Series in Contemporary Applied Mathematics
CAM 20

An Introduction to Applied Matrix Analysis

应用矩阵分析导论

金小庆　黄锡荣

高等教育出版社

Series in Contemporary Applied Mathematics CAM 20

An Introduction to Applied Matrix Analysis

应用矩阵分析导论

YINGYONG JUZHEN FENXI DAOLUN

金小庆 黄锡荣

澳门大学

高等教育出版社·北京

Authors

Xiao-Qing JIN
Department of Mathematics
University of Macau

Seak-Weng Vong
Department of Mathematics
University of Macau

图书在版编目（C I P）数据

应用矩阵分析导论 = An Introduction to Applied Matrix Analysis：英文 / 金小庆，黄锡荣著．-- 北京：高等教育出版社，2016.4
ISBN 978-7-04-044994-5

Ⅰ. ①应… Ⅱ. ①金… ②黄… Ⅲ. ①矩阵分析 – 英文 Ⅳ. ① O151.21

中国版本图书馆 CIP 数据核字（2016）第 052656 号

策划编辑 王丽萍　　责任编辑 王丽萍　　封面设计 张 楠　　责任印制 韩 刚

出版发行	高等教育出版社	网　　址	http://www.hep.edu.cn
社　　址	北京市西城区德外大街4号		http://www.hep.com.cn
邮政编码	100120	网上订购	http://www.hepmall.com.cn
印　　刷	涿州市星河印刷有限公司		http://www.hepmall.com
开　　本	787mm × 960mm 1/16		http://www.hepmall.cn
印　　张	9		
字　　数	180 千字	版　　次	2016 年 4 月第 1 版
购书热线	010-58581118	印　　次	2016 年 4 月第 1 次印刷
咨询电话	400-810-0598	定　　价	49.00 元

物 料 号　44994-00

An Introduction to Applied Matrix Analysis

应用矩阵分析导论

Series in Contemporary Applied Mathematics CAM

3 Hanji Shang, Alain Tosseti, eds., *Actuarial Science: Theory and Practice* (2002)
4 Alexandre Ern, Weiping Liu, eds., *Mathematical Problems in Environmental Science and Engineering* (2002)
5 Haïm Brezis, Tatsien Li, eds., *Ginzburg-Landau Vortices* (2005)
6 Tatsien Li, Pingwen Zhang, eds., *Frontiers and Prospects of Contemporary Applied Mathematics* (2006)
7 Deguan Wang, Christian Duquennoi, Alexandre Ern, eds., *Mathematical Methods for Surface and Subsurface Hydrosystems* (2007)
8 Rolf Jeltsch, Tatsien Li, Ian H. Sloan, eds., *Some Topics in Industrial and Applied Mathematics* (2007)
9 Philippe G. Ciarlet, Tatsien Li, eds., *Differential Geometry: Theory and Applications* (2008)
10 Tatsien Li, Pingwen Zhang, eds., *Industrial and Applied Mathematics in China* (2009)
11 Zhien Ma, Yicang Zhou, Jianhong Wu, eds., *Modeling and Dynamics of Infectious Diseases* (2009)
12 Tomas Y. Hou, Chun Liu, Jianguo Liu, eds., *Multi-Scale Phenomena in Complex Fluids: Modeling, Analysis and Numerical Simulation* (2010)
13 Guiqiang Chen, Tatsien Li, Chun Liu, eds., *Nonlinear Conservation Laws, Fluid Systems and Related Topics* (2010)
14 A. Damlamian, S. Jaffard, eds., *Wavelet Methods in Mathematical Analysis and Engineering* (2010)
15 Tatsien Li, Yuejun Peng, Bopeng Rao, eds., *Some Problems on Nonlinear Hyperbolic Equations and Applications* (2010)
16 A. Damlamian, B. Miara, Tatsien Li, eds., *Multiscale Problems: Theory, Numerical Approximation and Applications* (2011)
17 Tatsien Li, Song Jiang, eds., *Hyperbolic Problems: Theory, Numerics and Applications*, Vol. 1 (2012)
18 Tatsien Li, Song Jiang, eds., *Hyperbolic Problems: Theory, Numerics and Applications*, Vol. 2 (2012)
19 Zhaojun Bai, Weiguo Gao, Yangfeng Su, *Matrix Functions and Matrix Equations* (2015)
20 Xiao-Qing Jin, Seak-Weng Vong, *An Introduction to Applied Matrix Analysis* (2016)

To Our Families

Preface

This book was written for a one-semester course in applied matrix theory for senior undergraduate students. It can also be used for a one-semester graduate course. The book is accessible to students who, in various disciplines, have basic knowledge in linear algebra, calculus, and numerical analysis. It is self-contained. Some recent developments in matrix theory are also contained in the book.

