高等学校试用教材

刘光中　编

凸分析与极值问题

高等教育出版社

高等学校试用教材

凸分析与极值问题

刘光中　编

高等教育出版社

内 容 提 要

本书是一本关于凸集和凸函数理论的入门书，包括凸集、凸函数、凸函数的微分、极值问题的最优性条件和广义凸函数等内容。

本书是经高等工科院校应用数学教材委员会评选、推荐的教材，适用于数学、应用数学等专业的本科生及研究生。

高等学校试用教材

凸分析与极值问题

刘光中 编

*

高等教育出版社出版

新华书店北京科技发行所发行

北京印刷一厂印装

*

开本850×1168 1/32 印张10.25 字数250 000

1991年5月第1版 1991年5月第1次印刷

印数 0001— 1 520

ISBN 7-04-002662-7/O·1012

定价4.45元

前　言

凸分析作为数学的一个比较年青的分支，是在五十年代以后随着数学规划、最优控制理论、数理经济学等应用数学学科的兴起而发展起来的．在这方面的先驱和作出重要贡献的有W.Fenchel, J.Moreau, R.T.Rockafellar, A.Brosted 等人．现在，凸分析的理论已在很多领域发挥着越来越大的影响．例如，在最优化理论中，凸性曾经作为方法和假设．然而，最近十多年来在诸如数学规划理论这样一些分支中，凸性代表了处理具有广泛应用的最优化问题的自然的结构．利用凸性的条件就可以避免关于连续性和可微性这样一些很强的限制．

本书的目的，是为应用数学系高年级学生、研究生和应用数学工作者提供一本在极值问题中起关键作用的凸集和凸函数理论的入门书．从 1981 年起，王荫清教授和我在成都科技大学应用数学系为运筹专业的本科生、研究生讲授“凸分析”课程．经过多次教学实践，将所编讲义加以修改、整理和补充而编成目前这本书．它的主要内容包括凸集的结构，凸函数的运算、连续性，可微性，凸函数的 Fenchel 形式的对偶理论，不等式理论，约束极值问题及 Lagrange 乘子理论等．考虑到最优化理论的某些发展趋势，最后用一章讨论了广义凸函数研究中的一些进展．

阅读本书要求具备线性代数、解析几何、实分析、泛函分析初步及点集拓扑方面的知识．当然，从纯数学的观点看，本书中的一部分材料是比较初等的．这主要是考虑到为了使需要运筹学知识的经济工作者、工程技术人员及其他非数学工作者也能学懂大部分内容．

考虑到凸分析的应用和叙述的简洁性，本书的讨论只限制在 n 维欧氏空间，虽然很多结果在无穷维抽象空间中仍然是正确的，

本书不包括一般极值问题和最优控制问题的讨论，也不涉及属于凸分析范畴的不动点理论，及凸多值映射理论.

在形成本书的过程中得到王荫清教授的帮助和成都科技大学应用数学系运筹与控制教研室各位老师的支持，得到上海交通大学应用数学系胡毓达教授的热情指导，在此一并表示感谢．本书的部分内容参考了 Rockafellar R. T. 著«*Convex analysis*»及 Пшеничный Б. Н. 著«*Выпуклый анализ и экстремальные задачи*»两书．由于水平有限，书中谬误之处敬请批评指正.

刘 光 中

于成都科技大学应用数学系

1988年12月

目　录

第一章 凸 集

§1 基本概念与记号

本书中所指的数都是实数，R 表示实数系，R^n 表示 n 维实向量空间，R^n 中的元素用 $x=(\xi_1,\cdots,\xi_n)$ 来表示，其中 ξ_i 表示 x 的第 i 个分量. 为了方便有时也将 $x\in R^n$ 看成空间中的一点. $\theta=(0,\cdots,0)$ 代表零向量，它对应坐标原点.

$x\geqslant\theta$ 表示 $\xi_i\geqslant 0, i=1,\cdots,n$.

$x\neq\theta$ 表示至少存在一个 i，$\xi_i\neq 0$.

$x=\theta$ 表示 $\xi_i=0, i=1,\cdots,n$.

设 $x=(\xi_1,\cdots,\xi_n)$，$y=(\eta_1,\cdots,\eta_n)$，则

$x+y=(\xi_1+\eta_1,\cdots,\xi_n+\eta_n)$.

$cx=(c\xi_1,\cdots,c\xi_n)$，其中 $c\in R$.

$x=y$ 表示，$\xi_i=\eta_i, i=1,\cdots,n$.

