研究生教学用书（英文版）

经济数学

Math in Economics

王苏生 编著

中国人民大学出版社

研究生教学用书（英文版）

经济数学

Math in Economics

王苏生 编著

中国人民大学出版社

·北京·

图书在版编目（CIP）数据

经济数学/王苏生编著.
北京：中国人民大学出版社，2008
研究生教学用书（英文版）
ISBN 978-7-300-09635-3

Ⅰ. 经…
Ⅱ. 王…
Ⅲ. 经济数学—研究生—教材—英文
Ⅳ. F224.0

中国版本图书馆 CIP 数据核字（2008）第 131461 号

研究生教学用书（英文版）
经济数学
王苏生　编著

出版发行	中国人民大学出版社		
社　　址	北京中关村大街 31 号	**邮政编码**	100080
电　　话	010－62511242（总编室）		010－62511398（质管部）
	010－82501766（邮购部）		010－62514148（门市部）
	010－62515195（发行公司）		010－62515275（盗版举报）
网　　址	http://www.crup.com.cn		
	http://www.ttrnet.com（人大教研网）		
经　　销	新华书店		
印　　刷	北京东君印刷有限公司		
规　　格	170 mm×228 mm　16 开本	**版　　次**	2008 年 9 月第 1 版
印　　张	13 插页 1	**印　　次**	2008 年 9 月第 1 次印刷
字　　数	239 000	**定　　价**	23.00 元

内容简介

当今社会，商学院的学生也必须拥有基本的数学知识，特别是优化理论。而这一点对经济、会计、财务、管理和市场学系的学生尤为重要。鉴于商学院研究生的不同学习背景，在他们进入研究生初期，院方一般会为他们提供一门速成数学课程。本书正是这类课程的合适教材，旨在为商学院的大学高年级学生、研究生和博士生提供基本的数学基础。内容涵盖矩阵分析、基本数学概念、一般最优化、动态最优化、常微分方程等。

第一章主要阐述线性代数中的矩阵分析部分，对矩阵理论进行了广泛深入的讨论。第二章讲述数学分析中的基本概念，包括集合、序列、收敛性、连续性、可微性、齐性、反函数和不动点等。第三章叙述一般优化理论，诸如定性矩阵、凹性、拟凹性、无条件和有条件最优化问题、Lagrange 定理、Kuhn-Tucker 定理和包络定理。第四章主要涉及动态规划，集中讨论离散时间随机优化和连续时间非随机优化，介绍了几种常见的动态规划算法，如反向递推法、Lagrange 方法、Hamilton 方法、Bellman 方程、Euler 方程、横截性条件和相图等。第五章讲述一阶常微分方程和线性方程系统、线性方程的标准解法和拉普拉斯变换。最后，第六章叙述常差分方程和线性差分方程，包含迭代法和 z-变换。

市面上已有数本为商学院学生编写的数学教材，本书已囊括这类教材的必要内容，但重点阐述优化理论。作者根据自己多年的教学经验，有意识地简略某些细节，目的在于鼓励教师和学生根据自己的需要，在课堂上详细教授和学习。本书也可以作为已具有基本数学知识人士的参考书。练习和题解将以 PDF 文件格式刊登在如下网页上：www. bm. ust. hk/～sswang/mathbook/，另外有关本书的错漏和补充资料也随时在上述网址更新。

作者简介

王苏生，江苏省苏州人。1985 年毕业于南开大学，获数学硕士学位，并留校任数学系讲师。后留学加拿大多伦多大学并于 1991 年获经济学博士学位。1991—1993 年曾任加拿大康科迪亚大学经济系助理教授。1993 年至今在香港科技大学经济系从事信息经济、组织理论、企业治理领域的教学与研究工作。在国际一流刊物上发表过多篇高质量的学术文章，并著有《微观经济学理论》等著作。曾两次荣获香港科技大学商学院（Franklin）教学奖，现被教育部聘为南京大学长江学者讲座教授。

Preface

The target. This book covers math at the senior undergraduate, Master's and Ph. D levels for students in business schools and economics departments. It concisely covers main math knowledge and tools useful for business and economics studies, including matrix analysis, basic math concepts, general optimization, dynamic optimization, and ordinary differential equations. Basic math tools, particularly optimization tools, are essential for students in a business school, especially for students in economics, accounting, finance, management, and marketing. It is a standard practice nowadays that a graduate program in a business school requires a short and intense course in math just before or immediately after the students enter the program. This book intends to be used in such a course.

