流形优化的一些例子（一）

北极甜虾（南半球版）

2026.1.6

Manifold Optimization Notes

传感器网络定位	矩阵极端特征值
字典学习	主成分分析

1 基于方向的传感器网络定位：仿射子流形

问题 1: 基于方向的传感器网络定位

假设空间中有 n 个传感器，它们的位置 $t_1, \ldots, t_n \in \mathbb{R}^d$ 是未知的，假设有一个图其中的每一条表 $(i, j) \in G$ 代表传感器 i 和 j 之间有观测，并且我们能得到一个带有噪声的方向观测值

$$v_{ij} \approx \frac{t_i - t_j}{\|t_i - t_j\|_2}.$$

目标是根据传感器之间的方向观测值来估计它们的位置。

约束条件的推导　注意到，由于我们只知道传感器两两之间的方位，所以这个问题的解不唯一，必须通过添加约束来固定解：

- **平移模糊性 (Translation Ambiguity)：**平移整个传感器网络的位置不会改变它们之间的相对方向，因此我们可以假设所有传感器的中心在原点
$$\sum_{i=1}^{n} t_i = 0. \tag{1}$$
- **尺度模糊性 (Scale Ambiguity)：**将所有位置乘以一个缩放因子 $\alpha > 0$ 得到 αt_i，其相对方向
$$\frac{\alpha t_i - \alpha t_j}{\|\alpha t_i - \alpha t_j\|_2} = \frac{t_i - t_j}{\|t_i - t_j\|_2}$$
保持不变。因此我们可以假设
$$\sum_{(i,j)\in G} \langle t_i - t_j, v_{ij}\rangle = 1. \tag{2}$$
如果一组位置 $t_1, \ldots, t_n$ 满足该条件，那么对于缩放后的位置 αt_i，该项和将变为 α. 因此，只有当 $\alpha = 1$ 时才满足，从而唯一确定了尺度。

目标函数的推导　为了衡量估计值 $\hat{t}_i$ 与测量值 v_{ij} 的兼容性，我们利用正交投影的概念。设向量 $w_{ij} = \hat{t}_i - \hat{t}_j$ 是我们要估计的方向向量。如果它与测量方向 v_{ij} 完全对齐，那么 w_{ij} 在 v_{ij} 方向上的投影应该等于 w_{ij} 本身。注意 w_{ij} 在单位向量 v_{ij} 上的投影为 $\langle w_{ij}, v_{ij}\rangle v_{ij}$. 垂直误差向量（即 w_{ij} 偏离 v_{ij} 方向的部分）为

$$w_{ij} - \langle w, v_{ij}\rangle v_{ij} = (\hat{t}_i - \hat{t}_j) - \langle \hat{t}_i - \hat{t}_j, v_{ij}\rangle v_{ij}.$$

当方向完全匹配时，这个向量的模长为 0; 匹配越差，模长越大。因此，总的目标函数设定为所有观测边上的误差平方和：

$$\min_{\hat{t}_1,\dots,\hat{t}_n\in\mathbb{R}^d} \sum_{(i,j)\in G} \|(\hat{t}_i - \hat{t}_j) - \langle \hat{t}_i - \hat{t}_j, v_{ij}\rangle v_{ij}\|^2$$

且 $\hat{t}_1,\dots,\hat{t}_n \in \mathbb{R}^d$ 同时满足上述两个约束条件 (1), (2).

流形优化的视角　我们可以将所有变量 $\hat{t}_i$ 组合成一个矩阵 $\hat{T} = [\hat{t}_1, \hat{t}_2, \dots, \hat{t}_n] \in \mathbb{R}^{d\times n}$.

- **搜索空间：** 由于约束条件 $\sum \hat{t}_i = 0$ 和 $\sum\langle \hat{t}_i - \hat{t}_j, v_{ij}\rangle = 1$ 都是关于变量的线性等式约束，因此该问题的搜索空间是 $\mathbb{R}^{d\times n}$ 的一个仿射子空间 (Affine Subspace).
- **流形性质：** 仿射子空间是一类特殊的流形，称为线性流形 (Linear Manifold)，它也是 $\mathbb{R}^{d\times n}$ 的一个嵌入子流形 (Embedded Submanifold).

