

# 浅谈高等代数在最优化问题的应用

苏 勇

苏州科技大学数学科学学院, 江苏 苏州

收稿日期: 2022年10月18日; 录用日期: 2022年11月16日; 发布日期: 2022年11月23日

## 摘 要

高等代数是大学数学专业开设的三大专业课之一, 其内容可以概括为两个方面: 一是特殊概念的抽象化; 二是概念之间的严密逻辑推理。与初等代数相比, 高等代数中的相关概念和知识更具一般性和抽象性。在课程学习过程中, 如何理解、应用概念是教学难点。本文首先介绍高等代数中矩阵的特征值和特征向量等相关知识, 然后用他们解决工程学和经济学中的最优化问题。从而让学生更好的理解、掌握和应用高等代数中的相关概念和知识。

## 关键词

矩阵, 特征值, 特征向量, 最优化问题

# The Application of Advanced Algebra in Optimization Problem

Yong Su

School of Mathematics, Suzhou University of Science and Technology, Suzhou Jiangsu

Received: Oct. 18<sup>th</sup>, 2022; accepted: Nov. 16<sup>th</sup>, 2022; published: Nov. 23<sup>rd</sup>, 2022

## Abstract

Advanced algebra is one of three major courses for mathematics students, whose contents include mainly: the abstraction of the certain concepts and logical reasoning among concepts. Compared with elementary algebra, the concepts and knowledge of advanced algebra are much more general and abstract. In its study, the understanding and applications of the concepts are difficult points. This paper mainly introduces the applications of advanced algebra to optimization problem, and guides students how to understand the concepts from advanced algebra and how to use them.

## Keywords

Matrix, Eigenvalue, Eigenvector, Optimization Problem

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

最优化理论是一门在尽可能节省时间、人力和物力的前提之下获取最佳效果的理论，其在人类实践活动中应用十分普遍。在决策、经济学以及工程学中，下述问题十分常见：在诸多限制约束下，选择使关注的指标获得最优的方案。如，在既定约束下，选择尺寸，使材料的体重最轻，这是结构设计问题；在已有资源情况下，恰当分配用户，以获得最大的生产效益，这是资源分配问题；在物资需求和装载能力有限情况下，安排运输路线及运输量，来控制运输成本，使其达到最低值，这是运输方案问题；按照顾客的需求，如何订购原料和零件、安排投产日程和数量，使得成本最低，利润最高，这是生产计划问题。这些问题的本质是：在既定约束限制下，设计使目标指标取得最优的方案，属于约束最优化问题[1]。拉格朗日(Lagrange)乘数法是求解这类约束最优化问题的有效方法[2]。

本文考虑一类特殊的约束最优化问题：目标函数是实二次型的约束最优化问题。它们可以利用特征值和特征向量的相关知识解决。纯代数的方法求解这类问题比利用拉格朗日乘数法求解更为简单。我们将证明：决策变量取最大特征值对应的单位特征向量时，目标函数取得最大值，该最大值为二次型矩阵的最大特征值；决策变量取最小特征值对应的单位特征向量时，目标函数取得最小值，该最小值为二次型矩阵的最小特征值。随后，我们将通过一个例子来解释上面结论。通过该例子学生可以更好的理解矩阵特征值和特征向量的这两个概念的本质，及它们在分析线性空间方面的重要意义：线性空间可划分为不变子空间(即特征空间)的直和；直观地说“大空间”可分解为“小空间”直和；高维空间的研究转化为低维空间的研究。本文结构如下：第2节介绍一些高等代数中矩阵特征值和特征向量方面的必要基础知识；第3节主要借助用高等代数中的矩阵特征值和特征向量相关知识来求解约束最优化问题；第4节为总结与展望。

## 2. 预备知识

在下文中， $\mathbf{A}$ 表示 $n$ 阶矩阵； $\mathbf{E}$ 表示 $n$ 阶实单位矩阵； $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 表示对角线元素分别为 $\lambda_1, \lambda_2, \dots, \lambda_n$ 的实对角矩阵； $\mathbf{x}$ 表示实 $n$ 维实列向量； $\mathbf{x}^T$ 表示列向量 $\mathbf{x}$ 的转置。

### 2.1. 特征值、特征向量

定义1 [3] 设 $\mathbf{A}$ 为 $n$ 阶矩阵， $\mathbf{x}$ 为非零向量。若存在数 $\lambda$ 使得 $\mathbf{Ax} = \lambda\mathbf{x}$ ，则称 $\lambda$ 为 $\mathbf{A}$ 的特征值， $\mathbf{x}$ 为对应于 $\lambda$ 的特征向量。

矩阵是描述线性变换的工具。线性变换与矩阵乘法相对应，将给定向量变换成新向量；在变换过程中，原向量旋转和伸缩，向量的方向和长度都可能改变。向量是某个矩阵的特征向量表明：在相应的变换过程中，向量不旋转；仅以特征值为比例进行伸缩。

