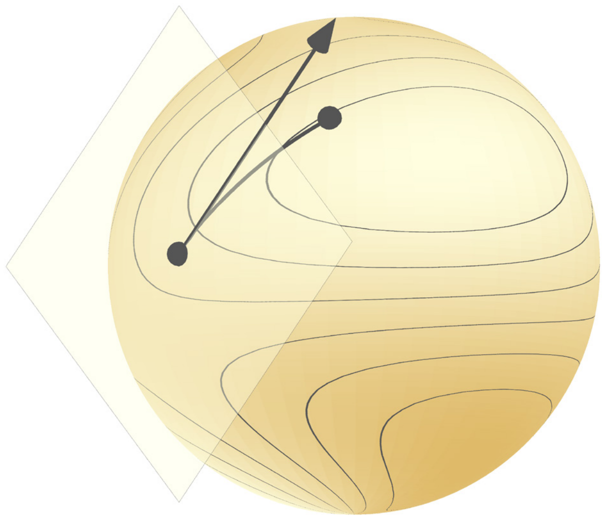
105

Riemannian metrics and gradients

Spring 2023

Optimization on manifolds, MATH 512 @ EPFL

Instructor: Nicolas Boumal



Remember Euclidean gradients

$\mathcal{E}$ is a Euclidean space, that is, a linear space with inner product $\langle\cdot,\cdot\rangle$.

Say $\bar{f}: \mathcal{E} \to \mathbf{R}$ is smooth. What is its gradient?

$\operatorname{grad} \bar{f} : \mathcal{E} \to \mathcal{E}$

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

Example: if $\mathcal{E} = \mathbf{R}^d$ and $\langle u, v\rangle = u^\top v$, then:

$$\operatorname{grad} \bar{f}(x)_i = e_i^\top \operatorname{grad} \bar{f}(x) = \langle e_i, \operatorname{grad} \bar{f}(x) \rangle = \mathrm{D}\bar{f}(x)[e_i]$$

$$= \frac{\partial \bar{f}}{\partial x_i}(x).$$

e_i is the ith col. of I_d

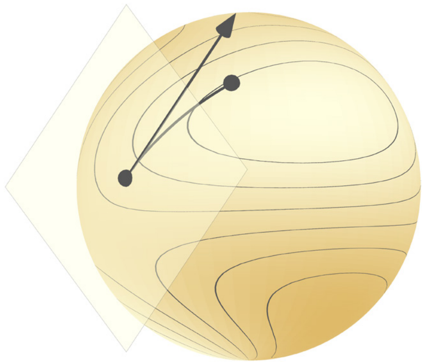
Make each tangent space Euclidean, nicely

Def.: An inner product on $\mathrm{T}_x\mathcal{M}$ is a map $\langle\cdot,\cdot\rangle_x : \mathrm{T}_x\mathcal{M} \times \mathrm{T}_x\mathcal{M} \to \mathbf{R}$ which is bilinear, symmetric and positive definite.
The associated norm on $\mathrm{T}_x\mathcal{M}$ is $\|v\|_x = \sqrt{\langle v, v\rangle_x}$.

Def.: A metric on $\mathcal{M}$ is a choice of inner product $\langle\cdot,\cdot\rangle_x$ for each x.

We would like the metric to "vary smoothly with x".

What does that mean?



Riemannian manifolds

Def.: A metric $\langle \cdot, \cdot \rangle_x$ on $\mathcal{M}$ is a Riemannian metric if it varies smoothly with x, in the following sense:

If V, W are two smooth vector fields on $\mathcal{M}$,
then $x \mapsto \langle V(x), W(x) \rangle_x$ is a smooth function on $\mathcal{M}$.

Def.: A Riemannian manifold is a manifold with a Riemannian metric.

Example 0: A Euclidean space $\mathcal{E}$ with its inner product $\langle \cdot, \cdot \rangle$ is a Riemannian manifold.

$T_x \mathcal{E} = \mathcal{E}$

Example 1: $S^{d-1} = \{x \in \mathbf{R}^d : x^\top x = 1\}$. Embedded in $\mathbf{R}^d$.

$$T_x S^{d-1} = \{ v \in \mathbb{R}^d : x^T v = 0 \} \subseteq \mathbb{R}^d$$

We could define $\langle u, v \rangle_x = u^T G(x) v$, where G is a smooth function of x s.t. $G(x) \succ 0$ for all x.

In particular, we could define:

$$\langle u, v \rangle_x = u^T v. \quad \text{(round metric)}$$

Equip $\mathbb{R}^d$ w/ $\langle u, v \rangle = u^T v$; then $\langle \cdot, \cdot \rangle_x$ is just the restriction of $\langle \cdot, \cdot \rangle$ to the tangent spaces of the sphere.