2 矩阵的极端特征值

问题 2: 对称矩阵的特征值

假设我们有一个实对称矩阵 $A \in \mathbb{R}^{n\times n}$. 根据谱定理，$A$ 拥有 n 个实特征值 $\lambda_1 \leqslant \cdots \leqslant \lambda_n$，以及对应的标准正交特征向量 $v_1, \dots, v_n \in \mathbb{R}^n$. 我们的目标是找到其中的一个极端特征对（比如最小的特征值 λ_1 及其对应的特征向量 v_1）。

Rayleigh 商　为了解决这个问题，通常会定义 Rayleigh 商

$$r(x) = \frac{\langle x, Ax\rangle}{\langle x, x\rangle}, \quad x \in \mathbb{R}^n \setminus \{0\}.$$

Rayleigh 商的一个关键特性是：它的最大值和最小值分别对应于矩阵 A 的最大特征值 λ_n 和最小特征值 λ_1.

尺度不变性与球面约束　注意到 Rayleigh 商具有**尺度不变性**：对于任何非零实数 $\alpha \neq 0$，都有 $r(\alpha x) = r(x)$。这意味着我们可以在不改变函数值的情况下任意固定 x 的长度。

为了计算方便，我们将 x 限制在单位向量集合中，即满足 $\|x\|_2^2 = \langle x, x\rangle = 1$。这个集合构成了 $\mathbb{R}^n$ 中的**单位球面**：

$$S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}.$$

此时，寻找最小特征值的问题可以简化为在球面上的受限优化问题：

$$\min_{x \in S^{n-1}} \langle x, Ax\rangle.$$

流形优化的视角　从流形几何的角度来看，该问题具有以下性质：

- **搜索空间：** 该问题的搜索空间不再是整个欧几里得空间 $\mathbb{R}^n$，而是单位球面 S^{n-1}。
- **流形性质：** 单位球面 S^{n-1} 是 $\mathbb{R}^n$ 的一个**嵌入子流形**。这是流形优化中最简单且非平凡的例子。

扩展：奇异值与乘积流形　类似的方法可以扩展到计算矩阵 $M \in \mathbb{R}^{m\times n}$ 的最大奇异值及其对应的左、右奇异向量：

$$\max_{x \in S^{m-1}, y \in S^{n-1}} \langle x, My\rangle. \tag{3}$$

- **搜索空间：** 此问题的搜索空间是两个球面的笛卡尔积 $S^{m-1} \times S^{n-1}$。
- **重要结论：** **流形的乘积仍然是流形**。这一性质对于构建涉及多变量的复杂流形结构非常有用。

3 字典学习：球面乘积

问题 3: 字典学习 (Dictionary Learning)

给定一组数据（如图像块）$Y \in \mathbb{R}^{d \times k}$，字典学习的目标是学习一个字典 $D = [b_1, \ldots, b_n] \in \mathbb{R}^{d \times n}$，其中每个列向量 b_i 称为原子 (Atoms)。我们希望每个数据点 y_i 都能表示为这些原子的稀疏线性组合，即 $Y \approx DC$。为了使系数矩阵 C 的量级有意义，通常约束每个原子具有单位范数。

目标函数的构建 字典学习通常通过最小化重构误差和表示稀疏度来建模：

$$\begin{aligned} \min_{D,C} \quad & \|Y - DC\|_F^2 + \lambda \|C\|_0 \\ \text{s.t.} \quad & \|b_i\|_2 = 1, \quad i = 1, \ldots, n. \end{aligned}$$

