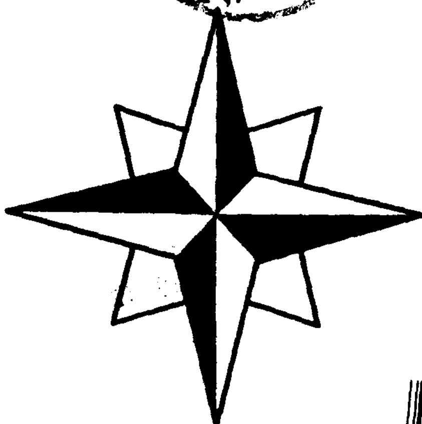
最优控制的要点·例题·习题

徐湘元　编解

华南理工大学出版社

最优控制的要点·例题·习题

徐湘元　编解



华南理工大学出版社
·广州·

内 容 简 介

本书是《最优控制理论与应用》课程的教学参考书．全书涉及确定性最优控制理论与应用的主要内容，分七章，每章包括有理论要点、例题的分析与解答、习题三部分．理论要点概括了该章的主要概念、定理和公式，有利于理清头绪，抓住重点；例题有 120 道，较全面地介绍解题的思路和步骤，阐述分析问题和解决问题的基本方法，有利于拓宽思路，掌握技巧；习题共有 90 多道，并附有部分习题的参考答案，供读者练习和参考．

本书可作为高等学校控制类和管理类专业的本科生、研究生的教学参考用书，也可供科研单位及工矿企业工程技术人员参考，尤其适用于最优控制的自学者．

图书在版编目（CIP）数据

最优控制的要点·例题·习题 / 徐湘元编解. — 广州：华南理工大学出版社，1997.6
ISBN 7-5623-1111-0

I. 最…
II. 徐…
III. 最佳控制-高等学校-教学参考资料
IV. TP13

华南理工大学出版社出版发行
（广州五山　邮编 510641）
责任编辑　赖淑华
各地新华书店经销
华南理工大学印刷厂印装
开本:787×1092　1/16　印张:11.75　字数:282 千
1997 年 6 月第 1 版　1998 年 4 月第 2 次印刷
印数:1001—2000 册
定价:15.00 元

前　言

作为现代控制理论重要组成部分的最优控制理论，自60年代以来，不仅在工程技术中得到了广泛而成功的应用，而且在社会、经济、军事、管理等方面受到了高度的重视并取得了卓有成效的实际效果。目前，除了理工科控制类的大学生、研究生和科技人员把它作为必修的专业课程外，其他专业或领域的科技人员也把它作为选修课程。

作为学习《最优控制理论与应用》的教材，国内虽有一些，但是，有关辅导方面的教学参考用书却极为罕见。由于最优控制理论涉及较多的数学基础和物理概念，令不少读者，尤其是自学者，感到困难和吃力，以致学习效果不佳。表现为：或概念不清，或支离破碎，或难以学以致用。他们迫切需要一本能加深理解、巩固概念、熟练应用的参考书。为此，本人编写了这本《最优控制的要点·例题·习题》，以应读者参考和提高之需。

本书为《最优控制理论与应用》课程的配套教学参考书，全书分为七章，每章包括有理论要点、题目和题解。理论要点部分归纳性地叙述各章的主要概念、重要定理和公式，一为解题提供依据，二为理顺关系，突出重点，便于记忆和掌握。题目部分包括例题的题目和习题的题目，选择时考虑了广泛性、典型性，并且有一定的深度，习题供练习用。题解部分包含有例题题目的分析与解答，以及部分习题的参考答案。本书在结构上采用例题题目和例题解答分离的做法，一是让读者看了例题题目后，在空间上留有思考的余地，二是例题既可作为解题示例阅读，又可作为习题练习，以满足不同读者的需要。

本书以实际应用为出发点，没有迂回于深奥的数学推理之中，而着重于用最优控制基本理论分析和解决实际问题。书中有许多题目的解法不止一种，例题中的解法仅起抛砖引玉的作用，读者可通过阅读和练习，达到加深理解、开阔思路、掌握技巧、提高能力的目的。

考虑到篇幅、计算工具、使用的广泛性等方面的问题，本书不涉及最优控制中的一些专门问题，如数值方法、准最优、奇异最优等。

本书在编写过程中，先后得到了广东工业大学符曦教授和华南理工大学毛宗源教授的关心与鼓励。初稿完成后，符老师认真审阅了全书(主审)，并且提出了宝贵的意见与建议，在此，特致谢意。另外，广州舰艇学院李忠老师审校了部分题解，华南理工大学应用数学系94级研究生张永东同学验核了除首尾两章以外的内容，广东工业大学电气工程系提供了出版资助，自控教研室给予了许多方便，华南理工大学出版社的同志为该书的早日出版做了许多具体工作，在此一并致谢。

由于水平所限，经验不足，书中难免有错误或不完善之处，恳请读者批评指正。

徐湘元

1996年12月于广州

目　　录

第 1 章　数学基础

1.1　理论要点

1.1.1　对数量变量的导数

1. 一个 n 维函数向量 $f(t)=\begin{bmatrix}f_1(t) & f_2(t) & \cdots & f_n(t)\end{bmatrix}^{\mathrm{T}}$ 对数量自变量 t 的导数为

