to \mathcal{M}$ 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

$$(f \circ c)''(t) = \frac{\mathrm{d}}{\mathrm{d}t} \left\langle \operatorname{grad} f(c(t)), c'(t) \right\rangle_{c(t)} = \left\langle \operatorname{Hess} f(c(t))[c'(t)], c'(t) \right\rangle_{c(t)},$$

where we also used that $c''(t) = 0$ since c is a geodesic.

If $\operatorname{Hess} f(x)$ 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

$$\left\langle \operatorname{Hess} f(c(0))[c'(0)], c'(0) \right\rangle_{c(0)} \geq 0$$

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 $c(t) = \operatorname{Exp}_x(tv)$ 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 $\langle \operatorname{Hess} f(x)[v], v \rangle_x \geq 0$, which confirms $\operatorname{Hess} f(x)$ 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 $\operatorname{Hess} f(x)$ 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 $S = (-1,1) \subset \mathbb{R}$. ☐

11.4 *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

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_1, \ldots, x_n > 0\} \tag{11.9}$$

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 and interesting notion of geodesic convexity. 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 $\varphi\colon \mathbb{R}^n_+ \to \mathbb{R}^n$:

$$\varphi(x) = \log(x) = (\log(x_1), \ldots, \log(x_n))^\top. \tag{11.10}$$

This is a diffeomorphism between the manifolds $\mathbb{R}^n_+$ and $\mathbb{R}^n$ because it is smooth and invertible, and $\varphi^{-1}\colon \mathbb{R}^n \to \mathbb{R}^n_+$, defined by

$$\varphi^{-1}(y) = \exp(y) = (e^{y_1}, \ldots, e^{y_n})^\top, \tag{11.11}$$

is smooth. Note also these expressions for the differential of φ and its inverse, both maps from $\mathbb{R}^n$ to $\mathbb{R}^n$:

$$\mathrm{D}\varphi(x)[u] = \left(\frac{u_1}{x_1}, \ldots, \frac{u_n}{x_n}\right)^\top, \quad (\mathrm{D}\varphi(x))^{-1}[z] \quad = (x_1 z_1, \ldots, x_n z_n)^\top.$$

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:

This metric is different also from the one we imposed on the relative interior of the simplex in Exercise 3.55, 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).

$$\langle u, v \rangle_x^+ \triangleq \langle \mathrm{D}\varphi(x)[u], \mathrm{D}\varphi(x)[v] \rangle = \sum_{i=1}^n \frac{u_i v_i}{x_i^2}. \tag{11.12}$$

Since $\mathbb{R}^n_+$ now has two distinct Riemannian geometries, we let

$$\mathcal{M} = (\mathbb{R}^n_+, \langle \cdot, \cdot \rangle^+) \tag{11.13}$$

denote the Riemannian manifold obtained with the pullback metric, to avoid ambiguities.

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

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. In fact, the metric at x is given by the Hessian of the log-barrier function $-\sum_{i=1}^n \log(x_i)$.

$$c(t) = \varphi^{-1}(y + tz) = \exp(y + tz) = \left(e^{y_1 + tz_1}, \ldots, e^{y_n + tz_n}\right), \tag{11.14}$$

for some $y, z \in \mathbb{R}^n$. These are defined for all t, hence $\mathcal{M}$ is complete.

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:

$$c(t) = \exp(\log(x) + t(\log(x') - \log(x))). \tag{11.15}$$

We are now in a good position to study geodesic convexity on $\mathcal{M}$.

Proposition 11.20. *A set $S \subseteq \mathbb{R}^n_+$ is geodesically convex on $\mathcal{M}$ if and only if $C = \log(S)$ is convex in $\mathbb{R}^n$.*

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 (11.15) 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 11.21. *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

$$f(c(t)) \le (1-t)f(x) + tf(x') = (1-t)g(y) + tg(y'),$$

where $x = \exp(y)$, $x' = \exp(y')$, and $c(t)$ is the geodesic uniquely specified by (11.15). 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

$$f(x) = x_1^{a_1} \cdots x_n^{a_n} \tag{11.16}$$

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 $\mathbb{R}^n$—indeed, it is the composition of a linear (hence convex) function of y with a convex, nondecreasing function (see also Exercise 11.16).