At the beginning of the book, we introduce some basic symbols and notations which will be used throughout the book. We study and review several important topics in linear algebra [1; 6; 27; 43], for instance, quadratic forms and symmetric positive definite matrices, complex inner product spaces, Hermitian and unitary matrices, the Kronecker product and sum, etc., which are essential for the development of later chapters.

It is well-known that the essential notions of distance and size in linear vector spaces are captured by norms. We therefore introduce vector and matrix norms in Chapter 2 and study their properties before we develop a perturbation and error analysis. We study effects of perturbation and error on numerical solutions of linear systems. Error analysis on floating point operations is also discussed briefly.

In Chapter 3 we study linear least squares (LS) problems:

$$\min_{\mathbf{y}\in\mathbb{R}^n} \|\mathbf{b} - A\mathbf{y}\|_2$$

where the matrix $A \in \mathbb{R}^{m\times n}$ with $m \geq n$ and the vector $\mathbf{b} \in \mathbb{R}^m$ are given. We introduce some well-known orthogonal transformations and the QR factorization for constructing efficient algorithms for solving LS problems.

We introduce the Moore-Penrose generalized inverse $A^\dagger$ in Chapter 4 and study some basic properties of this inverse [6; 42]. The Moore-Penrose generalized inverse was used in Chapter 3 for the solution of LS problems and will be used in Chapter 6 for constructing the generalized superoptimal preconditioner. *We show that $A^\dagger\mathbf{b}$ is a minimizer to the least squares problem:*

$$\min_{\mathbf{y}\in\mathbb{C}^n} \|\mathbf{b} - A\mathbf{y}\|_2$$

where the matrix $A \in \mathbb{C}^{m\times n}$ and the vector $\mathbf{b} \in \mathbb{C}^m$ are given. Finally, we discuss other generalized inverses related to the Moore-Penrose generalized inverse.

In Chapter 5 we study the conjugate gradient (CG) method. The CG method proposed by Hestenes and Stiefel [25] in 1952 is one of the best known iterative methods for solving any symmetric positive definite linear system $A\mathbf{x} = \mathbf{b}$ [24; 30; 38]. The method is a realization of an orthogonal projection technique onto the Krylov subspace $\mathcal{K}(A, \mathbf{r}_0, k)$ where $\mathbf{r}_0 = \mathbf{b} - A\mathbf{x}_0$ with a given initial vector $\mathbf{x}_0$. Preconditioning technique is also discussed briefly.

In Chapter 6 we introduce two popular preconditioners: the optimal preconditioner $c_U(A)$ proposed by Chan [12] in 1988 and the superoptimal preconditioner $t_U(A)$ proposed by Tyrtyshnikov [39] in 1992, respectively. The optimal preconditioner is studied from an operator viewpoint and a generalized superoptimal preconditioner is constructed by using the Moore-Penrose generalized inverse. A spectral relationship between the optimal preconditioned matrix and the superoptimal preconditioned matrix is also discussed [29].

We propose two optimal preconditioners for different functions of matrices in Chapter 7. More precisely, let f be a function of matrices from $\mathbb{C}^{n\times n}$ to $\mathbb{C}^{n\times n}$. Given $A \in \mathbb{C}^{n\times n}$, there are two choices of constructing optimal preconditioners for $f(A)$: $c_U(f(A))$ and $f(c_U(A))$. Both of them are called optimal preconditioners of $f(A)$ [32]. We study properties of both $c_U(f(A))$ and $f(c_U(A))$ for different functions of matrices, for instance, the matrix exponential, the matrix cosine and matrix sine, and the matrix logarithm, respectively.

