A GENERAL REVIEW OF SINGULAR VALUE DECOMPOSITION (SVD) AND ITS APPROACHES FOR CONTINUOUS ALGEBRAIC LYAPUNOV EQUATION (CALE)

M.S THESIS

(Branch- Pure Mathematics)

A thesis
submitted to the Department of Mathematics ,
University of Chittagong, Bangladesh, for the partial fulfillment
of the requirement for the degree of
Master of Science in Mathematics

FARHANA AKTER
EXAMINATION ROLL NO : 07203094
REGISTRATION NO : 07203094
SESSION : 2010-2011
YEAR OF EXAMINATION : 2011
HELD IN: 2012

Chittagong
November, 2013

Department of Mathematics
University of Chittagong
Chittagong - 4331
Bangladesh.

DEDICATION

DEDICATED TO
MY BELOVED PARENTS

CONTENTS

ACKNOWLEDGEMENT

At first I would like to express my gratitude to the Almighty Allah for He has provided me strength and patience to complete this thesis paper.

I would like to express my immense indebtedness, cordial thanks and deep appreciation and best regards to my honourable supervisor Professor Dr. Ganesh Chandra Ray, Department of Mathematics, University of Chittagong for his constructive suggestion, wise advice, scholastic guidance, constant inspiration and encouraging me to sustain my strength, overall sparing his valuable time for me to ensure a successful completion of this thesis.

I am also grateful to Professor Dr. Mohammad Ashraful Islam, Chairman, Department of Mathematics, for his keen interest, inspiration and encouragement to carry out my thesis work.

I would like to thank Professor Dr. A. K. Azad former Chairman, Department of Mathematics, for his valuable suggestions and indispensable guidence to complete my paper.I humbly render my gratitude to all of my venerable teachers of my department, Professor Dr. Mosleh Uddin Ahmed, Professor Dr. Nil Ratan Bhattacherjee, Professor Dr. Talky Bhattacharjee,Professor Milon Kanti Dhar, Professor Shafiul Alam Tarafder, Associate Professor Dr.Aman Ullah, Assistant Professor Mohammad Forkan, Department of Mathematics for their inspiration and valuable suggesstions to continue my thesis work. Especially,I humbly render my indebtedness to Assistant Professor Abu Shahed Mohammad Moinuddin , Department of Mathematics, for his generous help instead of my continuous pestering.

Finally,my parents and all of my family members showed their unlimited patience and extended their helping hand for the completion of my thesis. I am really too much indebted to them for having a lots of pain to provide me a study- friendly environment.

Chittagong **AUTHORESS**
November, 2013

ABSTRACT

In engineering theory for stability analysis Continuous Algebraic Lyapunov Equation (CALE) is an important one. In this dissertation we extend the set of Hurwitz matrix with a viable tool Singular Value Decomposition (SVD) to determine the upper solution bounds for CALE which are widely applicable methods of linear algebra, especially in stability theory. In theory for practical purposes one can estimate a stability margin by using some available control bounds. But these upper bounds for CALE are valid under some restrictive condition which are inapplicable. To make them applicable this thesis is an attempt to extend the solutions of upper bounds illustrated with some numerical examples. The following summarizes the contents of this thesis:

Chapter-1 contains basic outlines for vectors, matrix theory in total from linear algebra.

Chapter-2 deals with numerical matrix eigenvalue problems with a discussion on generalized eigenvalue problems.

Chapter-3 describes different matrix factorization specially LU, QR and Singular Value Decomposition in a detail.

Chapter-4 discusses about the computational tool SVD with its wide range of applications.

Finally, in **chapter- 5,** we are trying to discuss about the basic concepts of Lyapunov equations and for practical purposes we extend Hurwitz matrices with SVD to find upper solution bounds for Continuous Algebraic Lyapunov Equations (CALE) which are numerically applicable.

INTRODUCTION

One of the most useful and important tools to emerge from linear algebra is the singular value decomposition (SVD).Of many useful decompositions, singular value decomposition is a factorization of a matrix A into the product $U\Sigma V^T$ of a unitary matrix U ,a diagonal matix Σ, and another unitary matrix V^T -has assumed a special role. There are several reasons –firstly, it is achieved by unitary matrices makes it an ideal vehicle for geometry of $n-$ space. Secondly, it is stable : small perturbation in A correspond to small perturbation in Σ and conversely. Thirdly, the diagonality of Σ makes it easy to determine when A is near to a rank-degenerate matrix. The SVD has a long and fascinating history. Though its existence is accredited to five mathematicians in particular: Eugenio Beltrami (1835-1899), Camille Jordan (1838-1921), James Joseph Sylvester (1814-1897), Erhard Schmidt (1876-1959), and Hermann Weyl (1885-1955).One could argue that it could never have been developed without the contributions of Gauss and Cauchy, both who contributed to the field of Linear Algebra in its naissance. In 1823 Gauss published his famous elimination algorithm for reducing matrices and from there was able to derive the inverse of a matrix. One obtains the inverse by a process called *eliminatio indefinita* (or "general elimination"):

$$y = Ax \rightarrow x = By$$

Soon after, mathematicians such as Cauchy, Jacobi, and Weierstrass played a vital role to set the stage for the singular value decomposition algorithm by contributing to various aspects of linear algebra such as the properties of eigenvalues and eigenvectors, an algorithm for the diagonalization of a symmetric matrix, and the canonical forms for pairs of bilinear functions which is contemporarily known as "the generalized eigenvalue problem". With the collective contributions of these mathematicians singular value decomposition is derived by Beltrami, Jordan, Sylvester, Schmidt, and Weyl, whose contributions are as follows:

1.Beltrami : Eugenio Beltrami is called the progenitor of Singular value decomposition. In 1873 he first derived SVD algorithm. Beltrami's contribution first published in *Journal of Mathematics* for university students which is an attempt to persuade students to study bilinear forms. However, though he was the first to discover singular value decomposition for a real, square and nonsingular matrix having distinct singular values, his derivation did not cover degeneracies.

2.Jordan : Camille Jordan can rightly be called the co discoverer of the singular value decomposition. Although he published his derivation a year after Beltrami, it is clear that his work is independent. In fact, his publication ''Memorie sur les forms bilineaires ''treats three problems, of which the reduction of a bilinear form to a diagonal form by orthogonal substitutions is the simplest.

Jordan's derivation: Jordan starts with the form

$$P = x^T Ay$$

and seeks the maximum and minimum of P subject to

$$\|x\|^2 = \|y\|^2 = 1 \qquad \text{.....(2.1)}$$

The maximum is determined by the equation

$$0 = dP = dx^T Ay + x^T Ady \qquad \text{.....(2.2)}$$

which must be satisfied for all dx and dy that satisfy

$$dx^T x = 0 \text{ and } dy^T y = 0 \qquad \text{.....(2.3)}$$

Jordan then asserts that equation (2.2) will therefore be a combination of equation (2.3) from which we get,

$$Ay = \sigma x \qquad \text{.....(2.4)}$$

and

$$x^T A = \tau y^T \qquad \text{.....(2.5)}$$

From (2.4) it follows that the maximum is

$$x^T(Ay) = \sigma x^T x = \sigma$$

Similarly the maximum is also τ ,so that $\sigma = \tau$

Jordan now observes that σ is determined by the vanishing determinant

$$D = \begin{pmatrix} -\sigma I & A \\ A^T & -\sigma I \end{pmatrix}$$

of the system of (2.4)-(2.5).He shows that his determinant contains only even powers of σ .Now let σ_1 be a root of the equation $D = 0$,and let (2.4) and (2.5) be satisfied by $x = u$ and $y = v$,where $\|u\|^2 = \|v\|^2 = 1$.

Let,

$$\hat{U} = (uU_*) \quad \text{and} \quad \hat{V} = (vV_*)$$

be orthogonal, and let

$$x = \hat{U}\hat{x} \text{ and } \quad y = \hat{V}\hat{y} \quad .$$

With these substitution, let

$$P = \hat{x}^T \hat{A}\hat{y}.$$

In this system, P attains its maximum for $\hat{x} = \hat{y} = e_1$,where $e_1 = (1,0,.....,0)^T$.

Moreover, at the maximum we have

$$\hat{A}\hat{y} = \sigma_1 \hat{x} \quad \text{and} \quad \hat{x}^T \hat{A} = \sigma_1 \hat{y}^T \ ,$$

which implies that

$$\hat{A} = \begin{pmatrix} \sigma & 0 \\ 0 & A_1 \end{pmatrix} .$$

Thus with $\xi_1 = \hat{x}_1$ and $\eta_1 = \hat{y}_1$, P assumes the form

$$\sigma_1 \xi_1 \eta_1 + P_1 \ ,$$

where P_1 is independent of ξ_1 and η_1.Then Jordan applies the reduction inductively to P_1 to arrive at the canonical form

$$P = \xi^T \Sigma \eta \ .$$

Finally, Jordan notes that when the roots of the characteristic equation $D = 0$ are simple, the columns of U and V can be directly from (2.1) , (2.4) and (2.5) .

Results: Jordan proceeds from problem to solution with economy and elegance.His approach of using a partial solution of the problem to reduce it to one of smaller size-deflation, which is the modern term. Avoids the degeneracies that complicate Beltrami's approach. Deflation is now a widely used theoretical and algorithmic tool. Another consequence of Jordan's approach is the variational characterization the largest singular values as the maximum of a function.

3.Sylvester: Mathematician James Joseph Sylvester, who similarly discovered SVD independently, although he did so almost a decade and a half after Beltrami. Sylvester proposed an iterative approach that reduced a quadratic form to a diagonal form. The fact that Sylvester ended up sending a note to *Comptes Rendus* attempting to lay claim for a discovery made by Beltrami over a decade before suggests he had no idea that its existence was already asserted. Part of his algorithm, labeled "Infinitesimal Iteration" offers an alternative way to diagonalize a matrix by making infinitesimal orthogonal substitutions on the off-diagonal components such that any of the components raised to powers higher than one can be ignored . He then asserts that an infinite iteration over these transformations will bring the off-diagonals to zero.

4.Schmidt: Schmidt was another important pioneer of the Singular value decomposition. He approached it from integral equation. He

demonstrated the usefulness of the SVD in deriving an optimal, low-rank approximation of a larger problem during his study of integral equations with unsymmetrical kernels . In his approach, Schmidt also derived the approximation theorem, which is useful when one needs to approximate a matrix with another matrix of a specific rank, the solution of which can be derived by the singular value decomposition of the original matrix. The approximation theorem is often incorrectly labeled as the Eckhart-Young theorem since it was proved by the duo in 1936; however, Schmidt had actually proven the theorem almost thirty years before.

Schmidt's Derivation : Schmidt begins with a kernel $A(s,t)$ that is continuous and symmetric on $[a,b]\times[a,b]$. A continuous, nonvanishing function satisfying

$$\varphi(s)=\lambda\int_a^b A(s,t)\varphi(t)dt$$

is said to be an eigenfunction of A corresponding to eigenvalue λ .Schmidt then establishes the following facts:

(i)The kernel A has at least one eigenfunction.

(ii)The eigenvalues and their eigenfunctions are real.

(iii)Each eigenvalue of A has at most a finite number of linearly independent eigen functions.

(iv)The kernel A has a complete, orthonormal system of eigenfunctions; that is,a sequence $\varphi_1(s),\varphi_2(s),\ldots\ldots\ldots$ of orthonormal eigenfunctions such that every eigenfunction can be expressed as a linear combination of a finite number of the $\varphi_j(s)$.

(v)The eigenvalues satisfy

$$\int_a^b\int_a^b(A(s,t))^2dsdt\geq\sum_i\frac{1}{\lambda_i^2}$$

which implies that the sequence of eigenvalues is unbounded.

Results: Schmidt's two contributions to the singular value decomposition are its generalization to function spaces and his approximation theorem. An important difference in Schmidt's version of the decomposition is the treatment of null vectors of A.

5. Hermann Weyl: In the history of SVD another contributing mathematician was Hermann Weyl, whose contribution to the theorem of the singular value decomposition was to develop a general perturbation theory and use it to give an elegant proof of the approximation theorem.

Weyl's Derivation: The heart of Weyl's development is a lemma concerning the singular values of a perturbed matrix. Specifically, if $B_k = XY^T$, then

$$\sigma_1(A - B_k) \geq \sigma_{k+1}(A) \qquad \text{.....(5.1)}$$

where $\sigma_i(\cdot)$ denotes the ith singular value

Proof: Since Y has k columns, then

$$v = \gamma_1 v_1 + \gamma_2 v_2 + \text{.....} + \gamma_{k+1} v_{k+1}$$

of the first $k+1$ columns of V such that $Y^T v = 0$. We may assume that $\|v\| = 1$, or equivalently that $\gamma_1{}^2 + \gamma_2{}^2 + \text{.....} + \gamma_{k+1}{}^2 = 1$. It follows that

$$\begin{aligned}
\sigma_1^2(A - B) &\geq v^T (A - B)^T (A - B) v \\
&= v^T (A^T A) v \\
&= \gamma_1{}^2 \sigma_1^2 + \gamma_2{}^2 \sigma_2^2 + \text{.....} + \gamma_{k+1}{}^2 \sigma_{k+1}^2 \\
&\geq \sigma_{k+1}^2 .
\end{aligned}$$

Weyl then proves two theorems. The first states that if $A = A' + A''$, then

$$\sigma_{i+j-1} \leq \sigma_i{}' + \sigma_j{}'' \qquad \text{.....(5.2)}$$

where σ' and σ'' are the singular values of A' and A'' .Weyl begins by establishing (5.2) for $i = j = 1$:

$$\sigma_1 = u_1^T A v_1 = u_1^T A' v_1 + u_1^T A'' v_1 \leq \sigma_1' + \sigma_1''$$

To establish the result, let A'_{i-1} and A''_{j-1} .Then $\sigma_1(A' - A'_{i-1}) = \sigma_i(A')$ and $\sigma_1(A'' - A''_{j-1}) = \sigma_j(A'')$ .From (5.1) it follows that

$$\begin{aligned} \sigma_i' + \sigma_j'' &= \sigma_1(A' - A'_{i-1}) + \sigma_1(A'' - A''_{j-1}) \\ &\geq \sigma_1(A - A'_{i-1} - A''_{j-1}) \\ &\geq \sigma_{i+j-1} \end{aligned}$$

which proves the theorem.

The second theorem is a corollary of the first. Set $A' = A - B_k$ and $A'' = B_k$, where B_k has rank k .Since $\sigma_{k+1}(B_k) = 0$,setting $j = k+1$ in (5.2)

$$\sigma_i(A - B_k) \geq \sigma_{k+i}, \qquad i = 1,2,.....$$

As a result we obtain

$$\|A - B_k\|^2 \geq \sigma_{k+1}^2 + ... + \sigma_n^2 \quad .$$

This inequality is equivalent to (5.2) and establishes the approximation theorem.

To know more about SVD and its application we have followed so many books and papers, many of which are cited in this dissertation as references. However, in these books and paper mentioned the calculations are not given in detail .We have carried out most of the calculations in detail.

CHAPTER ONE

A REVIEW OF SOME BASIC CONCEPTS OF LINEAR ALGEBRA

This chapter contains some important definitions, propositions, and theorems from Linear Algebra and Matrix theory. For the sake of consciences, profound discussion of a topic and proofs of the theorems are omitted, since detailed are available in the references listed at the end of the thesis.

1.1 Vectors

1.1.1 Definition

In general, an ordered set of real numbers is called a vector, the numbers themselves are called the components of a vector. A vector v having n components has the form

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

A vector in this form is called column vector and its transpose

$$v^T = (v_1, v_2, ..., v_n)^T$$

is known as row vector.

1.1.2 Vector Space R^n

The set of all n-vectors (that is each vector having n-components) denoted by R^n, is called the vector space R^n. In R^n space all the elements known as scalar.

1.1.3 Operations on Vectors

Consider, two vectors u, v in R^n.Here

$$u = (a_1, a_2, ..., a_n)$$

$$v = (b_1, b_2, ..., b_n)$$

Then their sum,

$$u + v = (a_1 + b_1, a_2 + b_2, ..., a_n + b_n)$$

The scalar product of the vector u with a real number k, written ku, is the vector obtained by multiplying each component of u by k. That is,

$$ku = k(a_1, a_2, ..., a_n)$$

$$= (ka_1, ka_2, ..., ka_n)$$

Here $u + v$ and ku is also a vector in R^n.

1.2 Matrices

1.2.1 Basic concepts

A collection of n vectors in R^n arranged in a rectangular array of m rows and n columns is called a matrix. A matrix A has the form

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

It is denoted by $A=(a_{ij})_{m\times n}$

or simply, $A=(a_{ij})$, where

$i=1,2,.....,m$ and $j=1,2,.....,n$

A is said to be of order $m\times n$. The set of all $m\times n$ matrices denoted by $R^{m\times n}$.

Example:

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

Here A is a matrix of 2×2 matrix .

1.2.2 Definitions

Square Matrix : A matrix A having the same number of rows as columns is called a square matrix.

Example:

$A=\begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix}$ is a square matrix of 2×2 order.

Row Matrix: The matrix with only one row is called a row matrix.

Example: $A=[1 \quad 2 \quad 3]$ is of order 1×3.

Column Matrix: The matrix with only one column is called a column matrix.

Example:

$P = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$ is a column matrix of 3×1 order.

Null Matrix: The matrix with all elements equal to zero is called a null matrix.

Example:

$M = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is a null matrix of 3×3 order.

Rectangular Matrix: Let $A = (a_{ij})$ is a matrix. If $m \neq n$ and $m \neq 0$, $n \neq 0$, then $A = (a_{ij})_{m \times n}$ is called a rectangular matrix of order $m \times n$.

Example:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$$

is a rectangular matrix of order 3×4.

Diagonal Matrix: The matrix A is called a diagonal matrix if it is a square matrix and whose each non-diagonal elements are all zero and the diagonal elements are non-zero.

Example:

$A=\begin{bmatrix}1 & 0 & 0\\ 0 & 4 & 0\\ 0 & 0 & 5\end{bmatrix}$ is a diagonal matrix of order 3×3.

Scalar Matrix: Any diagonal matrix A is called a scalar matrix if and only if every diagonal elements are equal.

Example:

$A=\begin{bmatrix}a & 0 & 0\\ 0 & a & 0\\ 0 & 0 & a\end{bmatrix}$ is a scalar matrix of 3×3 order.

Identity Matrix: The diagonal A is called an identity matrix or unit matrix if and only if all the diagonal elements are equal to 1.

It is denoted by I_n or simply I.

Example:

$A=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ is an Identity matrix of order 2×2.

Upper triangular Matrix: Any square matrix A is said to be an upper triangular matrix if $a_{ij}=0$, for $i>j$.

Example:

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & \cdots & a_{1n}\\ 0 & a_{22} & \cdots & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \cdots & a_{nn}\end{pmatrix}$$

is an upper triangular matrix.

Lower triangular Matrix: Any square matrix A is said to be an lower triangular matrix if $a_{ij} = 0$ for $i < j$.

Example:

$$A=\begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{pmatrix}$$

is a lower triangular matrix.

Tridiagonal Matrix: Any matrix of the type

$$A=\begin{pmatrix} a_{11} & a_{12} & 0 & \cdots & \cdots & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{22} & 0 & \cdots & \cdots & \cdots & a_{2n} \\ 0 & a_{32} & a_{33} & a_{34} & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & a_{(n-1)n} \\ 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & a_{nn} \end{pmatrix}$$

is called a tridiagonal matrix

Example:

$$A=\begin{pmatrix} 1 & 2 & 0 & 0 \\ 3 & 1 & 4 & 0 \\ 0 & 5 & 2 & 7 \\ 0 & 0 & 6 & 4 \end{pmatrix}$$

is a tridiagonal matrix.

Principal diagonal and Trace : The elements a_{ij} $(i=j)$ in a square matrix form the principal diagonal and their sum $a_{11}+a_{22}+a_{33}+.....+a_{nn}$ is called the trace of that square matrix.

Hessen berg Matrix: A square matrix A is upper Hessenberg if $a_{ij}=0$ for $i>\mathrm{j}+1$.The transpose of an upper Hessen berg matrix is a lower Hessen berg matrix if $a_{ij}=0$ for $j>\mathrm{i}+1$.

Here $\begin{pmatrix} * & * & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ \vdots & & & * \\ * & \cdots & \cdots & * \\ * & * & \cdots & * \end{pmatrix}$ $\qquad$ $\begin{pmatrix} * & \cdots & * & * \\ * & \cdots & * & * \\ * & & \vdots & \vdots \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & * & * \end{pmatrix}$

Lower Hessen berg $\qquad$ Upper Hessen berg

Symmetric Matrix : Any square matrix A is said to symmetric matrix if $a_{ij}=a_{ji}\ \forall\ i$ and j.

Example:

$A=\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ is a Symmetric matrix.

Skew Symmetric Matrix : Any square matrix A is said to be skew symmetric matrix if $a_{ij}=-a_{ji}$, $\forall\ i$ and j.

Example:

$A=\begin{pmatrix} a & h & g \\ -h & b & f \\ -g & -f & c \end{pmatrix}$ is a skew symmetric matrix.