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 11.15).

Definition 11.22. *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 exactly one of the coefficients c_k is positive, f is called a monomial.

By Exercise 11.13, this implies 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 11.21, 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

$$\min_{x \in \mathbb{R}^n_+} f(x) \qquad \text{subject to} \qquad f_i(x) \leq 1, \qquad i = 1, \ldots, m, \qquad (11.17)$$
$$g_j(x) = 1, \qquad j = 1, \ldots, p,$$

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 tutorial by Boyd, Kim, Vandenberghe and Hassibi for the more standard construction of this class of problems and a list of applications [BKVH07]. This is also discussed under the lens of geodesic convexity in [Rap97, Ch. 10].

Exercise 11.23. *Show that $\langle \cdot, \cdot \rangle^+$ is indeed a Riemannian metric on $\mathbb{R}^n_+$. Then, show that the Riemannian connection on $\mathcal{M}$ is given by*

$$\nabla_u V = \mathrm{D}\varphi(x)^{-1}\left[\mathrm{D}\tilde{V}(\varphi(x))[\tilde{u}]\right]$$

for all $u \in \mathrm{T}_x\mathcal{M}$ and $V \in \mathfrak{X}(\mathcal{M})$, with $\tilde{u} = \mathrm{D}\varphi(x)[u]$ a vector in $\mathbb{R}^n$ and $\tilde{V}$ a smooth vector field in $\mathbb{R}^n$ defined by $\tilde{V}(\varphi(x)) = \mathrm{D}\varphi(x)[V(x)]$. From there, deduce an expression for $\frac{\mathrm{D}}{\mathrm{d}t}$ on $\mathcal{M}$, and conclude that geodesics of $\mathcal{M}$ correspond to geodesics of $\mathbb{R}^n$ as specified by (11.15).

11.5 *Positive definite matrices*

Consider the set of symmetric, positive definite matrices of size n:

$$\mathrm{Sym}(n)^+ = \{X \in \mathrm{Sym}(n) : X \succ 0\}. \tag{11.18}$$

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,

See Section 4.7 for questions related to the computation of matrix functions and their differentials.

$$\varphi(X) = \log(X). \tag{11.19}$$

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 $\mathrm{T}_X\mathrm{Sym}(n)^+ = \mathrm{Sym}(n)$:

$$\langle U, V \rangle_X^{\log} \triangleq \langle \mathrm{D}\log(X)[U], \mathrm{D}\log(X)[V] \rangle. \tag{11.20}$$

This is called the *Log-Euclidean metric*; it was 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:

- The unique (and minimizing) geodesic connecting $X, X' \in \mathrm{Sym}(n)^+$ with respect to the Log-Euclidean metric is
$$c(t) = \exp(\log(X) + t(\log(X') - \log(X))). \tag{11.21}$$
- A set $S \subseteq \mathrm{Sym}(n)^+$ is geodesically convex in that metric if and only if $\log(S)$ is convex in $\mathrm{Sym}(n)$.
- 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—Riemannian 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:

The central expression ensures the inputs to $\langle \cdot, \cdot \rangle$ are symmetric matrices. The metric at X is given by the Hessian of the log-barrier function $-\log(\det(X))$.

$$\langle U, V \rangle_X^{\mathrm{aff}} = \left\langle X^{-1/2} U X^{-1/2}, X^{-1/2} V X^{-1/2} \right\rangle = \mathrm{Tr}(X^{-1} U X^{-1} V). \tag{11.22}$$

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:

$$\left\langle MUM^\top, MVM^\top \right\rangle^{\text{aff}}_{MXM^\top} = \langle U, V\rangle^{\text{aff}}_X. \tag{11.23}$$

One concrete consequence is that if $c\colon [0,1] \to \mathrm{Sym}(n)^+$ is a smooth curve, then the length of $c(t)$ 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]:

$$\mathrm{Exp}_X(tV) = c(t) = X^{1/2}\exp\left(tX^{-1/2}VX^{-1/2}\right)X^{1/2}. \tag{11.24}$$

This is defined for all t, thus the manifold is complete. In order to ensure $c(1) = X'$ (another positive definite matrix), set V to be

$$\mathrm{Log}_X(X') = X^{1/2}\log(X^{-1/2}X'X^{-1/2})X^{1/2}. \tag{11.25}$$

This provides the initial velocity at X of the unique geodesic segment connecting X and X'. It follows that

$$\mathrm{dist}(X, X')^2 = \left\langle \mathrm{Log}_X(X'), \mathrm{Log}_X(X')\right\rangle^{\text{aff}}_X = \|\log(X^{-1/2}X'X^{-1/2})\|_\mathrm{F}^2, \tag{11.26}$$

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).