其中，$\|C\|_0$ 表示 C 中非零元素的数量。在实际计算中，为了处理不连续性，通常将其松弛为 $\|C\|_1$ 或进一步平滑化。

斜流形 (Oblique Manifold) 从流形优化的视角看，字典 D 的约束条件定义了一个特殊的几何结构：

- **构造方式：** 既然 D 的每一列都在单位球面上，那么 D 的搜索空间就是 n 个单位球面 S^{d-1} 的笛卡尔积。
- **数学定义：** 这一乘积空间被称为 **斜流形**，记为 $\mathrm{OB}(d, n)$：

  $$\mathrm{OB}(d, n) = (S^{d-1})^n = \{X \in \mathbb{R}^{d \times n} : \mathrm{diag}(X^\top X) = \mathbf{1}\}.$$

  其中 $\mathbf{1}$ 是全 1 向量，$\mathrm{diag}(\cdot)$ 提取矩阵的对角线。
- **流形性质：** $\mathrm{OB}(d, n)$ 是 $\mathbb{R}^{d \times n}$ 的一个嵌入子流形。由于流形本身的**非凸性**，即使代价函数是凸的，在斜流形上寻找全局最小值依然是一个非凸优化问题。

重要结论 该例子再次印证了流形优化的一个重要性质：**流形的乘积依然是流形**。这允许我们将复杂的约束分解为多个简单流形的组合。

4 主成分分析：Stiefel 与 Grassmann 流形

问题 4: 主成分分析 (PCA)

给定中心化数据矩阵 $X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^{d\times n}$，PCA 的目标是找到 k 个标准正交的向量 $u_1, \ldots, u_k \in \mathbb{R}^d$，使得数据投影到这些向量张成的子空间时，拥有最大的"方差"：

$$\max_{u_1,\ldots,u_k\in\mathbb{R}^d} \sum_{i=1}^{k} \alpha_i \langle XX^\top u_i, u_i\rangle$$

其中 $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_k > 0$ 为主成分的方差权重。

Stiefel 流形 (Stiefel Manifold) 当我们优化 k 个权重各不相同的正交主成分时，搜索空间由有标准正交列的矩阵组成。

- **定义：** $\mathrm{St}(d,k) = \{U \in \mathbb{R}^{d\times k} : U^\top U = I_k\}$。
- **性质：** 它是 $\mathbb{R}^{d\times k}$ 的一个**嵌入子流形 (Embedded Submanifold)**。
- **PCA 的表述：** 当我们需要寻找前 k 个主成分，且各主成分权重不同时，则对应在 $\mathrm{St}(d,k)$ 上优化：

$$\max_{U\in\mathrm{St}(d,k)} \langle XX^\top U, UD\rangle, \quad D = \mathrm{diag}(\alpha_1, \ldots, \alpha_k).$$

Grassmann 流形 (Grassmann Manifold) 如果主成分的顺序不重要 ($\alpha_1 = \cdots = \alpha_k = 1$)，只关心它们张成的 k 维子空间，则引入等价关系 $\sim$：

- **等价关系：** $U \sim V \iff V = UQ$，其中 $Q \in O(k)$（正交群）。

- **定义：** Grassmann 流形 $Gr(d, k)$ 是所有等价类 $[U]$ 的集合，即商空间 $St(d, k)/O(k)$。

- **性质：**它是一个**商流形 (Quotient Manifold)**。它将具有对称性（旋转不变性）的搜索空间简化为一个更紧凑的几何结构。

- **PCA 的子空间表述：**寻找最优 k 维特征空间可写为：

$$\max_{[U]\in Gr(d,k)} \langle XX^\top U, U \rangle.$$

意义 从流形视角看，Stiefel 流形处理“带顺序的正交基”，而 Grassmann 流形处理“无顺序的子空间”。这种抽象使得我们可以利用问题的固有几何对称性来提高算法的理论完备性和数值稳定性。当权重 α_i 部分相等、部分不等时（例如 $\alpha_1 = \alpha_2 > \alpha_3 = \alpha_4$），我们进入了一个介于 Stiefel 流形和 Grassmann 流形之间的几何结构，这在数学上被称为 Flag 流形。这种情况下的对称性（旋转不变性）会被部分打破。

参考文献

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