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t}=\begin{bmatrix}\frac{\mathrm{d}f_1(t)}{\mathrm{d}t} & \frac{\mathrm{d}f_2(t)}{\mathrm{d}t} & \cdots & \frac{\mathrm{d}f_n(t)}{\mathrm{d}t}\end{bmatrix}^{\mathrm{T}}$$

2. 一个 $n\times m$ 矩阵函数

$$\boldsymbol{F}(t)=\begin{pmatrix}f_{11}(t) & f_{12}(t) & \cdots & f_{1m}(t)\\ f_{21}(t) & f_{22}(t) & \cdots & f_{2m}(t)\\ \vdots & & & \\ f_{n1}(t) & f_{n2}(t) & \cdots & f_{nm}(t)\end{pmatrix}=\left(f_{ij}(t)\right)_{n\times m}$$

对数量自变量 t 的导数为

$$\frac{\mathrm{d}\boldsymbol{F}(t)}{\mathrm{d}t}=\left[\frac{\mathrm{d}f_{ij}(t)}{\mathrm{d}t}\right]_{n\times m}$$

性质：设 $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$ 均为函数矩阵，$\lambda=\lambda(t)$ 为数量函数，则有

（1）$\dfrac{\mathrm{d}(\boldsymbol{A}\pm\boldsymbol{B})}{\mathrm{d}t}=\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\pm\dfrac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}$

（2）$\dfrac{\mathrm{d}(\lambda\boldsymbol{A})}{\mathrm{d}t}=\dfrac{\mathrm{d}\lambda}{\mathrm{d}t}+\lambda\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}$

（3）$\dfrac{\mathrm{d}(\boldsymbol{AC})}{\mathrm{d}t}=\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\boldsymbol{C}+\boldsymbol{A}\dfrac{\mathrm{d}\boldsymbol{C}}{\mathrm{d}t}$

注：若 $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$ 改为向量，则求导公式不变。

1.1.2　对向量的导数

1. 设 $f=f(\boldsymbol{x})=f(x_1\ \ x_2\ \ \cdots\ \ x_n)$ 是以向量 $\boldsymbol{x}$ 为自变量的数量函数，$\boldsymbol{x}$ 为 n 维列向量，则

$$\frac{\mathrm{d} f(\boldsymbol{x})}{\mathrm{d} \boldsymbol{x}}=\left[\frac{\partial f(\boldsymbol{x})}{\partial x_1} \quad \frac{\partial f(\boldsymbol{x})}{\partial x_2} \quad \cdots \quad \frac{\partial f(\boldsymbol{x})}{\partial x_n}\right]^{\mathrm{T}}$$

它实际上是多元函数的导数或梯度，记为

$$\nabla f(\boldsymbol{x})=\operatorname{grad} f(\boldsymbol{x})=\frac{\mathrm{d} f}{\mathrm{d} \boldsymbol{x}}$$

2. 设 $\boldsymbol{f}=\boldsymbol{f}(\boldsymbol{x})=\left[f_1(t) \quad f_2(t) \quad \cdots \quad f_m(t)\right]^{\mathrm{T}}$ 为 m 维向量函数，且 $\boldsymbol{x}$ 为 n 维列变量，则

$$\frac{\mathrm{d} \boldsymbol{f}(\boldsymbol{x})}{\mathrm{d} \boldsymbol{x}}=\frac{\mathrm{d} \boldsymbol{f}(\boldsymbol{x})}{\mathrm{d} \boldsymbol{x}^{\mathrm{T}}}=\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}=\left(\frac{\partial f_i(\boldsymbol{x})}{\partial \boldsymbol{x}_j}\right)_{m \times n}$$

$$\frac{\mathrm{d} \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x})}{\mathrm{d} \boldsymbol{x}}=\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_2} \\ \vdots & & & \\ \frac{\partial f_1}{\partial x_n} & \frac{\partial f_2}{\partial x_n} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}=\left(\frac{\partial f_j(\boldsymbol{x})}{\partial x_i}\right)_{n \times m}=\left(\frac{\mathrm{d} \boldsymbol{f}(\boldsymbol{x})}{\mathrm{d} \boldsymbol{x}^{\mathrm{T}}}\right)^{\mathrm{T}}$$

性质：设

$$\boldsymbol{a}(\boldsymbol{x})=\left[a_1(\boldsymbol{x}) \quad a_2(\boldsymbol{x}) \quad \cdots \quad a_m(\boldsymbol{x})\right]^{\mathrm{T}},\ \boldsymbol{x}=\left[x_1 \quad x_2 \quad \cdots \quad x_n\right]^{\mathrm{T}}$$

$\boldsymbol{b}(\boldsymbol{x})=\left[b_1(\boldsymbol{x}) \quad b_2(\boldsymbol{x}) \quad \cdots \quad b_m(\boldsymbol{x})\right]^{\mathrm{T}}$，$\lambda=\lambda(\boldsymbol{x})$ 为数量函数，则有

(1) $$\frac{\mathrm{d}(\boldsymbol{a} \pm \boldsymbol{b})}{\mathrm{d} \boldsymbol{x}}=\frac{\mathrm{d} \boldsymbol{a}}{\mathrm{d} \boldsymbol{x}} \pm \frac{\mathrm{d} \boldsymbol{b}}{\mathrm{d} \boldsymbol{x}}$$