Example 2: $\mathcal{M} = \{x \in \mathbf{R}^d : x_1, \dots, x_d > 0 \text{ and } x_1 + \cdots + x_d = 1\}$.

Embedded in $\mathbf{R}^d$.

relative interior of the simplex.

$$\mathcal{M} = \{x \in \mathbb{R}^d : x > 0\} \cap \{x \in \mathbb{R}^d : \mathbb{1}^T x - 1 = 0\}$$

$$\mathbb{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^d$$

$$h(x) = \mathbb{1}^T x - 1 \; ; \quad Dh(x)[\dot{x}] = \mathbb{1}^T \dot{x}$$

$$T_x \mathcal{M} = \ker Dh(x) = \ker(\mathbb{1}^T) = (\operatorname{im} \mathbb{1})^\perp = \{v \in \mathbb{R}^d : v_1 + \cdots + v_d = 0\}$$

Fisher-Rao metric: $\langle u, v \rangle_x = \sum_i \frac{u_i v_i}{x_i}$ this is a Riemannian metric for $\mathcal{M}$.

Riemannian submanifolds

If $\mathcal{M}$ is embedded in a Euclidean space $\mathcal{E}$ with inner product $\langle \cdot, \cdot \rangle$ then it is natural to restrict $\langle \cdot, \cdot \rangle_x$ to each tangent space to define:

$$\langle u, v \rangle_x = \langle u, v \rangle$$

Fact: This is a Riemannian metric.

Def.: With *this* metric, $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$.

!! A Riemannian submanifold is not just a submanifold equipped with "a" Riemannian metric.

M is a Riemannian manifold

Riemannian gradients

Def.: The Riemannian gradient of a smooth $f: \mathcal{M} \to \mathbf{R}$ is the vector field $\mathrm{grad} f$ defined by:

$$\forall (x, v) \in \mathrm{T}\mathcal{M}, \qquad \langle \underbrace{\mathrm{grad} f(x)}_{\in \mathrm{T}_x \mathcal{M}}, \overset{\in \mathrm{T}_x \mathcal{M}}{v} \rangle_x = \mathrm{D} f(x)[v].$$

$= (f \circ c)'(0)$, $c(0) = x$, $c'(0) = v$

$= \mathrm{D}\bar{f}(x)[v]$

$\bar{f}|_{\mathcal{M}} = f$

Fact: $\mathrm{grad} f$ is a well-defined smooth vector field.

How to compute $\mathrm{grad} f(x)$?

1. From the definition.
2. Through a retraction.
3. Via a smooth extension (especially good for Riemannian submanifolds).

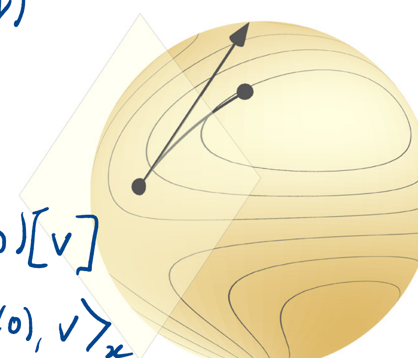
Computing gradients through a retraction

$R : TM \to M \;;\quad (x, v) \mapsto R_x(v)$

$R_x : T_xM \to M \;;\qquad f : M \to \mathbb{R} \;;\qquad \overbrace{f \circ R_x}^{\text{pullback of } f \text{ to } T_xM \text{ through } R.} : \underbrace{T_xM}_{\text{Euclidean space with } \langle \cdot, \cdot \rangle_x} \to \mathbb{R}$

Fact: $\mathrm{grad} f(x) = \mathrm{grad}(f \circ R_x)(0)$

$$
\begin{aligned}
\langle \mathrm{grad} f(x), v \rangle_x &= Df(x)[v] && \forall\, v \in T_xM \\
&= (f \circ c)'(0) && \text{let } c(t) = R_x(tv) \\
&= \frac{d}{dt}\Big(f(R_x(tv)) \Big)\Big|_{t=0} \\
&= \frac{d}{dt}\Big((f \circ R_x)(tv) \Big)\Big|_{t=0} = D(f \circ R_x)(0)[v] \\
&= \langle \mathrm{grad}(f \circ R_x)(0), v \rangle_x
\end{aligned}
$$



Computing gradients through an extension

Say $\mathcal{M}$ is a Riemannian manifold embedded in a Euclidean space $\mathcal{E}$.