用 $\langle x,y\rangle$ 表示 x、y 的内积，定义为：

$$\langle x,y\rangle=\xi_1\eta_1+\cdots+\xi_n\eta_n. \tag{1}$$

具有这种内积的向量空间称为 n 维欧几里德(Euclid)空间，仍用 R^n 表示，简称 n 维欧氏空间.

R^n 中的向量 x 的欧氏模（或长度）定义为：

$$|x|=\langle x,x\rangle^{\frac{1}{2}}, \tag{2}$$

即 $|x|=\left(\sum\limits_{i=1}^{n}\xi_i^2\right)^{\frac{1}{2}}$ 显然，$|x|\geqslant 0$；当且仅当 $x=\theta$ 时，$|x|=0$.

设 $x,y\in R^n$，定义 x,y 的距离 $d(x,y)$ 为：

$$d(x,y)=|x-y|. \tag{3}$$

设 $x\in R^n$，$S\subset R^n$，定义 x 和集合 S 的距离 $d(x,S)$ 是：

$$d(x,S)=\inf\{|x-y|\mid y\in S\}.\tag{4}$$

定理 1.1 对于距离 $d(x,y)$，下面的结论成立：

1) $d(x,y)\geqslant 0$；当且仅当 $x=y$ 时，$d(x,y)=0$.

2) $d(x,z)\leqslant d(x,y)+d(y,z)$.

3) $d(x,y)=d(y,x)$.

4) $d(x,y+z)=d(x-y,z)$.

5) 当 $\lambda\geqslant 0$时，$d(\lambda x,\lambda y)=\lambda d(x,y)$.

证明略.

定理 1.2 对于欧氏模，下面的结论成立：

1) $|\langle x,y\rangle|\leqslant|x||y|$.（Cauchy 不等式） (5)

2) $|x+y|\leqslant|x|+|y|$.（三角形不等式） (6)

证明 1）如果 $y=\theta$，(5)式的结论显然成立，所以可设 $y\neq\theta$. 对于每个数 t，有

$$\langle x+ty,x+ty\rangle=\langle x,x\rangle+2t\langle x,y\rangle+t^2\langle y,y\rangle.\tag{7}$$

(7) 式右边是关于 t 的二次式，当

$$t=t_0=-\frac{\langle x,y\rangle}{\langle y,y\rangle}$$

时达到极小值.将 t_0 代入(7)得

$$0\leqslant|x+t_0y|^2=|x|^2-\frac{|\langle x,y\rangle|^2}{|y|^2},$$

即
$$|\langle x,y\rangle|^2\leqslant|x|^2|y|^2,$$

这和(5) 式是等价的.

2) $\langle x+y,x+y\rangle=\langle x,x\rangle+2\langle x,y\rangle+\langle y,y\rangle$,

由(5) 式得

$$|x+y|^2\leqslant|x|^2+2|x||y|+|y|^2=(|x|+|y|)^2,\tag{8}$$

这和(6)式是等价的. ∎

注 1)从证明过程知道(5)式中等号成立等价于$|x+t_0y|=0$，即 $x+t_0y=\theta$. 所以当 $y\neq\theta$ 时，$|\langle x,y\rangle|=|x||y|$的充分必要

条件是 x 是 y 的数量倍数. 如果 $\langle x, y\rangle=|x||y|$ 则 x 是 y 的非负数量倍数（反过来结论也成立).

2） $y\neq\theta$ 时,(6) 式中等号成立的充分必要条件是 x 是 y 的非负数量倍数.

R^n 中线性无关的 n 个元素的集合 $\{b_1,\cdots,b_n\}$ 称为 R^n 的一组基. 如果这个基还满足

$$<b_i,b_j>=\begin{cases}1 & i=j,\\ 0 & i\neq j,\end{cases} \quad i,j=1,\cdots,n$$

则称它是正交基. 单位坐标向量

$$\begin{aligned}e_1&=(1,0,\cdots,0),\\ e_2&=(0,1,\cdots,0),\\ &\vdots\\ e_n&=(0,0,\cdots,1),\end{aligned}$$

构成 R^n 的标准正交基. 显然, 对每一个 $x=(\xi_1,\cdots,\xi_n)$, 有

$$x=\sum_{i=1}^{n}\xi_i e_i. \tag{9}$$

例如, $x=(2,-1,3)=2e_1-e_2+3e_3$.

