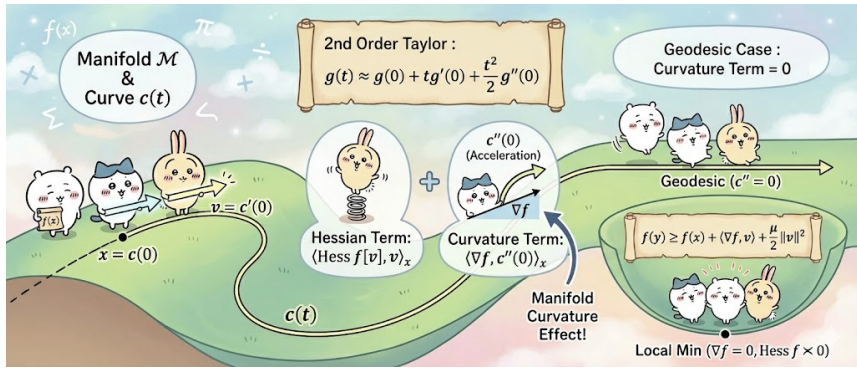


沿曲线的二阶泰勒展开

测地线上的二阶下界估计



1 沿曲线的二阶泰勒展开

在流形优化中，算法通常沿着流形上的曲线移动（例如通过收缩映射生成的曲线）。为了分析函数的下降性质，我们需要理解函数 f 沿着曲线 $c(t)$ 的变化规律。

定理 1: 函数沿曲线的二阶泰勒展开

令 $f: \mathcal{M} \rightarrow \mathbb{R}$ 是光滑函数， $c: I \rightarrow \mathcal{M}$ 是一条光滑曲线，满足 $c(0) = x$ 且 $c'(0) = v \in T_x \mathcal{M}$ 。函数 $g(t) = f(c(t))$ 在 $t = 0$ 处的二阶泰勒展开为

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

其中各项系数为：

$$\begin{aligned} g(0) &= f(x), \quad g'(0) = \langle \text{grad } f(x), v \rangle_x, \\ g''(0) &= \langle \text{Hess } f(x)[v], v \rangle_x + \langle \text{grad } f(x), c''(0) \rangle_x. \end{aligned}$$

这里 $c''(0)$ 是曲线 c 在 x 处的内蕴加速度。

证明. 一阶导数由梯度定义给出： $g'(t) = Df(c(t))[c'(t)] = \langle \text{grad } f(c(t)), c'(t) \rangle_{c(t)}$. 为了求二阶导数，利用黎曼联络 ∇ 诱导的沿曲线协变导数 $\frac{D}{dt}$ 对内积进行求导（利用度量相容性）：

$$\begin{aligned} g''(t) &= \frac{d}{dt} \langle \text{grad } f(c(t)), c'(t) \rangle_{c(t)} \\ &= \left\langle \frac{D}{dt} (\text{grad } f \circ c)(t), c'(t) \right\rangle_{c(t)} + \left\langle \text{grad } f(c(t)), \frac{D}{dt} c'(t) \right\rangle_{c(t)}. \end{aligned}$$

根据黎曼 Hessian 的定义，第一项为：

$$\frac{D}{dt} (\text{grad } f \circ c)(t) = \nabla_{c'(t)} \text{grad } f = \text{Hess } f(c(t))[c'(t)].$$

根据定义，第二项中的 $\frac{D}{dt} c'(t)$ 即为内蕴加速度 $c''(t)$ 。在 $t=0$ 处代入 $c(0) = x$ 和 $c'(0) = v$ 即得结论。 \square

注解 2: 对比欧几里得空间中的泰勒展开

与欧几里得空间中的泰勒展开相比，流形上的公式多出了 $\langle \text{grad } f(x), c''(0) \rangle_x$ 这一项。这揭示了以下几点：

1. **曲线弯曲的影响**：在流形上，即使速度 v 给定，曲线 c 为了留在流形内可能必须产生加速度。这一加速度与梯度的内积会直接影响函数值的二阶变化。

2. **测地线的情形**: 如果 c 是测地线, 则 $c''(0) = 0$, 展开式回归到类似于欧几里得空间的经典形式。
3. **临界点的情形**: 如果 x 是 f 的临界点 ($\text{grad } f(x) = 0$), 则加速度项消失, 二阶导数完全由 $\text{Hess } f(x)$ 决定。

这一结论对于后续理解收缩映射在优化算法中的二阶收敛性至关重要。

2 测地线上的二阶下界估计

推论 3: 函数沿曲线的均值定理型展开

令 $c : [0, 1] \rightarrow \mathcal{M}$ 是黎曼流形 \mathcal{M} 上的一条光滑曲线, 满足 $c(0) = x$ 且 $c(1) = y$. 那么存在 $t \in (0, 1)$, 使得

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

此外, 若 c 是测地线, 则其速率 $\|c'(t)\|_{c(t)}$ 为常数。

证明. 定义标量函数 $g : [0, 1] \rightarrow \mathbb{R}$ 为 $g(t) = f(c(t))$ 。根据单变量带拉格朗日余项的泰勒展开定理, 存在 $t \in (0, 1)$ 使得:

$$g(1) = g(0) + g'(0) + \frac{1}{2} g''(t).$$

根据定理 1 中的求导公式, 代入 $g(1), g(0), g'(0)$ 和 $g''(t)$ 即可得证。

对于测地线 c ，有 $c''(t) = \frac{D}{dt}c'(t) = 0$ 。计算速率平方的导数：

$$\frac{d}{dt}\|c'(t)\|_{c(t)}^2 = 2\left\langle \frac{D}{dt}c'(t), c'(t) \right\rangle_{c(t)} = 2\langle c''(t), c'(t) \rangle_{c(t)} = 0.$$

故速率为常数。 □

引理 4: 测地线上的二阶下界估计

令 $c : [0, 1] \rightarrow \mathcal{M}$ 是黎曼流形 \mathcal{M} 上的一条测地线，满足 $c(0) = x$ 且 $c(1) = y$ 。假设对于所有 $t \in [0, 1]$ ，黎曼 Hessian 满足 $\text{Hess } f(c(t)) \succeq \mu \cdot \text{Id}$ 。那么

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

其中 $v = c'(0)$ 是初始速度向量。

证明. 根据推论 3，存在 $t \in (0, 1)$ 满足展开式。由于 c 是测地线， $c''(t) = 0$ ，且速率 $\|c'(t)\|_{c(t)} = \|v\|_x$ 为常数。由 $\text{Hess } f(c(t)) \succeq \mu \cdot \text{Id}$ 知：

$$\langle \text{Hess } f(c(t))[c'(t)], c'(t) \rangle_{c(t)} \geq \mu \|c'(t)\|_{c(t)}^2 = \mu \|v\|_x^2.$$

代入展开式即得结论。 □

引理 4 是欧几里得空间中强凸函数性质在黎曼流形上的自然推广。在后续学习**测地凸性 (Geodesic Convexity)**时，这个不等式将起到核心作用。

参考文献

[Bou23] Nicolas Boumal. *An introduction to optimization on smooth manifolds*. Cambridge University Press, 2023.