

Rockafellar, R. Tyrrell. Convex analysis. Vol. 28.
Princeton university press, 1970. 盛宝怀译

屈彬

2019 年 7 月 4 日

1 基本概念

1.1 仿射集 (2019 年 7 月 2 日)

定义 1.1 (向量空间) R 表示实数系, 其中 R^n 表示由实 n 元组 $x = (\xi_1, \xi_2, \dots, \xi_n)$ 所组成的向量空间。

定义 1.2 (向量的内积) 记两向量 x 与 x^* 在实数系 R^n 中的内积为

$$\langle x, x^* \rangle = \xi_1 \xi_1^* + \xi_2 \xi_2^* + \dots + \xi_n \xi_n^*.$$

定义 1.3 (线性变换) 设向量空间 R^n 到向量空间 R^m 的线性变换为 $x \rightarrow Ax$, 其中 A 为一 $m \times n$ 阶实矩阵, 称 A 为线性变换 $R^n \rightarrow R^m$ 的变换矩阵。记 A^* 为 A 的转置矩阵, y^* 为一实 m 元组, 那么它们存在以下内积关系

$$\langle Ax, y^* \rangle = \langle x, A^* y^* \rangle.$$

定义 1.4 (直线) 设 x 和 y 为空间 R^n 中两个不同的点, 记通过 x 和 y 的直线为 $l(x, y)$, 那么 $l(x, y)$ 上所有的点都可以表示为 x 坐标和 y 坐标系数和为 1 的线性叠加, 即

$$l(x, y) \equiv (1 - \lambda)x + \lambda y, \lambda \in R.$$

定义 1.5 (仿射集) 令 M 为 R^n 中的一个子集, 若 $\forall x, y \in M \& \forall \lambda \in R$ 都有

$$(1 - \lambda)x + \lambda y \in M$$

即经过子集 M 中任意两点的直线完全处于 M 内, 则称 M 为 \mathbf{R}^n 中的仿射集 (affine set), 又称作仿射流形 (affine manifold)、仿射变量 (affine variety)、线形变量 (linear variety) 或称 M 是平坦的 (flat)。空集 \emptyset 和全集 \mathbf{R}^n 一定是空间 \mathbf{R}^n 中的仿射集。

定义 1.6 (线性子空间) 若 M 是 \mathbf{R}^n 的一个线性子空间 (简称子空间), 那么以下条件必须同时成立

1. $\mathbf{0} \in M$ 。
2. $\forall \mathbf{x} \in M, \lambda \in \mathbf{R}$ 满足 $\lambda \mathbf{x} \in M$ (乘法封闭性)。
3. $\forall \mathbf{x}, \mathbf{y} \in M$ 满足 $\mathbf{x} + \mathbf{y} \in M$ (加法封闭性)。

定理 1.1 \mathbf{R}^n 中包含原点的仿射集是 \mathbf{R}^n 的子空间。

证明 1.1 若 M 为 \mathbf{R}^n 的一个包含原点的仿射集, 则 $\mathbf{0} \in M$ 。

下面证明 M 满足乘法封闭性。因为 M 是一个包含原点的仿射集, 那么对于原点 $\mathbf{0}$ 和 M 中任一点 \mathbf{x} , 有

$$(1 - \lambda)\mathbf{0} + \lambda\mathbf{x} \in M \Leftrightarrow \lambda\mathbf{x} \in M, \lambda \in \mathbf{R}$$

即 M 满足乘法封闭性。

下面证明 M 满足加法封闭性。设 \mathbf{x} 和 \mathbf{y} 是 M 中的任意两点, 因为 M 是一个仿射集, 所以满足

$$\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y} \in M \Leftrightarrow \frac{1}{2}(\mathbf{x} + \mathbf{y}) \in M$$

又因为 M 满足乘法封闭性, 所以

$$\mathbf{x} + \mathbf{y} \in M$$

即 M 满足加法封闭性。

综上, M 是 \mathbf{R}^n 的一个子空间。

定义 1.7 (仿射集的平行) 仿射集 M 与 L 平行是指存在 α 使得 $M = L + \alpha$ 。易证平行具有传递性，若仿射集 L_1 和 L_2 皆平行于 M ，那么 L_1 平行于 L_2 。

定理 1.2 \mathbf{R}^n 上的每个非空仿射集 M 一定平行于唯一的子空间 L ，其中 L 由下式给出。

$$L = M - M = \{x - y | x, y \in M\}$$

证明 1.2 首先证明 M 不能同时与两个不同的子空间平行。设 L_1 和 L_2 为 \mathbf{R}^n 的两个子空间，且 L_1 和 L_2 皆平行于 M ，则 L_1 与 L_2 互相平行，那么存在 α 使得

$$L_2 = L_1 + \alpha$$

因为 L_1 是一个子空间，所以 $0 \in L_1$ ，则 $\alpha \in L_2$ 。设 x 是 L_1 中任意一点，则 $x + \alpha \in L_2$ 。根据加法封闭性，得 $x \in L_2$ ，那么 $L_1 \subseteq L_2$ 。同理可证 $L_2 \subseteq L_1$ 。因为 $L_1 \subseteq L_2$ 且 $L_2 \subseteq L_1$ ，所以 $L_1 = L_2$ ，子空间唯一性得证。