定义2 [3] 设 $\mathbf{A}$ 为 $n$ 阶矩阵，则行列式 $|\lambda\mathbf{E} - \mathbf{A}|$ 称为矩阵 $\mathbf{A}$ 的特征多项式， $|\lambda\mathbf{E} - \mathbf{A}| = 0$ 称为 $\mathbf{A}$ 的特

征方程。

数  $\lambda$  是  $\mathbf{A}$  的特征值当且仅当方程

$$\lambda \mathbf{E} - \mathbf{A} = 0 \quad (*)$$

有非平凡解。矩阵  $\mathbf{A}$  的对应于  $\lambda$  的特征空间是矩阵  $\lambda \mathbf{E} - \mathbf{A}$  的零空间，即方程(\*)的全体解集构成的集合。

求矩阵  $\mathbf{A}$  的特征值即是求特征方程的根；对应于  $\lambda$  的特征向量是齐次方程(\*)的非零解。

## 2.2. 实二次型

18 世纪开始，几何学关注二次曲线方程和二次曲面方程，通过适当选取坐标轴可以简化这两类方程，将它们化为标准形。这是二次型的起源。

定义 3 [3] 称函数  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ：

$$f(x_1, x_2, \dots, x_n) = \mathbf{x}^T \mathbf{A} \mathbf{x}$$

为实二次型，此处， $\mathbf{A}$  表示  $n$  阶实对称矩阵，且称矩阵  $\mathbf{A}$  为关于二次型的矩阵。

标准二次型是只含有平方项的二次型，它的二次型矩阵是对角矩阵(详见文献[3])。借助特征方程的概念，可将二次型的化简。欧拉的著作出现了特征方程的思想；拉格朗日首次明确提出该概念。柯西证明了：通过同一个线性变换，可将  $n$  个变元的两个二次型同时化为标准形。

定理 1 [3] 设  $\mathbf{x}^T \mathbf{A} \mathbf{x}$  为实二次型， $\lambda_i (i=1, 2, \dots, n)$  是  $\mathbf{A}$  的全部特征值，列向量  $\alpha_1, \alpha_2, \dots, \alpha_n$  是  $\mathbf{A}$  分别属于  $\lambda_1, \lambda_2, \dots, \lambda_n$  的单位特征向量。令  $\mathbf{T} = (\alpha_1, \alpha_2, \dots, \alpha_n)$ ，则通过正交线性变换  $\mathbf{x} = \mathbf{T} \mathbf{y}$ ，可使实二次型  $\mathbf{x}^T \mathbf{A} \mathbf{x}$  化为标准型  $\sum_{i=1}^n \lambda_i y_i^2$ 。

## 3. 约束最优化问题

利用拉格朗日乘数法解决约束最优化问题，较为有效。纯代数的方法解决对一些特殊问题更加简单、有效。如，当目标函数为实二次型时，矩阵的特征值和特征向量方法更加有效、简单。这类问题的数学模性如下：

$$\begin{aligned} \max_{\mathbf{x}} f(\mathbf{x}) \\ \text{s.t. } g(\mathbf{x}) = 0 \end{aligned}$$

其中， $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  是决策变量；目标函数  $f(\mathbf{x})$  是实二次型  $\mathbf{x}^T \mathbf{A} \mathbf{x}$ ； $g(\mathbf{x}) = \mathbf{x}^T \mathbf{x} - 1$ ；记号 s.t. 是“subject to”的缩写，专指约束条件。

下面，我们将证明：目标函数的最大(小)值为二次型矩阵的最大(小)特征值；决策变量取最大特征值对应的单位特征向量和最小特征值对应的单位特征向量时，目标函数取得极值。

定理 2 设  $\mathbf{A}$  为  $n$  阶实对称矩阵，令

$$\begin{aligned} m &= \min \{ \mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x}^T \mathbf{x} = 1 \}, \\ M &= \max \{ \mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x}^T \mathbf{x} = 1 \}, \end{aligned}$$

则  $m$  和  $M$  分别是矩阵  $\mathbf{A}$  的最小和最大特征值。设  $\alpha$  是  $\mathbf{A}$  的属于  $M$  的单位特征向量，则  $\alpha^T \mathbf{A} \alpha = M$ 。设  $\beta$  是  $\mathbf{A}$  的属于  $m$  的单位特征向量，则  $\beta^T \mathbf{A} \beta = m$ 。

证明：不妨设  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ ，且列向量  $\alpha_1, \alpha_2, \dots, \alpha_n$  是  $\mathbf{A}$  分别属于  $\lambda_1, \lambda_2, \dots, \lambda_n$  的单位特征向量，则基于定理 1 知：存在变换  $\mathbf{x} = \mathbf{T} \mathbf{y}$ ，使得

$$\sum_{i=1}^n \lambda_i y_i^2 = \mathbf{y}^T \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{y}.$$