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. Expressions for the Riemannian connection, gradient, Hessian and parallel transport can be found in [SH15, §3]. 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.

11.6 *Notes and references*

We follow Rapcsák for the definitions of geodesic convexity [Rap91], in that we define a set to be geodesically convex if every two points in the set are connected by *some* geodesic remaining in the set.

Contrast this with the usual definition in Riemannian geometry, namely: a set S in $\mathcal{M}$ is geodesically convex (in the traditional sense) if all two points $x, y \in S$ are connected by a unique minimizing geodesic segment, and that segment lies in S [Lee18, p166]. Sakai calls such sets *strongly convex* [Sak96, §IV.5]. If a set is strongly convex, then it is also geodesically convex in the sense of Definition 11.2, but the converse is not true: consider the sphere as a counter-example. Any point $x \in \mathcal{M}$ admits a strongly convex neighborhood (hence also a geodesically convex one) in the form of a geodesic ball centered at x [Lee18, Thm. 6.17]: this is helpful to prove local convergence results in optimization.

Another notion is that of a *totally convex set S* as defined by Sakai and also Udrişte [Udr94]: for all two points $x, y \in S$, *all* geodesics of $\mathcal{M}$ which connect x to y must remain in S.

O'Neill [O'N83, Def. 5.5] defines $S \subseteq \mathcal{M}$ to be an *open* convex set of a Riemannian manifold $\mathcal{M}$ if it is open and it is a *normal neighborhood* [O'N83, p71] of each of its points. In particular, this implies that for all $x, y \in S$, there is a *unique* (not necessarily minimizing) geodesic connecting x to y in S. There may be other geodesics connecting x to y which do not remain in S.

In several papers about geodesic convexity, these differences are little or not addressed because they deal with *Hadamard manifolds*: on such manifolds, any two points are connected by exactly one geodesic (which is then necessarily minimizing). This simplifies matters significantly (in particular, all above definitions become equivalent), and also provides the most favorable setup for convexity. Euclidean spaces, $\mathbb{R}^n_+$ and $\mathrm{Sym}(n)^+$ with the metrics we discussed all have that property.

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]. Much 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.

More generally, 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 $f \colon S \to \mathbb{R}$ 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 connection 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.

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 [BHTVN19] for a discussion of that notion on Riemannian manifolds.

Perhaps the most famous algorithm for convex optimization in $\mathbb{R}^n$ is the accelerated gradient method (also known as the fast gradient method). There is theoretical 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] and of a continuous-time perspective in [AOBL19], and approaches based on estimate sequences [AS20, AOBL20].

Recently, interest in geodesic convexity over the cone of positive definite matrices is growing in relation to the Brascamp–Lieb constant and the operator scaling problem: see an introductory treatment of geodesic convexity for this topic by Vishnoi [Vis18].

Bibliography

[Abb84] E. Abbott. *Flatland: A Romance of Many Dimensions*. Seeley & Co., 1884.

[ABBC20] N. Agarwal, N. Boumal, B. Bullins, and C. Cartis. Adaptive regularization with cubics on manifolds. *Mathematical Programming*, 2020.

[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.

[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.

[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.

[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.

[AK06] P.-A. Absil and K. Kurdyka. On the stable equilibrium points of gradient systems. *Systems & Control Letters*, 55(7):573–577, July 2006.

[AM12] P.-A. Absil and J. Malick. Projection-like retractions on matrix manifolds. *SIAM Journal on Optimization*, 22(1):135–158, 2012.

[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.

[AMS08] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008.

[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.

[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.

[AOBL19] F. Alimisis, A. Orvieto, G. Bécigneul, and A. Lucchi. A continuous-time perspective for modeling acceleration in Riemannian optimization. *arXiv 1910.10782*, 2019.