The coverage. Chapter 1 focuses on the part of linear algebra that is essential for business studies: matrix analysis. It covers a wide range of subjects on matrices. Chapter 2 covers basic concepts in real analysis, including set, sequence, convergence, continuity, differentiation, homogeneity, inverse functions, and fixed points. Chapter 3 covers general topics in optimization, including definite matrices, concavity, quasi-concavity, unconstrained and constrained optimization, Lagrange theorem, Kuhn-Tucker theorem, and Envelope theorem. Chapter 4 covers dynamic optimization, focusing on discrete-time stochastic optimization and continuous-time deterministic optimization. It covers several popular approaches in dynamic optimization, including backward induction, Lagrange method, Hamilton method, Bellman equation, Euler equation, transversality conditions, and phase diagram. Chapter 5 covers first-order differential equations and linear equation systems. It introduces various ways in solving first-order differential equations, the standard method in solving linear equations, and the Laplace transformation. Finally, Chapter 6 covers first-order difference equations and linear difference equations. It introduces the iterative method and z-transformation.

Special features of this book. There are quite a few math books for business studies. In terms of its contents, this book covers the necessary topics for such a book. The emphasis of this book is optimization, which occupies more than half of the book. This book is intended to be concise with many details deliberately left out. Such a book is good for students who will learn most of the details in class. It is also good for instructors who would want to fill in the details by themselves in their own words. This book can also serve as a reference book for those who have already learned much of the materials.

Supporting materials. Exercises and solutions are available in PDF files at www. bm. ust. hk/~sswang/math-book/. Errors are inevitable and corrections will be posted there. I may also provide additional materials such as new sections and chapters at the site.

Contents

Chapter 1 Linear Algebra

We will focus on the part of linear algebra that is most useful for business studies; specifically, we will focus on matrices. We assume that the readers have some preliminary knowledge of linear algebra.

1.1 Vector

A *scalar* is either a real number $a \in \mathbb{R}$ or a complex number $z \in \mathbb{C}$, where $\mathbb{R}$ is the set of all real numbers and $\mathbb{C}$ is the set of all complex numbers. A *vector* is an ordered sequence of numbers:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{1.1}$$

The numbers x_i are called entries, coefficients or elements. The entries can generally be complex numbers. However, except when eigenvalues are involved, we typically have real entries and denote an n-dimensional vector as $x \in \mathbb{R}^n$. For convenience, we sometimes write it as (x_i) or $(x_i)_n$. In business studies, for example, a vector can represent a financial asset, with its entries representing features of the financial asset.

A special vector is the *zero vector*: $0 \in \mathbb{R}^n$. Given a vector in (1.1), we can define its *transpose* and denote it as x' or x^T, with

$$x' = x^T \equiv (x_1, x_2, \cdots, x_n)$$

We typically write a vector as a vertical/column vector; its horizontal/row version is x'. For a vector $x \in \mathbb{R}^n$, we can define its length as $\| x \|$, where

$$\| x \| \equiv \sqrt{x_1^2 + \cdots + x_n^2}$$

and call it the *norm*. Then, we can define the distance of any two points x and y in $\mathbb{R}^n$ by the norm $\| x - y \|$.