In 2005, Böttcher and Wenzel [7] proposed the following conjecture: for any real matrices $X, Y \in \mathbb{R}^{n\times n}$, the following inequality may hold

$$\|XY - YX\|_F \le \sqrt{2}\, \|X\|_F\|Y\|_F$$

where $\|\cdot\|_F$ is the Frobenius norm. In the final chapter, we prove the conjecture in an elementary way and study its related problems.

In writing the book, we have been influenced and helped by many people. In particular, we appreciate some helpful and detailed comments of several people who have taken the time to read the preliminary manuscript that was the precursor of this book: Professor Z. J. Bai of School of Mathematical Sciences, Xiamen University; Professor Z. L. Xu of Department of Mathematics, Shanghai Maritime University; and our former PhD student Dr. Z. Zhao of Department of Mathematics, Hangzhou Dianzi University. Special thanks go to the most important institution in the authors' life: University of Macau for providing a wonderful intellectual atmosphere for writing this book. Last but not least, we should mention that our families have given us much needed encouragement, patience, and endless love which are essential to the completion of the book. Writing of the book is supported by the research grants MYRG098(Y2-L3)-FST13-JXQ and MYRG098(Y3-L3)-FST13-JXQ from University of Macau; the research grant 010/2015/A from FDCT of Macao.

Contents

Preface vii

1. Introduction and Review 1

1.1 Basic symbols 1
1.2 Quadratic forms and positive definite matrices 2
1.2.1 Quadratic forms 2
1.2.2 Problems involving quadratic forms 3
1.2.3 Positive definite matrix 5
1.2.4 Other methods to determine the positive definiteness 5
1.3 Theorems for eigenvalues of symmetric matrices 6
1.4 Complex inner product spaces 10
1.5 Hermitian, unitary, and normal matrices 11
1.6 Kronecker product and Kronecker sum 13

2. Norms and Perturbation Analysis 17

2.1 Vector norms 17
2.2 Matrix norms 18
2.3 Perturbation analysis for linear systems 27
2.4 Error on floating point numbers 32

3. Least Squares Problems 37

3.1 Solution of LS problems 37
3.2 Perturbation analysis for LS problems 41
3.3 Orthogonal transformations 42
3.3.1 Householder reflections 43
3.3.2 Givens rotations 44

3.4 An algorithm based on QR factorization 45
3.4.1 QR factorization 45
3.4.2 A practical algorithm for LS problems 47

4. Generalized Inverses 51

4.1 Moore-Penrose generalized inverse 51
4.2 Basic properties 53
4.3 Relation to LS problems 54
4.4 Other generalized inverses 56

5. Conjugate Gradient Method 61

5.1 Steepest descent method 61
5.1.1 Steepest descent method 62
5.1.2 Convergence rate 63
5.2 Conjugate gradient method 66
5.2.1 Conjugate gradient method 66
5.2.2 Basic properties 67
5.2.3 Practical CG method 70
5.3 Preconditioning technique 72

6. Optimal and Superoptimal Preconditioners 75

6.1 Introduction to optimal preconditioner 75
6.1.1 Circulant matrix 75
6.1.2 Optimal preconditioner 77
6.2 Linear operator c_U 80
6.2.1 Algebraic properties 80
6.2.2 Geometric properties 82
6.3 Stability 85
6.4 Superoptimal preconditioner 87
6.5 Spectral relation of preconditioned matrices 94

7. Optimal Preconditioners for Functions of Matrices 99

7.1 Optimal preconditioners for matrix exponential 99
7.2 Optimal preconditioners for matrix cosine and matrix sine 104
7.3 Optimal preconditioners for matrix logarithm 107

8. Böttcher-Wenzel Conjecture and Related Problems 111

8.1 Introduction to Böttcher-Wenzel conjecture 111

8.2 The proof of Böttcher-Wenzel conjecture 112
8.3 Maximal pairs of the inequality 116
8.4 Other related problems 121
8.4.1 The use of other norms in the inequality 121
8.4.2 The sharpening of the inequality 122
8.4.3 The extension to other products similar to the commutator 123

Bibliography 125

Index 129

Chapter 1

Introduction and Review

We study and review several important topics in linear algebra [1; 6; 27; 43] which are essential for the development of later chapters.

1.1 Basic symbols

We use the following symbols throughout this book.