Idempotent Matrix: The matrix A is called an Idempotent matrix if it is a square matrix and also if $A^2 = A$.

Example:

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

Now,

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

$$= \begin{pmatrix} 4-2 & 2-1 \\ -4+2 & -2+1 \end{pmatrix}$$

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

$$= A$$

Therefore, A is an Idempotent matrix.

Involutory Matrix: The matrix A is called an Involutory matrix if it is a square matrix and also if $A^2 = I$.

Example:

$$A = \begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix}$$

Now,

$$A^2 = \begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= I$$

Singular Matrix: Any square matrix A is called a singular matrix if $|A| = 0 \ \ (Det\ A = 0)$.

Example:

Let, $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

$$|A| = 1 \times 4 - 2 \times 2 \ = 0$$

$\therefore A$ is a singular matrix.

Non-Singular Matrix: Any square matrix A is said to be non-singular if $|A| \neq 0 \ \ (Det\ A \neq 0)$.

Example:

Let, $A = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}$

$$|A| = 2 \times 3 - 4 \times 1$$

$$= 2 \neq 0$$

$\therefore A$ is a nonsingular matrix.

1.2.3 Operations on matrices

Addition of matrices

If A and B are $m\times n$ matrices the sum of A and B is defined to be the $m\times n$ matrix $A+B$, obtained by adding corresponding entries.

That is

$[A+B]_{ij}=[A]_{ij}+[B]_{ij}$,for each i and j.

For Example: Consider two matrix:

$$A=\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix} \text{and } B=\begin{bmatrix}2 & 3\\ 4 & 2\end{bmatrix}$$

Then,

$$A+B=\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}+\begin{bmatrix}2 & 3\\ 4 & 2\end{bmatrix}$$

$$=\begin{bmatrix}1+2 & 2+3\\ 2+4 & 3+2\end{bmatrix}$$

$$=\begin{bmatrix}3 & 5\\ 6 & 5\end{bmatrix}$$

Here $A+B$ is also a 2×2 matrix.

Properties of matrix addition

For $m\times n$ matrices $A,B,$ and C the following properties are hold:

(i)Closure Property : $A+B$ is again an $m\times n$ matrix

(ii)Associative Property: $(A+B)+C=A+(B+C)$

(iii)Commutative Property: $A + B = B + A$

(iv)Additive Identity: The $m \times n$ matrix 0 consisting of all zeros has the property that $A + 0 = A$.

(v)Additive Inverse : The $m \times n$ matrix $(-A)$ has the property that $A + (-A) = 0$.

Scalar Multiplication

The product of a scalar α times a matrix A, denoted by αA or $A\alpha$ is defined to be the matrix obtained by multiplying each entry of A by α .

For example:

Consider a matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

and a scalar $\alpha = 2$ .Then,

$$\alpha A = 2 \times \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}, \text{ is also a matrix.}$$

Properties of scalar multiplication

For $m \times n$ matrices A and B and for scalars α and β, the following properties hold.

(i) Closure Property: αA is again an $m \times n$ matrix.

(ii)Associative Property: $(\alpha\beta)A = \alpha(\beta A)$

(iii)Distributive Property: (a) $\alpha(A+B)=\alpha A+\alpha B$

Scalar Multiplication is distributed over matrix addition.

(b) $(\alpha+\beta)=\alpha A+\beta A$

Scalar multiplication is distributed over scalar addition.

(iv)Identity Property: $1\times A=A$, Here 1 is an identity element under scalar multiplication.

Transpose of a matrix : The transpose of a matrix A is obtained by interchanging the rows and columns of A and it is denoted by A^t.

Example:

Let, $A=\begin{bmatrix}1 & 2 & 3\\ 4 & 6 & 7\end{bmatrix}$

$$A^t=\begin{bmatrix}1 & 4\\ 2 & 6\\ 3 & 7\end{bmatrix}$$

When $A^t=A$, then A is also called transpose of A.

Determinant of a matrix: For every square matrix A, there a unique scalar is called determinant of A, which is denoted by $\det(A)$ or $|A|$.

Example:

Let, $A=\begin{bmatrix}2 & 4\\ 1 & 3\end{bmatrix}$

$$|A|=2\times 3-4\times 1$$

$$=2$$

The inverse of a matrix: An $n \times n$ matrix A is said to invertible if there exists an $n \times n$ matrix B such that

$$AB = BA = I$$

Properties of inverse of a matrix

For $n \times n$ matrix A, the followings are equivalent:

(i) A is nonsingular.

(ii) $\det(A)$ is nonzero.

(iii) $rank(A) = rank(A^T) = n$

(iv) $N(A) = 0$

(v) A^{-1} exists.

(vi) A has linearly independent rows and columns.

(vii)The eigen values of A are nonzero.

Example:

Let, $A = \begin{pmatrix} -1 & 2 & 3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{pmatrix}$

$\therefore \quad |A| = \begin{pmatrix} -1 & 2 & 3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{pmatrix}$

$= -1(5-0) - 2(10-0) - 3(-4-4)$

$= -1 \neq 0$

i.e A is a non-singular matrix.

Now

$$(A,I) = \begin{pmatrix} -1 & 2 & -3 & \dots & 1 & 0 & 0 \\ 2 & 1 & 0 & \dots & 0 & 1 & 0 \\ 4 & -2 & 5 & \dots & 0 & 0 & 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} -1 & 2 & -3 & \dots & 1 & 0 & 0 \\ 2 & 1 & 0 & \dots & 0 & 1 & 0 \\ 4 & -2 & 5 & \dots & 0 & 0 & 1 \end{pmatrix}; R_2' = R_2 + 2R_1,\ R_3' = R_3 + 4R_1$$

$$\sim \begin{pmatrix} -1 & 2 & -3 & \dots & 1 & 0 & 0 \\ 0 & 5 & -6 & \dots & 2 & 1 & 0 \\ 0 & 6 & -7 & \dots & 4 & 0 & 1 \end{pmatrix}; R_3' = 5R_3 - 6R_2$$

$$\sim \begin{pmatrix} -5 & 0 & -3 & \dots & 1 & -2 & 0 \\ 0 & 5 & -6 & \dots & 2 & 1 & 0 \\ 0 & 0 & 1 & \dots & 8 & -6 & 5 \end{pmatrix}; R_1' = 5R_1 - 2R_2$$

$$\sim \begin{pmatrix} -5 & 0 & 0 & \dots & 25 & -20 & 15 \\ 0 & 5 & 0 & \dots & 50 & 35 & 30 \\ 0 & 0 & 1 & \dots & 8 & -6 & 5 \end{pmatrix}; R_1' = R_1 + 3R_3,\ R_2' = R_2 + 6R_3$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & \dots & -5 & 4 & -3 \\ 0 & 1 & 0 & \dots & 10 & -7 & 6 \\ 0 & 0 & 1 & \dots & 8 & -6 & 5 \end{pmatrix}; R_1' = -\frac{1}{5}R_1,\ R_2' = \frac{1}{5}R_2$$

$$= (I,B)$$

$$\therefore\ B = A^{-1} = \begin{pmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{pmatrix}$$

Test for the existence of the Inverse:

Here, $A=\begin{pmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{pmatrix}$

and $B=\begin{pmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{pmatrix}$

$\therefore \quad AB=\begin{pmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{pmatrix}\begin{pmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{pmatrix}$

$=\begin{pmatrix} 5+20-24 & -4-14+18 & 3+12-15 \\ -10+10+0 & 8-7+0 & -6+6+0 \\ -20-20+40 & 16+14-30 & -12-12+25 \end{pmatrix}$

$=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$=I_3$

Here B is the required inverse of A.

Conjugate of a matrix: Let A be any matrix of complex number. Then the matrix $\overline{A}$ is obtained from A by replacing each element by its conjugate, is called the conjugate of A.

Here $\overline{A}$ is read as conjugate of A.

Example:

Let $A=\begin{pmatrix} a+ib & c-id \\ 2+i & 4i \end{pmatrix}$

$$\overline{A}=\begin{pmatrix} a-ib & c+id \\ 2-i & -4i \end{pmatrix}$$

Adjoint of a square matrix: Let $A=(a_{ij})_{n\times n}$ be any $n\times n$ matrix. The transpose of the matrix $B=(A_{ij})_{n\times n}$. where A_{ij} denotes the co-factor of the elements a_{ij} in the $|A|$, is called the adjoint of the matrix A and is denoted by the symbol $Adj\ A$.

Example:

Let,

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

Then

$$Adj\ A=\begin{pmatrix} A_{11} & A_{12} & \cdots & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & \cdots & A_{nn} \end{pmatrix}^{T}$$

$$= \begin{pmatrix} A_{11} & A_{21} & \cdots & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & \cdots & A_{nn} \end{pmatrix}$$

Range and Null space: For every $m \times n$ matrix A, there are two important associated subspaces: the r**ange** of A, denoted by $R(A)$,and the **null space** of A, denoted by $N(A)$, defined as follows:

R(A)={b $\mid$ b=Ax for some x}

N(A)={x $\mid$ Ax=0}

The dimension of $N(A)$ is called the **nullity** of A and denoted by **null** (A)**.**

Rank of a Matrix: The rank of a matrix is defined as the number of independent rows or which is the same as the number of independent columns it contains. It is denoted by $R(A)$.

Example:

Let,

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

$$R(A) = 2$$

Properties of Rank of a Matrix

Let A be an $m \times n$ matrix. Then

(i) $rank\,(A) = rank\,(A^T)$

(ii) $rank\;(A) + null\,(A) = n$

(iii) $rank\;(AB) \geq rank\,(A) + rank\;(B) - n$, where B is $n \times p$.

(iv) $rank\;(BA) = rank\,(A) = rank\;(AC)$, where B and C are nonsingular matrices of appropriate order.

(v) $rank\;(AB) \leq \min\{rank\,(A), rank\;(B)\}$

(vi) $rank\;(A+B) \leq rank\,(A) + rank\;(B)$

1.2.4 Some Special Matrices

Similar Matrix : Two matrix A and B are called similar if there exists a nonsingular matrix T such that

$$T^{-1}AT = B$$

Unitary Matrix: A complex square matrix U is unitary if $UU^* = U^*U = I$, where $U^* = (\overline{U})^T$

Orthogonal Matrix: A real square matrix O is orthogonal if $OO^T = O^TO = I$. If U is an $n \times k$ matrix such that $U \times U = 1_k$ then U is said to orthogonal.

Companion Matrix: An unreduced upper Hessen berg matrix of the form

$$C = \begin{pmatrix} 0 & 0 & \cdots & \cdots & c_1 \\ 1 & 0 & \cdots & \cdots & c_2 \\ 0 & 1 & \cdots & \cdots & \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & c_n \end{pmatrix}$$

is called an upper companion matrix. The transpose of an upper companion matrix is a lower companion matrix.

Non derogatory Matrix: A matrix A is non derogatory if and only if it is similar to a companion matrix of its characteristic polynomial. That is A is a non derogatory matrix if and only if there exists a nonsingular matrix T such that $T^{-1}AT$ is a companion matrix.

Positive Definite Matrix : A real symmetric matrix A is Positive definite (Positive semi definite) if $x^T Ax > 0$ (≥ 0) for every nonzero x .

Similarly, a complex Hermitian matrix A is Positive definite (Positive semi definite) if $x^* Ax > 0$ (≥ 0) for every nonzero complex vector x.

Positive semi-definite matrices : A matrix is said to be positive semi-definite when it can be obtained as the product of a matrix by its transpose. This implies that a positive semi-definite matrix is always symmetric. So, formally, the matrix A is positive semi-definite if it can be obtained as:

$$A = XX^T$$

for a certain matrix X (containing real numbers). Positive semi definite matrices of special relevance for multivariate analysis

positive semi-definite matrices include correlation matrices, covariance and, cross-product matrices. The important properties of a positive semi-definite matrix is that its eigen values are always positive or null and that its eigenvectors are pair wise orthogonal when their eigen values are different.

Block Matrices : A matrix whose each entry is a matrix is called a block matrix. A block diagonal matrix is a diagonal matrix whose each entry is a matrix. A block triangular matrix is similarly defined.

A block Matrix

$$A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

Projector Matrix: A projector or projection matrix P is a square matrix that satisfies $P^2=P$, such a matrix is also known as idempotent matrix. If P is a projector, $I-P$ is also a projector because

$$(I-P)^2=1-2P+P^2=1-2P+P=1-P:$$

$I-P$ is called complementary projector to P.

Orthogonal Projection : Let S be a subspace of R^n.Then an $m\times n$ matrix P having the properties

(i) $R(P)=S$

(ii) $P^T=P(P\,is\,symmetric)$

(iii) $P^2=P(P\,is\,Idempotent)$

is called the orthogonal projection onto S or, simply, the projection matrix. We denote the orthogonal projection P onto S by P_S.

Stable Matrix : A square matrix A is said to be a stable matrix if every eigen value of matrix A has negative real part. The matrix is called positive stable if every eigen value has positive real part. In control theory an eigenvalue with negative real part is called a stable eigenvalue.

Discrete Stable Matrix: A matrix A having all its eigen values inside the unit circle is called a discrete stable matrix or a convergent matrix or Schur matrix.

Stability Margin : Let $\lambda_1, \lambda_2, \ldots\ldots\ldots, \lambda_n$ be the eigen values of A.Then the distance $\min\{-\mathrm{Re}(\lambda_i) : i = 1, \ldots\ldots\ldots, n\}$ to the imaginary axis is called the stability margin of A.

Hurwitz matrix and the Hurwitz stability criterion

Namely, given a real polynomial

$$p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n$$

the $n \times n$ square matrix

$$H(p) := \begin{pmatrix} a_1 & a_3 & a_5 & a_7 & \cdots & 0 \\ 1 & a_2 & a_4 & a_6 & \cdots & 0 \\ 0 & a_1 & a_3 & a_5 & \cdots & 0 \\ 0 & 1 & a_2 & a_4 & \cdots & 0 \\ 0 & 0 & a_1 & a_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a_n \end{pmatrix}$$

is called **Hurwitz matrix** corresponding to the polynomial p. It was established by Adolf Hurwitz in 1895 that a real polynomial is

stable (that is, all its roots have strictly negative real part) if and only if all the leading principal minors of the matrix $H(p)$: are positive

$$\Delta_1(p) = |a_1| = a_1 \rangle 0$$
$$\Delta_2(p) = \begin{vmatrix} a_1 & a_3 \\ 1 & a_2 \end{vmatrix} = a_2 a_1 - a_0 a_3 \rangle 0$$
$$\Delta_3(p) = \begin{vmatrix} a_1 & a_3 & a_5 \\ 1 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix} = a_3\Delta_2 - a_1(a_1 a_4 - a_0 a_5) \rangle 0$$

and so on. The minors $\Delta_k(p)$ are called the Hurwitz determinants.

Hurwitz stable matrices : In engineering and stability theory, a square matrix A is called stable matrix (or sometimes Hurwitz matrix) if every eigenvalue of A has strictly negative real part, that is,

$$\text{Re}[\lambda_i] \langle 0$$

for each eigenvalue λ_i. A is also called a stability matrix, because then the differential equation

$$\dot{x} = Ax$$

is asymptotically stable, that is, $x(t) \to 0$ as $t \to \infty$ Hurwitz matrix is named after Adolf Hurwitz. If $G(s)$ is a (matrix-valued) transfer function, then G is called Hurwitz if the poles of all elements of G have negative real part. Note that it is not necessary that $G(s)$ for a specific argument S, be a Hurwitz matrix it need not even be square. The connection is that if A is a Hurwitz matrix, then the dynamical system

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t)$$

has a Hurwitz transfer function.

Transfer Function: The transfer function of a linear dynamical system is the ratio of the Laplace transform of its output to the Laplace transform of its input. In systems theory, the Laplace transform is called the "frequency domain" representation of the system.

Consider a canonical dynamical system

$$\dot{x}(t) = Ax(t)$$
$$y(t) = Cx(t)$$

with input $u : R \to R_n$, output $y : R \to R_m$ and state $x : R \to R_p$, and (A, B, C, D) are constant matrices of conformable sizes.

The frequency domain representation is

$$y(s)$$

and thus the transfer function matrix is $D + C(sI - A)$.

In the case of single-input-single-output systems $m = n = 1$, the transfer function is commonly expressed as a rational function of S:

$$H(s)$$

The values z_i are called the zeros of $H(s)$, and the values p_i are called the poles. If any of the poles has positive real part, then the transfer function is termed unstable; if all of the poles have strictly negative real part, it is stable.

1.3 Vector Norms and Matrix Norms

Let X be a vector space. A real valued function $\|.\|: X \to R$ is said to be a norm on X satisfies the following properties:

(i) $\|x\| \geq 0$ and $\|x\| = 0$ iff $x = 0$

(ii) $\|x + y\| \leq \|x\| + \|y\|$,

(iii) $\|\alpha x\| = |\alpha| \|x\|$,

For any $x \in X$ and $y \in X$ and $\alpha \in R$. Let $x \in C^n$ .Then we define the vector p -norm of x as

$$\|x\|_p = \left(\sum_{i=1}^{n} \|x_i\|^p \right)^{\frac{1}{p}}; \quad \text{for } 1 \leq P \leq \infty$$

In particular when $p = 1, 2, \infty$ we have

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

Let $A = (a_{ij}) \in C^{m \times n}$,then the matrix norm induced by a vector p -norm is defined as

$$\|A\|_P = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$$

The matrix norm induced by vector p-norms are sometimes called induced p-norms. In particular the induced matrix 1-norm and 2-norm can be computed by

$$\|A\|_1 = \max_{1 \le j \le n} \|a_j\|_1 \quad ; \; a_j \text{ is the } j^{\text{th}} \text{ column of } A,$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A \times A)}$$

The most important matrix norm which is not induced by a vector norm is Frobenius norm defined by

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}}.$$

CHAPTER TWO

INTRODUCTION TO MATRIX EIGEN VALUE PROBLEMS

The Eigen value problem is a problem of considerable theoretical interest and wide ranging application. Eigen values and eigenvectors were introduced in connection with the convergence of iterative methods for approximating the solution to a linear solution. But now it is also crucial in solving system of differential equation. This chapter is devoted to study of numerical matrix Eigen value problems. These problems are very important practical problems and arise in a variety of application, including engineering, statistics and economics.

2.1 Eigenvalues and Eigenvectors

Let A be an $m \times n$ matrix. Then λ is an eigen value of A if there exists a nonzero vector x such that-

$$Ax = \lambda x$$

or, $(A - \lambda I)x = 0$ ……..(2.1.1)

The vector x is a right eigenvector of A associated with the eigenvalue λ. (λ, x) will be called an eigen pair of A.

The vector y given by

$$y^T A = \lambda y^T$$

is called a left eigenvector of A associated with eigenvalue λ. In generally we mention right eigenvector just an eigenvector.

The homogenous system $(A - \lambda I)\, x = 0$ has a nontrivial solution if and only if

$$\det\left(A - \lambda I\right) = 0$$

Here $\det(A-\lambda I)$ is a polynomial in λ of degree n and is called the characteristics polynomial of A. Thus the n eigen values of A are the n roots of its characteristics polynomial.

The sum of the eigen values of matrix A is called the trace of A. It is denoted by trace(A) or Tr(A)

Mathematical Example

Let

$$A = \begin{pmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{pmatrix}$$

be an 3×3 square matrix set of all eigen vectors x for A is defined as those vectors which, when multiplied by A, result in a simple scaling λ of x. Thus

$$Ax = \lambda x$$

If we restrict ourselves to real eigen values, only effect of the matrix on the eigenvectors will be change their length, and possibly reverse their direction. So multiplying the right hand side by the identity matrix I, we have

$$Ax = (\lambda I)x$$

$$\text{or, } (A-\lambda I)x = 0$$

Now for non-trivial solutions, we require the determinant $\det(A-\lambda I)$ which is called the characteristics polynomial of A to be zero. i.e.

$$\det\left(\begin{pmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) = 0$$

$$\text{or, } \det\begin{pmatrix} 2-\lambda & 2 & -7 \\ 2 & 1-\lambda & 2 \\ 0 & 1 & -3-\lambda \end{pmatrix} = 0$$

$$\text{or, } (2-\lambda)\{(1-\lambda)\times(-3-\lambda)-2\}-2\{2\times(-3-\lambda)\}-7(-2)=0$$

$$\text{or, } (2-\lambda)(\lambda^2-4\lambda-3-2)-2(-6-2\lambda)+14=0$$

$$\therefore \lambda = 1, 3, -4$$

These are the eigenvalues of the matrix of A.

For finding eigenvector $v(x,y,z)$ such that $(A-\lambda I)v=0$

$$\text{or, } \begin{pmatrix} 2-\lambda & 2 & -7 \\ 2 & 1-\lambda & 2 \\ 0 & 1 & -3-\lambda \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

For $\lambda = 1$:

$$\begin{pmatrix} 1 & 2 & -7 \\ 2 & 0 & 2 \\ 0 & 1 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{matrix} x+2y-7z=0 \\ 2x+2z=0 \\ y-4z=0 \end{matrix} \sim \begin{matrix} x+2y-7z=0 \\ x+z=0 \\ y-4z=0 \end{matrix} \sim \begin{matrix} 2y-8z=0 \\ x+z=0 \\ y-4z=0 \end{matrix} \sim \begin{matrix} y-4z=0 \\ x+z=0 \end{matrix}$$

Here is only one free variable. Considering $z=s$,we get $x=-s$, $y=4s$ .Thus we get the corresponding eigenvector $v_1 = \begin{pmatrix} -s \\ 4s \\ s \end{pmatrix}$

If we choose $s=1$, then $z=1$, $y=4, x=-1$: $v_1 = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$

$\therefore$ $v_1 = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$ is the eigen vector for $\lambda = 1$.