下面证明子空间 $L = M - M$ 。设 x 是 M 中任意一点，那么仿射集 $M - x$ 平行于 M 且 $M - x$ 包含 0 。那么根据定理 1.1， $M - x$ 是 \mathbf{R}^n 中的一个子空间，记作 L ，则 $L = M - x$ 。因为 x 是 M 中的任意一点，无论选择什么样的 x ，都会有 $L = M - x$ ，所以 $L = M - M$ 成立。

综上，定理 1.2 得证。

定义 1.8 (仿射集的维数) 非空仿射集的维数等于与其平行的子空间的维数。特别地，定义空集 \emptyset 的维数为 -1 。维数为 0 、 1 、 2 的仿射集分别被称为点、线和平面。空间 \mathbf{R}^n 中的 $n-1$ 维仿射集被称为超平面（比空间少一维）。

定义 1.9 (正交) 若向量 x 与向量 y 的内积为 0 ，即 $\langle x, y \rangle = 0$ ，则称 x 与 y 正交，记作 $x \perp y$ 。

定义 1.10 (正交补 (orthogonal complement)) 给定 \mathbf{R}^n 的子空间 L ，若对于一切 $y \in L$ ，都有 $x \perp y$ ，这些由 x 构成的集合称为 L 的正交补，记作 L^\perp 。易证，空间 \mathbf{R}^n 中子空间 L 的维数与其正交补 L^\perp 的维数之和等于空间的维数，即

$$\dim(L) + \dim(L^\perp) = n.$$

另外，子空间的正交补的正交补是其本身，即

$$L = (L^\perp)^\perp$$

借助正交补，可通过一维子空间刻画超平面。设 \mathbf{b} 是一个非零向量，那么集合 $\{\mathbf{x} | \mathbf{x} \perp \mathbf{b}\}$ 表示一个 $n-1$ 维的子空间（也表示一个经过原点的超平面），该子空间经过平移变换可表示任何与之平行的超平面，形式如下：

$$\{\mathbf{x} | \mathbf{x} \perp \mathbf{b}\} + \mathbf{a} = \{\mathbf{x} + \mathbf{a} | \langle \mathbf{x}, \mathbf{b} \rangle = 0\}.$$

定理 1.3 (超平面的表示) 给定 $\beta \in \mathbf{R}$ 以及非零 $\mathbf{b} \in \mathbf{R}^n$ ，集合

$$H = \{\mathbf{x} | \langle \mathbf{x}, \mathbf{b} \rangle = \beta\}$$

可表示 \mathbf{R}^n 中的一个超平面。其中，向量 \mathbf{b} 正交于超平面 H ，其余每个与 H 正交的向量为 \mathbf{b} 的整倍数。

证明 1.3 略。

定理 1.4 (仿射集表示) 空间 \mathbf{R}^n 中的仿射集可通过一个 n 元线性方程组表示。给定 $\mathbf{b} \in \mathbf{R}^n$ 及 $m \times n$ 阶 ($m \leq n$) 实矩阵 B ，集合

$$M = \{\mathbf{x} \in \mathbf{R}^n | B\mathbf{x} = \mathbf{b}\}$$

为 \mathbf{R}^n 中的一个仿射集。

证明 1.4 略。书中的证明不够严谨，应该仔细考虑 $m > n$ ，即方程组为超定方程组的情形。

定理 1.5 (子空间的表示) 设 L 为 \mathbf{R}^n 中的一个子空间， $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ 为 L^\perp 的一组基，那么

$$L = (L^\perp)^\perp = \{\mathbf{x} | \mathbf{x} \perp \mathbf{b}_1, \dots, \mathbf{x} \perp \mathbf{b}_m\} = \{\mathbf{x} | \langle \mathbf{x}, \mathbf{b}_i \rangle = 0, i = 1, \dots, m\} = \{\mathbf{x} | B\mathbf{x} = \mathbf{0}\}$$

其中， B 为一个 $m \times n$ 实矩阵，它的行向量为 $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ 。