- Let $\mathbb{N}$ denote the set of natural numbers, $\mathbb{Z}$ denote the set of integers, $\mathbb{R}$ denote the set of real numbers, $\mathbb{C}$ denote the set of complex numbers, and $\mathbf{i} \equiv \sqrt{-1}$.
- Let $\mathbb{R}^n$ denote the linear vector space of real n-vectors and $\mathbb{C}^n$ denote the linear vector space of complex n-vectors. Vectors will almost always be column vectors.
- Let $\mathbb{R}^{m\times n}$ denote the linear vector space of $m \times n$ real matrices and $\mathbb{C}^{m\times n}$ denote the linear vector space of $m \times n$ complex matrices.
- The symbol $\mathbf{0}$ denotes the zero matrix or the zero vector with appropriate size.
- We use the upper case letters such as A, B, C, Δ, Λ, etc. to denote matrices. We use the bold lower case letters such as $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, etc. to denote vectors, and use the lower case letters such as x, y, z, α, β, etc. to denote scalars.
- The symbol a_{ij} denotes the (i, j)th entry in a matrix A.
- For any matrix A, let A^T denote the transpose of A, A^* denote the conjugate transpose of A, and $A^\dagger$ denote the Moore-Penrose generalized inverse of A.
- Let $\mathrm{rank}(A)$ denote the rank of a matrix A.
- Let $\mathrm{tr}(A)$ denote the trace of a square matrix A.
- Let $\det(A)$ denote the determinant of a square matrix A. The

matrix A is a nonsingular matrix if $\det(A) \neq 0$, otherwise A is a singular matrix when $\det(A) = 0$.

- Let $\text{diag}(a_{11}, a_{22}, \ldots, a_{nn})$ denote the $n \times n$ diagonal matrix:

$$\text{diag}(a_{11}, a_{22}, \ldots, a_{nn}) = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}.$$

- The symbol I_n denotes the $n \times n$ identity matrix, for instance,

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

  The symbol $\mathbf{e}_i$ denotes the ith unit vector, i.e., the ith column vector of I_n. Sometimes, we use the symbol I to denote the identity matrix with appropriate size if there is no confusion.
- Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ (or $\mathbb{C}^n$). We use $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ to denote the linear vector space of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$.
- Let $\dim(S)$ denote the dimension of a linear vector space S.

1.2 Quadratic forms and positive definite matrices

In this section we study functions in which the terms are squares of variables or products of two variables. Such functions arise in a variety of applications, including geometry, vibrations of mechanical systems, statistics, and electrical engineering.

1.2.1 *Quadratic forms*

For an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where b, $a_i \in \mathbb{R}$ for $i = 1, 2, \ldots, n$, the expression on the left of this equation is a "linear form", in which all variables occur to the first power. Now, we are concerned with "quadratic forms" which are functions of the form

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2 + (\text{all possible terms of form } 2a_kx_ix_j \text{ for } i < j).$$

For example, a quadratic form in two variables x_1 and x_2 is

$$a_1x_1^2 + a_2x_2^2 + 2a_3x_1x_2 \tag{1.1}$$

and a quadratic form in three variables x_1, x_2, and x_3 is

$$a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + 2a_4x_1x_2 + 2a_5x_1x_3 + 2a_6x_2x_3. \tag{1.2}$$

The terms in a quadratic form that involve products of different variables are called the cross-product terms. Note that (1.1) can be written in matrix form as

$$[x_1,\ x_2] \begin{bmatrix} a_1 & a_3 \\ a_3 & a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{1.3}$$

and (1.2) can be written as

$$[x_1,\ x_2,\ x_3] \begin{bmatrix} a_1 & a_4 & a_5 \\ a_4 & a_2 & a_6 \\ a_5 & a_6 & a_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \tag{1.4}$$

The products in (1.3) and (1.4) are both of the form $\mathbf{x}^T A\mathbf{x}$ where $\mathbf{x}$ is the column vector of variables and A is a symmetric matrix whose diagonal entries are the coefficients of the squared terms and whose entries off the main diagonal are half the coefficients of the cross-product terms. By using the Euclidean inner product, we can write

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T(A\mathbf{x}) = \langle A\mathbf{x}, \mathbf{x}\rangle = \langle \mathbf{x}, A\mathbf{x}\rangle. \tag{1.5}$$

We recall that the Euclidean inner product is defined by

$$\langle \mathbf{x}, \mathbf{y}\rangle \equiv \sum_{i=1}^{n} x_i y_i$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$, $\mathbf{y} = [y_1, y_2, \ldots, y_n]^T \in \mathbb{R}^n$. Then the Euclidean norm of $\mathbf{x}$ is defined by

$$\|\mathbf{x}\| \equiv \langle \mathbf{x}, \mathbf{x}\rangle^{1/2} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}.$$

1.2.2 *Problems involving quadratic forms*

Let $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$. The following are some important mathematical problems relating to quadratic forms.

