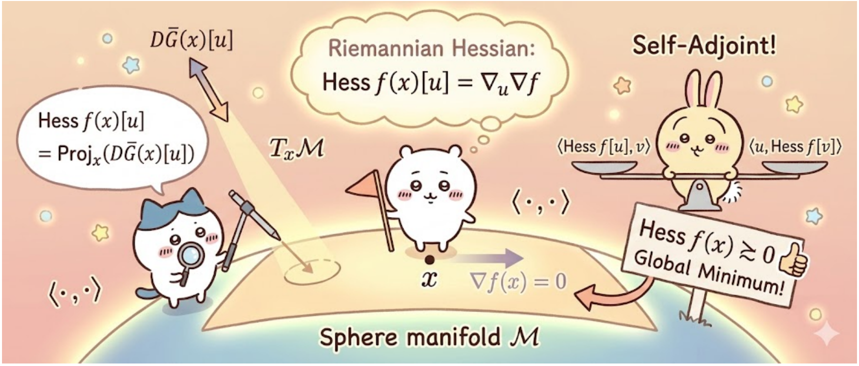
# 曲线上的向量场微分

北极甜虾 (南半球版)

2026 年 1 月 19 日

Manifold Optimization Notes

黎曼 Hessian 的定义 | 例子：球面上的二次型




## 1 黎曼 Hessian 的定义

在欧几里得空间优化中，Hessian 矩阵（二阶导数）对于分析临界点的性质（极大值, 极小值或鞍点）以及构造牛顿法等二阶优化算法至关重要。现在我们可以利用利用黎曼联络 $\nabla$ 来定义一种流形上的二阶导数的概念, 之后会学习用它来设计流形上的二阶优化算法。

**定义 1: 黎曼 Hessian**

设 $f : \mathcal{M} \to \mathbb{R}$ 是黎曼流形 $\mathcal{M}$ 上的光滑函数，$\nabla$ 是其黎曼联络。函数 $f$ 在点 $x \in \mathcal{M}$ 处的黎曼 Hessian 是一个线性映射 $\operatorname{Hess} f(x) : T_x\mathcal{M} \to T_x\mathcal{M}$，定义为

$$\operatorname{Hess} f(x)[u] = \nabla_u \operatorname{grad} f, \quad \forall u \in T_x\mathcal{M}.$$

等价地，$\operatorname{Hess} f : \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$，$\operatorname{Hess} f[U] = \nabla_U \operatorname{grad} f$.

黎曼 Hessian 的自伴随性保证了它拥有实特征值，这是研究凸性和极值判定的基础。这与欧氏空间中 Hessian 矩阵是对称矩阵这一事实是完全对应的。

**命题 2: 黎曼 Hessian 的自伴随性**

黎曼 Hessian 是自伴随的:

$$\langle \operatorname{Hess} f(x)[u], v\rangle_x = \langle u, \operatorname{Hess} f(x)[v]\rangle_x, \quad \forall x \in \mathcal{M}, u, v \in T_x\mathcal{M},$$

或者等价地,

$$\langle \operatorname{Hess} f[U], V\rangle = \langle U, \operatorname{Hess} f[V]\rangle.$$

证明. 用度量相容性，我们有

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

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

根据李括号的定义, 以及对称性，我们有

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

证明完毕。 □

直接计算黎曼 Hessian 可能比较繁琐，但对于嵌入在欧几里得空间中的子流形, 我们可以利用正交投影算子得到一个更易于计算的公式。以下是上次笔记中命题 1 的推论。

**推论 3: 嵌入子流形的黎曼 Hessian**

设 $\mathcal{M}$ 是 $\mathcal{E}$ 的黎曼子流形，对于一个光滑函数 $f: \mathcal{M} \to \mathbb{R}$，如果 $\bar{G}$ 是黎曼梯度场 $\operatorname{grad} f$ 在 $\mathcal{E}$ 的光滑延拓，那么

$$\operatorname{Hess} f(x)[u] = \operatorname{Proj}_x(D\bar{G}(x)[u]), \quad \forall x \in \mathcal{M}, u, v \in T_x\mathcal{M}.$$