联系定理1.4, 仿射集 M 可通过子空间 L 表示, 存在 $\mathbf{a} \in \mathbf{R}^n$ 使得

$$M = L + \mathbf{a} = \{\mathbf{x} | \mathbf{B}\mathbf{x} = \mathbf{b}\}$$

其中, $\mathbf{b} = \mathbf{B}\mathbf{a}$ 。

推论 1.1 \mathbf{R}^n 中的每个仿射集都是有限个超平面的交集。

定理1.4中仿射集的表示也可以写成

$$M = \{\mathbf{x} | \mathbf{B}\mathbf{x} = \mathbf{b}\} = \{\mathbf{x} | \langle \mathbf{x}, \mathbf{b}_1 \rangle = \beta_1, \langle \mathbf{x}, \mathbf{b}_2 \rangle = \beta_2, \dots, \langle \mathbf{x}, \mathbf{b}_m \rangle = \beta_m\}$$

其中, \mathbf{b}_i 为 \mathbf{B} 的第 i 个行 (向量), β_i 为 \mathbf{b} 的第 i 个分量, 而根据定理1.3

$$H_i = \{\mathbf{x} | \langle \mathbf{x}, \mathbf{b}_i \rangle = \beta_i\}$$

表示第 i 个超平面, 则

$$M = \bigcap_{i=1}^m H_i$$

因为任意一个 m 维的仿射集也是 $m+1$ 维空间的超平面, 所以仿射集与仿射集的交集依然是仿射集。

定义 1.11 (仿射包) 给定 $S \subset \mathbf{R}^n$ (S 是一个集合), 存在唯一一个包含 S 的最小的仿射集 (例如, 满足 $S \subset M$ 的所有仿射集的交集)。这个仿射集被称为 S 的仿射包 (afine hull), 记作 $\text{aff}S$ 。 $\text{aff}S$ 可以由所有形式为 $\lambda_1 \mathbf{x}_1 + \dots + \lambda_m \mathbf{x}_m$ 向量组成, 其中 $\mathbf{x}_i \in S$ 且 $\lambda_1 + \dots + \lambda_m = 1$ 。

定义 1.12 (仿射无关) $m+1$ 个向量 $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_m$ 的集合若使得 $\text{aff}\{\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_m\}$ 的维数为 m , 则称它们是仿射无关 (affinely independent) 的。反之, 若 $\text{aff}\{\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_m\}$ 的维数小于 m , 则称它们是仿射相关的。

其中, 以 \mathbf{b}_0 的模作为 $\text{aff}\{\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_m\}$ 到原点的距离, 那么 \mathbf{b}_0 满足

$$\text{aff}\{\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_m\} = L + \mathbf{b}_0$$

其中

$$L = \text{aff}\{\mathbf{0}, \mathbf{b}_1 - \mathbf{b}_0, \dots, \mathbf{b}_m - \mathbf{b}_0\}$$

是唯一平行于 $\text{aff}\{\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_m\}$ 的子空间。 L 的维数为 m 当且仅当向量 $\mathbf{b}_1 - \mathbf{b}_0, \dots, \mathbf{b}_m - \mathbf{b}_0$ 是线性无关的。

结合定理1.4和定理1.5, 若仿射集 M 中的向量 $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_m$ 是仿射无关的, 则 M 中的任一向量 \mathbf{x} 具有以下表达形式

$$\mathbf{x} = \lambda_1(\mathbf{b}_1 - \mathbf{b}_0) + \dots + \lambda_m(\mathbf{b}_m - \mathbf{b}_0) + \mathbf{b}_0 = \lambda_1 \mathbf{b}_1 + \dots + \lambda_m \mathbf{b}_m + \mathbf{b}_0 - \sum_{i=1}^m \lambda_i \mathbf{b}_0$$

即

$$\mathbf{x} = \sum_{i=0}^m \lambda_i \mathbf{b}_i, \lambda_0 = -\sum_{i=1}^m \lambda_i \Leftrightarrow \sum_{i=0}^m \lambda_i = 0$$

其中, 称 $\lambda_0, \lambda_1, \dots, \lambda_m$ 为 M 中的重心坐标系统 (*barycentric coordinate system*)。

定义 1.13 (仿射变换) 给定一个从 \mathbf{R}^n 到 \mathbf{R}^m 的变换 $T: \mathbf{x} \rightarrow T(\mathbf{x})$, 若变换前后的满足系数保持, 即对于每个 \mathbf{R}^n 中的任意两点 \mathbf{x} 和 \mathbf{y} 及 $\lambda \in \mathbf{R}$, 都有

$$T((1-\lambda)\mathbf{x} + \lambda\mathbf{y}) = (1-\lambda)T\mathbf{x} + \lambda T\mathbf{y}$$

则称 T 为仿射变换。

仿射变换的形式形如 $T(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{a}$, 其中 \mathbf{A} 为线性变换, $\mathbf{a} \in \mathbf{R}^m$ 且 $\mathbf{a} = T(\mathbf{0})$ 。