(i) Find the maximum and minimum values of the quadratic form $\mathbf{x}^T A\mathbf{x}$ if $\mathbf{x}$ is subject to the constraint $\|\mathbf{x}\| = 1$ where $\|\mathbf{x}\|$ is the Euclidean norm of $\mathbf{x}$.

(ii) What conditions must A satisfy in order for a quadratic form to satisfy the inequality $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$?

Theorem 1.1. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix whose eigenvalues in decreasing order are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. If $\mathbf{x}$ is subject to the constraint $\|\mathbf{x}\| = 1$, then:*

(a) $\lambda_1 \geq \mathbf{x}^T A\mathbf{x} \geq \lambda_n$.
(b) $\mathbf{x}^T A\mathbf{x} = \lambda_1$ *if* $\mathbf{x}$ *is an eigenvector of* A *corresponding to* λ_1 *and* $\mathbf{x}^T A\mathbf{x} = \lambda_n$ *if* $\mathbf{x}$ *is an eigenvector of* A *corresponding to* λ_n.

Proof. We only prove (a). Since A is symmetric, there is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A. Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is such a basis where $\mathbf{v}_k$ is an eigenvector corresponding to the eigenvalue λ_k for $k = 1, 2, \ldots, n$. If $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, then for any $\mathbf{x}$ in $\mathbb{R}^n$,

$$\mathbf{x} = \langle \mathbf{x}, \mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{x}, \mathbf{v}_2\rangle \mathbf{v}_2 + \cdots + \langle \mathbf{x}, \mathbf{v}_n\rangle \mathbf{v}_n.$$

Thus,

$$\begin{aligned} A\mathbf{x} =& \langle \mathbf{x}, \mathbf{v}_1\rangle A\mathbf{v}_1 + \langle \mathbf{x}, \mathbf{v}_2\rangle A\mathbf{v}_2 + \cdots + \langle \mathbf{x}, \mathbf{v}_n\rangle A\mathbf{v}_n \\ =& \langle \mathbf{x}, \mathbf{v}_1\rangle \lambda_1\mathbf{v}_1 + \langle \mathbf{x}, \mathbf{v}_2\rangle \lambda_2\mathbf{v}_2 + \cdots + \langle \mathbf{x}, \mathbf{v}_n\rangle \lambda_n\mathbf{v}_n \\ =& \lambda_1\langle \mathbf{x}, \mathbf{v}_1\rangle \mathbf{v}_1 + \lambda_2\langle \mathbf{x}, \mathbf{v}_2\rangle \mathbf{v}_2 + \cdots + \lambda_n\langle \mathbf{x}, \mathbf{v}_n\rangle \mathbf{v}_n. \end{aligned}$$

It follows that the coordinate vectors for $\mathbf{x}$ and $A\mathbf{x}$ relative to the basis S are

$$(\mathbf{x})_S = [\langle \mathbf{x}, \mathbf{v}_1\rangle, \langle \mathbf{x}, \mathbf{v}_2\rangle, \ldots, \langle \mathbf{x}, \mathbf{v}_n\rangle]^T,$$

and

$$(A\mathbf{x})_S = [\lambda_1\langle \mathbf{x}, \mathbf{v}_1\rangle, \lambda_2\langle \mathbf{x}, \mathbf{v}_2\rangle, \ldots, \lambda_n\langle \mathbf{x}, \mathbf{v}_n\rangle]^T.$$