For $\lambda = 3$:

$$\begin{pmatrix} -1 & 2 & -7 \\ 2 & -2 & 2 \\ 0 & 1 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{array}{l} -x+2y-7z=0 \\ 2x-2y+2z=0 \\ y-6z=0 \end{array} \sim \begin{array}{l} -x+2y-7z=0 \\ x-y+z=0 \\ y-6z=0 \end{array} \sim \begin{array}{l} y-6z=0 \\ x-y+z=0 \\ y-6z=0 \end{array} \sim \begin{array}{l} y-6z=0 \\ x-y+z=0 \end{array}$$

Here is only one free variable. Considering $z = s$,we get $y = 6s$, $x = 5s$.

Thus we get the corresponding eigenvector $v_2 = \begin{pmatrix} 5s \\ 6s \\ s \end{pmatrix}$.

If we choose $s = 1$, then $z = 1$, $y = 6$ and $x = 5$: $v_2 = \begin{pmatrix} 5 \\ 6 \\ 1 \end{pmatrix}$.

$\therefore\ v_2 = \begin{pmatrix} 5 \\ 6 \\ 1 \end{pmatrix}$ is the eigen vector for $\lambda = 3$.

For $\lambda = -4$:

$$\begin{pmatrix} 6 & 2 & -7 \\ 2 & 5 & 2 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{array}{l} 6x+2y-7z=0 \\ 2x+5y+2z=0 \\ y+z=0 \end{array} \sim \begin{array}{l} 6x+2y-7z=0 \\ 6x+15y+6z=0 \\ y+z=0 \end{array} \sim \begin{array}{l} -13y-13z=0 \\ 6x+15y+6z=0 \\ y+z=0 \end{array} \sim \begin{array}{l} y+z=0 \\ 6x+15y+6z=0 \end{array}$$

Here is only one free variable. Considering $y = 2s$,we get $x = -3s$, $z = -2s$. Thus we get the corresponding eigenvector $v_3 = \begin{pmatrix} -3s \\ 2s \\ -2s \end{pmatrix}$.

If we choose $s = 1$, then we get $y = 2$, $x = -3$ and $z = -2$: $v_3 = \begin{pmatrix} -3 \\ 2 \\ -2 \end{pmatrix}$.

$\therefore\ v_3 = \begin{pmatrix} -3 \\ 2 \\ -2 \end{pmatrix}$ is the eigen vector for $\lambda = -4$.

Now for **repeated eigenvalues** again we assume

$$A = \begin{pmatrix} 7 & 0 & -3 \\ -9 & -2 & 3 \\ 18 & 0 & -8 \end{pmatrix}$$

For eigenvalues of A we solve

$\det(A - \lambda I) = 0$

or, $\begin{vmatrix} 7-\lambda & 0 & -3 \\ -9 & -2-\lambda & 3 \\ 18 & 0 & -8-\lambda \end{vmatrix} = (-2-\lambda)(-1)^4 \begin{vmatrix} 7-\lambda & -3 \\ 18 & -8-\lambda \end{vmatrix}$

$$= -(2+\lambda)[(7-\lambda)(-8-\lambda)+54]$$
$$= -(2+\lambda)(\lambda^2+\lambda-2)$$
$$= -(2+\lambda)^2(\lambda-1)$$

Thus we get, $\lambda_{1,2} = -2$ and $\lambda_3 = 1$.

These are the eigenvalues of the matrix of A . A has two eigenvalues $\lambda_{1,2} = -2$ (repeated) and $\lambda_3 = 1$.We might say $\lambda_2 = -2$ as it is a repeated eigenvalue , since as a root -2 has a multiplicity of two. For finding eigenvector $v(x,y,z)$ such that $(A - \lambda I)v = 0$

$$\begin{pmatrix} 7-\lambda & 0 & -3 \\ -9 & -2-\lambda & 3 \\ 18 & 0 & -8-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

For $\lambda = -2$:

$$\begin{pmatrix} 9 & 0 & -3 \\ -9 & 0 & 3 \\ 18 & 0 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{matrix} 9x & - & 3z = 0 \\ -9x & + & 3z = 0 \\ 18x & - & 6z = 0 \end{matrix} \sim \begin{matrix} 9x & - & 3z = 0 \\ -9x & + & 3z = 0 \\ 18x & - & 6z = 0 \end{matrix} \sim \begin{matrix} 9x & - & 3z = 0 \\ 18x & - & 6z = 0 \end{matrix} \sim 3x = z$$

Here we have found that y can be chosen arbitrarily, and independently of x and z, whereas z cannot be chosen independently. Considering, $x = s + t$,we get, $z = 3s$, $y = t$. Thus we get the corresponding eigenvector $v_1 = \begin{pmatrix} s+t \\ t \\ 3s \end{pmatrix}$.

When $s = 1,\ t = 0,$ we get $x = 1$, $z = 3,\ y = 0$: $v_1 = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$.

Again , when $s = 0$, $t = 1,$ we get $x = 1$, $z = 0,\ y = 1$: $v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

Thus we get two linearly independent eigenvectors.

$v_1 = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ are two linearly independent eigenvectors for $\lambda = -2$.

For $\lambda = 1$:

$$\begin{pmatrix} 6 & 0 & -3 \\ -9 & -3 & 3 \\ 18 & 0 & -9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{matrix} 6x & - & 3z = 0 \\ -9x & -3y+ & 3z = 0 \\ 18x & - & 6z = 0 \end{matrix} \sim \begin{matrix} 6x & - & 3z = 0 \\ -9x & -3y+ & 3z = 0 \end{matrix} \sim \begin{matrix} 2x & = & z \\ -x & = & y \end{matrix}$$

Here is only one free variable. Considering $x = s$,we get $z = 2s$, $y = -s$

Thus we get the corresponding eigenvector $v_3 = \begin{pmatrix} s \\ -s \\ 2s \end{pmatrix}$.

If we choose $s=1$, then we get $x=1, z=2$ and $y=-1 : v_3 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$.

$v_3 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ is the eigen vector for $\lambda = 1$.

Physical example

As the earth rotates, every arrow pointing outward from the center of the earth also rotates, except those arrows which are parallel to the axis of rotation. Consider the transformation of the earth after one hour of rotation: An arrow from the center of the earth to the geographic south pole would be an eigenvector of this transformation, but an arrow from the center of earth to anywhere on the equator would not be an eigenvector. Since the arrow pointing at the poles not stretched by the rotation of the earth, its eigenvalue is1.

2.2 Properties of eigenvalues & eigenvectors

(a) The eigenvalues of a real symmetric matrix are real.

(b) Eigenvectors of a symmetric matrix are orthogonal, but only for distinct eigenvalues.

(c) If A be non-singular matrix, then the eigenvalues of A^{-1} are reciprocals of the eigenvalues of A .

(d) Two matrices A & B are said to be similar if there exists a non-singular matrix P such that $B = P^{-1}AP$ and similar matrices have the same eigenvalues.

(e) If A and B are two square matrices then AB and BA have the same eigenvalues. Also $A^{-1}B$ and $B^{-1}A$ have the same eigenvalues when $|A| \neq 0$.

(f) If eigenvalues of a matrix A are $\lambda_1, \lambda_2, \lambda_3, \ldots\ldots\ldots\ldots, \lambda_n$ then the eigenvalues of the matrix A^k will be $\lambda_1^k, \lambda_2^k, \lambda_3^k, \ldots\ldots\ldots\ldots, \lambda_n^k$

(g) A set of non-zero eigenvectors belonging to distinct eigenvalues are linearly independent.

(h) If X_i is an eigenvector of a given matrix, then scalar multiple of this is also an eigenvector.

(i) The matrix A and A^T (transpose of A) have the same eigenvalues.

(j) The eigenvalues of a triangular matrix are exactly the diagonal elements of the matrix.

(k) A symmetric matrix A is positive definite if and only if all the eigenvalues of A are positive.

(l) Normalization of an eigenvector can be done in two ways. One method is to divide all the elements of a vector by the largest element so that vectors have unity as the largest element. In the second method, each element is divided by the sum of the squares of the elements of the vector in which case vectors have unit length.

(m) The absolute value of a determinant $\det(A)$ is the product of the absolute values of the eigenvalues of matrix A.

(n) $\lambda = 0$ is an eigenvalue of A if A is a singular (non-invertible) matrix.

2.3 Spectrum and spectral radius

The eigenvectors X_i corresponding to λ_i; (i=1, 2, 3, ... , n) for the square matrix A is given by

$$AX_i = \lambda_i X_i$$

the eigenvalues λ_i may be either distinct or repeated. The set of all eigenvalues λ of a matrix A is called the spectrum of A and the largest of $|\lambda_i|$ is called the spectral radius $p(A))$ of a matrix A is defined by

$$p(A) = \max|\lambda_i|$$

where λ is an eigenvalue of A.

Example: Let the square matrix be

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

to compute the eigenvalues of A, consider

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) \\ &= \det\left(\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \\ &= \det\begin{pmatrix} 1-\lambda & 0 \\ 2 & 1-\lambda \end{pmatrix} \\ &= \lambda^2 - 3\lambda + 2 \end{aligned}$$

The eigenvalues of A are the solution of

$$\begin{aligned} p(\lambda) &= 0 \\ \text{or, } \lambda^2 - 3\lambda + 2 &= 0 \\ \text{or, } (\lambda - 2)(\lambda - -1) &= 0 \\ \text{or, } \lambda &= 1,2 \\ \therefore \lambda &= 1,\ 2 \end{aligned}$$

Here the set $\{1,2\}$ of matrix A is called spectrum of A and the spectral radius is $p(\lambda)=\max\{1,2\}=2$

$\therefore p(\lambda)=2$

2.4 Basic Theorems

Theorem-1: Any similarity transformation PAP^{-1} applied to A leaves the eigenvalues of the matrix unchanged.

Proof: Let λ be an eigenvalue of A and X be the associated eigenvector. Then

$AX=\lambda X$ …….. (2.4.1)

Multiplying both sides of equation (2.4.1) by P, we have

$PAX=\lambda PX$ ……. (2.4.2)

Let $Y=PX$. So that $X=P^{-1}Y$.

Substituting this in equation (2.4.2), then we get

$PAP^{-1}Y=\lambda Y$

Thus X is an eigenvalue of PAP^{-1} and Y is the associated eigenvector.

Cay1ey-Hamilton Theorem

Theorem-2: Let $P_n(\lambda)$ be the characteristics polynomial of an arbitrary $n\times n$ square matrix A. Then A satisfies its own characteristics equation and so is a solution of the matrix polynomial equation $P_n(A)=0$

.

Proof: For simplicity, we only prove the theorem for real symmetric matrices, though it is true for every $n\times n$ matrix. If A is a real $n\times n$ symmetric matrix, then $A=PAP^{-1}$. Let the characteristic polynomial of A be

$P_n(\lambda)=(-1)\{\lambda^n+c_1\lambda^{n-1}+c_2\lambda^{n-2}+.....+c_{n-1}\lambda+c_nI\}$

Then replacing λ by A converts $P_n(\lambda)$ to the matrix polynomial

$$P_n(A)=(-1)^n\left\{A^n+c_1A^{n-1}+c_2A^{n-2}+.....+c_{n-1}A+c_nI\right\}$$

But $A=PAP^{-1}$, therefore

$$P_n(A)=(-1)^n\left[P\left\{A^n+c_1A^{n-1}+c_2A^{n-2}+.....+c_{n-1}A+c_nI\right\}P^{-1}\right]$$

The i-th row of the matrix polynomial

$$\left\{A^n+c_1A^{n-1}+c_2A^{n-2}+.....+c_{n-1}A+c_nI\right\}$$

is simply,

$$\lambda_i^n+c_1\lambda_i^{n-1}+c_2\lambda_i^{n-2}+.....+c_{n-1}\lambda_i+c_nI$$

But this $P_n(\lambda_i)$, and it must vanish for $i=1,2,3,....., n$ because λ_i is an eigenvalue of A. Thus

$$\left\{A^n+c_1A^{n-1}+c_2A^{n-2}+.....+c_{n-1}A+c_nI\right\}$$

Showing that

$P_n(A)=0$. Hence the result proved.

Verification of the Cayley-Hamilton theorem

Consider a 3×3 square matrix

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

For this matrix,

$$(A-\lambda I)=\begin{pmatrix}1&2&1\\2&4&2\\1&2&1\end{pmatrix}-\lambda\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

$$=\begin{pmatrix}1-\lambda&2&1\\2&4-\lambda&2\\1&2&1-\lambda\end{pmatrix}$$

Hence

$$\begin{aligned}\det(A-\lambda I) &= (1-\lambda)\{(4-\lambda)\times(1-\lambda)\}-2\{2\times(1-\lambda)-2\}+1\{4-(4-\lambda)\}\\ &= (1-\lambda)(4-5\lambda+\lambda^2-4)-2(2-2\lambda-2)+1(4-4+\lambda)\\ &= (1-\lambda)(-5\lambda+\lambda^2)+4\lambda+\lambda\\ &= \lambda^2-5\lambda-\lambda^3+5\lambda^2+5\lambda\\ &= \lambda^3-6\lambda^2\end{aligned}$$

The characteristic equation for A is,

$$\lambda^3-6\lambda^2=0 \qquad \text{.....(2.5.1)}$$

And the square of A is,

$$A^2=\begin{pmatrix}6 & 12 & 6\\ 12 & 24 & 12\\ 6 & 12 & 6\end{pmatrix}$$

Therefore, $A^3=\begin{pmatrix}36 & 72 & 36\\ 72 & 144 & 72\\ 36 & 72 & 36\end{pmatrix}$

Now substituting A for λ in equation (2.5.1), we obtain

$$\begin{aligned}A^3-6A^2 &= \begin{pmatrix}36 & 72 & 36\\ 72 & 144 & 72\\ 36 & 72 & 36\end{pmatrix}-6\begin{pmatrix}6 & 12 & 6\\ 12 & 24 & 12\\ 6 & 12 & 6\end{pmatrix}\\ &= \begin{pmatrix}36 & 72 & 36\\ 72 & 144 & 72\\ 36 & 72 & 36\end{pmatrix}-\begin{pmatrix}36 & 72 & 36\\ 72 & 144 & 72\\ 36 & 72 & 36\end{pmatrix}\\ &= \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\end{aligned}$$

i.e. every matrix is non-zero of its characteristics polynomial.

2.5 Power Method for finding eigenvalues

The Power method is an iterative approach that can be employed to determine the largest or dominant eigenvalue with the largest magnitude. The largest eigenvalue of a matrix is also called the principal eigenvalue. One of the simplest methods for finding the largest eigenvalue and eigenvector of a matrix is the Power Method, also called the Vector Iteration Method. The method fails if there is no dominant eigenvalue.

2.5.1 Definition of Dominant eigenvalues

The Power method, can be used when

(i) the $n \times n$ matrix of A has n linearly independent eigenvectors

(ii) the eigenvalues can be ordered in magnitude as

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots\ldots\ldots\ldots \geq |\lambda_n| \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.5.1)$$

$\uparrow$

Note the strict inequality

When this ordering can be done, $|\lambda_1|$ is called the dominant eigen value of A . The eigenvectors corresponding to λ_1 are called dominant eigenvectors of A .

Example: Let A be the matrix whose eigenvalues are 2, 5, 0, -7 and -2. Then the dominant eigenvalue of A is given by the following way:

$$|-7| > |5| > |2| \geq |-2| > |0|$$

Therefore, A has a dominant eigenvalue of $\lambda_1 = 7$.

2.5.2 The role of Dominant Eigen values and Eigenvectors in Dynamic Systems

Consider the homogenous discrete time system

$x_{k+1} = Ax_k, \quad k = 0,1,2.....$

Let λ_1 be the dominant eigenvalue of A that is $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq ..... \geq |\lambda_n|$ where $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of A. Suppose A has a set of independent eigenvectors: $v_1, v_2, ..., v_n$ Then the state x_k at any time $k > 0$ is given by

$$x_k = \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + ... + \alpha_n \lambda_n^k v_n$$

where $x_0 = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$. Because $\lambda_1^k > \lambda_i^k \; i = 1,2,3,....., n$, it follows that for large values of k.

$$|\alpha_1 \lambda_1^k| > |\alpha_i \lambda_i^k|, \qquad i = 2,3..... .., n$$

provided that $\alpha_1 \neq 0$. This means that for large values of k the state vector x_k will approach the direction of the vector v_1 corresponding to the dominant eigenvalue λ_1. Furthermore, the rate at which the state vector approaches v_1 is determined by the ratio of the second to the first dominant eigenvalue: $\left[\lambda_2 / \lambda_1\right]$

In the case $\alpha_1 = 0$ the second dominant eigenvalue λ_2 and the corresponding eigenvector assume the role of the first dominant eigenpair. Similar result holds for the continuous time system

$$\dot{x}(t) = Ax(t)$$

2.6 Procedure of Power method

The Power method for approximating eigenvalues is iterative. First we assume that the matrix A has a dominant eigenvalue with corresponding dominant eigenvectors. Then we choose an initial approximation of one of the dominant eigenvectors of A. This initial

approximation must be a nonzero vector in R^n. Finally we form the sequence given by

$$X_1 = AX_0$$

$$X_2 = AX_1 = A(AX_0) = A^2X_0$$

$$X_3 = AX_2 = A(A^2X_0) = A^3X_0$$

$$\vdots$$

$$\vdots$$

$$X_k = AX_{k-1} = A(A^{k-1}X_0) = A^kX_0$$

By properly scaling this above sequence for large powers of k, we will see that we obtain a good approximation of the dominant eigenvector of A.

Example: Finding the largest eigenvalue and the corresponding eigenvector for the matrix

$$A = \begin{pmatrix} 5 & 2 & 4 \\ -3 & 6 & 2 \\ 3 & -3 & 1 \end{pmatrix}$$

starting with

$$X_0 = (1 \quad 0 \quad 0)^T$$

Solution:

$$X_1 = AX_0 = \begin{pmatrix} 5 & 2 & 4 \\ -3 & 6 & 2 \\ 3 & -3 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \\ 3 \end{pmatrix}$$

$$= 5\begin{pmatrix} 1 \\ -0.6 \\ 0.6 \end{pmatrix}$$

$$X_2 = AX_1 = A^2X_0 = \begin{pmatrix} 31 \\ -27 \\ 27 \end{pmatrix} = 31\begin{pmatrix} 1 \\ -0.871 \\ 0.871 \end{pmatrix}$$

$$X_3 = AX_2 = A^3X_0 = \begin{pmatrix} 209 \\ -201 \\ 201 \end{pmatrix} = 209\begin{pmatrix} 1 \\ -0.962 \\ 0.962 \end{pmatrix}$$

$$X_4 = AX_3 = A^4X_0 = \begin{pmatrix} 1447 \\ -1431 \\ 1431 \end{pmatrix} = 1447\begin{pmatrix} 1 \\ -0.977 \\ 0.977 \end{pmatrix}$$

$$X_5 = AX_4 = A^5X_0 = \begin{pmatrix} 10097 \\ -10065 \\ 10065 \end{pmatrix} = 10097\begin{pmatrix} 1 \\ -0.997 \\ 0.997 \end{pmatrix}$$

$$X_6 = AX_5 = A^6X_0 = \begin{pmatrix} 70615 \\ -70551 \\ 70551 \end{pmatrix} = 70615\begin{pmatrix} 1 \\ -0.999 \\ 0.999 \end{pmatrix}$$

$$X_7 = AX_6 = A^7X_0 = \begin{pmatrix} 494177 \\ -494049 \\ 494049 \end{pmatrix} = 494177\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

The ratio of $\frac{X_7}{X_6}$ is

$$\frac{X_7}{X_6} = \frac{494177}{70615} \approx 6.998 \approx 7(\text{approx})$$

Therefore the largest eigenvalue is 7 and the corresponding eigenvector is

$$\begin{pmatrix} 1 & -1 & 1 \end{pmatrix}^T$$

2.7 The importance of largest and smallest eigenvalues

In several applications all we needed to compute were a few largest and smallest eigenvalues and the corresponding eigenvectors. For example in buckling it is the smallest eigen value that is the most important one.

In vibration analysis of structures, a common engineering practice is just to compute the a few smallest eigenvalues (frequencies) and the corresponding eigenvectors (modes), because it has been in practice that the larger eigenvalues and eigenvectors contribute very little to the total response of the system. The same remarks hold in the case of control problems modeled by a system of second-order differential equations arising in the finite element-generated reduced-order model of large flexible space structure (Inman 1989).

In statistical applications, such as in principal component analysis, only the first few largest eigenvalues are computed. There are other applications where only the dominant and subdominant eigenvalues and the corresponding eigenvectors play an important role.