Inner product on $\mathrm{T}_x\mathcal{M}$: $\langle\cdot,\cdot\rangle_x$ Inner product on $\mathcal{E}$: $\langle\cdot,\cdot\rangle_{\mathcal{E}}$

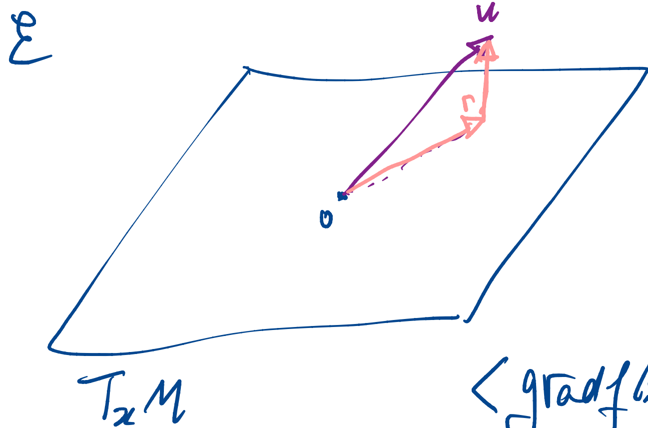
Given $f\colon \mathcal{M} \to \mathbf{R}$ smooth at x, pick a smooth extension $\bar{f}$ around x.

Can we relate $\operatorname{grad} f(x)$ and $\operatorname{grad} \bar{f}(x)$?

$f = \bar{f}|_{\mathcal{M}}$

$\forall v \in \mathrm{T}_x\mathcal{M}$

$$\langle \operatorname{grad} f(x), v \rangle_x = \mathrm{D}f(x)[v] = \mathrm{D}\bar{f}(x)[v] = \langle \operatorname{grad} \bar{f}(x), v \rangle_{\mathcal{E}}$$



$$\operatorname{grad} \bar{f}(x) = \underbrace{\operatorname{grad} \bar{f}(x)_{\parallel}}_{\in T_x M} + \underbrace{\operatorname{grad} \bar{f}(x)_{\perp}}_{\perp T_x M \text{ w.r.t. } \langle \cdot, \cdot \rangle_{\mathcal{E}}}$$

$$\langle \operatorname{grad} f(x), v \rangle_x = \langle \operatorname{grad} \bar{f}(x)_{\parallel}, v \rangle_{\mathcal{E}} + \underbrace{\langle \operatorname{grad} \bar{f}(x)_{\perp}, v \rangle_{\mathcal{E}}}_{= 0 \ \forall v \in T_x M}$$

Gradients on Riemannian submanifolds

$\mathcal{M}$ is an embedded submanifold of a Euclidean space $\mathcal{E}$ with $\langle\cdot,\cdot\rangle$.
$f: \mathcal{M} \to \mathbf{R}$ is smooth and $\bar{f}$ is a smooth extension of f.

Theorem: If $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$, then:

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

$\mathrm{Proj}_x : \mathcal{E} \to \mathcal{E}$ is the orthogonal projector to $T_x\mathcal{M}$.

w.r.t. the Euclidean inner product on $\mathcal{E}$.

Orthogonal projectors to tangent spaces

$\mathcal{M}$ is an embedded submanifold of a Euclidean space $\mathcal{E}$ with $\langle\cdot,\cdot\rangle$.

Def.: For each $x \in \mathcal{M}$, the orthogonal projector to $\mathrm{T}_x\mathcal{M}$ is the linear map $\mathrm{Proj}_x : \mathcal{E} \to \mathcal{E}$ which satisfies:

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

≡ Symmetric

Fact: Proj_x is self-adjoint, that is, $\langle u, \mathrm{Proj}_x(v)\rangle = \langle \mathrm{Proj}_x(u), v\rangle$ for all u, v in $\mathcal{E}$.

Example 0: $\mathcal{M}$ is open in $\mathcal{E}$.

$T_x\mathcal{M} = \mathcal{E}$; $\langle u, v \rangle_x = \langle u, v \rangle_{\mathcal{E}}$

$f: \mathcal{M} \to \mathbb{R}$ is smooth; it is its own smooth extension. Also: $\mathrm{Proj}_x = \mathrm{Identity}$

Example 1: $\mathcal{M} = \{x \in \mathbf{R}^d : x^\top x = 1\}$ with $\mathcal{E} = \mathbf{R}^d$ and $\langle u, v \rangle = u^\top v$.

Equip $\mathcal{M}$ w/ the Riemannian submanifold metric.

$T_x\mathcal{M} = \{v \in \mathbb{R}^d : x^T v = 0\} = (\mathrm{im}\, x)^\perp$

$\mathrm{Proj}_x(v) = v - \alpha x$; should have: $x^T(v - \alpha x) = 0$

$$x^T(v-\alpha x) = x^T v - \alpha x^T x = x^T v - \alpha$$

$$= v - (x^T v)x$$

$$= (I - xx^T)v.$$