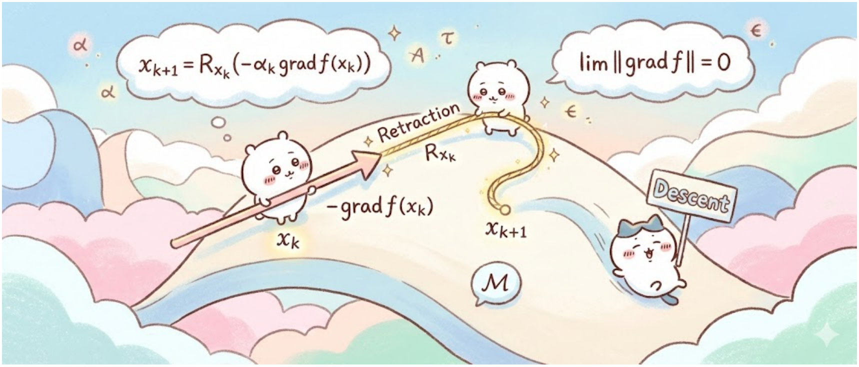
一阶优化算法：黎曼梯度下降

北极甜虾 (南半球版)

2026 年 1 月 25 日

Manifold Optimization Notes



1 算法动机

在传统的欧几里得优化中，我们沿着负梯度方向 $-\nabla f(x)$ 走一小步。但在流形 $\mathcal{M}$ 上，直接应用 $x - \alpha\nabla f(x)$ 会面临以下问题：

1. **可行性问题**：更新后的点通常会脱离流形（例如，在球面上加一个切向量，新点就不在球面上了）。

2. **方向问题**：欧几里得梯度 $\nabla f(x)$ 往往不在切空间 $T_x\mathcal{M}$ 内，无法直接作为流形上的“方向”。

黎曼梯度下降通过引入我们之前学习过的**黎曼梯度**来解决方向问题，引入**收缩映射 (Retraction)** 来解决可行性问题。

2 迭代公式与算法

定义 1: 黎曼梯度下降（Riemannian Gradient Descent, RGD)

给定点 $x_k \in \mathcal{M}$，黎曼梯度下降的迭代定义为:

$$x_{k+1} = R_{x_k}(-\alpha_k \operatorname{grad} f(x_k))$$

其中:

- $\operatorname{grad} f(x_k) \in T_{x_k}\mathcal{M}$ 是黎曼梯度，它指明了流形上局部下降最快的切方向。
- R_{x_k} 是收缩映射，它将切空间中的一步收缩到流形表面，并保持一阶近似性质。
- $\alpha_k > 0$ 是步长 (Step Size)。

线搜索 (line-searching) 是寻找合适步长的方法, 通过近似地去最小化下列函数得到:

$$g(t) = f(R_{x_k}(-t \cdot \operatorname{grad} f(x_k))).$$

这里，$g(t)$ 的值表示在 x_k 处进行一次步长为 t 的 RGD 迭代后的目标函数值。

注解 2: 步长 α_k 的选择

- **固定步长:** $\alpha_k = \alpha > 0, \forall k \in \mathbb{N}.$
- **最优线搜索 (Optimal Line-searching):**

$$\alpha_k = \arg\min_t g(t) = f(R_{x_k}(-t \cdot \operatorname{grad} f(x_k)))$$

- **回溯线搜索 (Backtracking Line-searching):** 这种方法通过寻找满足 **Armijo–Goldstein 条件（简称 Armijo 条件)** 的步长来平衡计算效率与收敛速度。给定参数 $\bar{\alpha} > 0$（初始步长），$\tau \in (0,1)$（缩减因子）和 $r \in (0,1)$（下降常数）。我们寻找最小的非负整数 m，使得 $\alpha_k = \bar{\alpha}\tau^m$ 满足 **Armijo 条件**：

$$f(R_{x_k}(-\alpha_k \operatorname{grad} f(x_k))) \leqslant f(x_k) - r\alpha_k \|\operatorname{grad} f(x_k)\|_{x_k}^2$$

其中，右侧的项代表了我们期望的最少下降量。

带回溯线搜索的黎曼梯度下降

1. **输入**：目标函数 f，收缩映射 R，初始点 $x_0 \in \mathcal{M}$，梯度阈值 $\epsilon > 0$.