2.8 Diagonalization

A square matrix A is said to be diagonalizable if A is similar to a diagonal matrix, i.e. if $A = PDP^{-1}$ where P is invertible and D is a diagonal matrix.
In general,

$$A\ [v_1, v_2, \ldots\ldots\ v_n] = [v_1, v_2, \ldots\ldots\ v_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

and if $[v_1, v_2, \ldots\ldots v_n]$ is invertible, A equals

$$[v_1, v_2, \ldots\ldots\ v_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} [v_1, v_2, \ldots\ldots\ v_n]^{-1}$$

The Diagonalization Theorem

Therom: An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the

diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

Example: Diagonalize the following matrix,

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

Step 1: Find the eigenvalues of A

$$\det(A-\lambda I)=0$$

$$\text{or, } \begin{pmatrix} 2-\lambda & 0 & 0 \\ 1 & 2-\lambda & 1 \\ -1 & 0 & 1-\lambda \end{pmatrix}=0$$

$$\text{or, } (2-\lambda)^2(1-\lambda)=0$$

Thus, eigenvalues of A are $\lambda=1,2$

Step 2: Find three linearly independent eigenvectors of A:
By solving $(A-\lambda I)X=0$, for each value of λ, we obtain the following:

Basis for $\lambda=1$: $v_1=\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

Basis for $\lambda=2$: $v_2=\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $v_3=\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

Step 3: Construct P from the eigenvectors

$$p=\begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Step 4: Construct D from the corresponding eigenvalues.

$$D=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Step 5: Verifying, AP= PD

$$AP=\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}=\begin{bmatrix} 0 & 0 & -2 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

$$PD=\begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}=\begin{bmatrix} 0 & 0 & -2 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

Therefore , $A = PDP^{-1}$ that is A is diagonalizable.

2.9 Eigen Decomposition

Let P be a matrix of eigenvectors of a given square matrix A and D be a diagonal matrix with the corresponding eigenvalues on the diagonal. Then, P is a square matrix, A can be written as an eigen decomposition

$$A = PDP^{-1}$$

where D is a diagonal matrix. Furthermore, if A is symmetric, then the columns of P are orthogonal vectors.

Mainly, matrix diagonalization is known as eigen decomposition.

2.9.1 Generalized Eigenvalue Problem

Let A and B be two $n \times n$ matrix.Then $Ax = \lambda Bx$, where λ is called an eigenvalue and the vector x is an eigenvector ,is known as generalized eigenvalue problem.

Matrix Pencil

The matrix $A - \lambda B$ is called a matrix pencil. Also known as eigen pencil. It is denoted by (A, B). The pair (A, B) is called regular if $\det(A - \lambda B)$ is not identically zero.Otherwise the pair (A, B) is called singular.

The Eigenvalue of a Regular Pencil

Case-1: Let (A, B) be a regular pencil.If B is nonsingular,then all the eigenvalues of the pair (A, B) are finite and are the same as of AB^{-1} or of BA^{-1}.

Proof : Since B is nonsingular, so we say that B is invertible. Then from the definition of generalized eigenvalue problem we have $Ax = \lambda Bx$ or equivalently $B^{-1}AX = \lambda X$ .Thus λ is an eigenvalue of.Also $AB^{-1} = BB^{-1}AB^{-1}$ so AB^{-1} and BA^{-1} are similar. Therefore, they have the same eigenvalues.

Case-2: If B is singular, then the degree of $p(\lambda) = \det(A - \lambda B)$ is less than n.

The Eigenvalue of a Singular Pencil

If (A, B) is a singular pencil, then, because $\det(A - \lambda B)$ vanishes identically, any number λ can be an eigenvalue of (A, B).

2.9.2 The Symmetric Definite Generalized Eigenvalue Problem

If A and B are real symmetric matrices and at least one them is positive definite. , furthermore, if B is positive definite, then the generalized eigenvalue problem

$$Ax = \lambda Bx$$

is called the symmetric definite generalized eigenvalue problem.

Many practical applications give rise to symmetric definite generalized eigenvalue problem of the form

$Ax = \lambda Bx$ and $ABx = \lambda x$

when B is positive definite we can write

$$ABx = \lambda x$$

in the form

$$\left(A - \lambda B^{-1}\right)Bx = 0$$

where B^{-1} is also positive definite. Then the problem $Ax = \lambda Bx$ and $ABx = \lambda x$ are same.

Many engineering applications give rise to generalized eigenvalue problems. A majority of eigenvalue problem arising in mechanical vibration are generalized eigenvalue problems. For example, the eigenproblems for vibrations of structures such as buildings and bridges are the so-called symmetric definite eigenvalue problems for the mass and stiffness matrices. Most important applications of the symmetric definite generalized eigenvalue problem is to find the solution of second-order differential equations in mechanical vibration .

CHAPTER THREE

MATRIX FACTORIZATION

One of the most fruitful ideas in the theory of matrices is that of a matrix factorization or matrix decomposition or canonical form. The theoretical utility of matrix decomposition has long been appreciated. More recently, they have become the mainstay of numerical linear algebra, where they serve as computational platforms from which a variety of problems can be solved. Three important matrix factorizations: LU, QR and the Singular Value Decomposition(SVD) and their applications are shortly describes in this chapter.

3.1 LU Factorization

In linear algebra LU decomposition factors a matrix as the product of a lower triangular matrix and upper triangular matrix. The LU decomposition can be viewed as the matrix form of Gaussian elimination. This decomposition was introduced by the mathematician Alan Turing. To solve a system of linear equation by LU decomposition we have to know about Gaussian elimination.

3.1.1 Gaussian Elimination

The most general system of n linear equations in n unknown can be written as

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
\vdots \quad\quad \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad \vdots & \\
\vdots \quad\quad \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad \vdots & \\
a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n &= b_n
\end{aligned}
$$

The a_{ij} and the b_i are known constants, and the x_i are the variables. This system can be expressed very compactly in matrix notation as $Ax = b$, where A is the $n \times n$ matrix

$$
\begin{bmatrix}
a_{11} & a_{12} & a_{13} & \cdots & \cdots & a_{1n} \\
a_{21} & a_{22} & a_{23} & \cdots & \cdots & a_{2n} \\
a_{31} & a_{32} & a_{33} & \cdots & \cdots & a_{3n} \\
\vdots & \vdots & \vdots & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & \cdots & \cdots & \vdots \\
a_{n1} & a_{n2} & a_{n3} & \cdots & \cdots & a_{nn}
\end{bmatrix}
$$

and x and b are the n-dimensional column vectors $\begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}^T$ and $\begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_n \end{bmatrix}^T$, respectively. A is called the coefficient matrix, x is the solution vector and b is the right-hand side vector for the system.

We will focus on the solution technique known as Gaussian elimination with back substitution.

Reviewing the Basics

The first step in the solution of a linear system of equations is to gather all the information needed to compute the solution (that is, the coefficients and the right-hand sides) into one structure, known as the augmented matrix for the system. For a system of n equations in n unknowns, the augmented matrix will have $n \times (n+1)$

dimensions. The first n columns are the coefficient matrix, A , for the system. The right-hand side vector, b , from the last column. For the general system of linear equations, the augmented matrix is

$$\left[\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & \cdots & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \cdots & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \cdots & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \vdots & \cdots & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \cdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & \cdots & a_{nn} & b_n \end{array}\right]$$

It is customary to use a vertical line to separate the two portions, coefficient and right-hand side, of the augmented matrix. The objective of Gaussian elimination is to transform the coefficient portion of the augmented matrix into upper triangular form. The transformation of the coefficient portion of the augmented matrix is carried out through the systematic application of three elementary row operations . The three operations, and the notation we will use to signify each, are

Rule-1: Any two rows can be interchanged. The notation $R_i \leftrightarrow R_j$ indicates that row i was interchanged with row j.

Rule-2: Any row can be multiplied by a nonzero constant. The notation $r_i \leftarrow m R_i$ indicates that row i was multiplied by m.

Rule-3: Any multiple of one row can be added to another row. The notation $r_i \leftarrow R_i + m R_j$ indicates that m times row j was added to row i.

To illustrate the Gaussian elimination process, consider the system,

$$\begin{aligned} x_1 &+ x_2 + x_3 + x_4 = 1 \\ x_1 &+ x_2 + 2x_3 + 3x_4 = 2 \\ -x_1 & \quad + 2x_3 + x_4 = 1 \\ 3x_1 &+ 2x_2 - x_3 \quad = 1. \end{aligned}$$

we begin by placing the pivot in the first row, first column of the augmented matrix. In the matrices shown below, the location of the pivot is indicated by angled braces, $\langle\ \rangle$. The pivot serves as a reference location for organizing subsequent calculations. The goal is to replace each element below the pivot, within the pivot column, with a zero. This can be done by performing **Rule-3** on the rows below the pivot row, each time adding an appropriate multiple of the pivot row. The required multiple, m is determined by the formula

$$m = -\frac{element\ \ to\ be\ replaced\ by\ \ zero}{element\ \ in\ \ pivot}$$

For the problem at hand, the multipliers for the second, third, and fourth rows are -1,+1, and -3 respectively. The result of carrying out the corresponding row operations is

$$\left[\begin{array}{cccc|c} \langle 1\rangle & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 3 & 2 \\ -1 & 0 & 2 & 1 & 1 \\ 3 & 2 & -1 & 0 & 1 \end{array}\right] \begin{array}{l} r_2 \leftarrow R_2 - R_1 \\ r_3 \leftarrow R_3 + R_1 \\ r_4 \leftarrow R_4 - 3R_1 \\ \rightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 1 & 3 & 2 & 2 \\ 0 & -1 & -4 & -3 & -2 \end{array}\right]$$

Having completed one elimination pass through the matrix (generating zero in one column), the pivot is moved down one row and to the right one column to set up for the next pass. At this point, the pivot element is zero. This problem can be by passed by locating a row below the pivot row which has a nonzero entry in the pivot column. Provided the original coefficient matrix was nonsingular, it will always be possible to find such a row. The current pivot row and the selected row are then interchanged. Here, we choose to

interchange rows 2 and 3. Adding the new pivot row to the fourth row completes the second elimination pass.

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & \langle 0 \rangle & 1 & 2 & 1 \\ 0 & 1 & 3 & 2 & 2 \\ 0 & -1 & -4 & -3 & -2 \end{array}\right] \begin{array}{c} R_2 \leftrightarrow R_3 \\ \rightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & \langle 1 \rangle & 3 & 2 & 2 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & -1 & -4 & -3 & -2 \end{array}\right]$$

$$\begin{array}{c} r_4 \leftarrow R_4 + R_3 \\ \rightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 2 & 2 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & -1 & -1 & 0 \end{array}\right]$$

For the third, and in this case final, pass through the matrix, the pivot is moved to the third row, third column. As a general rule, the number of elimination passes is always one less than the number of equations. By adding the third row to the fourth row, the transformation of the coefficient portion of the augment matrix to upper triangular form is complete:

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 2 & 2 \\ 0 & 0 & \langle 1 \rangle & 2 & 1 \\ 0 & 0 & -1 & -1 & 0 \end{array}\right] \begin{array}{c} r_4 \leftarrow R_4 + R_3 \\ \rightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 2 & 2 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

To obtain the solution to the system, we are now in position to perform back substitution. The equation corresponding to the bottom row of the transformed augmented matrix contains just one variable and can be solved directly. Here, we find $x_4 = 1$. This value is then substituted into the equation corresponding to the next to last row to give $x_3 + 2(1) = 1$ or $x_3 = -1$. Continuing to work back up the matrix, the values for x_4 and x_3 are substituted into the second equation, yielding $x_2 + 3(-1) + 2(1) = 2$ or $x_2 = 3$. Finally, from the first equation we

find $x_1 + 3 - 1 + 1 = 1$ or $x_1 = -2$. Collecting these four values, the solution vector is

$$X = \begin{bmatrix} -2 & 3 & -1 & 1 \end{bmatrix}^T .$$

3.1.2 General procedure of LU decomposition

Let a matrix A can be factored or decomposed into a unit lower triangular matrix L and a upper triangular matrix U ,thus we will get $LU = A$,which is termed as LU decomposition of A .A typical procedure of LU decomposition is as follows:

(a)To get the matrix U , we have to use row operations until an upper triangular matrix is formed. By using Gaussian elimination we can obtain the upper triangular matrix U .

(b)To get a unit lower matrix L , we have to start with the identity matrix and use the following rules.

- Any row operations that involves adding a multipleof one row to another. As $R_i + kR_j$,put the value $-k$ in the i^{th} row , j^{th} column of the identity matrix.
- Any row operations that involves getting a leading one on the main diagonal. As kR_i, put the value $1/k$ in the position of the identity matrix where the leading one occurs.

An example to find an LU decomposition of the following matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -4 & 6 \\ 3 & -9 & -3 \end{pmatrix}$$

(i)Use Gaussian Elimination to get the upper triangular matrix U .

$$\begin{pmatrix}1&2&3\\2&-4&6\\3&-9&-3\end{pmatrix}\begin{matrix}R_2-2R_1\\R_3-3R_1\\\rightarrow\end{matrix}\begin{pmatrix}1&2&3\\0&-8&0\\0&-15&-12\end{pmatrix}\begin{matrix}\frac{-1}{8}R_2\\\rightarrow\end{matrix}\begin{pmatrix}1&2&3\\0&1&0\\0&-15&-12\end{pmatrix}\begin{matrix}R_3+15R_2\\\rightarrow\end{matrix}\begin{pmatrix}1&2&3\\0&1&0\\0&0&-12\end{pmatrix}$$

$$\begin{matrix}\frac{-1}{12}R_3\\\rightarrow\end{matrix}\begin{pmatrix}1&2&3\\0&1&0\\0&0&1\end{pmatrix}.$$

(ii) Form the lower triangular matrix L by using the rules mentioned above for the row operations involved to get U .

→Start with the identity matrix $\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$.

Row operations:

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{matrix}R_2-2R_1\\R_3-3R_1\\\rightarrow\end{matrix}\begin{pmatrix}1&0&0\\2&1&0\\3&0&1\end{pmatrix}\begin{matrix}\frac{-1}{8}R_1\\\rightarrow\end{matrix}\begin{pmatrix}1&0&0\\2&-8&0\\3&0&1\end{pmatrix}\begin{matrix}R_3+15R_2\\\rightarrow\end{matrix}\begin{pmatrix}1&0&0\\2&-8&0\\3&-15&1\end{pmatrix}$$

$$\begin{matrix}\frac{-1}{6}R_3\\\rightarrow\end{matrix}\begin{pmatrix}1&0&0\\2&-8&0\\3&-15&-12\end{pmatrix}$$

Thus an LU decomposition is given by

$$\begin{pmatrix}1&2&3\\2&-4&6\\3&-9&-3\end{pmatrix}=\begin{pmatrix}1&0&0\\2&-8&0\\3&-15&-12\end{pmatrix}\begin{pmatrix}1&2&3\\0&1&0\\0&0&1\end{pmatrix}$$

Now we show how an LU decomposition can be used to solve a system of linear equations.

Steps to solve a system of using an LU decomposition:

(a) Set up the equation $Ax = b$.

(b) Find an LU decomposition for A . This will yield the equation $LUx = b$.

(c) Let $Ux = y$. Then solve the equation $Ly = b$ for y .

(d) Take the values for y and solve the equation $Ux = y$ for x . This will give the solution to the system $Ax = b$.

Example: Consider system of linear equations:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 5 \\ 2x_1 - 4x_2 + 6x_3 &= 18 \\ 3x_1 - 9x_2 - 3x_3 &= 6 \end{aligned}$$

Set the equations as $Ax = b$:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & -4 & 6 \\ 3 & -9 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 18 \\ 6 \end{pmatrix}$$

This will yield the equation, $(LU)x = b$:

From the above we get, $\begin{pmatrix} 1 & 0 & 0 \\ 2 & -8 & 0 \\ 3 & -15 & -12 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 18 \\ 6 \end{pmatrix}$

and $y = Ux \rightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

Then we solve $Ly = b$ for y : $\begin{pmatrix} 1 & 0 & 0 \\ 2 & -8 & 0 \\ 3 & -15 & -12 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 18 \\ 6 \end{pmatrix}$

Now solving for y gives the following values:

$$\begin{aligned} y_1 &= 5 \\ 2y_1 - 8y_2 &= 18 \\ 3y_1 - 15y_2 - 12y_3 &= 6 \end{aligned} \quad \rightarrow \quad \begin{aligned} y_1 &= 5 \\ y_2 &= -1 \\ y_3 &= 2 \end{aligned}$$

Take the values for y and solve the equation $y = Ux$ for x. This will give the solution to the system $Ax = b$:

$$\begin{pmatrix} 5 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \rightarrow \begin{aligned} x_1 + 2x_2 + 3x_3 &= 5 \\ x_2 &= -1 \\ x_3 &= 2 \end{aligned} \rightarrow \begin{aligned} x_1 &= 1 \\ x_2 &= -1 \\ x_3 &= 2 \end{aligned}$$

Therefore, the solution to the system is $x_1 = 1$, $x_2 = -1$ and $x_3 = 2$.

3.2 QR Decomposition

QR decomposition is one of the most frequently used methods to complete eigen analysis of a non-symmetric matrix, despite the fact that its convergence is not ensured. This is the method that handles these shorts of problems in a unique way .In numerical linear algebra QR algorithm is an eigenvalue algorithm: that is introduced by John G.F. Francis and Vera N.Kublanovskaya in the late 1950's.

3.2.1 General procedure of QR Decomposition

The basic idea is to perform a QR decomposition, writing the matrix as a product of an orthogonal matrix and an upper triangular matrix, multiply the factors in the other order, and iterate. Before this method the matrix A must be reduced to a combination of tri-diagonal and an upper triangular matrix with one more diagonal, which is a Hessen berg matrix of the form

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & \cdots & A_{1n} \\ 0 & A_{22} & \cdots & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & A_{nn} \end{pmatrix}$$

In the first step, the matrix A is transformed into a Hessen berg matrix by using Givens or Householder transformations. In the second step, this Hessen berg matrix is subject to the iterative process called chasing.

Given matrix $A = A^{(0)}$. The QR algorithm constructs the sequence of matrices $\{A^{(i)}\}$ as follows: for $i = 0,1,2,.....$

• factor $A^{(i)}$ into the product $Q^{(i)} R^{(i)}$,where $Q^{(i)}$ is an orthogonal matrix, so $\left[Q^{(i)}\right]^{-1} = \left[Q^{(i)}\right]^{T}$ and $R^{(i)}$ is an upper triangular matrix; and

• let us compute $A^{(i+1)} = Q^{(i)} R^{(i)}$

From the relation $A^{(i)} = Q^{(i)} R^{(i)}$,it follows that $Q^{(i)^T} A^{(i)} = R^{(i)}$,since $Q^{(i)}$ is an orthogonal matrix. The calculation in the second step is then equivalent to $A^{(i+1)} = R^{(i)} Q^{(i)} = Q^{(i)^T} A^{(i)} Q^{(i)}$. Hence, each iteration performs a similarity transformation with an orthogonal matrix, which implies that the eigen values of $A^{(i+1)}$ are identical to those of $A^{(i)}$.

Example: In general,

$$\begin{cases} A^i = Q^i R^i \\ A^{i+1} = R^i Q^i , \end{cases}$$

Putting $i = 1,2,3,.....$ then we have $A = QR$ factorization.

Let , $A = A^0 = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{pmatrix} = QR$

$$= \begin{pmatrix} -0.3333 & -.05788 & -0.7442 \\ -0.6667 & -0.4134 & 0.6202 \\ -0.6667 & 0.7029 & -0.2481 \end{pmatrix} \begin{pmatrix} -3 & -1.3333 & -0.3333 \\ 0 & -2.6874 & 2.3980 \\ 0 & 0 & -0.3721 \end{pmatrix}$$

$$A^{(1)} = RQ = \begin{pmatrix} -3 & -1.3333 & -0.3333 \\ 0 & -2.6874 & 2.3980 \\ 0 & 0 & -0.3721 \end{pmatrix} \begin{pmatrix} -0.3333 & -.05788 & -0.7442 \\ -0.6667 & -0.4134 & 0.6202 \\ -0.6667 & 0.7029 & -0.2481 \end{pmatrix}$$

$$= \begin{pmatrix} 2.1111 & 2.0585 & 1.4884 \\ 0.1929 & 2.7966 & -2.2615 \\ 0.2481 & -0.2615 & -0.0923 \end{pmatrix}$$

Similarly, by computing we get the iterations:

$$A^{(2)} = \begin{pmatrix} 2.4636 & 1.8104 & -1.3865 \\ -0.0310 & 3.0527 & 1.7694 \\ 0.0616 & -0.1047 & -0.5161 \end{pmatrix}$$

$$A^{(3)} = \begin{pmatrix} 2.4056 & 1.8691 & 1.3930 \\ 0.0056 & 2.9892 & -1.9230 \\ 0.0099 & -0.0191 & -0.3948 \end{pmatrix}$$

$$A^{(4)} = \begin{pmatrix} 2.4157 & 1.8579 & -1.3937 \\ -0.0010 & 3.0021 & 1.8930 \\ 0.0017 & -0.0038 & -0.4178 \end{pmatrix}$$

$$A^{(5)} = \begin{pmatrix} 2.4140 & 1.8600 & -1.3933 \\ -0.0002 & 2.9996 & -1.8982 \\ 0.0003 & -0.0007 & -0.4136 \end{pmatrix}$$

$$A^{(6)} = \begin{pmatrix} 2.4143 & 1.8597 & 1.3934 \\ -0.0000 & 3.0001 & -1.8974 \\ 0.0001 & -0.0001 & -0.4143 \end{pmatrix}$$

$$A^{(7)} = \begin{pmatrix} 2.4142 & 1.8597 & 1.3934 \\ -0.0000 & 3.0000 & -1.8974 \\ 0.0000 & 0.0000 & -0.4142 \end{pmatrix}$$

Hence the eigen values are $\lambda_1 = 2.4142,\ \lambda_2 = 3.0000, \lambda_3 = -0.4142$.