因为  $\mathbf{T}$  是正交矩阵, 所以  $\mathbf{T}^T \mathbf{T} = \mathbf{E}$ 。因此,  $\mathbf{x}^T \mathbf{x} = (\mathbf{T}\mathbf{y})^T \mathbf{T}\mathbf{y} = \mathbf{y}^T \mathbf{y}$ 。故,

$$M = \max \{ \mathbf{y}^T \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{y} \mid \mathbf{y}^T \mathbf{y} = 1 \}.$$

于是,

$$\mathbf{y}^T \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{y} = \sum_{i=1}^n \lambda_i y_i^2 \leq \lambda_1 \sum_{i=1}^n y_i^2 = \lambda_1.$$

因此,  $M \leq \lambda_1$ 。另一方面, 对于  $\mathbf{e}_1 = (1, 0, \dots, 0)$ , 有  $\mathbf{e}_1^T \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{e}_1 = \lambda_1$ 。于是,  $M = \lambda_1$ 。简单计算可知:  $\mathbf{T}\mathbf{e}_1 = \boldsymbol{\alpha}_1$ 。于是,

$$\boldsymbol{\alpha}_1^T \mathbf{A} \boldsymbol{\alpha}_1 = \mathbf{e}_1^T \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{e}_1 = \lambda_1 = M.$$

类似地, 可以证明关于最小特征值的情况。

例 1: 一县政府计划制定下一年度的公共工作计划, 拟修  $x$  公里路,  $y$  公顷的公园, 如何分配资源使得效用最大, 这是政府部门须要考虑的问题。与仅开始一个项目相比, 两个项目同时进行效用更高, 但  $x$  和  $y$  必须满足限制曲线:  $4x^2 + 9y^2 \leq 36$  (\*\*). 可行集中的点(满足不等式(\*\*)的点  $(x, y)$ )表示一个可能的公共工作计划。资源利用率取得最大的地方在限制曲线边界  $4x^2 + 9y^2 = 36$  上。为了选择该年度的公共计划, 县政府考虑居民的意见, 经济学家给出无差别曲线  $q(x, y) = xy$  作为效用函数。求公共工作计划, 使  $q$  最大。

解: 定义  $x_1 = x/3, x_2 = y/2$ , 取  $\mathbf{x} = (x_1, x_2)^T$ , 则原问题转化为: 在约束条件  $\mathbf{x}^T \mathbf{x} = 1$  下, 求目标函数  $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$  的最大值问题, 其中

$$\mathbf{A} = \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix}.$$

矩阵  $\mathbf{A}$  的特征值为  $\pm 3$ ; 在  $x_1 = x_2 = \frac{1}{\sqrt{2}}$  处,  $Q(\mathbf{x})$  取得最大值为 3。

## 4. 结论

约束最优化问题在工程、经济、决策等领域中有重要的应用。拉格朗日(Lagrange)乘数法是求解约束最优化问题的有效方法[2]。本文考虑一类特殊的约束最优化问题: 目标函数是实二次型。利用特征值和特征向量的相关知识可以解决这类约束最优化问题, 该解法比拉格朗日乘数法更为简单。关于高等代数在约束最优化问题中的进一步应用, 可参考文献[4][5]。在高等代数的教学过程中, 如何理解和应用概念是教学的难点。本文通过例 1 来解释如何理解矩阵特征值和特征向量概念的本质, 以及它们在分析线性空间方面的重要意义: 线性空间可划分为不变子空间(即特征空间)的直和; 直观地说“大空间”可分解为“小空间”直和; 高维空间的研究转化为低维空间的研究。然而, 在教学过程中, 如何更加直观的讲述高等代数中的概念, 如何应用相关理论知识, 是须要进一步探索的问题。

## 基金项目

国家自然科学基金(11801220), 江苏省自然科学基金(BK20180590), 苏州科技大学天平学院校级教育教学改革研究课题(2022TJGB-11)。

---

## 参考文献

- [1] 刘浩洋, 户将, 李勇锋, 文再文. 最优化: 建模、算法与理论[M]. 北京: 高等教育出版社, 2020.
- [2] 郭进峰, 李炜玲, 沈菁华. 高等数学[M]. 北京: 高等教育出版社, 2020.
- [3] 邱森. 高等代数[M]. 武汉: 武汉大学出版社, 2012.
- [4] Gilbert, S. (2004) Linear Algebra and Its Applications. Brooks Cole, Belmont.
- [5] Gilbert, S. (2021) Introduction to Linear Algebra. Wellesley Cambridge Press, Wellesley.