2. **参数**：初始步长 $\bar{\alpha} > 0$，收缩因子 $\tau \in (0,1)$，下降常数 $r \in (0,1)$。

3. **循环**：对于 $k = 0, 1, 2, \ldots$ 执行：

   - **确定搜索方向：** $\eta_k = -\operatorname{grad} f(x_k)$.
   - **回溯步长**：寻找最小的 $m \in \mathbb{N}$，使得 $\alpha_k = \bar{\alpha}\tau^m$ 满足：

   $$f(R_{x_k}(\alpha_k \eta_k)) \leqslant f(x_k) - r\alpha_k \|\operatorname{grad} f(x_k)\|_{x_k}^2$$

   - **更新：** $x_{k+1} = R_{x_k}(\alpha_k \eta_k)$.
   - **停止准则**：若 $\|\operatorname{grad} f(x_{k+1})\|_{x_{k+1}} < \epsilon$ 则停止。

3 简化版本的收敛性定理

我们这次笔记中先证明一个简化版本的收敛性定理，通过假设算法会输出一个充分下降的序列，随后的笔记中我们证明通过一些正则条件和合适的步长可以满足该条件。

定理 3: 全局收敛性（收敛速率）

如果下列条件满足：

- **有下界：** 存在 $f_{\text{low}} \in \mathbb{R}$ 使得 $f(x) \geqslant f_{\text{low}}$，$\forall x \in \mathcal{M}$.
- **充分下降：** 存在 $c > 0$ 使得序列 $\{x_k\}_{k\in\mathbb{N}}$ 满足：

$$f(x_{k+1}) - f(x_k) \leqslant -c\|\operatorname{grad} f(x_k)\|_{x_k}^2, \quad \forall k \in \mathbb{N}.$$

那么对于任意整数 $K \geqslant 1$，我们都有

$$\min_{0\leqslant k\leqslant K-1} \|\operatorname{grad} f(x_k)\|_{x_k} \leqslant \sqrt{\frac{f(x_0) - f_{\text{low}}}{cK}}.$$

证明. 根据充分下降条件，对于任意 $k \in \mathbb{N}$，都有

$$c\|\operatorname{grad} f(x_k)\|_{x_k}^2 \leqslant f(x_k) - f(x_{k+1})$$

我们将这个不等式从 $k = 0$ 到 $K - 1$ 进行裂项相消：

$$\begin{aligned} c\sum_{k=0}^{K-1} \|\operatorname{grad} f(x_k)\|_{x_k}^2 &\leqslant \sum_{k=0}^{K-1} (f(x_k) - f(x_{k+1})) \\ &= f(x_0) - f(x_K) \end{aligned}$$

由于假设 f 有下界 f_{low}，因此对于任何 K，都有 $f(x_K) \geqslant f_{\text{low}}$。

由此得

$$c\sum_{k=0}^{K-1}\|\operatorname{grad} f(x_k)\|_{x_k}^2 \leqslant f(x_0)-f_{\text{low}}$$

注意到和式中的每一项都大于等于最小项，即 $\sum_{k=0}^{K-1}\|\dots\|^2 \geqslant K\cdot\min_{0\leqslant k\leqslant K-1}\|\dots\|^2$，代入上式：

$$cK\min_{0\leqslant k\leqslant K-1}\|\operatorname{grad} f(x_k)\|_{x_k}^2 \leqslant f(x_0)-f_{\text{low}}$$

两边除以 cK 并开平方根，即得

$$\min_{0\leqslant k\leqslant K-1}\|\operatorname{grad} f(x_k)\|_{x_k} \leqslant \sqrt{\frac{f(x_0)-f_{\text{low}}}{cK}}$$

□

推论 4: 全局收敛性（极限版本）

在定理 3 的同样条件下，

$$\lim_{k\to\infty}\|\operatorname{grad} f(x_k)\|_{x_k} = 0.$$

并且序列 $\{x_k\}_{k\in\mathbb{N}}$ 中所有的聚点都是 f 的临界点。

证明. 由于序列 $\{f(x_k)\}_{k\in\mathbb{N}}$ 是单调递减且有下界的，它必定收敛到一个有限值。因此级数 $\sum_{k=0}^{\infty}(f(x_k)-f(x_{k+1}))$ 是收敛的。根据前面的求和不等式可知，正项级数 $\sum_{k=0}^{\infty}\|\operatorname{grad} f(x_k)\|_{x_k}^2$ 也必定收敛。由级数收敛的必要条件可知，一般项必须趋于零，所以得到

$$\lim_{k\to\infty}\|\operatorname{grad} f(x_k)\|_{x_k} = 0$$

由于梯度函数和范数是连续的，这也意味着序列的所有聚点 x_* 都满足 $\|\operatorname{grad} f(x_*)\|_{x_*}=0$，即它们都是临界点。 □

参考文献

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