3.3 Singular Value Decomposition (SVD)

The theoretical utility of matrix decomposition has long been appreciated. More recently, they have become the mainstay of

numerical linear algebra. One of the most useful and important tools to emerge from linear algebra is the singular value decomposition (SVD). The SVD has a long and fascinating history. Though its existence is accredited to five mathematicians in particular: Eugenio Beltrami (1835-1899), Camille Jordan (1838-1921), James Joseph Sylvester (1814-1897), Erhard Schmidt (1876-1959), and Hermann Weyl (1885-1955).

3.3.1 General procedure of SVD

The singular value decomposition factors a matrix A into the product $U\Sigma V^T$ of a unitary matrix U,a diagonal matix Σ,and another unitary matrix V^T. The SVD algorithm can be described as follows :

(i) Calculate AA^T and A^TA.

(ii) Find the eigenvalues and Σ.

(iii) Find U and V

(iv) Concatenate $U\Sigma V^T$

The singular value decomposition of a matrix A takes the form

$$A = U \Sigma V^T$$

Example:

Let, $A=\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$

$$A^T=\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$$

Then $AA^T=\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}$

$$A^TA=\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}=\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

Then we have to derive the eigen values for both AA^T and A^TA

$$\left|AA^T - \lambda I\right| = 0$$

$$\left|A^T A - \lambda I\right| = 0$$

Thus for AA^T ,we get the eigen values are $\lambda_1 = 8,\ \lambda_2 = 2$.

The singular values $\sigma_1,.....,\sigma_n$ are equivalent to $\sqrt{\lambda_1},.....,\sqrt{\lambda_n}$ respectively such that

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{8}$$

$$\sigma_2 = \sqrt{\lambda_2} = \sqrt{2}$$

Thus, $\Sigma = \begin{pmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{pmatrix}$.

To find U ,whose columns are the unit eigenvectors of AA^T :

$$\left(AA^T - \lambda I\right)x = 0$$

$$\text{or, } \left| \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix} - \lambda_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| x_1 = 0$$

$$\text{or, } \left| \begin{pmatrix} 8-\lambda_1 & 0 \\ 0 & 2-\lambda_1 \end{pmatrix} \right| x_1 = 0$$

$$\text{or, } \left| \begin{pmatrix} 0 & 0 \\ 0 & -6 \end{pmatrix} \right| x_1 = 0$$

$$\therefore\ x_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Again, $\left(AA^T - \lambda I\right)x = 0$

$$\text{or, } \left| \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix} - \lambda_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| x_2 = 0$$

$$\text{or, } \left| \begin{pmatrix} 8-\lambda_2 & 0 \\ 0 & 2-\lambda_2 \end{pmatrix} \right| x_2 = 0$$

or, $\left| \begin{pmatrix} 0 & 0 \\ 0 & -6 \end{pmatrix} \right| x_2 = 0$

$\therefore \; x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Turn x_1 and x_2 into orthonormal unit vectors, such that

$u_1 = \frac{x_1}{\|x_1\|} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ and $u_2 = \frac{x_2}{\|x_2\|} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

This two vectors form U

$$U = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

To find the unit eigenvectors of V we can do the followings:

$$E_i = \ker \begin{pmatrix} A^T A_{11} - \lambda_i & A^T A_{12} \\ A^T A_{22} & A^T A_{22} - \lambda_i \end{pmatrix}$$

We get, $E_1 = \ker \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix} = span \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$E_2 = \ker \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} = span \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

We need orthonormal basis so we have to multiply the vectors by reciprocals of length, we get

$$\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Thus we get, $V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

With all of these values we get the final equation as:

$$A = U \; \Sigma \; V^T$$

$$\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Among so many matrix decompositions, SVD is a reliable and widely used computational technique. There are so many reasons – firstly, it is achieved by unitary matrices makes it an ideal vehicle for geometry of $n-$ space. Secondly, it is stable: small perturbation in A correspond to small perturbation in $\sum$and conversely. Thirdly, the diagonality of $\sum$makes it easy to determine when A is near to a rank-degenerate matrix.

CHAPTER FOUR

SINGULAR VALUE DECOMPOSITION

The SVD has a long and fascinating history. It has become a computationally viable tool for solving a wide variety of problems arising in many practical applications. The crux of using the SVD in these applications is in the fact these applications require knowledge of the rank of matrix, approximations of a matrix using matrices of lower rank, the orthonormal bases for the row and column spaces of a matrix as well as for their orthogonal complements and orthogonal projections onto these subspaces. The SVD is very effective for these computations.

4.1 Basic Concepts

More generally, any $m \times n$ matrix can be factored as

$$A = U \sum V^T$$

where U is an $m \times m$ orthogonal matrix whose columns are the eigenvectors of AA^T. V is an $n \times n$ orthogonal matrix whose columns are the eigenvectors of $A^T A$ and $\sum$ is an $m \times n$ diagonal matrix of the form

$$\sum = \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & 0 & \\ & & \sigma_r & & \\ & 0 & & \ddots & \\ & & & & 0 \end{pmatrix}$$

with $\sigma_1 \geq \sigma_2 \geq \ldots\ldots\ldots\ldots\ldots \geq \sigma_r > 0$ and r = rank (A).

In the above $\sigma_1, \sigma_2, \sigma_3, \ldots\ldots\ldots\ldots\ldots, \sigma_r$ are the square roots of the eigenvalues of $A^T A$. They are called singular values of A. This

decomposition is called the singular value decomposition or SVD of A .The SVD is closely related eigen decomposition of AA^T .

Reduced SVD

The equations with right singular values $\{v_j\}$ and left singular values $\{u_j\}$ can be written as

$$Av_j = \sigma_j u_j, \quad 1 \le j \le n \qquad \text{……..(4.1.1)}$$

The collection of vector equation can be expressed as a matrix equation

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix}$$

or more compactly $AV = \hat{U}\hat{\Sigma}$ .In this matrix equation $\hat{\Sigma}$ is an $n \times n$ diagonal matrix with positive real entries, $\hat{U}$ is an $m \times n$ matrix with orthonormal columns and V is an $n \times n$ matrix with orthonormal columns. Thus V is unitary we can multiply on the right by V^T to obtain

$$A = \hat{U}\hat{\Sigma}V^T \qquad \text{……(4.1.2)}$$

This factorization of A is called a reduced singular value decomposition or reduced svd of A.

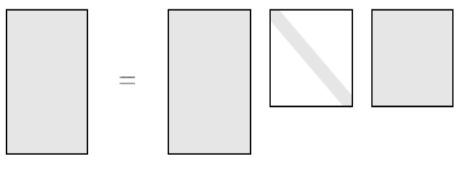

Fig-Reduced SVD $(m \ge n)$

Full SVD

To get full svd the columns of U are n orthonormal vectors in m-dimensional space C^m,unless $m=n$ they do not form a basis of C^m, nor $\hat{U}$ a unitary matrix. By adjoining an additional $m-n$ orthonormal columns $\hat{U}$ can be extended to a unitary matrix, call the result U.For the product to remain unaltered, the last m-n columns of U should be multiplied by zero. Let Σ be the $m\times n$ matrix consisting of $\hat{\Sigma}$ in the upper $n\times n$ block together with $m-n$ rows of zeros below.

Now we have a new factorization ,the full SVD of A

$$A=U\Sigma V^T \qquad (m\geq n) \qquad \text{.....(4.1.3)}$$

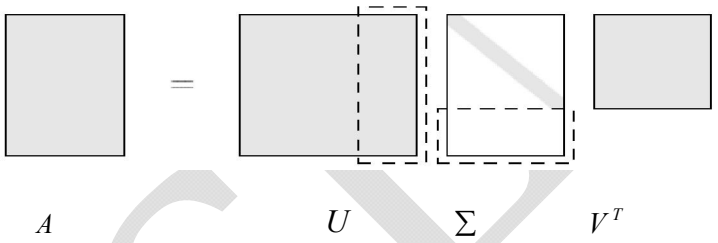

Fig: Full SVD

4.2 A geometric interpretation of the SVD

One way to understand how A deforms space is to consider its action on the unit sphere in R^n: An arbitrary element x of this unit sphere can be represented by $x=x_1v_1+x_2v_2+.....+x_nv_n$ with $\sum_1^n x_i^2=1$ .The image is $Ax=\sigma_1x_1u_1+.....+\sigma_kx_ku_k$. Letting, $y_i=x_i\sigma_i$, we see the image of the unit sphere consists of the vectors $y_1u_1+y_2u_2+.....+y_ku_k$;
where

$$\frac{y_1^2}{\sigma_1^2}+\frac{y_2^2}{\sigma_2^2}+.....+\frac{y_k^2}{\sigma_k^2}=\sum_1^k x_i^2$$

$$\leq 1$$

If A has full column rank, so that $k = n$; the inequality is actually a strict equality. Otherwise, some of the x_i are missing on the right, and the sum can be anything from 0 to 1. This shows that A maps the unit sphere of R^n to a $k-$dimensional ellipsoid with semi-axes in the directions u_i and with the magnitudes σ_i. If $k = n$ the image is just the surface of the ellipsoid, otherwise it is the solid ellipsoid.

In summary, we can visualize the effect A as follows: it first collapses $n-k$ dimensions of the domain, then distorts the remaining dimensions, stretching and squeezing the unit $k-$sphere into an ellipsoid and finally embeds the ellipsoid in R^m: This is illustrated for $n = m = 3$ and $k = 2$ in following Fig:

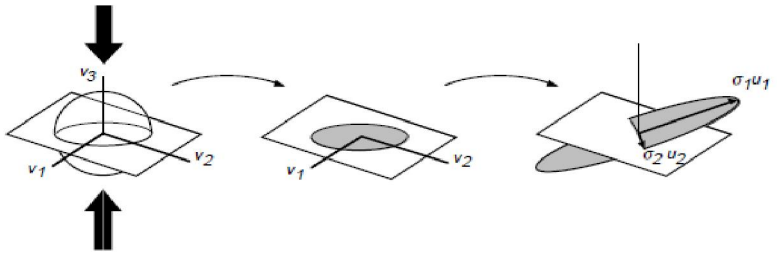

Figure :How A deforms $\mathbf{R^n}$

As an immediate consequence, we see that $\|A\|$ the operator norm of A, defined as the maximum value of $|Av|$ for v on the unit sphere, is simply σ_1, the largest singular value of A. Put another way, we have the inequality $|Ax| \le \sigma_1 |x|$ for all $x \in R^n$, with equality only when x is a multiple of v_1.

4.3 The Singular Value Decomposition Theorem

Statement: Let A be a real $m \times n$ matrix. Then there exist orthogonal matrices U and V such that

$$U^T AV = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} = \Sigma$$

where Σ_1 is a nonsingular diagonal matrix. The diagonal entries of Σ are all nonnegative and can be arranged in non increasing order. The number of nonzero diagonal entries of Σ equals the rank of A.

Proof: Consider the matrix $A^T A$.It is an $n \times n$ symmetric positive semi definite matrix; therefore its eigen values are nonnegative. Denote the eigen values of $A^T A$ by $\lambda_1 = \sigma_1^2, \lambda_2 = \sigma_2^2, \ldots\ldots\ldots\ldots\ldots, \lambda_n = \sigma_n^2$. Assume that these have been ordered such that $\sigma_1 \geq \sigma_2 \geq \ldots\ldots\ldots\ldots\ldots \geq \sigma_r > 0$ and

$$\sigma_{r+1} = \sigma_{r+2} = \ldots\ldots\ldots = \sigma_n = 0$$

We know that a symmetric matrix has a set of orthonormal eigenvectors. Denote the set of orthonormal eigenvectors of $A^T A$ corresponding to λ_1 through λ_n by $v_1, v_2, \ldots\ldots v_n$; that is v_1 through v_n are orthonormal and satisfy

$$A^T A v_i = \sigma_i^2 v_i \qquad i = 1,2,\ldots\ldots, n$$

Then,

$$v_i^T A^T A v_i = \sigma_i^2 \neq 0, \qquad i = 1,2,\ldots\ldots, r \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.3.2)$$

and

$$v_j^T A^T A v_j = 0, \qquad i = 1,2,\ldots\ldots, r, j \neq i \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.3.3)$$

Write

$$V_1 = (v_1, v_2, \ldots\ldots\ldots, v_r)$$

$$V_2 = (v_{r+1}, v_{r+2}, \ldots\ldots\ldots, v_n)$$

where v_1 through v_r are the eigenvectors associated with the nonzero eigenvalues λ_1 through λ_r, and $v_{r+1}, v_{r+2}, \ldots\ldots\ldots, v_n$ correspond to the zero eigen values. Then

$$V_2^T A^T A V_2 = V_2^T A^T A (v_{r+1}, v_{r+2}, \ldots\ldots\ldots, v_n) = V_2^T (0,0,\ldots\ldots\ldots\ldots\ldots, 0) = 0$$

This implies that $AV_2 = 0$, or

$$AV_k = 0, \qquad k = r+1, r+2, \ldots . n \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.3.4)$$

Define now a set of nonzero vectors $\{u_i\}$ by

$$u_i = \frac{1}{\sigma_i} A v_i \qquad i = 1,2,\ldots\ldots, r \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.3.5)$$

The u_i's, $i = 1,2,\ldots\ldots, r$, then form an orthonormal set, because

$$u_i^T u_j = \frac{1}{\sigma_i}(\Sigma v_i)^T \ \frac{1}{\sigma_j}(\Sigma v_j)$$

$$= \frac{1}{\sigma_i \sigma_j}\left(v_i^T A^T A v_j\right) \qquad \ldots\ldots\ldots\ldots\ldots\ldots (4.3.6)$$

$$= \begin{cases} 0 & when\ i \neq j \\ 1 & when\ i = j \end{cases}$$

Define $U_1 = (u_1, u_2, \ldots\ldots\ldots, u_r)$ and $U_2 = (u_{r+1}, u_{r+2}, \ldots\ldots\ldots, u_m)$ such that $U = (U_1, U_2)$ is orthonormal. Then, for any $k > r$, we have

$u_k^T A v_i = \Sigma_i u_k^T u_i = 0$, $\quad i = 1, \ldots\ldots\ldots, r$ (by orthogonality of the vectors of U) and

$u_k^T A v_i = 0$, $\quad i = r+1, \ldots\ldots\ldots, n$

Let

$V = (V_1, V_2)$, Then

$$U^T AV = \begin{pmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_m^T \end{pmatrix} A(v_1, v_2, \ldots . v_n) = \begin{pmatrix} \frac{1}{\sigma_1} v_1^T A^T \\ \frac{1}{\sigma_2} v_2^T A^T \\ \vdots \\ \frac{1}{\sigma_r} v_r^T A^T \\ u_{r+1}^T \\ \vdots \\ u_m^T \end{pmatrix} A(v_1, v_2, \ldots\ldots v_n) \qquad \ldots\ldots\ldots\ldots (4.3.7)$$

$$= \begin{pmatrix} \frac{1}{\sigma_1}\sigma_1^2 & & & 0 \\ & \frac{1}{\sigma_2}\sigma_2^2 & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_r}\sigma_r^2 \\ 0 & & & 0 \end{pmatrix} = \Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix}$$

where $\Sigma_1 = diag\ (\sigma_1, \sigma_2, ... \sigma_r)$.

The statement about the rank is obvious, because rank (A) = rank $(U\Sigma V^T)$ = rank$(\Sigma) = r$.The decomposition $A = U\Sigma V^T$ is known as the singular value decomposition (SVD) of A.

Note:

We can assume that $m \geq n;$ because if m < n, we consider the SVD of A^T, and if the SVD of A^T is $U\Sigma V^T$,then the SVD of A is $V\Sigma^T U^T$. Here singular values are in non increasing order. Thus $\sigma_1 = \sigma_{\max}$ is the largest singular value and $\sigma_n = \sigma_{\min}$ the smallest singular value and $\sigma(A)$ denote the set of singular values of A.

4.4 Uniqueness of Singular Value Decomposition

There are $k = \min(m,n)$ singular values of A.Let r be the rank of A. Then there are r positive singular values. These are the positive square roots of the nonzero eigenvalues of $A^T A$ (or AA^T).The remaining $(k-r)$,if $r < k$, singular values are zero. Thus, the singular values are unique. However, singular vectors are not unique. For example, if A has a multiple singular value $s > 0$,then the corresponding columns of the matrix V can be chosen as any orthonormal basis of the space spanned by the eigenvectors associated with the multiple eigen value $\lambda = s^2$ of $A^T A$. Therefore, the singular value decomposition is “almost unique”. There are two sources of ambiguity. The first is in the orientation of the singular vectors. One can flip any right singular vector, provided that the corresponding left singular vector is flipped as well, and still obtain a valid SVD. Singular vectors must be flipped in pairs (a left vector

and its corresponding right vector) because the singular values are required to be nonnegative. This is a trivial ambiguity. If desired, it can be removed by imposing, for instance, that the first nonzero entry of every left singular value be positive. The second source of ambiguity is deeper. If the matrix A maps a hyper sphere into another hyper sphere, the axes of the latter are not defined. For instance, the identity matrix has an infinity of SVDs, all of the form

$$I = UIU^T$$

where U is any orthogonal matrix of suitable size. More generally, whenever two or more singular values coincide, the subspaces identified by the corresponding left and right singular vectors are unique, but any ortho normal basis can be chosen within, say, the right subspace and yield, together with the corresponding left singular vectors, a valid SVD. Except for these ambiguities, the SVD is unique.

Even in the general case, the singular values of a matrix A are the lengths of the semi-axes of the hyper ellipse E defined by

$$E = \{Ax : \|x\| = I\}$$

The SVD reveals a great deal about the structure of a matrix. If we define r by $\sigma_1 \geq ....\geq \sigma_r \geq \sigma_{r+1} = ...= 0$ that is, if is the smallest nonzero singular value of A, then

$$rank\,(A) = r$$
$$null\,(A) = span\,(v_{r+1}, v_{r+2}, ..., v_r)$$
$$range\,(A) = span\,(u_1, u_2, ..., u_n)$$

The sizes of the matrices in the SVD are as follows: U is $m \times m$, and Σ is $m \times n$, and V is $m \times n$.
Thus, Σ has the same shape and size as A, while U and V are square. However, if $m > n$, the bottom $(m-n) \times n$ block of Σ is zero, so that the last $(m-n)$ columns of U are multiplied by zero.

Similarly, if $m < n$, the rightmost $m \times (n-m)$ block of Σ is zero, and this multiplies the last $(n-m)$ rows of V . This suggests a "small,"

equivalent version of the SVD. If $p=\min(m,n)$, we can define $U_p = U(:;1:p)$, $\Sigma_p = \Sigma(1:p;1:p)$, and $V_p = V(:;1:p)$, and write

$$A = U_P \Sigma_P V_P^T$$

where U_p is $m \times p$, Σ_p is $p \times p$, and V_p is $n \times p$.

Moreover, if $p-r$ singular values are zero, we can let $U_r = U(:;1:r)$, $\Sigma_r = \Sigma(1:r;1:r)$, and $V_r = V(:;1:r)$, then we have

$$A = U_r \Sigma_r V_r^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

which is an even smaller, minimal, SVD.

Finally, both the 2-norm and the Frobenius norm

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}$$

and

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

are neatly characterized in terms of the SVD:

$$\|A\|_F^2 = \sigma_1^2 + \ldots\ldots\ldots\ldots + \sigma_P^2$$

$$\|A\|_2 = \sigma_1$$

4.5 Relationship between the SVD and EVD

The SVD theorem shows the fact that the singular values Of A are the nonnegative square roots of eigenvalues of $A = AA^T$.The following theorem shows that how the SVD of A is related to the eigen decompositions of AA^T and $A^T A$.

Theorem

Let $A = U\Sigma V^T$ be the singular value decomposition of m× n matrix A $(m \geq n)$.Let r be the rank of the matrix .Then

1. $V^T\left(A^T A\right) V = diag\left(\sigma_1^2, \sigma_2^2, \ldots\ldots\ldots\ldots, \sigma_r^2, 0, \ldots 0\right)_{(n \times n)}$
2. $U^T\left(AA^T\right) U = diag\left(\sigma_1^2, \sigma_2^2, \ldots\ldots\ldots\ldots, \sigma_r^2, 0, \ldots 0\right)_{(m \times m)}$

Proof: We get,

$$A^TA = \left(V\sum^T U^T\right)\left(U\sum V^T\right)$$

$$= V\sum^T \sum V^T$$

$$= V\sum' V^T \qquad \text{.....(4.5.1)}$$

where $\sum'$ is an n×n diagonal matrix with $\sigma_1^2,.....,\sigma_r^2,0,.....0$ as its diagonal entries.