(2) $$\frac{\mathrm{d}(\lambda \boldsymbol{a}^{\mathrm{T}})}{\mathrm{d} \boldsymbol{x}}=\frac{\mathrm{d} \lambda}{\mathrm{d} \boldsymbol{x}} \boldsymbol{a}^{\mathrm{T}}+\lambda \frac{\mathrm{d} \boldsymbol{a}^{\mathrm{T}}}{\mathrm{d} \boldsymbol{x}}$$

(3) $$\frac{\mathrm{d}(\boldsymbol{a}^{\mathrm{T}} \boldsymbol{b})}{\mathrm{d} \boldsymbol{x}}=\frac{\mathrm{d} \boldsymbol{a}^{\mathrm{T}}}{\mathrm{d} \boldsymbol{x}} \boldsymbol{b}+\frac{\mathrm{d} \boldsymbol{b}^{\mathrm{T}}}{\mathrm{d} \boldsymbol{x}} \boldsymbol{a}$$

实用等式

$$\frac{\mathrm{d} \boldsymbol{x}}{\mathrm{d} \boldsymbol{x}^{\mathrm{T}}}=\boldsymbol{I} \quad , \quad \frac{\mathrm{d} \boldsymbol{x}^{\mathrm{T}}}{\mathrm{d} \boldsymbol{x}}=\boldsymbol{I}$$

其中，$\boldsymbol{I}$ 为单位矩阵。

3. 设 $\boldsymbol{F}(\boldsymbol{x})$ 为 $m \times l$ 函数矩阵，$\boldsymbol{x}$ 为 n 维向量变量，即

$$
\boldsymbol{F}(\boldsymbol{x})=\begin{pmatrix} f_{11}(\boldsymbol{x}) & f_{12}(\boldsymbol{x}) & \cdots & f_{1l}(\boldsymbol{x}) \\ f_{21}(\boldsymbol{x}) & f_{22}(\boldsymbol{x}) & \cdots & f_{2l}(\boldsymbol{x}) \\ \vdots & & & \\ f_{m1}(\boldsymbol{x}) & f_{m2}(\boldsymbol{x}) & \cdots & f_{ml}(\boldsymbol{x}) \end{pmatrix}, \quad \boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}
$$

定义

$$
\frac{\mathrm{d}\boldsymbol{F}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}=\begin{pmatrix} \dfrac{\partial \boldsymbol{F}}{\partial x_1} \\ \dfrac{\partial \boldsymbol{F}}{\partial x_2} \\ \vdots \\ \dfrac{\partial \boldsymbol{F}}{\partial x_n} \end{pmatrix}_{nm\times l}, \quad \frac{\mathrm{d}\boldsymbol{F}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}=\begin{pmatrix} \dfrac{\partial \boldsymbol{F}}{\partial x_1} & \dfrac{\partial \boldsymbol{F}}{\partial x_2} & \cdots & \dfrac{\partial \boldsymbol{F}}{\partial x_n} \end{pmatrix}_{m\times ln}
$$

其中

$$
\frac{\partial \boldsymbol{F}}{\partial x_i}=\begin{pmatrix} \dfrac{\partial f_{11}(\boldsymbol{x})}{\partial x_i} & \dfrac{\partial f_{12}(\boldsymbol{x})}{\partial x_i} & \cdots & \dfrac{\partial f_{1l}(\boldsymbol{x})}{\partial x_i} \\ \dfrac{\partial f_{21}(\boldsymbol{x})}{\partial x_i} & \dfrac{\partial f_{22}(\boldsymbol{x})}{\partial x_i} & \cdots & \dfrac{\partial f_{2l}(\boldsymbol{x})}{\partial x_i} \\ \vdots & & & \\ \dfrac{\partial f_{m1}(\boldsymbol{x})}{\partial x_i} & \dfrac{\partial f_{m2}(\boldsymbol{x})}{\partial x_i} & \cdots & \dfrac{\partial f_{ml}(\boldsymbol{x})}{\partial x_i} \end{pmatrix}
$$

注：(1) 一般$\dfrac{\mathrm{d}\boldsymbol{F}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}} \neq \dfrac{\mathrm{d}\boldsymbol{F}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}$；

(2) 2 中性质(1)、性质(2)仍适用，但性质(3)一般不成立。

1.1.3 对矩阵变量的导数

1. 设$f=f(\boldsymbol{X})$为数量函数，$\boldsymbol{X}$为$n\times m$矩阵，则

$$
\frac{\mathrm{d}f(\boldsymbol{X})}{\mathrm{d}\boldsymbol{X}}=\begin{pmatrix} \dfrac{\partial f}{\partial x_{11}} & \dfrac{\partial f}{\partial x_{12}} & \cdots & \dfrac{\partial f}{\partial x_{1m}} \\ \dfrac{\partial f}{\partial x_{21}} & \dfrac{\partial f}{\partial x_{22}} & \cdots & \dfrac{\partial f}{\partial x_{2m}} \\ \vdots & & & \\ \dfrac{\partial f}{\partial x_{n1}} & \dfrac{\partial f}{\partial x_{n2}} & \cdots & \dfrac{\partial f}{\partial x_{nm}} \end{pmatrix}=\left(\frac{\partial f(\boldsymbol{X})}{\partial x_{ij}}\right)_{n\times m}
$$