若用同一个符号 A 表示 $m\times n$ 实矩阵和相应的从 R^n 到 R^m 的线性变换: $x\longrightarrow Ax$, 用 A^* 表示 A 的转置矩阵和相应的从 R^m 到 R^n 的伴随线性变换, 则对于任意的 $x\in R^n$, $y^*\in R^m$, 恒等式

$$\langle Ax,y^*\rangle\doteq\langle x,A^*y^*\rangle \tag{10}$$

成立.（表示向量时, 符号"$*$"没有运算意义. 为遵守矩阵乘法规则, 全部向量均为列向量).

设 $x_1,y_1\in R^n$, $x_1\neq y_1$, 集合

$$\{x\,|\,x=(1-t)x_1+ty_1,t\in R\} \tag{11}$$

称为过 x_1 和 y_1 的直线. 如果设 $z=y_1-x_1$, 则该直线亦可表示为

$$\{x \mid x = x_1 + tz, t \in R\}. \tag{12}$$

连结 x_1 和 y_1 的线段定义为集合

$$\overline{x_1 y_1} = \{x \mid x = (1-t)x_1 + ty_1, t \in [0,1]\}. \tag{13}$$

如果 $t = \frac{1}{2}$，则 x 表示连结 x_1 和 y_1 的线段的中点，$t = \frac{1}{3}$，$t = \frac{2}{3}$ 分别对应该线段的两个三等分点(图 1)。

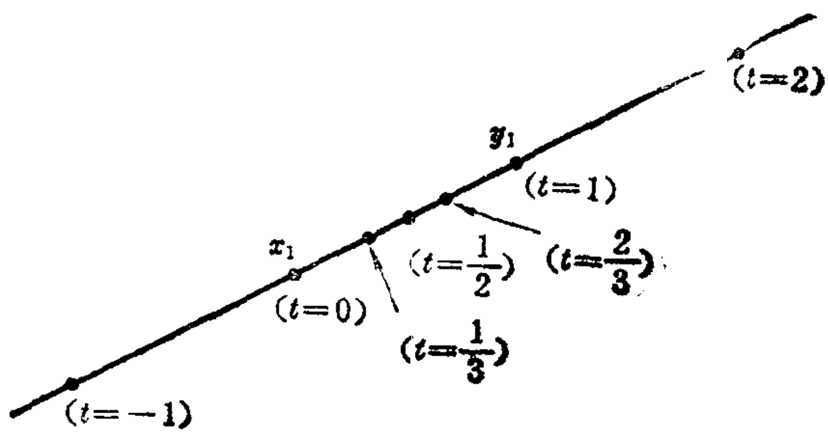

图 1

§ 2 R^n 中的仿射结构

这一节叙述 R^n 中的仿射理论的主要结果. 为了使叙述更加自然，首先列出 R^n 中线性结构的一些基本结论. 从这些结果以及后面将要讨论的 R^n 中的凸结构，我们就可以认识它们之间的必然联系了. 因为我们已经假定读者熟悉线性代数的有关内容，故大多数定理的证明均略去.

1. R^n 中的线性结构.

定义 2.1 设 $L \subset R^n$，如果对于 $\forall x, y \in L, \lambda_1, \lambda_2 \in R$，均有 $\lambda_1 x + \lambda_2 y \in L$，则称 L 是 R^n 中的子空间.

设 $x_1, \cdots, x_p$ 是 R^n 中的向量，$\lambda_1, \cdots \lambda_p \in R$，形如 $x = \lambda_1 x_1 + \cdots + \lambda_p x_p$ 的向量 x 称为 $x_1, \cdots, x_p$ 的线性组合. 所以定义 2.1 表示子空间 L 中的两个向量的任意线性组合属于 L. 事实上，下面的结果成立.

定理 2.1 设$L\subset R^n$,则L是R^n中的子空间的等价条件是L的元素的任意线性组合仍属于L.

证明略.

从线性代数知道,R^n中的子空间族的交是子空间. 所以,对于R^n中的任意集合M,存在包含M的最小子空间,这个子空间就是包含M的全体子空间的交. 我们称这个子空间是由M张成的子空间,或由M生成的子空间,或M的线性包,用span$M$表示。

定理 2.2 设$M\subset R^n$,则span M是M的向量的全体线性组合构成的集合.

证明略。

定义 2.2 设$\{x_1,\cdots,x_p\}$是R^n中的向量组,$\lambda_1,\cdots,\lambda_p\in R$,如果仅当$\lambda_1=\cdots=\lambda_p=0$时,有

$$\lambda_1x_1+\cdots+\lambda_px_p=\theta,$$

则称向量组$\{x_1,\cdots,x_p\}$线性无关,不是线性无关的向量组称为线性相关.