Thus,

$$V^TA^TAV = \sum'$$

$$= diag\left(\sigma_1^2,.....,\sigma_r^2,0,.....0\right)_{(n\times n)} \qquad \text{.....(4.5.2)}$$

Similarly,

$$AA^T = U\sum\sum^T U^T \qquad \text{..... (4.5.3)}$$

$$= diag\left(\sigma_1^2,.....,\sigma_r^2,0,.....0\right)_{(m\times m)}.$$

Note:

i) The right singular vectors $v_1,v_2,.....v_n$ are the eigenvectors of the matrix A^TA.

ii) The left singular vectors $u_1,u_2,.....u_m$ are the eigenvectors of the matrix AA^T.

iii) $\sigma_1^2,.....,\sigma_r^2$ are the nonzero eigenvalues of both A^TA and AA^T.

Example

Finding the singular value decomposition of

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

Let us find $A^TA = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$

Now the eigenvalues of A^TA are 2 and 8 and corresponding unit eigenvectors are

$$v_1 = \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

respectively $\sigma_1 = \sqrt{2}$ and $\sigma_2 = \sqrt{8} = 2\sqrt{2}$ .We have

$$Av_1 = \sigma_1 u_1 v_1^T v_1 = \sigma_1 u_1 = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix}, \text{ so } \quad u_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$Av_2 = \sigma_2 u_2 v_2^T v_2 = \sigma_1 u_2 = \begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}, \text{ so } \quad u_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The SVD of A is therefore,

$$A = U \Sigma V^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

We can also compute the eigenvectors u_1 and u_2 directly from AA^T .

4.6 Applications of the SVD

The SVD has become an effective tool in handling various important problems arising in a wide variety of applications areas, such as control theory, signal and image processing, identification and estimation, speech synthesis, pattern recognition, time series analysis, electrical network theory and biomedical engineering.

The aspects of control theory and identification problems requiring use of the SVD including problems on controllability and observability , realization of state-space models, the H-infinity control, balancing, robust feedback stabilization, model reduction and so on.

In signal and speech processing the SVD can be regarded as a filter that produces an estimate of a signal from noisy data. For example, when a person speaks, speech is absent about 50% of the time. Thus, if there are background noises coming from a fan, a vibrating machine, and so on, then these disturbances dominate in the microphone signal when speech is absent. In such a situation the ratio of speech signal to the background noise has to be enhanced, and the SVD can be used effectively to do so.

In image-processing applications the SVD is routinely used for image compression, to remove noise in picture, and so on.In biomedical engineering the SVD plays an important role in obtaining a meaningful fetal ECG from that of the mother.

Furthermore, the SVD is the most effective tool in solving least squares and the generalized least-sqaure problems and to determine the rank of a matrix. As well as SVD is the most reliable way to determine the rank deficiency of a matrix, orthonomal bases for the low and column spaces of a matrix and also for orthogonal complements and projections.

4.6.1 Image Processing(Fetal ECG)

Consider a problem of taking the electrocardiogram of a fetus by placing the cutaneous electrodes at the heart and the abdomen of the pregnant mother. The maternal ECG (MECG) clearly disturbs the observation of the fetal ECG(FECG),because the contributions of the maternal heart signals are much stronger than those of the fetal heart. The objective then will be to detect the FECG while simultaneously suppressing the MECG respective to noise.

Let us suppose there are p measurement signals $m_1(t),.....,m_p(t)$. Let these measurements be arranged in a vector called the measurement vector, $m(t)$:

$$m(t)=\left(m_1(t), m_2(t),.....,m_p(t)\right)^T \qquad \ldots\ldots\ldots\ldots\ldots \quad (4.6.1.1)$$

Let there be r source signals $s_1(t), s_2(t)\ldots\ldots\ldots\ldots, s_r(t)$ arranged in the source signal vector, $s(t)$:

$$s(t) = (s_1(t), s_2(t), \ldots\ldots\ldots, s_r(t))^T \qquad \ldots\ldots\ldots\ldots\ldots\ldots(4.6.1.2)$$

Obviously, the measurement signals are corrupted by an additive noise signal, and there exists a relationship between the measurement signals. It can be assumed that this relationship is linear and indeed each measurement signal $m_i(t)$ can be written as a linear combination of r source signals $s_i(t)$ and additive noise signal $n_i(t)$. This leads to the following equations:

$$\begin{aligned} m_1(t) &= t_{11}\, s_1(t) + t_{12}\, s_2(t) + \ldots\ldots\ldots + t_{1r}\, s_r(t) + n_1(t) \\ m_2(t) &= t_{21}\, s_1(t) + t_{22}\, s_2(t) + \ldots\ldots\ldots + t_{2r}\, s_r(t) + n_2(t) \\ &\vdots \\ &\vdots \\ m_p(t) &= t_{p1}\, s_1(t) + t_{p2}\, s_2(t) + \ldots\ldots\ldots + t_{pr}\, s_r(t) + n_r(t) \end{aligned} \qquad \ldots\ldots\ldots\ldots(4.6.1.3)$$

or ,

$$m(t) = Ts(t) + n(t) \qquad \ldots\ldots\ldots\ldots(4.6.1.4)$$

where, $T = (t_{ij})$ and $n(t) = (n_1(t), n_2(t), \ldots\ldots\ldots, n_r(t))^T$ $\ldots\ldots\ldots\ldots(4.6.1.5)$

The matrix T is called the transfer matrix and depends upon the geometry of the body, the positions of the electrodes and sources, and the conductivities of the body tissues.

Now, the problem is to get an estimate of the source signals $s(t)$ knowing only $m(t)$ and, from that estimate, separate out the estimate of fetal source signals.

Let each measurement consist of q samples. Then the measurement can be stored in a matrix M of order $p \times q$. The matrix equation form of (4.6.1.4) is $M = TS + N$.

We now show that the SVD of M can be used to get an estimates of the source signals. Let

$$M = U \sum V^T \qquad \ldots\ldots\ldots\ldots..(4.6.1.6)$$

be the SVD of M. Then $p \times q$ matrix $\hat{S}$ defined by

$$\hat{S} = U^T M \qquad \ldots\ldots\ldots\ldots..(4.6.1.7)$$

will contain p estimates source signals. Next, we need to get the estimates of the fetal source signals from $\hat{S}$; let this be called $\hat{S}_F$.

Partition the matrix of singular values Σ of M as follows:

$$\Sigma = \begin{pmatrix} \Sigma_M & 0 & 0 \\ 0 & \Sigma_F & 0 \\ 0 & 0 & \Sigma_0 \end{pmatrix} \qquad \text{.....(4.6.1.8)}$$

where Σ_M contains r_m large singular values associated with the maternal heart, Σ_F contains r_j singular values, those smaller ones associated with the fetal heart, and Σ_0 contains the remaining singular values associated with noise, and so on.

Let $U = (U_M, U_F, U_0)$ be a comfortable partitioning of U. Then obviously, we have

$$\hat{S} = U^T M = \begin{pmatrix} U_M^T \\ U_F^T \\ U_0^T \end{pmatrix} M \qquad \text{.....(4.6.1.9)}$$

$$= \begin{pmatrix} U_M^T \\ U_F^T \\ U_0^T \end{pmatrix} = \begin{pmatrix} \hat{S}_M \\ \hat{S}_F \\ \hat{S}_0 \end{pmatrix}$$

Thus $\hat{S}_F = U_F^T M$.

Once $\hat{S}_F$ is determined , we can also construct a matrix F containing only the contributions of fetus in each measured signal, as follows:

$$F = U_F \hat{S}_F = \sum_{i=r_m+1}^{r_m+r_j} \sigma_i u_i v_i^T$$

where u_i and v_i are the ith column of U and V and σ_i is the ith singular value of M.The signal in $\hat{S}_F$ are called fetal signals.

4.6.2 The Rank the Rank-Deficiency of a Matrix

Finding the matrix of an $m \times n$ matrix and, in particular, determining the non singularity of a square matrix are very important tasks that frequently arise in linear algebra and many important applications. The most obvious and the least expensive way of determining the rank of a matrix is , to triangularize the matrix using Gaussian elimination and then to find the rank of the reduced upper triangular matrix. Unfortunately, however, this is not for a very reliable approach in floating- point arithmetic. In practice, it is more important as we will see, to determine if the given matrix is near to a matrix of a certain rank and, in particular, to know if a matrix of full rank is near a rank-deficient matrix. The Gaussian elimination method, which uses elementary transformations, may transform a rank-deficient matrix rank into one having full rank, due to numerical round-off error. The most reliable way to determine the rank and nearness to rank-deficiency is to use the SVD.

Let us suppose that A has rank r, that is, $\sigma_1 \geq \sigma_2 \geq \ldots\ldots\ldots \geq \sigma_n > 0$. Then by using the SVD we can find the distance of matrix A from the nearest matrix of lower rank which is $k < r$.

Low Rank Approximations

The most widespread application of the singular value decomposition is the detection of rank degeneracy. If A is of rank k, then

$$\sigma_k > 0 = \sigma_{k+1} = \ldots\ldots\ldots + \sigma_n .$$

Thus if A has small singular values, then A is near a matrix of defective rank. Specifically, set $\Sigma_k = diag(\sigma_1, \ldots\ldots, \sigma_k, 0, \ldots\ldots 0)$ and

$$A_k = U \begin{pmatrix} \Sigma_k \\ 0 \end{pmatrix} V^T .$$

Then A_k has rank not greater than k and

$\|A_k - A\|_F^2 = \sigma_{k+1}^2 + \ldots\ldots\ldots + \sigma_n^2$.

The above construction shows that small singular values are a sufficient condition fort a rank degeneracy.

Again, let B be any matrix of rank not greater than k,let the singular values of B be denoted by $\phi_1 \geq \phi_2 \geq \ldots\ldots\ldots \geq \phi_n$. Then $\phi_{k+1} = \phi_{k+2} = \ldots\ldots\ldots = \phi_n = 0$

By Mirsky's theorem

$$\|B - A\|_F^2 \geq \sum_{i=1}^{n} |\phi_i - \sigma_i|^2 \geq \sigma_{k+1}^2 + \ldots\ldots\ldots + \sigma_n^2 \geq \|A_k - A\|_F^2 .$$

Thus if A is near a matrix B of rank k,then the sum of squares of the k smallest singular values of A is not greater than $\|B - A\|_F^2$.

Theorem: (Schmidt)

"The matrix A_k is a matrix of rank k that is nearest A in the frobenius norm."

Or

Let A and A_k be two matrix. Then

(i) A_k has rank k.

(ii) The distance of A_k from A: $\|A - A_k\|_2 = \sigma_{k+1}$.

(iii) Out of all matrices of rank k, A_k is closest to A ;that is

$$\min_{rank(B)=k} \|A - B\|_2 = \|A - A_k\|_2$$

Proof of (i): Since, $\sigma_1 \geq \sigma_2 \geq \ldots\ldots\ldots \geq \sigma_n$.

It is obvious that rank $(A_k) = k$; because

rank (A_k) =rank $(U \Sigma_k V^T)$

$= \text{rank } (\Sigma_k)$

$= k$.

Thus A_k has rank k.

Proof of (ii): Since $A = U\Sigma V^T$ and $A_k = U\Sigma_k V^T$

and $A - A_k = U(\Sigma - \Sigma_k)V^T$.

So, we have $\|A - A_k\|_2 = \|U(\Sigma - \Sigma_k)V^T\|_2$

$= \|(\Sigma - \Sigma_k)\|_2$

$= \sigma_{k+1}$.

Thus the distance between A and A_k is σ_{k+1}.

Proof of (iii): To prove (iii), we have to show that if $B \in R^{m\times n}$ is any matrix of rank k,then

$\|A - B\|_2 \geq \sigma_{k+1}$;that is A_k is closest to A among all other matrices of rank k.Since B has rank k, the full null space of B , $N(B)$ has dimension $n-k$. Consider now the space $S = span\{v_1, v_2, ..., v_{k+1}\}$, where v_1 through v_{k+1} are the right singular vectors of A .Since $N(B)$ and S are both subspaces of R^n and the sum of their dimensions is greater than n ,their intersection must be nonempty. Let z be a unit vector lying in this intersection. Then, since $z \in span\{v_1, v_2, ..., v_{k+1}\}$,there exists scalars such that $z = c_1 v_1 + c_2 v_2 + ... + c_k v_{k+1}$.

Furthermore, because $v_1, v_2, ..., v_{k+1}$ are orthonormal. We must have $|c_1|^2 + |c_2|^2 + ... + |c_{k+1}|^2 = 1$. Because $z \in N(B)$,we have $B_Z = 0$.So

$$(A - B)z = Az = \sum_{i=1}^{k+1} c_i A v_i = \sum_{i=1}^{k+1} \sigma_i c_i v_i$$

Because $u_1, u_2, \ldots\ldots, u_{k+1}$ are also orthonormal,

$$\|(A-B)z\|_2^2 = \sum_{i=1}^{k+1} |\sigma_i c_i|^2 \geq \sigma_{k+1}^2 \sum_{i=1}^{k+1} |c_i|^2 = \sum_{i=1}^{k+1} c_i A v_i = \sigma_{k+1}^2$$

Thus $\|A-B\|_2 \geq \dfrac{\|(A-B)z\|_2}{\|z\|_2} \geq \sigma_{k+1}^2$ $\quad [\because \|z\|_2 = 1]$.

4.6.3 Least Square Problem and The Pseudo inverse

The singular value decomposition is an effective tool to solve least square problems, both in the full rank and rank deficient cases. Consider the algebraic linear system

$$Ax = b$$

where A is a $m \times n$ matrix and b is a real vector. Then the linear least square problem is defined as follows to find a real n–vector x such that is $\|r\|_2 = \|Ax - b\|_2$ minimum.

Let, $A = U \Sigma V^T$ be the SVD of A.Then we have

$$\|r\|_2 = \|(U \Sigma V^T - b)\|_2$$

$$= \|U(\Sigma V^T x - U^T b)\|_2$$

$$= \|\Sigma y - b'\|_2$$

where $V^T x = y$ and $U^T b = b'$.Thus, the use of the SVD of A reduces the least-squares problem for a full matrix A to one with a diagonal matrix Σ:

The reduced problem is trivial to solve. We have

$$\|\Sigma y - b'\|_2 = \sum_{i=1}^{k} |\sigma_i y_i - b_i'|^2 + \sum_{i=k+1}^{m} |b_i'|^2$$

where k is the number of nonzero values of A .Thus, the vector

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

that minimizes $\|\Sigma y - b'\|_2$ is given by :

$$y_i = \begin{cases} \dfrac{b_i'}{\sigma_i} & if\ \sigma_i \neq 0 \\ arbitrary & if\ \sigma_i = 0 \end{cases}$$

When y is computed the solution can be recovered from $x = Vy$.

Because corresponding to each zero singular value σ_i , y_i can be set arbitrarily, in the rank-deficient case, we will have infinitely many solutions to the least-squares problem. In the full-rank case the least squares solution is unique.

The Pseudo inverse

One of the most important applications of the SVD is the solution of linear systems in the least squares sense. A linear system of the form

$$Ax = b$$

To solve $Ax = b$ for $m \times n$ matrix , we have to use $n \times m$ pseudo inverse matrix of the form

$$A^{\dagger} = U \Sigma^{\dagger} V^T .$$

The matrix $A^{\dagger} = \left(A^T A\right)^{-1} A^T$,when A is of order $m \times n (m \geq n)$ and has rank n, or the matrix $A^{\dagger} = A^T \left(AA^T\right)^{-1}$, when A is of order $m \times n$ $(m < n)$ and has rank m ,is called the pseudoinverse of A .The pseudoinverse is also reffered to as the Moore-Penorse generalized inverse of A .

Clearly, the definition of pseudo inverse is the ordinary definition of the inverse of square matrix A. When A is square and invertible

$$A^{\dagger} = \left(A^T A\right)^{-1} A^T = A^{-1}\left(A^T\right)^{-1} A^T = A^{-1}$$

Let , $A = U\Sigma V^T$ be SVD of A ,then it is easy to verify that $A^{\dagger} = U\Sigma^{\dagger} V^T$, where $\Sigma^{\dagger} = diag\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots\ldots\ldots\ldots\ldots, \frac{1}{\sigma_r}, 0, \ldots\ldots, 0\right) \in R^{n \times m}$, satisfying the following conditions:

$(i)\ AXA = A$
$(ii)\ XAX = X$
$(iii)\left(AX\right)^T = AX$
$(iv)\ \left(XA\right)^T = XA$

This expression $A^{\dagger} = U\Sigma^{\dagger} V^T$ for the pseudo inverse coincides with A^{-1} when A is non singular, because

$$\begin{aligned} A^{-1} &= \left(A^T A\right)^{-1} A^T \\ &= \left(V\Sigma^T U^T U\Sigma V^T\right)^{-1} V\Sigma^T U^T \\ &= V\Sigma^{-1}\left(\Sigma^T\right)^{-1} V^T V\Sigma^T U^T \\ &= V\Sigma^{-1} U^T, \text{ here } \Sigma^{\dagger} = \Sigma^{-1}. \end{aligned}$$

CHAPTER FIVE

UPPER SOLUTION BOUNDS FOR CALE:

A SINGULAR VALUE DECOMPOSITION APPROACH

The Continuous Algebraic Lyapunov equation (CALE) is one of the forms of Lyapunov equation. CALE is a fundamental matrix equation which plays an significant role in control theory, model reduction and stochastic analysis of dynamical systems. Mainly in control system solution bounds of CALE is very important to solve stability analysis problems .Although the bounds of the exact solution of the Lyapunov equation can be found numerically, the computational burden increases with the dimension of the system matrices. For some applications such as stability analysis, it is often not necessary to know the exact solution. In this chapter we discuss about basic terminology of Lyapunov equation and the upper bounds solution for Continuous Algebraic Lyapunov equation via SVD approach.

5.1 Preliminaries

The stability of a continuous time system is defined here with respect to an equilibrium state. An equilibrium state of the uncontrolled system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0 \qquad \text{.....(5.1.1)}$$

is the vector x_e satisfying

$$Ax_e = 0$$

Clearly $x_e = 0$ is an equilibrium state and it is the unique equilibrium state if and only only if A is nonsingular. An equilibrium state x_e is asymptotically stable if for any initial state, the state vector $x(t)$ approaches x_e as time increases.

The system (5.1.1) is asymptotically stable if and only if the equilibrium state is asymptotically stable. Thus, the system (5.1.1) is asymptotically stable if and only if $x(t) \to 0$ as $t \to \infty$.

Criterion of Continuous-Time Stability

A well-known criterion of asymptotic stability of a continuous-time system is given below :

"The system (.5.1.1) is asymptotically stable if and only if all the eigenvalues of the matrix A have negative real parts.''

Proof: The general solution of (5.1.1) is

$$x(t) = e^{At}x_0$$

Thus, $x(t) \to 0$ if and only if $e^{At} \to 0$ as $t \to \infty$ We will now show that this happens if and only if all the eigenvalues of A have negative real parts.

Let $X^{-1}AX = diag\,(J_1, J_2, J_3, \ldots\ldots J_k)$ be the Jordan canonical form of A. Then,

$$e^{At} = Xdiag\,(e^{J_1 t}, e^{J_2 t}, \ldots\ldots\ldots, e^{J_K t})X^{-1}$$

Let λ_i be the eigenvalue of A associated with J_i. Then $e^{J_i t} \to 0$ if and only if λ_i has a negative real part. Therefore, $e^{At} \to 0$ if and only if all the eigenvalues A of have negative real parts.

5.2 Lyapunov Equations

For continuous time stability the matrix equation:

$$XA + A^T X = -M$$

and its dual

$$AX + XA^T = -M$$

are called the continuous time Lyapunov equations.

Similarly,the matrix equation :

$$X - A^T XA = M$$

and its dual

$$X - AXA^T = M$$

are known as discrete time Lyapunov equations.

5.3 Continuous-Time Lyapunov Stability Theory

Before the advent of computers, finding the eigenvalues of a matrix A was an extremely difficult task. In 1892, the Russian mathematician A. Lyapunov (1857-1918) developed a historical stability criterion for nonlinear systems of equations. In the linear case, this criterion may be formulated in terms of the solution of a matrix equation.