性质：设$f(\boldsymbol{X})$、$g(\boldsymbol{X})$为数量函数，$\boldsymbol{X}$为矩阵变量，则

$$
\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}[f(\boldsymbol{X})\pm g(\boldsymbol{X})]=\frac{\mathrm{d}f(\boldsymbol{X})}{\mathrm{d}\boldsymbol{X}}\pm\frac{\mathrm{d}g(\boldsymbol{X})}{\mathrm{d}\boldsymbol{X}}
$$

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}[\lambda f(\boldsymbol{X})]=\lambda\frac{\mathrm{d}f(\boldsymbol{X})}{\mathrm{d}\boldsymbol{X}}\quad(\lambda\text{ 为常数})$$

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}[f(\boldsymbol{X})g(\boldsymbol{X})]=\frac{\mathrm{d}f(\boldsymbol{X})}{\mathrm{d}\boldsymbol{X}}g(\boldsymbol{X})+f(\boldsymbol{X})\frac{\mathrm{d}g(\boldsymbol{X})}{\mathrm{d}\boldsymbol{X}}$$

2. 设 $z(\boldsymbol{X})$ 为 l 维向量函数，$\boldsymbol{X}$ 为 $n\times m$ 变量矩阵，则有

$$\frac{\mathrm{d}z(\boldsymbol{X})}{\mathrm{d}\boldsymbol{X}}=\begin{pmatrix}\frac{\partial z}{\partial x_{11}} & \frac{\partial z}{\partial x_{12}} & \cdots & \frac{\partial z}{\partial x_{1m}}\\ \frac{\partial z}{\partial x_{21}} & \frac{\partial z}{\partial x_{22}} & \cdots & \frac{\partial z}{\partial x_{2m}}\\ \vdots & & & \\ \frac{\partial z}{\partial x_{n1}} & \frac{\partial z}{\partial x_{n2}} & \cdots & \frac{\partial z}{\partial x_{nm}}\end{pmatrix}=\left(\frac{\partial z(\boldsymbol{X})}{\partial x_{ij}}\right)_{ln\times m}$$

3. 设 $\boldsymbol{F}(\boldsymbol{X})$ 为 $l\times p$ 矩阵，$\boldsymbol{X}$ 为 $n\times m$ 矩阵变量，定义

$$\frac{\mathrm{d}\boldsymbol{F}(\boldsymbol{X})}{\mathrm{d}\boldsymbol{X}}=\begin{pmatrix}\frac{\partial F}{\partial x_{11}} & \frac{\partial F}{\partial x_{12}} & \cdots & \frac{\partial F}{\partial x_{1m}}\\ \frac{\partial F}{\partial x_{21}} & \frac{\partial F}{\partial x_{22}} & \cdots & \frac{\partial F}{\partial x_{2m}}\\ \vdots & & & \\ \frac{\partial F}{\partial x_{n1}} & \frac{\partial F}{\partial x_{n2}} & \cdots & \frac{\partial F}{\partial x_{nm}}\end{pmatrix}=\left(\frac{\partial \boldsymbol{F}(\boldsymbol{X})}{\partial x_{ij}}\right)_{ln\times pm}$$

其中

$$\frac{\mathrm{d}\boldsymbol{F}(\boldsymbol{X})}{\mathrm{d}x_{ij}}=\begin{pmatrix}\frac{\partial f_{11}}{\partial x_{ij}} & \frac{\partial f_{12}}{\partial x_{ij}} & \cdots & \frac{\partial f_{1p}}{\partial x_{ij}}\\ \frac{\partial f_{21}}{\partial x_{ij}} & \frac{\partial f_{22}}{\partial x_{ij}} & \cdots & \frac{\partial f_{2p}}{\partial x_{ij}}\\ \vdots & & & \\ \frac{\partial f_{l1}}{\partial x_{ij}} & \frac{\partial f_{l2}}{\partial x_{ij}} & \cdots & \frac{\partial f_{lp}}{\partial x_{ij}}\end{pmatrix}$$

1.1.4 复合函数的导数

1. 设 $f=f(\boldsymbol{x})$，$\boldsymbol{x}=\boldsymbol{x}(t)$（$f$ 为数量函数，$\boldsymbol{x}$ 为向量，t 为数量变量），则

$$\frac{\mathrm{d}f}{\mathrm{d}t}=\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$$

2. 设 $f=f(\boldsymbol{y}),\boldsymbol{y}=\boldsymbol{y}(\boldsymbol{x})$，则

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}=\frac{\mathrm{d}\boldsymbol{y}^{\mathrm{T}}}{\mathrm{d}\boldsymbol{x}}\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{y}}\quad,\quad\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}=\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{y}^{\mathrm{T}}}\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}$$

3. 设$z=z(y)$，$y=y(t)$，则

$$\frac{\mathrm{d}z}{\mathrm{d}t}=\frac{\mathrm{d}z}{\mathrm{d}y^{\mathrm{T}}}\frac{\mathrm{d}y}{\mathrm{d}t}$$