Given two vectors $a = (a_i)_n$ and $b = (b_i)_n$, we can define their *summation* $a + b$, *subtraction* $a - b$, and *multiplication* $a'b$, $\langle a, b \rangle$ or $a \cdot b$, where

$$a + b = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}, \quad a - b = \begin{pmatrix} a_1 - b_1 \\ \vdots \\ a_n - b_n \end{pmatrix}, \quad a \cdot b = \sum_{i=1}^{n} a_i b_i$$

We can also multiply a vector $a \in \mathbb{R}^n$ by a number $\lambda \in \mathbb{R}$:

$$\lambda a \equiv \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix}$$

These operations are intuitively shown in Figure 1.1, where

$$a \cdot b = \| a \| \cdot \| b \| \cdot \cos(\theta)$$

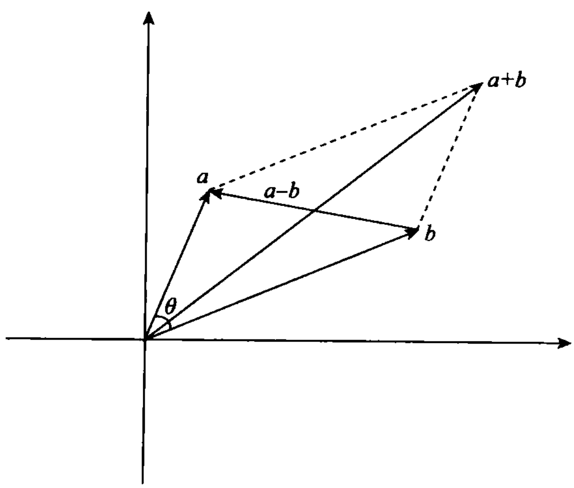

Figure 1.1 Graphic Illustration of Vector Operations

Proposition 1.1 For vectors $a, b, c \in \mathbb{R}^n$, we have

(a) Associative law of summation: $(a+b)+c=a+(b+c)$.

(b) Commutative law of summation: $a+b=b+a$.

(c) Commutative law of multiplication: $a \cdot b=b \cdot a$.

(d) Distributive law: $a \cdot (b+c)=a \cdot b+a \cdot c$.

A vector $\beta \in \mathbb{R}^n$ is a *linear combination* of vectors $\alpha_1, \cdots, \alpha_m \in \mathbb{R}^n$, if there exist $\lambda_1, \cdots, \lambda_m \in \mathbb{R}$ s. t. (such that)

$$\beta=\lambda_1\alpha_1+\cdots+\lambda_m\alpha_m$$

For example, a mutual fund is a linear combination of some basic assets. Define the *span* of a few vectors $\alpha_1, \cdots, \alpha_m$ as

$$span(\alpha_1, \cdots, \alpha_m)=\{\text{all linear combinations of vectors } \alpha_1, \cdots, \alpha_m\}$$

For example, a financial market can be considered as the span of some basic financial assets.

A few vectors $\alpha_1, \alpha_2, \cdots, \alpha_m \in \mathbb{R}^n$ are *linearly dependent* if there exist numbers $\lambda_1, \lambda_2, \cdots, \lambda_m \in \mathbb{R}$, not all 0, such that

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m=0$$

Vectors which are not linearly dependent are *linearly independent*. Hence, vectors $\alpha_1, \alpha_2, \cdots, \alpha_m$ are linearly independent if equation $c_1\alpha_1+c_2\alpha_2+\cdots+c_m\alpha_m=0$ holds only if $c_1=c_2=\cdots=c_m=0$. Linear independence means that any of the vectors cannot be a linear combination of the rest.

Example 1.1 Any two linearly independent vectors in $\mathbb{R}^2$ can span the whole space $\mathbb{R}^2$. This can be shown graphically. In general, $\mathbb{R}^n$ can be spanned by n linearly independent vectors.

1.2 Matrix

A *matrix* is an ordered sequence of column vectors or row vectors, which can be written in the following form:

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \tag{1.2}$$

The numbers a_{ij} are called entries, coefficients or elements. The entries can generally be complex numbers. However, except when eigenvalues are involved, we typically have real entries. We say that matrix A in (1.2) is of *dimension* $m \times n$ and denote $A \in \mathbb{R}^{m\times n}$. For convenience, we sometimes write it as (a_{ij}) or $(a_{ij})_{m\times n}$. In business, a matrix can represent a set of data of a few economic variables.

Example 1.2 The following is a linear equation system for variables $x_1, \cdots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= d_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= d_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= d_m \end{aligned}$$

This linear equation system can be written as $Ax = d$, where A is defined in (1.2) and $d = (d_i)_m$.