[AOBL20] F. Alimisis, A. Orvieto, G. Bécigneul, and A. Lucchi. Practical accelerated optimization on Riemannian manifolds. *arXiv 2002.04144*, 2020.

[AS20] K. Ahn and S. Sra. From Nesterov's estimate sequence to Riemannian acceleration. *arXiv 2001.08876*, 2020.

[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.

[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.

[BAC18] N. Boumal, P.-A. Absil, and C. Cartis. Global rates of convergence for nonconvex optimization on manifolds. *IMA Journal of Numerical Analysis*, 2018.

[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, 2020.

[Bar95] A.I. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete & Computational Geometry*, 13(1):189–202, 1995.

[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.

[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.

[BC70] F. Brickell and R.S. Clark. *Differentiable manifolds: an introduction*. Van Nostrand Reinhold, 1970.

[Ber19] R. Bergmann. Manopt.jl, 2019.

[Bes87] A.L. Besse. *Einstein Manifolds*. Springer Berlin Heidelberg, 1987.

[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.

[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.

[BH18] R. Bergmann and R. Herzog. Intrinsic formulation of KKT conditions and constraint qualifications on smooth manifolds. *arXiv preprint arXiv:1804.06214*, 2018.

[Bha07] R. Bhatia. *Positive definite matrices*. Princeton University Press, 2007.

[BHTVN19] R. Bergmann, R. Herzog, D. Tenbrinck, and J. Vidal-Núñez. Fenchel duality theory and a primal-dual algorithm on Riemannian manifolds. *arXiv 1908.02022*, 2019.

[BKVH07] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi. A tutorial on geometric programming. *Optimization and Engineering*, 8(1):67–127, 2007.

[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.

[BM05] S. Burer and R.D.C. Monteiro. Local minima and convergence in low-rank semidefinite programming. *Mathematical Programming*, 103(3):427–444, 2005.

[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.

[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.

[Bon13] S. Bonnabel. Stochastic gradient descent on Riemannian manifolds. *Automatic Control, IEEE Transactions on*, 58(9):2217–2229, 2013.

[BV19] P. Breiding and N. Vannieuwenhoven. The condition number of riemannian approximation problems, 2019.

[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.

[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.

[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.

[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.

[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.

[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.

[Cha06] I. Chavel. *Riemannian geometry: a modern introduction*, volume 108 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2006.

[Cif19] D. Cifuentes. Burer-Monteiro guarantees for general semidefinite programs. *arXiv 1904.07147*, 2019.

[CM19] D. Cifuentes and A. Moitra. Polynomial time guarantees for the Burer-Monteiro method. *arXiv 1912.01745*, 2019.

[CMSZ18] S. Chen, S. Ma, A.M.C. So, and T. Zhang. Proximal gradient method for nonsmooth optimization over the Stiefel manifold. *arXiv preprint arXiv:1811.00980*, 2018.

[Dar94] R.W.R. Darling. *Differential forms and connections*. Cambridge University Press, 1994.

[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.

[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.

[dCN95] J.X. da Cruz Neto. *Métodos Geodésicos na Programaçao Matemática*. PhD thesis, Ph.D. Thesis, COPPE/UFRJ, Rio de Janeiro, Brazil, 1995.

[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.

[DE99] L. Dieci and T. Eirola. On smooth decompositions of matrices. *SIAM Journal on Matrix Analysis and Applications*, 20(3):800–819, 1999.

[Deh95] J. Dehaene. *Continuous-time matrix algorithms, systolic algorithms and adaptive neural networks*. PhD thesis, Katholieke universiteit Leuven, 1995.

[DH94] E. Dudek and K. Holly. Nonlinear orthogonal projection. *Annales Polonici Mathematici*, 59(1):1–31, 1994.

[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.

[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.

[dOF18] 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. *arXiv:1810.11636*, 2018.

[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.

[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.

[Fep17] F. Feppon. Riemannian geometry of matrix manifolds for Lagrangian uncertainty quantification of stochastic fluid flows. Master's thesis, Massachusetts Institute of Technology, 2017.

[FL19] F. Feppon and P. Lermusiaux. The extrinsic geometry of dynamical systems tracking nonlinear matrix projections, 2019.