Consequently from the fact that $\|\mathbf{x}\| = 1$, we obtain

$$\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{v}_1\rangle^2 + \langle \mathbf{x}, \mathbf{v}_2\rangle^2 + \cdots + \langle \mathbf{x}, \mathbf{v}_n\rangle^2 = 1$$

and

$$\langle \mathbf{x}, A\mathbf{x}\rangle = \lambda_1\langle \mathbf{x}, \mathbf{v}_1\rangle^2 + \lambda_2\langle \mathbf{x}, \mathbf{v}_2\rangle^2 + \cdots + \lambda_n\langle \mathbf{x}, \mathbf{v}_n\rangle^2.$$

Using these two equations and (1.5), we can prove that $\mathbf{x}^T A\mathbf{x} \leq \lambda_1$ as follows:

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} =& \langle \mathbf{x}, A\mathbf{x}\rangle = \lambda_1\langle \mathbf{x}, \mathbf{v}_1\rangle^2 + \lambda_2\langle \mathbf{x}, \mathbf{v}_2\rangle^2 + \cdots + \lambda_n\langle \mathbf{x}, \mathbf{v}_n\rangle^2 \\ \leq& \lambda_1\langle \mathbf{x}, \mathbf{v}_1\rangle^2 + \lambda_1\langle \mathbf{x}, \mathbf{v}_2\rangle^2 + \cdots + \lambda_1\langle \mathbf{x}, \mathbf{v}_n\rangle^2 \\ =& \lambda_1\Big(\langle \mathbf{x}, \mathbf{v}_1\rangle^2 + \langle \mathbf{x}, \mathbf{v}_2\rangle^2 + \cdots + \langle \mathbf{x}, \mathbf{v}_n\rangle^2\Big) = \lambda_1. \end{aligned}$$

Similarly, we can show that $\mathbf{x}^T A\mathbf{x} \geq \lambda_n$. □

1.2.3 *Positive definite matrix*

We introduce the definition of the symmetric positive definite matrix.

Definition 1.1. A quadratic form $\mathbf{x}^T A\mathbf{x}$ is called positive definite (positive semidefinite) if $\mathbf{x}^T A\mathbf{x} > 0$ ($\mathbf{x}^T A\mathbf{x} \geq 0$) for any $\mathbf{x} \neq \mathbf{0}$, and a symmetric matrix A is called a positive definite (positive semidefinite) matrix if $\mathbf{x}^T A\mathbf{x}$ is a positive definite (positive semidefinite) quadratic form.

We remark that a symmetric matrix A is negative definite if $\mathbf{x}^T A\mathbf{x} < 0$ for any $\mathbf{x} \neq \mathbf{0}$.

Theorem 1.2. *A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is positive definite if and only if all eigenvalues of A are positive.*

Proof. Assume that A is positive definite and let λ be any eigenvalue of A. If $\mathbf{x}$ is an eigenvector of A corresponding to λ, then $\mathbf{x} \neq \mathbf{0}$ and $A\mathbf{x} = \lambda\mathbf{x}$, so

$$0 < \mathbf{x}^T A\mathbf{x} = \mathbf{x}^T \lambda\mathbf{x} = \lambda\mathbf{x}^T\mathbf{x} = \lambda\|\mathbf{x}\|^2$$

where $\|\mathbf{x}\|$ is the Euclidean norm of $\mathbf{x}$. Since $\|\mathbf{x}\|^2 > 0$ it follows that $\lambda > 0$.

Conversely, assume that all eigenvalues of A are positive. We must show that $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$. We can normalize $\mathbf{x}$ to obtain the vector $\mathbf{y} = \mathbf{x}/\|\mathbf{x}\|$ with the property $\|\mathbf{y}\| = 1$ if $\mathbf{x} \neq \mathbf{0}$. It now follows from Theorem 1.1 that

$$\mathbf{y}^T A\mathbf{y} \geq \lambda_n > 0$$

where λ_n is the smallest eigenvalue of A. Thus,

$$\mathbf{y}^T A\mathbf{y} = \left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right)^T A \left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) = \frac{1}{\|\mathbf{x}\|^2}\mathbf{x}^T A\mathbf{x} > 0$$

which implies

$$\mathbf{x}^T A\mathbf{x} > 0.$$

□

1.2.4 *Other methods to determine the positive definiteness*

Our next objective is to give some criteria that can be used to determine whether a symmetric matrix is positive definite without finding its eigenvalues. To do this it will be helpful to introduce some terminology. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$