A special matrix is the *zero matrix*: $0 \in \mathbb{R}^{m\times n}$, for which all the entries are zero. When $m = n$, we have a *square matrix*. A special square matrix is the *identity matrix*: $I_n \in \mathbb{R}^{n\times n}$, for which all the diagonal entries are 1 and the rest are zero.

Given two matrices A and B, we can define their *summation* $A+B$, *subtraction* $A-B$, and *multiplication* AB. A requirement on the operations is to have matching dimensions as shown in the following:

$$A_{m\times n} + B_{m\times n}, A_{m\times n} - B_{m\times n}, A_{m\times n}B_{n\times k}$$

For $A = (a_{ij})$ and $B = (b_{ij})$, with matching dimensions, the operations are defined by

$$A \pm B = \begin{pmatrix} a_{11} \pm b_{11} & \cdots & a_{1n} \pm b_{1n} \\ \vdots & & \vdots \\ a_{m1} \pm b_{m1} & \cdots & a_{mn} \pm b_{mn} \end{pmatrix}$$

$$AB = \begin{pmatrix} \sum_{t=1}^{n} a_{1t}b_{t1} & \cdots & \sum_{k=1}^{n} a_{1t}b_{tk} \\ \vdots & & \vdots \\ \sum_{t=1}^{n} a_{mt}b_{t1} & \cdots & \sum_{t=1}^{n} a_{mt}b_{tk} \end{pmatrix}$$

We can also multiply a matrix $A = (a_{ij})_{m\times n}$ by a number $\lambda \in \mathbb{R}$:

$$\lambda A \equiv \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix}$$

Vectors can be treated as one-column matrices. By this, the definition of matrix operations applies to vectors too. That is, the operations for one-column matrices are consistent with the operations defined for vectors.

Example 1.3 Given two vectors $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$, derive ab'.

Why do we define such matrix operations? The reason is that these definitions turn out to be very convenient in many applications. Take an example of the multiplication operation. Suppose that a vector $x \in \mathbb{R}^n$ is mapped linearly to another vector $y \in \mathbb{R}^n$ by $y = Ax$, where $A \in \mathbb{R}^{n\times n}$; and suppose that vector y is further mapped to $z \in \mathbb{R}^n$ by $z = By$, where $B \in \mathbb{R}^{n\times n}$. The question is: how can we map x directly to z? The answer is: we can find a matrix $C \in \mathbb{R}^{n\times n}$ such that $z = Cx$ and this matrix C turns out to be $C = BA$, where the multiplication operation is defined above.

Theorem 1.1 Whenever the matrix operations are feasible, we have

(a) Associative law of summation: $(A+B)+C = A+(B+C)$.

(b) Associative law of multiplication: $A(BC) = (AB)C$.

(c) Commutative law of summation: $A+B = B+A$.

(d) Distributive law: $A(B+C) = AB+AC$, $(B+C)A = BA+CA$.

Denote the *determinant* of a square matrix A as $|A|$. Why do we define the determinant this way? The reason is that, for the case of a 3×3 matrix $A = (\alpha_1, \alpha_2, \alpha_3)$, the absolute value of the determinant is the size of an object defined by the column vectors of the matrix (See Figure 1.2).

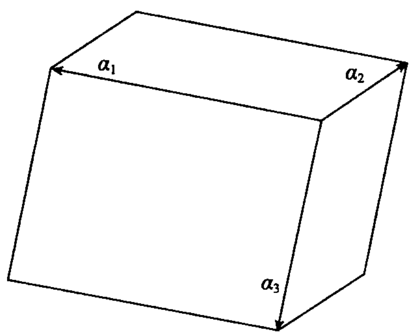

Figure 1.2 The Volume of a Parallelepiped

Theorem 1.2 For any $A, B \in \mathbb{R}^{n\times n}$, we have $|AB| = |A||B|$.

Given a square matrix $A = (a_{ij})_{n\times n}$, we call the following determinant the *minor* of a_{ij}:

$$M_{ij} \equiv \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

and call the following value the *cofactor* of a_{ij}:

$$C_{ij} \equiv (-1)^{i+j} M_{ij}$$

Denote

$$A^* \equiv \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix}^T$$

and call it the *adjoint* of A.