4. 设$z=z(y)$，$y=y(x)$，则

$$\frac{\mathrm{d}z}{\mathrm{d}x^{\mathrm{T}}}=\frac{\mathrm{d}z}{\mathrm{d}y^{\mathrm{T}}}\frac{\mathrm{d}y}{\mathrm{d}x^{\mathrm{T}}} \quad , \quad \frac{\mathrm{d}z^{\mathrm{T}}}{\mathrm{d}x}=\frac{\mathrm{d}y^{\mathrm{T}}}{\mathrm{d}x}\frac{\mathrm{d}z^{\mathrm{T}}}{\mathrm{d}y}$$

5. 设$z=z(x,y)$，$y=y(x)$，则

$$\frac{\mathrm{d}z}{\mathrm{d}x^{\mathrm{T}}}=\frac{\partial z}{\partial x^{\mathrm{T}}}+\frac{\partial z}{\partial y^{\mathrm{T}}}\frac{\mathrm{d}y}{\mathrm{d}x^{\mathrm{T}}} \quad , \quad \frac{\mathrm{d}z^{\mathrm{T}}}{\mathrm{d}x}=\frac{\partial z^{\mathrm{T}}}{\partial x}+\frac{\mathrm{d}y^{\mathrm{T}}}{\mathrm{d}x}\frac{\partial z^{\mathrm{T}}}{\partial y}$$

1.1.5 函数极值

1. 函数的无条件极值

n元函数$f(x)$，其中 $x=\begin{pmatrix}x_1 & x_2 & \cdots & x_n\end{pmatrix}^{\mathrm{T}}$ 在x^*点取得极小值的充要条件为

$$\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x^*}=\left.\begin{pmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n}\end{pmatrix}^{\mathrm{T}}\right|_{x^*}=0 \quad , \quad \left.\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2}\right|_{x^*}=A>0$$

其中

$$A=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n}\\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 x_n}\\ \vdots & \vdots & \cdots & \vdots\\ \frac{\partial^2 f}{\partial x_n x_1} & \frac{\partial^2 f}{\partial x_n x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{pmatrix}_{x^*}$$

为$f(x)$在x^*处的二阶偏导数矩阵，称为海赛（Hessian）矩阵，大于零表明为正定矩阵。同理，若A为负定矩阵，则函数有极大值。

2. 拉格朗日乘子法

设连续可微的标量函数为$J=f(x,u)$，约束条件为 $g(x,u)=0$，其中

$$x=[x_1,\ x_2,\ \cdots,\ x_n]^{\mathrm{T}}$$
$$u=[u_1,\ u_2,\ \cdots,\ u_n]^{\mathrm{T}}$$
$$g(x,u)=\begin{bmatrix}g_1(x,u) & g_2(x,u) & \cdots & g_n(x,u)\end{bmatrix}^{\mathrm{T}}$$

求使标量函数J为最小的控制向量u。

先设拉格朗日函数

$$L(x,u,\lambda)=f(x,u)+\lambda^{\mathrm{T}}g(x,u)$$

其中 $\lambda=\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix}^{\mathrm{T}}$ 称为拉格朗日乘子，则取得最小值的充分条件为

$$L_x=f_x+g_x^{\mathrm{T}}\lambda=0$$

$$L_u = f_u + g_u^{\mathrm{T}}\lambda = 0$$
$$L_\lambda = g(x,u) = 0$$
$$\begin{pmatrix} -g_x^{-1}g_u \\ I \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} L_{xx} & L_{xu} \\ L_{ux} & L_{uu} \end{pmatrix} \begin{pmatrix} -g_x^{-1}g_u \\ I \end{pmatrix} > 0$$

其中

$$L_x = \frac{\partial L}{\partial x}, \quad f_x = \frac{\partial f}{\partial x}, \quad L_u = \frac{\partial L}{\partial u}, \quad g_x = \frac{\partial g}{\partial x}, \quad g_u = \frac{\partial g}{\partial u}$$
$$L_{xx} = \frac{\partial^2 L}{\partial x^2}, \quad L_{xu} = \frac{\partial^2 L}{\partial x \partial u}, \quad L_{ux} = \frac{\partial^2 L}{\partial u \partial x}, \quad L_{uu} = \frac{\partial^2 L}{\partial u^2}$$

3. 库恩-塔克尔（Kuhn-Tucker）定理

设函数 $f(x)$ 连续可微，在连续可微的不等式

$$g_i(x) \leqslant 0, \quad i = 1, 2, \cdots, q$$

约束下，$f(x)$ 在 $x = x^*$ 处取得极小值的必要条件为

(1) $\lambda \geqslant 0, \quad g(x^*) \leqslant 0$

(2) $\lambda^{\mathrm{T}} g(x^*) = 0$

(3) $\dfrac{\partial L(x^*,\lambda)}{\partial x} = 0$

其中 $\lambda = \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_q \end{bmatrix}^{\mathrm{T}}$，$g(x) = \begin{bmatrix} g_1(x) & g_2(x) & \cdots & g_q(x) \end{bmatrix}^{\mathrm{T}}$，