[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.

[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.

[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.

[Gab82] D. Gabay. Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, 37(2):177–219, 1982.

[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.

[GHL04] S. Gallot, D. Hulin, and J. LaFontaine. *Riemannian geometry*. Springer Verlag, 2004.

[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.

[GW08] A. Griewank and A. Walther. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. Society for Industrial and Applied Mathematics, 2 edition, 2008.

[GZAL14] J. Goes, T. Zhang, R. Arora, and G. Lerman. Robust stochastic principal component analysis. In *Artificial Intelligence and Statistics*, pages 266–274, 2014.

[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.

[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.

[Hig08] N. Higham. *Functions of Matrices*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

[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.

[HM96] U. Helmke and J.B. Moore. *Optimization and dynamical systems*. Springer Science & Business Media, 1996.

[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.

[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.

[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.

[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.

[Hua13] Wen Huang. *Optimization algorithms on Riemannian manifolds with applications*. PhD thesis, Florida State University, 2013.

[HUL01] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis*. Grundlehren Text Editions. Springer-Verlag Berlin Heidelberg, 1 edition, 2001.

[IEE08] IEEE. Ieee standard for floating-point arithmetic. *IEEE Std 754-2008*, pages 1–70, 2008.

[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.

[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.

[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.

[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.

[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.

[KMU+19] V. Khrulkov, L. Mirvakhabova, E. Ustinova, I. Oseledets, and V. Lempitsky. Hyperbolic image embeddings. *arXiv 1904.02239*, 2019.

[KN63] S. Kobayashi and K. Nomizu. *Foundations of differential geometry, Vol. I*. John Wiley and Sons, 1963.

[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.

[KSV14] D. Kressner, M. Steinlechner, and B. Vandereycken. Low-rank tensor completion by Riemannian optimization. *BIT Numerical Mathematics*, 54(2):447–468, Jun 2014.

[Lag07] C. Lageman. Pointwise convergence of gradient-like systems. *Mathematische Nachrichten*, 280(13–14):1543–1558, October 2007.

[LB19] C. Liu and N. Boumal. Simple algorithms for optimization on Riemannian manifolds with constraints. *Applied Mathematics and Optimization*, 2019.

[Lee12] J.M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer-Verlag New York, 2nd edition, 2012.

[Lee18] J.M. Lee. *Introduction to Riemannian Manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 2018.

[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.

[Łoj65] S. Łojasiewicz. Ensembles semi-analytiques. *Lecture Notes IHES (Bures-sur-Yvette)*, 1965.

[Lue72] D.G. Luenberger. The gradient projection method along geodesics. *Management Science*, 18(11):620–631, 1972.

[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.

[Man02] J.H. Manton. Optimization algorithms exploiting unitary constraints. *IEEE Transactions on Signal Processing*, 50(3):635–650, March 2002.

[Mat96] R. Mathias. A chain rule for matrix functions and applications. *SIAM Journal on Matrix Analysis and Applications*, 17(3):610–620, 1996.

[Mic08] P.W. Michor. *Topics in differential geometry*, volume 93. American Mathematical Society, 2008.

[MMP18] L. Malagò, L. Montrucchio, and G. Pistone. Wasserstein Riemannian geometry of positive definite matrices, 2018.

[Moa03] M. Moakher. Means and averaging in the group of rotations. *SIAM Journal on Matrix Analysis and Applications*, 24(1):1–16, 2003.

[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.

[MS16] B. Mishra and R. Sepulchre. Riemannian preconditioning. *SIAM Journal on Optimization*, 26(1):635–660, 2016.

[MT11] M. McCoy and J.A. Tropp. Two proposals for robust PCA using semidefinite programming. *Electronic Journal of Statistics*, 5:1123–1160, 2011.

[MV13] B. Mishra and B. Vandereycken. A Riemannian approach to low-rank algebraic Riccati equations. *arXiv preprint arXiv:1312.4883*, 2013.

[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.

[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.

[NSAY+19] V.A. Nguyen, S. Shafieezadeh-Abadeh, M.-C. Yue, D. Kuhn, and W. Wiesemann. Calculating optimistic likelihoods using (geodesically) convex optimization, 2019.