向量组的线性无关性等价于这个向量组的任何一个向量均不是向量组其他向量的线性组合. 如果某一个向量x表示成向量组$\{x_1,\cdots,x_p\}$的线性组合的形式,即$x=\lambda_1x_1+\cdots+\lambda_px_p$,则当且仅当$x_1,\cdots,x_p$线性无关时,系数$\lambda_1,\cdots,\lambda_p$可以单值地确定.

设L是R^n中的子空间,$\{x_1,\cdots,x_p\}$是R^n中的一个线性无关向量组. 如果$L=\text{span}\{x_1,\cdots,x_p\}$,则称$\{x_1,\cdots,x_p\}$是$L$的一组基. L可以有无穷多个基,构成L的基的向量组中的元素并不是唯一的,但其向量的个数是相同的. L的维数$\dim L$等于L的基的元素的个数.

设$M\subset R^n$,$\dim(\text{span}M)=p$,则存在属于M的向量组成的线性无关组$\{x_1,\cdots,x_p\}$,即span$M$存在由M的向量组$\{x_1,\cdots,x_p\}$组成的基. 所以下面的定理成立:

理 2.3 设$M\subset R^n$，则在M中存在线性无关的向量组$\{x_1,\cdots,x_p\}$，使由形如$\sum\limits_{i=1}^{p}\lambda_i x_i$的线性组合的全体构成的集合就是spa$M$.

这个结论表明，为了得到 span$M$，只要取M中固定的线性无关向量 $x_1,\cdots,x_p$ 的全体线性组合就可以了. 此外，span$M$ 中每一个向量都可以唯一地表示成 $x_1,\cdots,x_p$ 的线性组合的形式.

定义 2.3 如果由 R^n 中的子空间L到 R^m 中的映射A保持线性组合，即

$$A\left(\sum_{i=1}^{p}\lambda_i x_i\right)=\sum_{i=1}^{p}\lambda_i A(x_i), \tag{1}$$

其中 $x_i\in L,\lambda_i\in R,i=1,\cdots,p$，则称$A$是线性变换.

称由 R^n 中的子空间 L_1 到 R^m 中的子空间 L_2 上的一对一线性变换为线性同构. 如果存在L_1到L_2上的线性同构，则 L_1 和 L_2 称为同构的子空间. 可以证明，当且仅当两个子空间维数相同时，这两个子空间线性同构.

2. R^n 中的仿射结构

如图 2，在 R^2 中，过原点的直线 l_1 是R^2 中的子空间. 不过原点的直线 l_2 则不是子空间，但它与 l_1 平行，是 l_1 的平移. 由子空间平移而形成的这类集合将是现在研究的对象.

定义 2.4 设$M\subset R^n$. 如果对于$\forall x$，$y\in M$，$\lambda\in R$，均有$(1-\lambda)x+\lambda y\in M$，则称$M$是 R^n中的仿射集.

有时也称仿射集为“仿射流形”、“仿射簇”、“线性流形”等.

从定义 2.4 可以看出，通过仿射集M中任意两点的直线仍然包含在M中.

例 1 空集$\varnothing$和空间 R^n 都是仿射集，R^n 中的点，直线和平面都是仿射集.