## 2 例子：球面上的二次型

**例子 4: 球面上的二次型**

考虑函数 $\bar{f}(x) = \frac{1}{2}x^\top Ax$ ($A \in \mathbb{R}^{d\times d}$ 是对称矩阵)，以及它在作为 $\mathbb{R}^d$ 的黎曼子流形的单位球面 $S^{d-1}$ 上的限制 $f = \bar{f}|_{S^{d-1}}$. 首先，我们已经确定了 $\bar{f}$ 的欧几里得梯度和 $f$ 的黎曼梯度：

$$\operatorname{grad} \bar{f}(x) = Ax,$$

$$\operatorname{grad} f(x) = \operatorname{Proj}_x(\operatorname{grad} \bar{f}(x)) = (I_d - xx^\top)Ax = Ax - (x^\top Ax)x$$

为了计算黎曼 Hessian，我们需要黎曼梯度场 $\operatorname{grad} f$ 在环境空间中的一个光滑延拓 $\bar{G}$. 通常情况下，黎曼梯度的表达式本身就提供了一个自然的选择，即令

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

根据乘积法则，$\bar{G}$ 在 $x$ 点处沿切向量 $u$ 方向的微分为

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

最后，将该结果正交投影到 $x$ 点处的切空间即可得到黎曼 Hessian（仅定义在切空间 $T_xS^{d-1}$ 上，而非整个 $\mathbb{R}^d$ 空间)：

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

**命题 5: 球面二次型的全局极小值判定, Exercise 5.18**

延续例子 4，考虑 $f(x) = \frac{1}{2}x^\top Ax$ 在单位球面 $S^{d-1}$ 上的限制，其中 $A$ 是对称矩阵。证明：如果 $x$ 满足 $\operatorname{grad} f(x) = 0$ 且 $\operatorname{Hess} f(x) \succeq 0$（半正定），则 $x$ 是 $f$ 在球面上的一个全局极小值点。即证明 $x$ 是 $A$ 的最小特征值对应的特征向量。

证明. 根据例子 4，球面上的黎曼梯度为

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

当 $\operatorname{grad} f(x) = 0$ 时，有 $Ax = (x^\top Ax)x$. 令 $\lambda = x^\top Ax$，则该式变为 $Ax = \lambda x$。由于 $x \in S^{d-1}$，这表明 $x$ 是 $A$ 的一个单位特征向量，其对应的特征值为 $\lambda$. 此时函数值为 $f(x) = \frac{1}{2}\lambda$.

同样根据例子 4，对于任何切向量 $u \in T_x S^{d-1}$（即满足 $x^\top u = 0$），黎曼 Hessian 作用于 $u$ 的结果为

$$\operatorname{Hess} f(x)[u] = Au - (x^\top Au)x - (x^\top Ax)u.$$

计算其二次型:

$$\begin{aligned}\langle u, \operatorname{Hess} f(x)[u]\rangle &= u^\top(Au - (x^\top Au)x - \lambda u) \\ &= u^\top Au - (x^\top Au)(u^\top x) - \lambda u^\top u.\end{aligned}$$

由于 $u$ 是切向量，满足 $u^\top x = 0$，因此中间项消失, 得到

$$\langle u, \operatorname{Hess} f(x)[u]\rangle = u^\top Au - \lambda\|u\|^2.$$

题目给定条件 $\operatorname{Hess} f(x) \succeq 0$，这意味着对于所有满足 $u \perp x$ 的 $u \neq 0$，都有

$$u^\top Au \geqslant \lambda\|u\|^2 \implies \frac{u^\top Au}{u^\top u} \geqslant \lambda.$$

设 $A$ 的特征值为 $\lambda_1 \leqslant \cdots \leqslant \lambda_d$. 根据 Rayleigh 商的性质，$\lambda$ 必须等于最小特征值 $\lambda_1$. 即 $x$ 是最小特征值对应的特征向量，对应的函数值 $f(x) = \frac{1}{2}\lambda_1$ 是 $f$ 在球面上的全局最小值。 □

# 参考文献

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