$L(x,\lambda) = f(x) + \lambda^{\mathrm{T}} g(x)$

1.1.6 二次型的正（负）定性

1. 定义

设 A 为二次型 $x^{\mathrm{T}}Ax$ 的 $n \times n$ 实对称矩阵（或为赫米特矩阵，下同）：

若当 $x \neq 0$ 时，$x^{\mathrm{T}}Ax > 0$，仅当 $x = 0$ 时，$x^{\mathrm{T}}Ax = 0$，则称 $x^{\mathrm{T}}Ax$ 正定，A 为正定矩阵。

若当 $x \neq 0$ 时，$x^{\mathrm{T}}Ax \geqslant 0$，当 $x = 0$ 时，$x^{\mathrm{T}}Ax = 0$，则称 $x^{\mathrm{T}}Ax$ 正半定，A 为正半定矩阵。

若当 $x \neq 0$ 时，$x^{\mathrm{T}}Ax < 0$，仅当 $x = 0$ 时，$x^{\mathrm{T}}Ax = 0$，则称 $x^{\mathrm{T}}Ax$ 负定，A 为负定矩阵。

若当 $x \neq 0$ 时，$x^{\mathrm{T}}Ax \leqslant 0$，当 $x = 0$ 时，$x^{\mathrm{T}}Ax = 0$，则称 $x^{\mathrm{T}}Ax$ 负半定，A 为负半定矩阵。

若 $x^{\mathrm{T}}Ax$ 可能为正或负，则称 $x^{\mathrm{T}}Ax$ 不定，A 为不定矩阵。

另外，当 A 的所有特征值为正数时，实对称矩阵 A 为正定矩阵，若 A 的所有特征值为非负，且至少有一特征值为零，则矩阵 A 为正半定矩阵。

2. 西尔威斯特（Sylvester）准则

二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正（负）定性的充分必要条件：

(1) 正定：$n\times n$ 实对称矩阵 $\boldsymbol{A}$ 的行列式为正，$\boldsymbol{A}$ 的行列式的顺序主子式（$\boldsymbol{A}$ 左上角的行列式）为正，即

$$a_{11}>0\ ,\ \begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}>0\ ,\ \begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}>0\ ,\ \cdots\ ,\ |A|>0$$

则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定，$\boldsymbol{A}$ 为正定矩阵。

(2) 负定：当 n 为偶数时，$\boldsymbol{A}$ 的行列式为正的，n 为奇数时，$\boldsymbol{A}$ 的行列式为负，并且偶阶顺序主子式为正，奇阶顺序主子式为负，即

$$a_{11}<0\ ,\ \begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}>0\ ,\ \begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}<0\ ,\ \cdots\ ,\ \begin{matrix}|A|>0\,(n\text{为偶数})\\|A|<0\,(n\text{为奇数})\end{matrix}$$

则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 负定，$\boldsymbol{A}$ 为负定矩阵。

(3) 正半定：设 $\boldsymbol{A}$ 为奇异矩阵，且所有主子式为非负的，即

$$a_{ii}\geqslant 0\ ,\ \begin{vmatrix}a_{ii}&a_{ij}\\a_{ji}&a_{jj}\end{vmatrix}\geqslant 0\ ,\ \begin{vmatrix}a_{ii}&a_{ij}&a_{ik}\\a_{ji}&a_{jj}&a_{jk}\\a_{ki}&a_{kj}&a_{kk}\end{vmatrix}\geqslant 0\ ,\ \cdots\ ,\ |A|=0$$

其中 $i<j<k$，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正半定，$\boldsymbol{A}$ 为正半定矩阵。

(4) 负半定：设 $\boldsymbol{A}$ 为奇异矩阵，并且所有偶阶主子式为非负的，所有奇阶主子式为非正的，即

$$a_{ii}\leqslant 0\ ,\ \begin{vmatrix}a_{ii}&a_{ij}\\a_{ji}&a_{jj}\end{vmatrix}\geqslant 0\ ,\ \begin{vmatrix}a_{ii}&a_{ij}&a_{ik}\\a_{ji}&a_{jj}&a_{jk}\\a_{ki}&a_{kj}&a_{kk}\end{vmatrix}\leqslant 0\ ,\ \cdots\ ,\ |A|=0$$

其中 $i<j<k$，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 负半定，$\boldsymbol{A}$ 为负半定矩阵。

1.2 题　目

1.2.1 例题题目

A-1-1 将标量函数 $f=x_1^2+4x_1x_3+2x_1x_2+2x_2^2+2x_2x_3+6x_3^2$ 写成 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 形式，并求 $\mathrm{d}f/\mathrm{d}\boldsymbol{x}$。

A-1-2 试证明

$$\frac{\mathrm{d}(\boldsymbol{A}\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}=\boldsymbol{A}$$

其中，$\boldsymbol{x}$为n维列向量，$\boldsymbol{A}$为$m\times n$常数矩阵。

A-1-3　求$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$对$\boldsymbol{x}^{\mathrm{T}}$的导数，其中，$\boldsymbol{x}$为$n$维列向量，$\boldsymbol{A}$为$n\times n$对称常数矩阵。

A-1-4　求 $\dfrac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}}(\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x})$，其中，$\boldsymbol{x}$为$n$维列向量，$\boldsymbol{\lambda}$为$n$维常数列向量，$\boldsymbol{A}$为$n\times n$常数矩阵。

