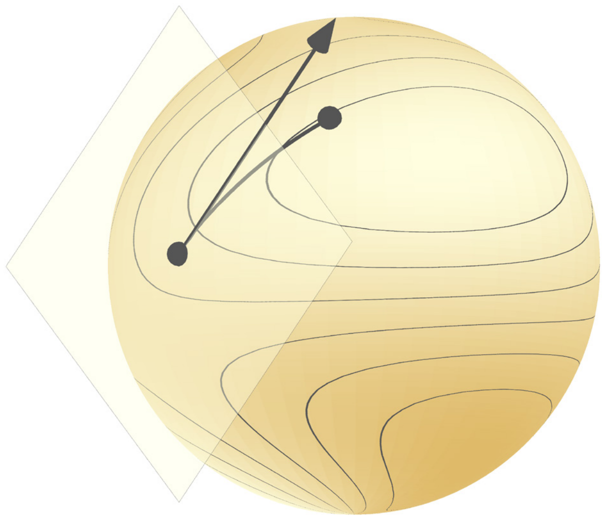
205

Taylor expansions and retractions, second order

Spring 2023

Optimization on manifolds, MATH 512 @ EPFL

Instructor: Nicolas Boumal



Function value along a curve, take two

Let $\mathcal{M}$ be a Riemannian manifold, with associated ∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$.

Consider smooth $f\colon \mathcal{M} \to \mathbf{R}$ and $c\colon \mathbf{R} \to \mathcal{M}$, with $c(0) = x$, $c'(0) = v$.

$$g(t) = f(c(t))$$

$$g(t) = g(0) + t\, g'(0) + \frac{t^2}{2} g''(0) + O(t^3)$$

$$g(0) = f(x);$$

$$g'(t) = Df(c(t))[c'(t)] = \langle \operatorname{grad} f(c(t)), c'(t) \rangle_{c(t)}$$

$$g''(t) = \frac{d}{dt} g'(t) = \frac{d}{dt} \langle \operatorname{grad} f(c(t)), c'(t) \rangle_{c(t)}$$

$\vdots$

by the product rule ④ for $\frac{D}{dt}$

$$= \left\langle \frac{D}{dt}(\operatorname{grad} f \circ c)(t),\ c'(t) \right\rangle_{c(t)} + \left\langle \operatorname{grad} f(c(t)),\ \underbrace{\frac{D}{dt} c'(t)}_{= c''(t)\ (\text{acceleration})} \right\rangle_{c(t)}$$

by the chain rule ③ for $\frac{D}{dt}$

$$= \left\langle \nabla_{c'(t)} \operatorname{grad} f,\ c'(t) \right\rangle_{c(t)} + \left\langle \operatorname{grad} f(c(t)),\ c''(t) \right\rangle_{c(t)}$$

$$= \left\langle \operatorname{Hess} f(c(t))[c'(t)],\ c'(t) \right\rangle_{c(t)} + \left\langle \operatorname{grad} f(c(t)),\ c''(t) \right\rangle_{c(t)}$$

$$g''(0) = \langle \operatorname{Hess} f(x)[v], v \rangle_x + \langle \operatorname{grad} f(x), c''(0) \rangle_x$$

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

Take away

Fact: Consider $f\colon \mathcal{M} \to \mathbf{R}$ smooth. Let $c\colon \mathbf{R} \to \mathcal{M}$ be a smooth curve with $c(0) = x$, $c'(0) = v$ and $c''(0) = w$. Then:

$$f\big(c(t)\big) = f(x) + t\langle \operatorname{grad} f(x), v\rangle_x + \frac{t^2}{2}\Big[\langle \operatorname{Hess} f(x)[v], v\rangle_x + \langle \operatorname{grad} f(x), w\rangle_x\Big] + O(t^3).$$

Second-order retractions

The factor $\langle \mathrm{grad} f(x), w \rangle_x$ vanishes if $\mathrm{grad} f(x) = 0$ or $c''(0) = 0$.

Def.: A retraction is second order if, for all $(x, v) \in \mathrm{T}\mathcal{M}$, the curve $c(t) = R_x(tv)$ satisfies $c''(0) = 0$.

Fact: If R is a second-order retraction, or if $\mathrm{grad} f(x) = 0$, then

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

Example: On the sphere S^{d-1},

the retraction $R_x(v) = \frac{x+v}{\|x+v\|}$ is second order.

$$c(t) = R_x(tv) = \frac{1}{\|x+tv\|}(x+tv) = \frac{1}{\sqrt{1+t^2\|v\|^2}}(x+tv)$$

$g(t) = 1 - \frac{\|v\|^2}{2}t^2 + O(t^3)$

$$c'(t) = g'(t)(x+tv) + g(t)v$$

$$\frac{d}{dt}c'(t) = g''(t)(x+tv) + 2g'(t)v \; ; \quad \frac{d}{dt}c'(0) = \underbrace{g''(0)}_{\neq 0}x + \cancel{\underbrace{2g'(0)}_{=0}v}$$

$$c''(0) = \frac{D}{dt}c'(0) = \mathrm{Proj}_x\left(\frac{d}{dt}c'(0)\right) = 0.$$

More generally, projection retractions are second order: see §5.12.

If R is a second-order retraction, or if $\mathrm{grad} f(x) = 0$, then
$f(R_x(s)) = f(x) + \langle \mathrm{grad} f(x), s \rangle_x + \frac{1}{2} \langle \mathrm{Hess} f(x)[s], s \rangle_x + O(\|s\|_x^3)$

Gradient and Hessian of pullbacks

For the pullback $f \circ R_x$, we already had $\mathrm{grad} f(x) = \mathrm{grad}(f \circ R_x)(0)$.

That holds for all retractions and for all $x \in \mathcal{M}$.

The expansion above yields a corollary that completes the picture:

Fact: If R is a second-order retraction, or if $\mathrm{grad} f(x) = 0$, then

$$\mathrm{Hess} f(x) = \mathrm{Hess}(f \circ R_x)(0)$$