Theorem 1.3 For any square matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, we have

(a) $\sum_{k=1}^{n} a_{ik} C_{ik} = \sum_{k=1}^{n} a_{kj} C_{kj} = |A|$, for any i, j;

(b) $\sum_{k=1}^{n} a_{ik} C_{jk} = \sum_{k=1}^{n} a_{ki} C_{kj} = 0$, for any $i, j, i \neq j$.

The first result in Theorem 1.3 comes directly from the definition of determinant and the second result can be easily derived from the first result.

For $A \in \mathbb{R}^{n\times n}$, if there is another $B \in \mathbb{R}^{n\times n}$ such that $AB = BA = I$, then A is said to be *invertible* or *nonsingular*, and denote B as A^{-1}. A^{-1} is called the *inverse matrix* of A. The inverse is unique.

Theorem 1.4 For $A \in \mathbb{R}^{n\times n}$, A is invertible if and only if $|A| \neq 0$. When A is invertible, the inverse is

$$A^{-1} = \frac{1}{|A|} A^*$$

Corollary 1.1 For a square matrix A,

(a) Its inverse matrix is unique.

(b) If there is a matrix B such that $AB=I$ or $BA=I$, then A is invertible and $B=A^{-1}$.

Example 1.4 Find the inverse of

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

We have $C_{11}=d, C_{12}=-c, C_{21}=-b, C_{22}=a$. Thus,

$$A^{-1}=\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Theorem 1.5 If $A,B\in\mathbb{R}^{n\times n}$ are invertible, then

(a) $(A^{-1})^{-1}=A$;

(b) $(AB)^{-1}=B^{-1}A^{-1}$.

Given matrix

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

denote the *transpose* of A as A' or A^{t}, where

$$A'=\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

That is, $(a_{ij})'_{m\times n}=(a_{ji})_{n\times m}$.

A matrix A is *symmetric* if $A'=A$. A symmetric matrix must be a square matrix.

The following theorem presents the key properties of the transpose operation.

Theorem 1.6 For the transpose operation, we have

(a) $(A')'=A$.

(b) $(A+B)'=A'+B'$.

(c) $(AB)'=B'A'$.

(d) $(cA)' = cA'$, for any $c \in \mathbb{R}$.

(e) $(A')^{-1} = (A^{-1})'$.

For a square matrix $A \in \mathbb{R}^{n\times n}$, denote

$$tr(A) \equiv a_{11} + a_{22} + \cdots + a_{nn}$$

and call it the *trace* of A. The trace applies only to square matrices. The trace is a second key number for a matrix, following the determinant. A third key number is the rank of a matrix, which will be defined later. The following theorem presents the key properties of the trace.

Theorem 1.7

(a) $tr(cA) = c \cdot tr(A)$, for any $c \in \mathbb{R}$.

(b) $tr(A') = tr(A)$.

(c) $tr(A+B) = tr(A) + tr(B)$.

(d) $tr(AB) = tr(BA)$.

(e) $tr(T^{-1}AT) = tr(A)$.

For a matrix $A \in \mathbb{R}^{m\times n}$, if the maximum number of linearly independent row vectors of A is r, then A is said to have *rank* r and is denoted as $rank(A) = r$. A *submatrix* of a given matrix consists of a rectangular array of entries lying in specified subsets of the rows and columns of the given matrix. These rows and columns need not be adjacent.

Theorem 1.8 For any $A \in \mathbb{R}^{n\times m}$,

$$\begin{aligned} rank(A) &= \text{the maximum number of independent row vectors} \\ &= \text{the maximum number of independent column vectors} \\ &= \text{the size of the largest invertible square submatrix of } A \end{aligned}$$

Example 1.5 Consider matrix $A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & -2 & 5 \\ 2 & -1 & 4 \end{pmatrix}$. Since the first two columns are proportional to each other, there are at most two linearly independent columns. Hence, $rank(A) \leqslant 2$. We can easily find a non-singular submatrix:

$$\begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = -2 \neq 0$$

indicating that the first and third columns are linearly independent. Hence,