A-1-5　求向量函数 $\boldsymbol{F}=\sin(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x})\cdot\boldsymbol{x}$ 对 $\boldsymbol{x}^{\mathrm{T}}$ 的导数。

A-1-6　求向量函数 $\boldsymbol{F}=\cos(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x})\cdot\boldsymbol{x}^{\mathrm{T}}$ 对$\boldsymbol{x}$的导数。

A-1-7　对于标量函数

$$H(\boldsymbol{x},\boldsymbol{u},\boldsymbol{\lambda})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}+\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u})$$

试求：$\dfrac{\partial H}{\partial \boldsymbol{x}}$、$\dfrac{\partial H}{\partial \boldsymbol{u}}$、$\dfrac{\partial H}{\partial \boldsymbol{\lambda}}$。其中，$\boldsymbol{x}$和$\boldsymbol{\lambda}$均为$n$维列向量，$\boldsymbol{u}$为$m$维列向量，$\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{Q}$和$\boldsymbol{R}$均为适当维数的常数矩阵。

A-1-8　已知标量函数$H[\boldsymbol{x}(t),\boldsymbol{\lambda}(t),\boldsymbol{u}(t),t]$，试求 $\mathrm{d}H/\mathrm{d}t$，其中，$\boldsymbol{x}(t)$、$\boldsymbol{u}(t)$、$\boldsymbol{\lambda}(t)$均为适当维数的列向量，t为实变量。

A-1-9　求标量函数 $f=(\boldsymbol{x}-\boldsymbol{a}-B\boldsymbol{z})^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{a}-B\boldsymbol{z})$ 对$\boldsymbol{x}$、$\boldsymbol{z}$的导数，其中$\boldsymbol{x}$、$\boldsymbol{a}$分别为n维变量和常数向量，$\boldsymbol{z}$为m维列向量，$\boldsymbol{B}$为$n\times m$常数矩阵。

A-1-10　试推证

$$\frac{\partial}{\partial X}\mathrm{Tr}AX^{\mathrm{T}}=A,\quad \frac{\partial}{\partial X}\mathrm{Tr}XX=2X^{\mathrm{T}}$$

其中，Tr为迹（Trace）符号，A和X均为矩阵。

A-1-11　设 $f_1(\boldsymbol{x},\boldsymbol{u}),f_2(\boldsymbol{x},\boldsymbol{u}),\cdots,f_n(\boldsymbol{x},\boldsymbol{u})$ 是连续可微函数，

$$\boldsymbol{f}=\begin{bmatrix}f_1 & f_2 & \cdots & f_n\end{bmatrix}^{\mathrm{T}}$$

为函数向量，试写出 $\boldsymbol{f}(\boldsymbol{x},\ \boldsymbol{u})$ 的一阶台劳展开式。

A-1-12　试求内接于椭圆、周长最大的矩形，即在约束

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

条件下，使$S=4(x+y)$取最大值。

A-1-13　试求给定椭圆内长方体的最大体积，即在约束

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

条件下，求$V=8xyz$ 的最大值。

A-1-14　试求函数 $f(x)=4-x_1^2-x_2^2$ 在约束 $g(x)=1-x_1^2+x_2=0$ 下的相对极大值和极小值。

A-1-15　求函数$f(x)=x_1^2+x_2^2$ 在不等式约束

$$(x_1-4)^2+x_2^2\leqslant 4\quad,\quad (x_1-1)^2+x_2^2\leqslant 4$$

条件下，$f(x)$取最小值。

A-1-16　试写出在不等式 $g(x)\leqslant 0$ 约束下，$f(x)$取极大的一阶必要条件，并以 A-1-15 题为例，求 $f(x)$ 的最大值。

A-1-17　试写出在不等式 $g(x)\geqslant 0$ 约束条件下，$f(x)$取极小的一阶必要条件。

A-1-18　应用拉格朗日乘子法求二次型函数

$$J(\boldsymbol{x},\boldsymbol{u})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}$$

在线性方程约束

$$f(\boldsymbol{x},\boldsymbol{u})=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}+\boldsymbol{c}=0$$

条件下，$J(\boldsymbol{x},\boldsymbol{u})$ 的极小点 $\boldsymbol{u}$，并证明满足必要条件的点是极小值点。其中 A 为 $n\times n$ 非异矩阵，$\boldsymbol{B}$ 为 $n\times m$ 矩阵，$\boldsymbol{R}$、$\boldsymbol{Q}$ 分别为 $m\times m$、$n\times n$ 正定对称常数矩阵，$\boldsymbol{u}$、$\boldsymbol{x}$ 和 $\boldsymbol{c}$ 分别为 m 维、n 维向量和 n 维常数向量。

A-1-19　检验下面矩阵 A 是否正半定：

$$A=\begin{pmatrix}1 & 2 & 1\\ 2 & 4 & 2\\ 1 & 2 & 0\end{pmatrix}$$

1.2.2 习题题目

B-1-1　求数量函数 $f=(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})$ 对向量 $\boldsymbol{x}^{\mathrm{T}}$ 的导数。

B-1-2　求数量函数 $f=(\boldsymbol{B}\boldsymbol{z}-\boldsymbol{x})^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{B}\boldsymbol{z}-\boldsymbol{x})$ 对矩阵 $\boldsymbol{R}$ 的导数($\boldsymbol{R}$ 是对称的)。

