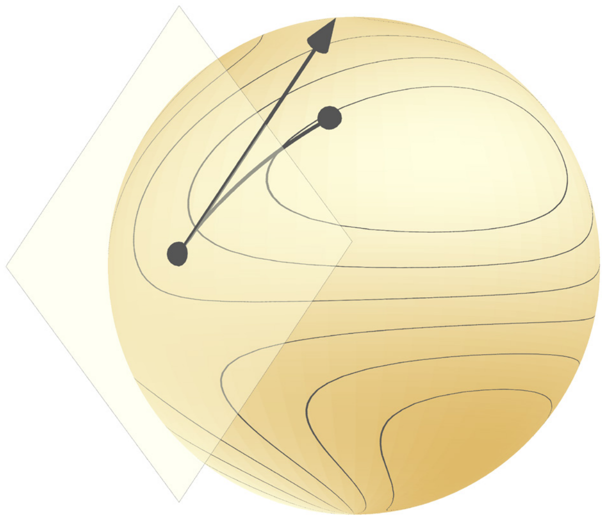
202

Optimality conditions, first order

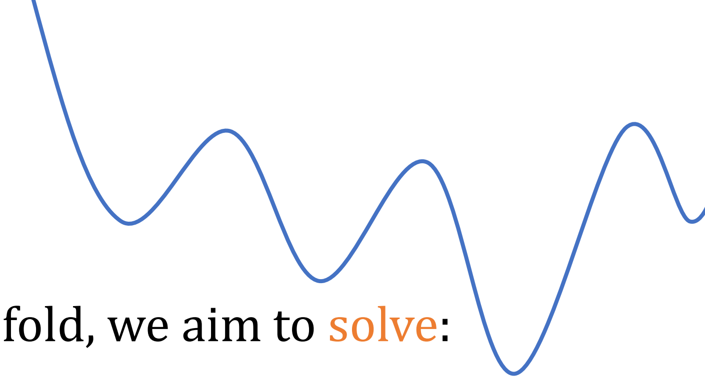
Spring 2023

Optimization on manifolds, MATH 512 @ EPFL

Instructor: Nicolas Boumal



Optimal points



Given a cost function $f : \mathcal{M} \to \mathbf{R}$ on a manifold, we aim to solve:

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

Def.: $x \in \mathcal{M}$ is a global minimum if $f(x) \leq f(y)$ for all $y \in \mathcal{M}$.

Def.: $x \in \mathcal{M}$ is a local minimum if there exists a ngbhd $\mathcal{U}$ of x on $\mathcal{M}$ such that $f(x) \leq f(y)$ for all $y \in \mathcal{U}$.

Critical points, first order

Def.: A point $x \in \mathcal{M}$ is critical or stationary for $f : \mathcal{M} \to \mathbf{R}$ if

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

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

Fact: If x is a local minimum, then it is critical.

Fact: On a Riemannian manifold, x is critical iff $\operatorname{grad} f(x) = 0$.

$$(f\circ c)'(0) \geq 0 \quad \forall c: \quad c(0)=x$$

Fact: On a Riemannian manifold, x is critical iff $\operatorname{grad} f(x) = 0$.

Proof. $(f\circ c)'(0) = Df(x)[v] = \langle \operatorname{grad} f(x), v\rangle_x$, $\quad c(0)=x$, $c'(0)=v$.

① If $\operatorname{grad} f(x) = 0$, then $(f\circ c)'(0) = 0 \geq 0 \quad \forall c$.

② If x is critical, then $\langle \operatorname{grad} f(x), v\rangle_x = (f\circ c)'(0) \geq 0 \quad \forall c$

$\equiv \quad \langle \operatorname{grad} f(x), v\rangle_x \geq 0 \quad \forall v \in T_x\mathcal{M}$

$\equiv \quad \langle \operatorname{grad} f(x), v\rangle_x = 0 \quad \forall v \in T_x\mathcal{M}$

$\equiv \quad \operatorname{grad} f(x) = 0.$ $\square$

Fact: If x is a local minimum, then it is critical.

Proof. Assume x is a local minimum.

There exists a ngbhd $\mathcal{U}$ s.t. $f(y) \geq f(x)$ for all $y \in \mathcal{U}$.

For contradiction, say there exists a smooth curve $c: \mathbf{R} \to \mathcal{M}$ with

$$c(0) = x \qquad \text{and} \qquad (f \circ c)'(0) < 0.$$

The function $g = f \circ c: \mathbf{R} \to \mathbf{R}$ is smooth by composition.

In particular, $g' = (f \circ c)'$ is continuous, and $g'(0) < 0$.

Thus, there exists $\bar{t} > 0$ such that $g'(\tau) < 0$ for all $\tau \in [0, \bar{t}]$, and so

$$f\big(c(t)\big) = g(t) = g(0) + \int_0^t g'(\tau)\mathrm{d}\tau < g(0) = f(x) \qquad \text{for all } t \in (0, \bar{t}].$$

But $I = c^{-1}(\mathcal{U})$ is open in $\mathbf{R}$ (by continuity of c), it contains 0 (since $c(0) \in \mathcal{U}$),

and it has no overlap with $(0, \bar{t}]$ (because $f\big(c(t)\big) \geq f(x)$ for $c(t) \in \mathcal{U}$).

That's impossible.