Lyapunov Stability Theorem

The linear system:

$$\dot{x}(t) = Ax(t)$$

is asymptotically stable if and only if for any symmetric positive definite matrix M,
there exists a unique symmetric positive definite matrix X satisfying the equation:

$$XA + A^T X = -M \qquad \text{.....(5.3.1)}$$

Proof : Let's define a matrix X by

$$X = \int_0^\infty e^{A^T t} M e^{At} dt \qquad \text{.....(5.3.2)}$$

Then, we show that when the system is asymptotic stable, X is a unique solution of the equation (5.3.1) and is symmetric positive definite. Using the expression of X in (5.3.1), we obtain

$$XA + A^T X = \int_0^\infty e^{A^T t} M e^{At} dt + \int_0^\infty A^T e^{A^T t} M e^{At} dt$$

$$= \int_0^\infty \frac{d}{dt}\left(e^{A^T t} M e^{At}\right) dt = \left[e^{A^T t} M e^{At}\right]_0^\infty$$

Since A is stable, $e^{A^T t} \to 0$ as $t \to \infty$. Thus, $XA + A^T X = -M$, showing that that X defined by (5.3.2) satisfies the equation (5.3.1).

To show that X is positive definite, we have to show that $u^T Xu > 0$ for any nonzero vector u .Using (5.3.2) we can write

$$u^T Xu = \int_0^\infty u^T e^{A^T t} M e^{At} u dt .$$

To prove that X is unique, assume that there are two solutions X_1 and X_2 of (5.3.1). Then,

$$A^T(X_1 - X_2) + (X_1 - X_2)A = 0,$$

which implies that

$$e^{A^T t}\left(A^T(X_1 - X_2) + (X_1 - X_2)A\right)e^{At} = 0$$

or

$$\frac{d}{dt}\left[e^{A^T t}(X_1 - X_2)e^{At}\right] = 0,$$

and hence $e^{A^T t}(X_1 - X_2)e^{At}$ is a constant matrix for all t.
Evaluating at $t = 0$ and $t \to \infty$ we conclude that $X_1 - X_2 = 0$.
We now prove the converse, that is, we prove that if X is a symmetric positive definite solution of the equation (5.3.1), then A is stable.
Let (λ, x) be an eigenpair of A. Then pre multiplying the equation (5.3.1) by x^* and post multiplying it by x, we obtain:

$$x^* XAx + x^* A^T Xx = \lambda x^* Xx + \bar{\lambda} x^* Xx = (\lambda + \bar{\lambda})x^* Xx = -x^* Mx$$

Since M and X are both symmetric positive definite, we have $\lambda + \bar{\lambda} < 0$ or $\mathbf{Re}(\lambda) < 0$.

5.4 Stability of Discrete-Time System

Consider the discrete time system:

$$x_{k+1} = Ax_k \qquad \text{.....(5.4.1)}$$

with initial value x_0.

The system (5.4.1) is asymptotically stable if and only if all the eigenvalues of A are inside the unit circle or all the eigenvalues of A have moduli less than 1.

Discrete-Time Lyapunov Stability Theory

In discrete system continuous time Lyapunov equations are replaced by discrete analogs. The discrete counterparts of the continuous-time Lyapunov equations are known as Stein equations in control theory.

Lyapunov Stability Theorem

The system

$$x_{k+1} = Ax_k$$

is asymptotically stable if and only if, for any positive definite matrix M, there exists a unique positive definite matrix X satisfying the discrete Lyapunov equation:

$$X - A^T XA = M \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5.4.2)$$

Proof : Let's define a matrix X by

$$X = \sum_{k=0}^{\infty} \left(A^T\right)^k MA^k \qquad \ldots\ldots\ldots\ldots\ldots\ldots (5.4.3)$$

Since A is discrete-stable, the infinite series on the right-hand side converges.
Furthermore, the matrix X is symmetric and positive definite. We now show that X is the unique solution of the equation. (5.4.2). Indeed,

$$X - A^T XA = \sum_{k=0}^{\infty} \left(A^T\right)^k MA^k - \sum_{k=1}^{\infty} \left(A^T\right)^k MA^k = M \ldots\ldots\ldots\ldots\ldots\ldots (5.4.4)$$

Thus, X defined by (5.4.3) satisfies the equation. (5.4.2).

To prove that X is unique, let's assume that there is another symmetric positive definite solution X_1 of (5.4.2).

Then,

$$X_1 - A^T X_1 A = M,$$

and

$$X = \sum_{k=0}^{\infty} \left(A^T\right)^k MA^k = \sum_{k=0}^{\infty} \left(A^T\right)^k \left(X_1 - A^T X_1 A\right) A^k,$$

$$= \sum_{k=0}^{\infty} \left(A^T\right)^k X_1 A^k - \sum_{k=1}^{\infty} \left(A^T\right)^k X_1 A^k = X_1 .$$

5.5 Lyapunov Functions

For the linear system :

$$\dot{x}(t) = Ax(t)$$

According to Lyapunov, to check stability of a system by finding some function $V(x)$, called the Lyapunov function, which for time invariant systems satisfies

$$V(x) > 0 \qquad V(0) = 0 \qquad \text{……..(5.5.1)}$$

$$\dot{V}(x) = \frac{dv}{dt} = \frac{\partial v}{\partial x}\frac{dx}{dt} < 0 \qquad \text{……..(5.5.2)}$$

For the linear system the linear functions can be chosen to be quadratic, that is

$$V(x) = x^T Xx \qquad X = X^T > 0 \quad \text{……..(5.5.3)}$$

where X is symmetric if $\dot{V}(x)$ is negative definite; which with the use of (5.3.1) this can be seen as follows:

$$\dot{V}(x) = \dot{x}^T Xx + x^T X\dot{x},$$

$$= x^T \left(A^T X + XA\right)x,$$

that is the system is asymptotically stable if the following condition is satisfied

$$A^T X + XA < 0$$

or, equivalently,

$$A^T X + XA = -M \;; M = M^T > 0 \qquad \text{……(5.5.4)}$$

Therefore,

$$\dot{V}(x) = x * (-M)x$$

Thus $\dot{V}(x)$ is negative definite if M is positive definite.

There is no general procedure for finding the Lyapunov functions for nonlinear systems, but for linear time invariant systems, the procedure comes down to the problem of solving a linear algebraic equation, called the Lyapunov algebraic equation. The matrix algebraic equation (5.5.4) is known as the Lyapunov algebraic equation.

5.6 Singular Value Decomposition Approch For CALE

The continuous algebraic Lyapunov equation (CALE) has been widely used in engineering theory. In practical application, especially for stability analysis upper bounds of CALE are desired. "**New estimates for solutions of Lyapunov equations**" by **Y.Fang, K.Loparo, and X.Feng** was published in March,1997.To this paper we can reach to solutions of upper bounds for continuous algebraic lyapunov equations. In section **5.6.1** we discuss about the above paper to get solutions of CALE .But in most cases the existing upper bounds are valid under some restrictive assumptions which are practically inapplicable. Afterwards, **Svetoslav G. Savoy and Ivan P. Popchev** were motivated by this fact that upper solution bounds for the continuous Lyapunov equation are valid under some very restrictive conditions, so an attempt "**New Upper Bounds for the CALE:A Singular Value Decomposition Approach**" is made in April, 2007 to extend the set of Hurwitz matrices for which such bounds are applicable. In section **5.6.2** it is shown that the matrix set for which solution bounds are available is only a subset of another stable matrices set. This helps to loosen the validity restriction and its easy for application. Extended matrices with upper bounds are illustrated here by examples via SVD approach.

5.6.1 Upper Bounds Solution for CALE with Hurwitz matrix

Consider, the continuous algebraic Lyapunov equation

$$A^T P + PA + Q = 0, \quad A \in H, \; Q > 0, \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

with respect to unknown matrix P.Here H is the set of Hurwitz matrices, the symmetric part of matrix A is,

$$A_S = \frac{1}{2}\left(A^T + A\right)$$

The eigen values of a $n \times n$ matrix A are denoted by $\lambda_1 \geq \lambda_2 \geq .....\geq \lambda_n$;

$\lambda_1(A_S) = \mu(A)$ is the matrix measure of A.The maximum and minimum singular values of matrix A are $\sigma_1(A)$and $\sigma_n(A)$ respectively. The maximum real part of the eigenvalues of matrix A is $\kappa(A)$ and I is the identity matrix. If $\mu(A) < 0$,the best known upper solution bounds are:

$$\lambda_1(p) \leq l_0 = \frac{1}{2}\lambda_1(-QA_S^{-1}) \quad .....(2)$$

$$t_r(p) \leq t_0 = \frac{1}{2}\sum_1^n \lambda_i(Q)/\lambda_i(A_S) \quad .....(3)$$

Let T be a non-singular matrix, denote $\tilde{A} = TAT^{-1}$ and define the set $H^- \equiv \{\tilde{A} : \mu[\tilde{A}] < 0\}$ .Then equation (1) can be written as –

$$\left.\begin{aligned} &\tilde{A}^T\tilde{P} + \tilde{P}\tilde{A} + \tilde{Q} = 0 \\ &\tilde{P} = T^{-T}PT^{-1} \\ &\tilde{Q} = T^{-T}QT^{-1} \end{aligned}\right\} \quad .....(4)$$

For any given matrix $A \in$ H and $A \notin$ H^{-1},There exists matrix T, such that $\tilde{A} \notin H^{-1}$, i.e $\tilde{A}_S < 0$.

Consider the positive scaler η,defined as-

$$\begin{aligned} \eta &= \frac{1}{2}\lambda_1\left\{-Q\left[\left(T^TTA\right)_S\right]\right\}^{-1} \\ &= \lambda_1\left[-Q\left\{\left(T^TTA\right)^T + T^TTA\right\}^{-1}\right] \\ &= \lambda_1[-Q\left(A^TT^TT + T^TTA\right)^{-1}] \qquad [\because \left(T^T\right)^T = T] \end{aligned}$$

$$= \lambda_1\left\{-Q\left(T^T T^{-T} A^T T^T T + T^T A T^{-1} T\right)^{-1}\right\}$$

$$= \lambda_1\left\{-Q\left[T^T T\left(T^{-1} A^T T^T + T A T^{-1}\right)\right]^{-1}\right\}$$

$$= \lambda_1[T^{-T} Q T^{-1} \times \left\{-\left(A^T + A\right)\right\}^{-1}]$$

$$= \frac{1}{2}\lambda_1\left\{\widetilde{Q}\left[-\widetilde{A}_S\right]^{-1}\right\} \qquad [\because \widetilde{Q} = T^{-T} Q T^{-1}]$$

$$= \frac{1}{2}\lambda_1\left[\left(-\widetilde{A}_S\right)^{\frac{1}{2}} \widetilde{Q}\left(-\widetilde{A}_S\right)^{-\frac{1}{2}}\right]$$

$$\text{or, } 2\eta = \lambda_1\left[\left(-\widetilde{A}_S\right)^{\frac{1}{2}} \widetilde{Q}\left(-\widetilde{A}_S\right)^{-\frac{1}{2}}\right]$$

Since, $2\eta I \geq \left(-\widetilde{A}_S\right)^{\frac{1}{2}} \widetilde{Q}\left(-\widetilde{A}_S\right)^{-\frac{1}{2}}$

then we get the inequality

$$\geq 2\eta\widetilde{A}_S + \widetilde{Q}$$

$$= 2\eta \cdot \frac{1}{2}\left(\widetilde{A}^T + \widetilde{A}\right) + \widetilde{Q} - \left(-\widetilde{A}^T \widetilde{P} - \widetilde{P}\widetilde{A} - \widetilde{Q}\right)$$

$$= \widetilde{A}^T\left(\eta I - \widetilde{P}\right) + \widetilde{A}\left(\eta I - \widetilde{P}\right)$$

Here, matrix $\widetilde{A}$ is Hurwitz and by Lyapunov Stability theory $\eta I - \widetilde{P}$ must be a positive semi definite matrix, which yields the upper matrix bound

$$P \leq \eta T^T T; \quad \eta = \frac{1}{2}\lambda_1\left\{-Q\left[\left(T^T A T\right)_S\right]^{-1}\right\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

and then we can obtain scalar bounds for $\lambda_1(p)$ and $t_r(p)$ easily.

The estimation problem for P has three important aspects: (i) restrictions on matrix *A*, (ii) computational burden and (iii) tightness of the bounds. Bounds based on equation (4) eliminate problem (i), but require the determination of matrix T, the selection of which to obtain the tightest bound is an open and difficult question.

Matrix T is obtained by some additional computational procedure and in this sense $\overline{p} = T^T T$, $\mu(\overline{P}A)$<0, is said to be an external Lyapunov matrix (ELM) for A. An internal Lyapunov matrix (ILM) is a matrix which can be defined entirely in terms of A. i.e if $A \in H^-$, then $\overline{P} = A^T A$ is an ILM for A.

Its an attempt to overcome to a certain extent the above mentioned difficulties concerning bouns based on ELM. This is closely related with the definition of an extension $\tilde{H}$ of the conservative set H^- in sense that, if $A \notin H^-$, but $A \in \tilde{H}$, there exists ILM for *A*. This will help to avoid the computation of an ELM.

5.6.2 Extension of Hurwitz matrices by using SVD

Using the singular value decomposition of the coefficient matrix in (1), it is always possible to present it as a product of two matrices as follows

$$A = FP_1 = P_2 F$$
$$F = UV^T$$
$$P_1^2 = A^T A, P_2^2 = AA^T$$

where $A = U\sum V^T$, $UU^T = VV^T = I$,and $\sum$ is a positive diagonal matrix containing singular values of A. Similarly the SVD of the transformed matrix results as follows-

$$\left.\begin{array}{l} \tilde{A} = TAT^{-1} = \tilde{U}\tilde{\sum}\tilde{V}^T = \tilde{F}\tilde{P}_1 = \tilde{P}_2\tilde{F} \\ \\ \tilde{F} = \tilde{U}\tilde{V}^T, \tilde{P}_1^2 = \tilde{A}^T\tilde{A}, \tilde{P}_2^2 = \tilde{A}\tilde{A}^T \end{array}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

With $\tilde{U}\tilde{U}^T = \tilde{V}\tilde{V}^T = I$ and $\tilde{\sum}$ is a positive diagonal matrix, Then we can define the matrix set

$$\tilde{H} \equiv \left\{\tilde{A} : \tilde{F} \in H\right\}$$

Theorem 1:For any matrix $A \in H$ and nonsingular *T,* one has

(a) $\lambda_1\left[\left(T^T TA\right)_S S^{-1}\right] \geq \kappa\left[\widetilde{F}\right]$,

(b) $H^- \in \widetilde{H}$,

(c) $\widetilde{A} \in \widetilde{H} \Leftrightarrow \mu(SA) \langle 0$; here we denote $S = T^T \widetilde{P}_1 T$.

Proof: From the representation of the SVD of the transformed matrix we get,

$$\widetilde{p}_1^{-\frac{1}{2}} \widetilde{A} \widetilde{p}_1^{-\frac{1}{2}} = \widetilde{p}_1^{-\frac{1}{2}} \widetilde{F} \widetilde{p}_1^{\frac{1}{2}} = X$$

Assertion (a) is proved applying the well known inequality $\mu(Y) \geq \kappa(Y)$ valid for any matrix *Y* to *X;* which results in

$$\begin{aligned}
\mu(X) &= \lambda_1\left[\widetilde{P}_1^{-\frac{1}{2}} T^{-T}\left(T^T TA\right)_S T^{-1} \widetilde{P}_1^{-\frac{1}{2}}\right] \\
&= \lambda_1\left[\left(T^T TA\right)_S S^{-1}\right] \\
&= \mu\left[\widetilde{P}_1^{-\frac{1}{2}} \widetilde{F} \widetilde{P}_1^{\frac{1}{2}}\right] \\
&\geq \kappa\left(\widetilde{F}\right),
\end{aligned}$$

since the eigenvalues of $\widetilde{F}$ are preserved under the nonsingular transformation. Assertion (b) follows immediately, i.e.,

$$\widetilde{A} \in H^- \Leftrightarrow \mu\left[\widetilde{A}\right] \langle 0 \Leftrightarrow \mu\left(T^T TA\right) \langle 0 \Leftrightarrow \mu(X) \langle 0 \Rightarrow \widetilde{F} \in H \Rightarrow \widetilde{A} \in H$$

Finally, $\widetilde{F}$ is unitary by definition and hence normal matrix, or

$$\widetilde{A} \in H^- \Leftrightarrow 0 \rangle \kappa\left(\widetilde{F}\right) = \mu\left(\widetilde{F}\right) = \mu\left(\widetilde{A}\widetilde{P}_1^{-1}\right) \Leftrightarrow \mu(SA) \langle 0,$$

which proves (c).

Corollary 1: For any T, such that $\overline{P} = T^T T$ is an ELM for A, $S = T^T \widetilde{P}_1 T$ is also a ELM for A.

Proof: It follows from assertions (b) and (c) in Theorem 1. Let $T^T T$ be an ELM for A i.e.,

$$\mu(T^T TA)\langle 0 \Leftrightarrow \mu(\widetilde{A})\langle 0 \Leftrightarrow \widetilde{A} \in H^-$$

$$\Rightarrow \widetilde{A} \in \widetilde{H} \Leftrightarrow \mu(SA) < 0.$$

Certainty The approach suggested for getting upper bounds for the solution of P in (1) is always theoretically applicable since the symmetric part of the transformed matrix $\widetilde{A}$ is negative definite. Here corollary 1 illustrates the important fact that the SVD approach does not introduce any conservatism concerning restrictions on the coefficient matrix for bounds validity. If $T^T T$ is an ELM for A, then the upper matrix bound in (5) becomes

$$P \le \eta S, \eta = \frac{1}{2}\lambda_1\left\{-Q[(SA)_S]^{-1}\right\}$$

Since the main purpose is to get ILM, let $T = I$,i.e $\widetilde{A} \equiv A, \widetilde{F} = F$, $S = \widetilde{P}_1 = P_1$. Then theorem (1) become:

Corollary 2: For any matrix $A \in H$, such that

(a) $\lambda_1(A_S P_1^{-1}) \ge \kappa[\widetilde{F}]$,

(b) $A \in H^- \Rightarrow A \in \widetilde{H}$,

(c) $A \in \widetilde{H} \Leftrightarrow \mu(P_1 A)\langle 0 \Leftrightarrow \mu(P_2^{-1} A)\langle 0$.

In other words , $A \in \widetilde{H}$, if and only if , P_1 and $P_2^{-1}$are ILMs for A. Here we denote $S_1 = (P_1 A)_S$ and $S_2 = (p_2^{-1} A)_S$.

Corollary 3: Let $A \in \widetilde{H}$ then the solution P in (1) has the following upper matrix bounds

$$\left.\begin{aligned} P \le \mu_1 P_1, \mu_1 = \frac{1}{2}\lambda_1(-QS_1^{-1}) \\ P \le \mu_2 P_2^{-1}, \mu_2 = \frac{1}{2}\lambda_1(-QS_2^{-1}) \end{aligned}\right\} \quad \text{.....(7)}$$

Lemma: If $A \in \widetilde{H}$, the maximum eigenvalue and the trace of P in (1) have the following upper bounds:

$$\left.\begin{aligned} \lambda_1(P) \le l_1 = \min\left[\mu_1\sigma_1(A), \mu_2\sigma_n^{-1}(A)\right], \\ t_r(P) \le t_1 = \min\left[t_1', t_2', t_1'', t_2''\right], \end{aligned}\right\} \quad \text{.....(8)}$$

$$\left.\begin{aligned} t_1' &= \frac{1}{2}\lambda_1\left(-QS_1^{-1}\right)t_r(P_1), \\ t_2' &= \frac{1}{2}\lambda_1\left(-QS_2^{-1}\right)t_r\left(P_2^{-1}\right), \\ t_1'' &= -\frac{1}{2}\lambda_1^{-1}\left(S_1P_1^{-2}\right)t_r\left(QP_1^{-1}\right), \\ t_2'' &= -\frac{1}{2}\lambda_1^{-1}\left(S_2P_2^{2}\right)t_r(QP_2), \end{aligned}\right\} \qquad \text{.....(9)}$$

Proof: Bounds l_1 and $t_i', i = 1,2....$ are obtained from the respective matrix bounds (7) for P. We know the SVD representation of $A = U\sum V^T = FP_1 = P_2F$, $F = UV^T$,the CALE (1) can be rewritten as

$$Q = -P_1F^TP - PFP_1 \Rightarrow QP_1^{-1} = -P_1F^TPP_1^{-1} - PF$$

$$Q = -F^TP_2P - PP_2F \Rightarrow QP_2 = -F^TP_2PP_2 - PP_2FP_2 .$$

Application of the t_r operator to both sides of the above equalities results in

$$\begin{aligned} t_r\left(QP_1^{-1}\right) &= -2t_r(PF) \\ &= -2t_r\left(P^{1/2}F_SP^{1/2}\right) \\ &\geq -2\mu(F)t_r(P) \ , \\ t_r(QP_2) &= -2t_r\left(PP_2FP_2\right), \\ &= -2t_r\left(P^{1/2}P_2F_SP_2P^{1/2}\right), \\ &\geq -2\mu(P_2FP_2)t_r(P). \end{aligned}$$

Since $F_S = P_1^{-1}S_1P_1^{-1} = S_2 \langle 0$, bounds $t_{1,2}''$ are proved.