B-1-3　已知 $\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}$，证明

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x})=2\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{A}\boldsymbol{x}$$

其中，$\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{B}$，但 $\boldsymbol{A}\neq\boldsymbol{A}^{\mathrm{T}}$。

B-1-4　求数量函数 $f=(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})$ 对向量 $\boldsymbol{b}$ 的导数。

B-1-5　已知 A 为 $n\times m$ 变元矩阵，$\boldsymbol{B}$ 为 $m\times n$ 常数矩阵，试证

$$\frac{\partial \mathrm{Tr}\boldsymbol{AB}}{\partial A}=\frac{\partial \mathrm{Tr}\boldsymbol{BA}}{\partial A}=\frac{\partial A^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}}{\partial A}=\frac{\partial \mathrm{Tr}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}}{\partial A}=\boldsymbol{B}^{\mathrm{T}}$$

其中 Tr 为迹（Trace）符号。

B-1-6　先将函数 $f=-3x_1^2+2x_1x_2+2x_1x_3-2x_2^2+4x_2x_3$ 写成 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的形式，然后判断其是否为负定。

B-1-7　试确定 $V(\boldsymbol{x})=-x_1^2+4x_2^2+x_3^2+2x_1x_2-6x_2x_3-2x_1x_3$ 是否为负定。

B-1-8　设多元函数为 $f(\boldsymbol{u})=2u_1^2+5u_2^2+u_3^2+2u_2u_3+2u_3u_1-6u_2+3$，求 $f(\boldsymbol{u})$ 的极值点及极小值。

B-1-9　试求函数 $f=x^2+4y^2$ 在约束 $x/3+y=1$ 条件下的极小值。

B-1-10 求满足约束条件 $g(x,u)=c-xu=0$，使函数

$$f=\frac{1}{2}\left(\frac{x^2}{a^2}+\frac{u^2}{b^2}\right)$$

最小的 f 值。其中 x、u 均为标量，a、b、c 均为正的常数。

B-1-11　检验下列 3×3 矩阵 A 是否为正定：

$$A=\begin{pmatrix}2 & 2 & -1\\ 2 & 6 & 0\\ -1 & 0 & 1\end{pmatrix}$$

1.3 题　解

1.3.1　例题题解

A-1-1

解：　$f=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$，其中

$$A=\begin{pmatrix}1 & 1 & 2\\ 1 & 2 & 1\\ 2 & 1 & 6\end{pmatrix}\quad,\qquad \boldsymbol{x}=\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$$

所以　$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}=\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}}(\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x})=A\boldsymbol{x}+A^{\mathrm{T}}\boldsymbol{x}=2A\boldsymbol{x}=\begin{pmatrix}2 & 2 & 4\\ 2 & 4 & 2\\ 4 & 2 & 12\end{pmatrix}\boldsymbol{x}$$

A-1-2

证明：令 $A=(\boldsymbol{a}_1\quad \boldsymbol{a}_2\quad\cdots\quad \boldsymbol{a}_m)^{\mathrm{T}}$，其中 $\boldsymbol{a}_i$ 为 n 维行向量，则

$$A\boldsymbol{x}=[\boldsymbol{a}_1\quad \boldsymbol{a}_2\quad\cdots\quad \boldsymbol{a}_m]^{\mathrm{T}}\boldsymbol{x}=[\boldsymbol{a}_1\boldsymbol{x}\quad \boldsymbol{a}_2\boldsymbol{x}\quad\cdots\quad \boldsymbol{a}_m\boldsymbol{x}]^{\mathrm{T}}$$

所以

$$\frac{\mathrm{d}(A\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}=\left(\frac{\mathrm{d}\boldsymbol{a}_1\boldsymbol{x}}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}\quad\frac{\mathrm{d}\boldsymbol{a}_2\boldsymbol{x}}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}\quad\cdots\quad\frac{\mathrm{d}\boldsymbol{a}_m\boldsymbol{x}}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}\right)^{\mathrm{T}}$$

对于 $\dfrac{\mathrm{d}(\boldsymbol{a}_i\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}$，有 $\dfrac{\mathrm{d}(\boldsymbol{a}_i\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}=\boldsymbol{a}_i\boldsymbol{I}$，其中 $\boldsymbol{I}$ 为单位矩阵。所以

$$\frac{\mathrm{d}(A\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}=(\boldsymbol{a}_1\boldsymbol{I}\quad \boldsymbol{a}_2\boldsymbol{I}\quad\cdots\quad \boldsymbol{a}_m\boldsymbol{I})^{\mathrm{T}}=A$$

A-1-3

解：$$\frac{\mathrm{d}(\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}=\left[\frac{\mathrm{d}(\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x})^{\mathrm{T}}}{\mathrm{d}\boldsymbol{x}}\right]^{\mathrm{T}}=\left[\frac{\mathrm{d}(\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}\right]^{\mathrm{T}}$$

$$=(A^{\mathrm{T}}\boldsymbol{x}+A\boldsymbol{x})^{\mathrm{T}}=\boldsymbol{x}^{\mathrm{T}}(A+A^{\mathrm{T}})$$