[NW06] J. Nocedal and S. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer Science & Business Media, 2 edition, 2006.

[O'N83] B. O'Neill. *Semi-Riemannian geometry: with applications to relativity*, volume 103. Academic Press, 1983.

[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.

[Pea94] B.A. Pearlmutter. Fast exact multiplication by the Hessian. *Neural Computation*, 6:147–160, 1994.

[Pet06] P. Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*. Springer, 2 edition, 2006.

[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.

[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.

[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.

[Qi11] C. Qi. *Numerical optimization methods on Riemannian manifolds*. PhD thesis, Florida State University, Tallahassee, FL, 2011.

[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.

[Rap91] T. Rapcsák. Geodesic convexity in nonlinear optimization. *Journal of Optimization Theory and Applications*, 69(1):169–183, Apr 1991.

[Rap97] T. Rapcsák. *Smooth Nonlinear Optimization in* $\mathbb{R}^n$, volume 19 of *Nonconvex Optimization and Its Applications*. Springer, 1997.

[Roc70] R.T. Rockafellar. *Convex analysis*. Princeton University Press, Princeton, NJ, 1970.

[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.

[Sak96] T. Sakai. *Riemannian geometry*, volume 149. American Mathematical Society, 1996.

[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.

[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.

[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.

[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.

[SI15] H. Sato and T. Iwai. A new, globally convergent Riemannian conjugate gradient method. *Optimization*, 64(4):1011–1031, 2015.

[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.

[Smi94] S.T. Smith. Optimization techniques on Riemannian manifolds. *Fields Institute Communications*, 3(3):113–135, 1994.

[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.

[Sra16] S. Sra. On the matrix square root via geometric optimization. *Electronic Journal of Linear Algebra*, 31(1):433–443, Jul 2016.

[Tak11] A. Takatsu. Wasserstein geometry of Gaussian measures. *Osaka Journal of Mathematics*, 48(4):1005–1026, 2011.

[TB97] L.N. Trefethen and D. Bau. *Numerical linear algebra*. Society for Industrial and Applied Mathematics, 1997.

[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.

[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.

[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.

[Udr94] C. Udrişte. *Convex functions and optimization methods on Riemannian manifolds*, volume 297 of *Mathematics and its applications*. Kluwer Academic Publishers, 1994.

[UV13] A. Uschmajew and B. Vandereycken. The geometry of algorithms using hierarchical tensors. *Linear Algebra and its Applications*, 439(1):133–166, July 2013.

[Van10] B. Vandereycken. Riemannian and multilevel optimization for rank-constrained matrix problems. *Faculty of Engineering, Katholieke Universiteit Leuven*, 2010.

[Van13] B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013.

[Vav91] S.A. Vavasis. *Nonlinear optimization: complexity issues*. Oxford University Press, Inc., 1991.

[Vis18] N.K. Vishnoi. Geodesic convex optimization: Differentiation on manifolds, geodesics, and convexity. *arXiv 1806.06373*, 2018.

[Wie12] A. Wiesel. Geodesic convexity and covariance estimation. *IEEE Transactions on Signal Processing*, 60(12):6182–6189, Dec 2012.

[WW18] I. Waldspurger and A. Waters. Rank optimality for the Burer-Monteiro factorization. *preprint arXiv:1812.03046*, 2018.

[WY13] Z. Wen and W. Yin. A feasible method for optimization with orthogonality constraints. *Mathematical Programming*, 142(1–2):397–434, 2013.

[YP06] N. Yurii and B.T. Polyak. Cubic regularization of Newton method and its global performance. *Mathematical Programming*, 108(1):177–205, 2006.

[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.

[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.

[ZRS16a] H. Zhang, S.J. Reddi, and S. Sra. Fast stochastic optimization on Riemannian manifolds. *arXiv preprint arXiv:1605.07147*, 2016.

[ZRS16b] 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.

[ZS16] H. Zhang and S. Sra. First-order methods for geodesically convex optimization. In *Conference on Learning Theory*, pages 1617–1638, 2016.

[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.

[ZZ18] J. Zhang and S. Zhang. A cubic regularized Newton's method over Riemannian manifolds. *arXiv preprint arXiv:1805.05565*, 2018.