The requirement $A \in \tilde{H}$ is less restrictive in comparison with the assumption that $A \in H^-$, due to the fact that $H^- \subseteq \tilde{H}$. Therefore, the presented bounds (7), (8) and (9) are less conservative with respect to the validity restrictions imposed on matrix A by the existing estimation approaches. The derived bounds (7), (8), and (9) are based on the SVD of the coefficient matrix A and in this sense they differ from all available bounds. Nevertheless, the only specific procedure consists in getting A in the form $A = U\sum V^T$. Once the

decomposition is done, one can easily compute the matrices involved in the respective bounds as follows:

$$P_1 = V \sum V^T, P_1^{-1} = V \sum^{-1} V^T, P_1^{-2} = V \sum^{-2} V^T,$$
$$P_2 = U \sum U^T, P_2^{-1} = U \sum^{-1} U^T, P_2^2 = AA^T, F = UV^T,$$
$$S_1 = (P_1 A)_S = V \sum V^T (UV^T + VU^T) V \sum V^T,$$
$$= V \sum V^T F_S V \sum V^T, S_2 = (P_2^{-1} A)_S = F_S.$$

Matrix F is a normal one and hence it is unitarily similar to a diagonal matrix Λ i.e., $F = W \Lambda W^T$. If Λ is Hurwitz, the proposed bounds are all valid. Then, the inverses of F_S, S_1 and S_2 are obtained by inversing diagonal matrices, i.e.,

$$F_S = 2W \Lambda_{\text{Re}} W^T, F_S^{-1} = \frac{1}{2} W \Lambda^{-1}{}_{\text{Re}} W^T,$$

$$S_1^{-1} = \frac{1}{2} V \sum^{-1} V^T W \Lambda^{-1}{}_{\text{Re}} W^T V \sum^{-1} V^{-1}, S_2^{-1} = F_S^{-1},$$

where A_{Re} denotes a diagonal matrix containing the real parts of the eigenvalues of F, while (2) and (5) require the computation of a possibly ill-conditioned general matrix. Since F is a normal matrix, the computation of W and Λ should not be a problem. In any case, this computation is easier than the computation of the eigenvalues of A which must be put in the form $A = C^T T C$ with C (unitary) and T (triangular) being complex matrices in the general case. Also, the computation of the spectrum of A is not required at all if the SVD is performed due to the following reasons. Matrix F must be Hurwitz in order to get valid bounds. If F is Hurwitz, this leads to

$$A = FP_1 \Rightarrow P_1 A = P_1 F P_1 \Rightarrow A^T P_1 + P_1 A = P_1 F_S P_1 \langle 0$$

which is possible only if A is Hurwitz.

The above results shows that the set of Hurwitz matrices for which there exist valid upper bounds for the solution of the CALE can be

further extended. The applicability for the proposed bounds for P is illustrated by several numerical examples.

Numerical Example :

Consider the unitary matrices

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}, \qquad V = \frac{1}{\sqrt{1+a^2}}\begin{pmatrix} 1 & a \\ -a & 1 \end{pmatrix}$$

where $a\rangle 0$ and the positive diagonal matrix $\Sigma = diag(\sigma,1)$. It is desired to investigate the influence of the parameters a and σ on the possibility to get bounds for p in (1), if the coefficient matrix is

$$\begin{aligned} A &= U \Sigma V^T \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} \times \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix} \times \frac{1}{\sqrt{1+a^2}}\begin{pmatrix} 1 & -a \\ a & 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2(1+a^2)}}\begin{pmatrix} -\alpha & 1 \\ -\alpha & -1 \end{pmatrix} \times \begin{pmatrix} 1 & -a \\ a & 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2(1+a^2)}}\begin{pmatrix} -\alpha + a & \alpha a + 1 \\ -\alpha - a & -1 + \alpha a \end{pmatrix} \end{aligned}$$

Denote $t = \left[\sqrt{2(1+a^2)}\right]^{-1}$. Then ,

$$A = t\begin{pmatrix} -\sigma + a & 1 + \sigma a \\ -\sigma - a & -1 + \sigma a \end{pmatrix}$$

For any $a\langle 1$ and any $\sigma\rangle 0$, $A \in H$ .The unitary part of A is

$$\begin{aligned} F &= UV^T \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} \times \frac{1}{\sqrt{1+a^2}}\begin{pmatrix} 1 & -a \\ a & 1 \end{pmatrix} \end{aligned}$$

$$= \frac{1}{\sqrt{2(1+a^2)}}\begin{pmatrix} -1+a & 1+a \\ -(1+a) & -1+a \end{pmatrix}$$

$$= t\begin{pmatrix} -1+a & 1+a \\ -(1+a) & -1+a \end{pmatrix}$$

$$F_S = \frac{1}{2}\left(F^T + F\right)$$

$$= \frac{1}{2}t\begin{pmatrix} -1+a & -(1+a) \\ 1+a & -1+a \end{pmatrix} + t\begin{pmatrix} -1+a & 1+a \\ -(1+a) & -1+a \end{pmatrix}$$

$$= \frac{1}{2}t\begin{pmatrix} -1+a-1+a & -1-a+1+a \\ 1+a-1-a & -1+a-1+a \end{pmatrix}$$

$$= \frac{1}{2}t\begin{pmatrix} -2+2a & 0 \\ 0 & -2+2a \end{pmatrix}$$

$$= \frac{1}{2}\times 2\times t\begin{pmatrix} -1+a & 0 \\ 0 & -1+a \end{pmatrix}$$

$$= t(-1+a)I\ ,$$

or $A \in H \Leftrightarrow A \in \tilde{H}$ in this case. If $A \in H^-$,then $a < \min(\sigma, \sigma^{-1})$,by necessity.

Consider the matrix bounds (7),with

$$S_1 = (P_1 A)_S = V\sum V^T\left(UV^T + VU^T\right)V\sum V^T,$$

$$= V\sum V^T F_S V\sum V^T,$$

$S_1 = t(-1+a)P_1^2$,

$S_2 = \left(P_2^{-1}A\right)_S = F_S = t(-1+a)I$, where

$P_1 = V\sum V^T$

$$= \frac{1}{\sqrt{1+a^2}}\begin{pmatrix} 1 & a \\ -a & 1 \end{pmatrix}\times\begin{pmatrix} \sigma & 0 \\ 0 & 1 \end{pmatrix}\times\frac{1}{\sqrt{1+a^2}}\begin{pmatrix} 1 & -a \\ a & 1 \end{pmatrix}$$

$$= \frac{1}{1+a^2}\begin{pmatrix} \sigma & a \\ -\sigma a & 1 \end{pmatrix} \times \begin{pmatrix} 1 & -a \\ a & 1 \end{pmatrix}$$

$$= \frac{1}{1+a^2}\begin{pmatrix} \sigma + a^2 & -\sigma a + a \\ -\sigma a + a & a^2\sigma + 1 \end{pmatrix}$$

$$= 2t^2\begin{pmatrix} \sigma + a^2 & -\sigma a + a \\ -\sigma a + a & a^2\sigma + 1 \end{pmatrix}$$

$$\therefore P_1^2 = 2t^2\begin{pmatrix} \sigma + a^2 & -\sigma a + a \\ -\sigma a + a & a^2\sigma + 1 \end{pmatrix} \times 2t^2\begin{pmatrix} \sigma + a^2 & -\sigma a + a \\ -\sigma a + a & a^2\sigma + 1 \end{pmatrix}$$

$$= 4t^2\begin{pmatrix} \sigma + a^2 & -\sigma a + a \\ -\sigma a + a & a^2\sigma + 1 \end{pmatrix} \times \begin{pmatrix} \sigma + a^2 & -\sigma a + a \\ -\sigma a + a & a^2\sigma + 1 \end{pmatrix}$$

$$P_2^{-1} = U \sum^{-1} U^T$$

$$= \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} \times \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & 1 \end{pmatrix} \times \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} \sigma^{-1} & 1 \\ -\sigma^{-1} & -1 \end{pmatrix} \times \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} \sigma^{-1} + 1 & \sigma^{-1} - 1 \\ \sigma^{-1} - 1 & \sigma^{-1} + 1 \end{pmatrix}$$

Now Putting, $a = 0.5$ and $\sigma = 2$.

Thus we get, $A = t\begin{pmatrix} -1.5 & 2 \\ -2.5 & 0 \end{pmatrix}$

$\mu(A)$= max eigenvalue of A_S. Since,

$\lambda_1(A_S) \rangle\ \lambda_2(A_S) \rangle \ldots\ldots\ldots\ldots\ldots\ldots \rangle \lambda_n(A_S)$

$\therefore \mu(A) = \lambda_1(A_S) \rangle\ \lambda_2(A_S) \rangle \ldots\ldots\ldots\ldots\ldots\ldots \rangle \lambda_n(A_S) \rangle 0$ .Thus $\mu(A) \rangle 0$,bounds (2) and (3) are not valid. Assume $Q = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.From the continuous algebraic lyapunov equation,

$A^T P + PA + Q = 0$; $A \in H$, $Q > 0$. Let, $P = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ .We have to find the value of P.

Now , $A^T P + PA + I = 0$

$$\text{or, } t\begin{pmatrix} -1.5 & 2.5 \\ 2 & 0 \end{pmatrix} \times \begin{pmatrix} x & y \\ z & w \end{pmatrix} + \begin{pmatrix} x & y \\ z & w \end{pmatrix} \times t\begin{pmatrix} -1.5 & 2.5 \\ 2 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 0$$

$$\text{or, } \begin{aligned} 1 - 3tx - 2.5ty - 2.5tz &= 0 \\ 2tx - 1.5ty - 2.5tw &= 0 \\ 2tx - 1.5tz - 2.5tw &= 0 \\ 1 + 2ty + 2tz &= 0 \end{aligned}$$

By solving this in **Mathematica 5** we get,

$$x = \frac{0.75}{t}, y = -\frac{0.25}{t}, z = \frac{0.75}{t}, w = -\frac{0.25}{t}$$

Thus we get, $P = \begin{pmatrix} \frac{0.75}{t} & -\frac{0.25}{t} \\ -\frac{0.25}{t} & \frac{0.75}{t} \end{pmatrix}$

$$= 0.25t^{-1}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

Trace : $tr(P) = 1.5t^{-1}$ and eigenvalues of P are $\frac{0.5}{t}, \frac{1}{t}$ .So we can write the largest eigenvalue of P as : $\lambda_1(P) = t^{-1}$. Since $A \in \tilde{H}$, bounds (8) and (9) are valid. In this case, $\mu_1 = \mu_2 = t^{-1}$.The matrix bound $\mu_2 P_2^{-1} = t^{-1}\left(AA^T\right)^{-\frac{1}{2}} = P$, which is evident from the fact that

$$A^T P_2^{-1} + P_2^{-1} A = -t\, I \quad \Rightarrow t^{-1} P_2^{-1} = P,$$

Since P is unique for any given matrices A and Q. Consider bounds (8) and (9).We get the maximum eigenvalue upper bound is

$$l_1 = \min[\mu_1 \sigma_1(A), \mu_2 \sigma_n^{-1}(A)]$$

Here we get the singular values of A are : $\{3.16228, 1.58114\}$
Therefore,

$$t^{-1} = 1.58114, \qquad 2t^{-1} = 3.16228$$

Thus by using **Mathematica 5** we get,

$$l_1 = \min(3.16228, 1.58114) = \min\left(2t^{-1}, t^{-1}\right) = 1.58114 = t^{-1} = \lambda_1(P)$$

Similarly, the trace upper bound is

$$t_1 = \min\left(3t^{-1}, 1.5t^{-1}, 1.5t^{-1}\right) = 1.5t^{-1} = t_r(P).$$

The proposed upper scalar bounds coincide with the exact respective solution parameters in this case.

APPENDIX

$$A = t \begin{pmatrix} -\sigma + a & 1 + \sigma a \\ -\sigma - a & -1 + \sigma a \end{pmatrix} /. \{\sigma \to 2, a \to 0.5\}$$

```
{{-1.5 t, 2. t}, {-2.5 t, 0. t}}
```

```
B = Transpose[A]
```

```
{{-1.5 t, -2.5 t}, {2. t, 0. t}}
```

$$P = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

```
{{x, y}, {z, w}}
```

$$B.P + P.A + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} == 0$$

```
{{1 - 3. t x - 2.5 t y - 2.5 t z, -2.5 t w + 2. t x - 1.5 t y}, {-2.5 t w + 2. t x - 1.5 t z, 1 + 0. t w + 2. t y + 2. t z}} == 0
```

$$M = \begin{pmatrix} -3\,t & -2.5\,t & -2.5\,t & 0 \\ 2\,t & -1.5\,t & 0 & -2.5\,t \\ 2\,t & 0 & -1.5\,t & -2.5\,t \\ 0 & 2\,t & 2\,t & 0 \end{pmatrix}$$

```
{{-3 t, -2.5 t, -2.5 t, 0}, {2 t, -1.5 t, 0, -2.5 t}, {2 t, 0, -1.5 t, -2.5 t}, {0, 2 t, 2 t, 0}}
```

$$T = \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \end{pmatrix};$$

```
LinearSolve[M, T] // MatrixForm
```

$$\begin{pmatrix} \frac{0.75}{t} \\ -\frac{0.25}{t} \\ -\frac{0.25}{t} \\ \frac{0.75}{t} \end{pmatrix}$$

$$\mathbf{K} = \begin{pmatrix} \frac{0.75`}{t} & -\frac{0.25`}{t} \\ -\frac{0.25`}{t} & \frac{0.75`}{t} \end{pmatrix}$$

$$\left\{\left\{\frac{0.75}{t}, -\frac{0.25}{t}\right\}, \left\{-\frac{0.25}{t}, \frac{0.75}{t}\right\}\right\}$$

Tr[K]

$$\frac{1.5}{t}$$

Eigenvalues[K]

$$\left\{\frac{0.5}{t}, \frac{1.}{t}\right\}$$

SingularValueList$\left[\begin{pmatrix} \mathbf{-1.5} & \mathbf{2} \\ \mathbf{-2.5} & \mathbf{0} \end{pmatrix}\right]$

{3.16228, 1.58114}

Inverse$\left[\begin{pmatrix} \mathbf{-1.5\,t} & \mathbf{2\,t} \\ \mathbf{-2.5\,t} & \mathbf{0} \end{pmatrix}\right]$

$$\left\{\left\{0, -\frac{0.4}{t}\right\}, \left\{\frac{0.5}{t}, -\frac{0.3}{t}\right\}\right\}$$

SingularValueList$\left[\begin{pmatrix} \mathbf{0} & \mathbf{-0.4} \\ \mathbf{0.5} & \mathbf{-0.3} \end{pmatrix}\right]$

{0.632456, 0.316228}

REFERENCES

[1]. **A. Ralston and Philip Rabinowitz**, *A First course in Numerical Analysis*, MCGRAW-HILL BOOK COMPANY, 2nd edition,1978

[2]. **S.Balchandra Rao and C.K.Shantha,** *Numerical methods with Programs in BASIC, FORTAN, and PASCAL,* University Press India Ltd.

[3]. **B.N.Datta,** *Numerical Linear Algebra and Applications,*2nd edition, Society for Industrial and Applied Mathematics.

[4]. **G.C.F. Francis,** *The QR Transformation; part I and II,* Computer Journal,V.4(1961-1962).

[5]. **A. R. Horn and C. R. Jhonson,** *Matrix Analysis,* Cambridge University Press,1985.

[6]. **R. Horn and C. Johnson**, *Topics in Matrix Analysis,* Cambridge University Press, Cambridge, 1991

[7]. **Vera N.Kublanovskaya**,*On Some Algorithms for the solution of the complete Eigenvalue Problem,* USSR Computational Mathematics and Mathematical Physics,V.1(3),1963.

[8]. **G.H. Golub and H. A. Van Der Vorst,** *Eigenvalue Computation in the 2Oth Century,* Journal of Computational and Applied Mathematics,V.123,2000.

[9]. **B.Noble,** *Applied Linear Algebra***,** Prentice-Hall, Inc, New Jersey.

[10]. **S.S. Sastry,** *Introductory Methods of Numerical Analysis,* Prentice-Hall of India Pvt.Ltd, New Delhi,2000.

[11]. **B.R.Bhatt and S.Chawraverty,** *Nummerical Analysis in Engineering.*

[12] **B.Bradie,** *A friendly Introduction to Numerical Analysis,*1st edition, Pearson Education,Inc,2007.

[13]. **S.Lipschutz,** *Schum's outline of Theory and Problems of Linear Algebra,* MCGRAW-HILL Companies, New York,1991.

[14]. **A.R. Vasistha and Vipin Vasistha,** *Numerical Analysis,* 4th edition,2004.

[15]. **J.H.Wilkinson,** *The Algebraic Eigenvalue Problem,* Claredon Press,Oxford.

[16]. **J.N.Sharma,** *Numerical Methods For Engineers and Scientists,* Narosa Publishing House,6th edition,2004.

[17]. **C.F. Gerald and P.O. Whealty,** *Applied Numerical Analysis,*Pearson Education,Asia,6th edition,2002.

[18]. **W. Kwon, Y. Moon, and S. Ahn**, *Bounds in algebraic Riccati and Lyapunov equations: A survey and some new results*, International Journal of Control, vol. 64, no. 3, pp. 377-389.

[19]. **G.W. Stewart**, *Introduction to Matrix Computations,* Academic Press, New York,1973.

[20]. **J. S. Pizer,** *Numerical Computing and Mathematical Analysis,* Chicago Science Research Associates,1975.

[21]. **L. Fox**, *An Introduction to Numerical Linear Analysis,* Oxford University Press, New York,1965.

[22]. **G.H. Golub and C.F.Van Loan,** *Matrix Computations,* Johns Hopkins University Press,3rd edition,1996.

[23]. **G.W. Stewart**, *Matrix Algorithms, Volumn 1:Basic Decompositions,* Society for Industrial and Applied Mathematics(SIAM),Philadelphia, 1998.

[24]. **B.N.Datta,** *Numerical Methods For Linear Control Systems,* Elsevier Science & Technology Books,2003.

[25]. **H. Anton and C. Rorres,** *Elementary Linear Algebra with Applications*, John Wiley, New York, 1987.

[26]. **Golub, Kahan, William**, *Calculating the singular values and pseudo-inverse of a matrix,* Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis 2 (2): 205–224, 1965.

[27]. **S.J. Leon,** *Linear Algebra with Applications*, 4th edition, Macmillan, New York, 1994.

[28]. **G. Strang,** *Linear Algebra and Its Applications,* 3rd edition, Academic Press, New York, 1988

[29]. **W.L. Brogan,** *Modern Control Theory,* 3rd edition, Prentice Hall, Englewood Cliffs, NJ, 1991.

[30]. **C.T. Chen**, *Linear Systems Theory and Design,* CBS College Publishing, New York, 1984.

[31]. **G.W. Stewart**,*On the early History of the Singular Value Decomposition, JSTOR*. 35.4 (1993): 551-566. Web. 7 Mar. 2013.

[32]. **S.H. Lehnigk ,** *Stability Theorems for Linear Motions with an Introduction to Lyapunov's Direct Method,* Prentice Hall, Englewood Cliffs, NJ, 1966.

[33]. **B.N. Datta,** *Stability and Inertia*, Lin. Alg. Appl., Vol. 302/303, pp. 563-600, 1999.

[34]. **D.Kressner**, *Numerical methods for general and structured eigenvalue problems*, Springer-Verlag , New York, 2005.

[35]. **G. E. Dullerud and F. G. Paganini**, *A course in robust control theory: A convex approach,* Springer-Verlag, New York, 2000.

[36]. **J. Distefano, A. Stubberud, and I. Williams,** *Feedback and control systems: continuous (analog) and discrete (digital)*, Schaums outline series, McGraw-Hill, New York, 1990.

[37]. **S. G. Savoy and I P. Popchev,** *New Upper Bounds for the CALE: A Singular Value Decomposition Approach,* International Journal of Control, Automation, and Systems, vol. 6, no. 2, pp. 288-294, April 2008.

[38]. **G.W. Stewart**, *Perturbation Theory For The Singular Value Decomposition,* UMIACS-TR-90-124, CS-TR-2539, September, 1990.

[39]. **T. Mori and I. Derese,** *A brief summary of the bounds on the solution of the algebraic matrix equations in control theory,* International Journal of Control, Vol. 39(2), pp. 247-256, February 1984.

[40]. **T. Mori and H. Kokame,** *On solution bounds for three types of Luapunov matrix equations: Continuous, discrete and unified equations,* IEEE Trans. on Automatic Control, vol. 47, no. 10, pp. 1767-1770, October 2002.

[41]. **S. Savov and I. Popchev**, *New upper estimates for the solution of the CALE,* IEEE Trans. on Automatic Control, vol. 49, no. 10, pp. 18411842, October 2004.

[42]. **Y. Fang, K. Loparo, and X. Feng,** *New estimates for solutions of Lyapunov equations*, IEEE Trans. on Automatic Control, vol. 42, no. 3, pp. 408-411, March 1997.

[43]. **N. Komaroff**, *Upper summation and product bounds for solution eigenvalues of the Lyapunov matrix equation,* IEEE Trans. on Automatic Control, vol. 37, no. 5, pp. 337-341, May 1992.

[44]. **Yan-Bin.Jia,** *Singular Value Decomposition,* Com S 477/577 Notes, 6th September, 2012.

[45]. **J.E LaSalle and S. Lefschetz ,** *Stability by Lyapunov's Direct Method with Applications,* Academic Press, New York, 1961.