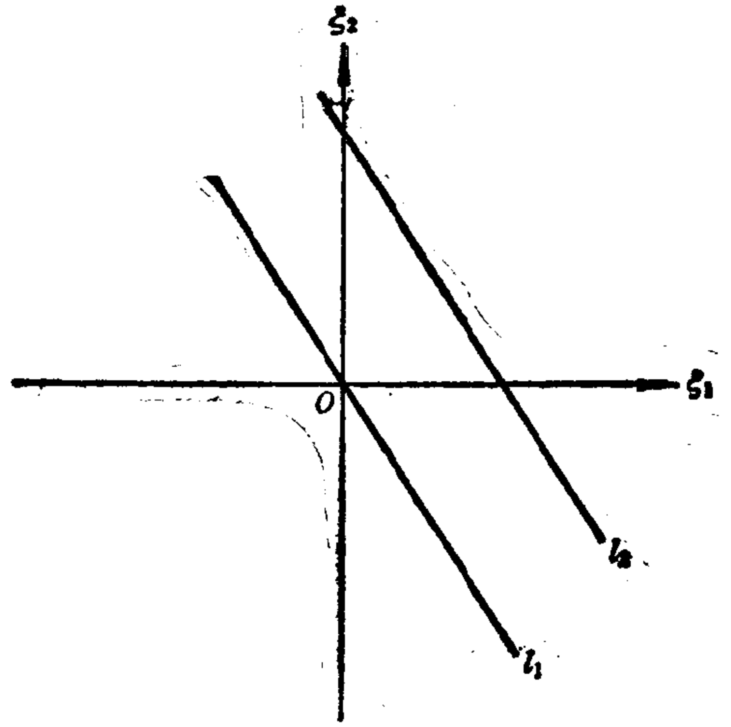

图 2

设 $x_1,\cdots,x_p$ 是 R^n 中的向量，$\lambda_1,\cdots,\lambda_p$ 是 R 中满足 $\sum\limits_{i=1}^{p}\lambda_i=1$ 的元素，称形如 $x=\lambda_1x_1+\cdots+\lambda_px_p$ 的向量 x 是 $x_1,\cdots,x_p$ 的仿射组合．所以定义 2.4 表示，仿射集 M 中两个向量的任意仿射组合仍然属于 M．事实上，下面的结果成立．

定理 2.4 设 $M\subset R^n$，则 M 是 R^n 中的仿射集的等价条件是 M 的元素的任意仿射组合仍然属于 M．

证明略．

容易证明，R^n 中的仿射集族的交仍是 R^n 中的仿射集．所以，对于 R^n 中的任意集合 M，存在着包含 M 的最小仿射集，这个仿射集就是包含 M 的全体仿射集的交。我们称这个仿射集是由 M 张成的仿射集，或由 M 生成的仿射集，或 M 的仿射包，用 $\mathrm{aff}M$ 表示．

例 2 如图 3，M 是 R^3 中的线段，H_1、H_2 是 R^3 中的平面，

M 在 H_1、H_2 的交线上，aff$M$ 就是 H_1、H_2 的交线。

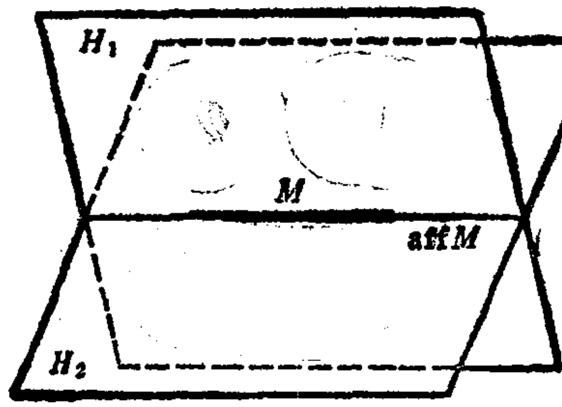

图 3

定理 2.5 设 $M\subset R^n$，则 aff M 是M的向量的全体仿射组合构成的集合.

证明留给读者.

定义 2.5 设$\{x_1,\cdots,x_p\}$是 R^n 中的向量组，$\lambda_1,\cdots,\lambda_p$ 是 R 中满足 $\lambda_1+\cdots+\lambda_p=0$ 的元素，如果仅当 $\lambda_1=\cdots=\lambda_p=0$ 时，有

$$\lambda_1x_1+\cdots+\lambda_px_p=\theta,$$

则称向量组$\{x_1,\cdots x_p\}$仿射无关；不是仿射无关的向量组称为仿射相关的.

向量组的仿射无关性等价于这个向量组的任一向量都不是其他向量的仿射组合. 如果某一个向量 x 表示成向量组$\{x_1,\cdots,x_p\}$的仿射组合的形式，即 $x=\lambda_1x_1+\cdots+\lambda_px_p$，$\lambda_1+\cdots+\lambda_p=1$，$\lambda_i\in R$，$i=1,\cdots,p$，则当且仅当 $x_1,\cdots,x_p$ 仿射无关时，系数 $\lambda_1,\cdots,\lambda_p$ 可以单值地确定.

定理 2.6 设向量组$\{x_1,\cdots,x_p\}\subset R^n$，则$\{x_1,\cdots,x_p\}$仿射无关的等价条件是每一组由 $p-1$ 个向量构成的向量组

$$\{x_1-x_i,\cdots,x_{i-1}-x_i,x_{i+1}-x_i,\cdots,x_p-x_i\},\quad i=1,\cdots,p$$

线性无关.

证明留给读者.

从定理 2.6 可知，与线性无关有联系的所有事实都可以应用于仿射无关. 例如，R^n 中的任意 $p+1$ 个仿射无关的点的集合可以扩张成 $n+1$ 个仿射无关的点的集合；等等.

定理 2.7 R^n 中多于 $n+1$ 个不同向量所组成的向量组必仿射相关.

证明留给读者.

设 M 是 R^n 中的仿射集，$\{x_1,\cdots,x_p\}$ 是 M 中仿射无关的向量组. 如果 $M=\mathrm{aff}\{x_1,\cdots,x_p\}$，则称 $\{x_1,\cdots,x_p\}$ 是 M 的仿射基. 相似地定义 M 的维数 $\dim M$ 等于 M 的仿射基的元素的个数减 1. 所以，当且仅当 M 的仿射基的元素个数为 p 时，$\dim M=p-1$. 当且仅当 $p=\dim M+1$ 时，M 的 p 个仿射无关的向量 $x_1,\cdots,x_p$ 组成它的仿射基.

当 $M=\varnothing$ 时，令 $\dim M=-1$. 一个点构成的仿射集是零维仿射集. 直线是一维仿射集，因为直线上不同的两个点 x_1、x_2 的集合仿射无关，故 $\mathrm{aff}\{x_1,x_2\}$ 的维数等于 1，而 $\mathrm{aff}\{x_1,x_2\}$ 正是过 x_1、x_2 的直线.

设 $M\subset R^n$，$\dim(\mathrm{aff}M)=p-1$，则存在属于 M 的向量组成的仿射无关组 $\{x_1,\cdots,x_p\}$，即在 $\mathrm{aff}M$ 中存在由 M 的向量 $x_1,\cdots,x_p$ 组成的仿射基. 所以下面的结论成立：

定理 2.8 设 $M\subset R^n$，则在 M 中存在仿射无关的向量组 $\{x_1,\cdots,x_p\}$，使形如 $\sum\limits_{i=1}^{p}\lambda_i x_i$ 的仿射组合的全体构成的集合就是 $\mathrm{aff}M$，其中 $\sum\limits_{i=1}^{p}\lambda_i=1$，$\lambda_i\in R$，$i=1,\cdots,p$.

这个结论表明，为了得到 $\mathrm{aff}M$，只要取 M 中固定的仿射无关向量 $x_1,\cdots,x_p$ 的全体仿射组合就可以了. 此外，$\mathrm{aff}M$ 中每一

个向量都可以唯一地表示成 $x_1,\cdots,x_p$ 的仿射组合的形式.

定义 2.6 如果由 R^n 中的仿射集M到 R^m 中的映射T保持仿射组合，即

$$T\left(\sum_{i=1}^{p}\lambda_i x_i\right)=\sum_{i=1}^{p}\lambda_i T(x_i), \tag{2}$$

其中 $\sum_{i=1}^{p}\lambda_i=1,\lambda_i\in R,i=1,\cdots,p$,则称$T$是仿射变换. 如果$m=1$,则仿射变换$T$也称为仿射函数.

可以证明,如果T是仿射变换,M是 R^n 中的仿射集,则$TM=\{Tx\,|\,x\in M\}$是 R^m 中的仿射集. 特别地，仿射变换保持仿射包，即

$$\mathrm{aff}(TM)=T(\mathrm{aff}M). \tag{3}$$

也可以证明,仿射变换将线段变为线段，平行线变为平行线，两平行线段的长度之比不因仿射变换而改变.

定理 2.9 仿射变换是形如 $Tx=Ax+a$ 的映射T,其中A是从 R^n 到 R^m 的线性变换,$a\in R^m$.

证明 设T是仿射变换,令 $a=T\theta,Ax=Tx-a$,容易验证 A 是从 R^n 到 R^m 的线性变换.

反之,设A是从 R^n 到 R^m 的线性变换,$Tx=Ax+a$，于是对于$\forall x,y\in R^n,\lambda\in R$,有

$$\begin{aligned}T[(1-\lambda)x+\lambda y]&=(1-\lambda)Ax+\lambda Ay+a\\&=(1-\lambda)Ax+(1-\lambda)a\\&\quad+\lambda Ay+\lambda a\\&=(1-\lambda)Tx+\lambda Ty.\end{aligned}$$

所以T是仿射变换. ▌

定理 2.10 设$\{x_0,x_1,\cdots,x_p\}$和$\{x_0',x_1',\cdots x_p'\}$是 R^n 中的两个仿射无关集,则存在从 R^n 到其自身的一对一仿射变换 T,使