CHAPTER 2

PRELIMINARIES

In this chapter, we give some definitions, notations and theorems that will be used in the later chapters. Throughout this thesis, our scalar field is the field of real numbers $\mathbb{R}$ and we let $\mathbb{N}$ denote the set of all natural numbers.

2.1 Sequences and Series

Definition 2.1.1 Let (x_n) be a sequence. We say that (x_n) *approaches the limit* L (as n approaches to infinity), if for any $\epsilon > 0$, there is a positive integer N_ϵ such that

$$|x_n - L| < \epsilon \quad \text{for all} \quad n \geq N_\epsilon. \tag{2.1}$$

We write $\lim_{n\to\infty} x_n = L$ or $x_n \to L$ as $n \to \infty$.

Definition 2.1.2 If (x_n) is a sequence having the limit L, we say that (x_n) is a *convergent* sequence. If (x_n) is not convergent, we say that (x_n) is *divergent*.

Definition 2.1.3 Let $\sum_{n=1}^{\infty} a_n$ be a series of scalar with partial sum $S_n = a_1 + a_2 + ... + a_n (n \in \mathbb{N})$, if the sequence (S_n) converge to S, we say that the series $\sum_{n=1}^{\infty} a_n$ *converges* to S. If (S_n) diverges, we say that $\sum_{n=1}^{\infty} a_n$ *diverges*.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่

2.2 Metric Spaces and Normed Spaces

Copyright © by Chiang Mai University
All rights reserved

Definition 2.2.1 A *metric space* is a pair (X, d), where X is a nonempty set and d a metric on X, that is $d : X \times X \to \mathbb{R}$ is a function satisfies the following conditions:

(M1) $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}, \mathbf{y} \in X$

(M2) $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in X$

(M3) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in X$

(M4) $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$.

Definition 2.2.2 A sequence $(\mathbf{x}_n)$ in metric space (X,d) is said to *converge* or to be *convergent* if there is an $\mathbf{x} \in X$ such that $\lim_{n\to\infty} d(\mathbf{x}_n, \mathbf{x}) = 0$, $\mathbf{x}$ is called *the limit* of $\mathbf{x}_n$ and we write $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$, or simply $\mathbf{x}_n \to \mathbf{x}$.

Definition 2.2.3 A sequence $(\mathbf{x}_n)$ in a metric space (X,d) is called *Cauchy sequence* if for every $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that

$$d(\mathbf{x}_n, \mathbf{x}_m) < \epsilon$$

for all $n, m \geq N_\epsilon$.

Definition 2.2.4 A metric space (X,d) is said to be *complete* if every Cauchy sequence in X is convergent.

Definition 2.2.5 Let X be a linear space (or vector space), a *norm* on X is a real - valued function $\|\cdot\|$ which satisfies the following conditions:

(N1) $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in X$

(N2) $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$

(N3) $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ for all $\alpha \in \mathbb{R}$ and $\mathbf{x} \in X$

(N4) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in X$.

A linear space X equipped with a norm $\|.\|$ is called a *normed linear space.* Every normed linear space gives rise to the metric $d(\mathbf{x}, \mathbf{y})$. It is called the metric induced by norm.

For $\mathbf{x}^0$ in X and $r > 0$ the set $B(\mathbf{x}^0, r) = \{\mathbf{x} \in X : \|\mathbf{x} - \mathbf{x}^0\| < r\}$ is called the *open ball* of radius r and center $\mathbf{x}^0$. Correspondingly, the set $\overline{B(\mathbf{x}^0, r)} = \{\mathbf{x} \in X : \|\mathbf{x} - \mathbf{x}^0\| \leq r\}$ is called the *closed ball.*

Definition 2.2.6 Let $\mathbf{x} \in \mathbb{R}^n$ (n=1,2,...). For $\mathbf{x}=(x_1, \ldots, x_n) \in \mathbb{R}^n$, let

$$\begin{aligned} \|\mathbf{x}\|_1 &= \Sigma_{i=1}^n |x_i|, \\ \|\mathbf{x}\|_\infty &= \max\{|x_i|, \ldots, |x_n|\}, \\ \|\mathbf{x}\|_p &= (\Sigma_{i=1}^n |x_i|^p)^{\frac{1}{p}}, (p > 1). \end{aligned}$$

Definition 2.2.7 A *Banach space* is a complete normed linear space.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Definition 2.2.8 A subset D of a linear space X is called *convex* if $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in D$ for all $\mathbf{x}, \mathbf{y} \in D$ and all $\lambda \in [0,1]$.

2.3 Matrices and Linear Transformations

The set of $m \times n$ matrices with element in $\mathbb{R}$ is denoted $\mathbb{R}^{m\times n}$.
A matrix $A \in \mathbb{R}^{m\times n}$ is *square* if $m = n$, *rectangular* otherwise.
The element of a matrix $A \in \mathbb{R}^{m\times n}$ are denoted by $a_{i,j}$ or $A[i,j]$.
The matrix A is *diagonal* if $A[i,j] = 0$ for $i \neq j$.
An $m \times n$ diagonal matrix A is denoted $A = diag(a_{11}, a_{22}, \ldots, a_{pp})$ where $p = min(m,n)$. Given a matrix $A \in \mathbb{R}^{m\times n}$, its *transpose* is the matrix $A^T \in \mathbb{R}^{n\times m}$ with $A^T[i,j] = A[j,i]$.
A square matrix is *symmetric* if $A = A^T$ and is *orthogonal* if $A^T = A^{-1}$.

Definition 2.3.1 If $F : V \to W$ is the function from the vector space V into the vector space W, then F is called a *linear transformation* if

(a) $F(\mathbf{u}+\mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v})$,$\forall \mathbf{u}, \mathbf{v} \in V$

(b)$F(k\mathbf{u}) = kF(\mathbf{u})$,$\forall \mathbf{u} \in V$ and all scalars k.

2.4 Newton's method

We now want to determine zeros of a function of n variables; i.e., we want to solve equations of the form

$$\mathbf{f}(\mathbf{x}) = 0, \tag{2.2}$$

where $\mathbf{f} : D \to \mathbb{R}^n$ is a continuously differentiable function defined on some open subset $D \subset \mathbb{R}^n$. We begin by considering a function of one variable. Let $\mathbf{x}^0$ be an approximation to a zero of the function $\mathbf{f}$. In a neighborhood of $\mathbf{x}^0$, by Taylor's formula we have that

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= \mathbf{f}(\mathbf{x}^0) + \mathbf{f}'(\mathbf{x}^0)(\mathbf{x}-\mathbf{x}^0) + \frac{\mathbf{f}''(\mathbf{x}^0)(\mathbf{x}-\mathbf{x}^0)^2}{2!} + \ldots \\ &\approx \mathbf{f}(\mathbf{x}^0) + \mathbf{f}'(\mathbf{x}^0)(\mathbf{x}-\mathbf{x}^0) = \mathbf{g}(\mathbf{x}). \qquad (2.3) \end{aligned}$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Therefore, we may consider the zero of the affine linear function $\mathbf{g}$ as a new approximation to the zero of $\mathbf{f}$ and denote it by $\mathbf{x}^1$.
From the linear equation

$$\mathbf{f}(\mathbf{x}^0) + \mathbf{f}'(\mathbf{x}^0)(\mathbf{x}^1 - \mathbf{x}^0) = 0 \tag{2.4}$$

we immediately obtain

$$\mathbf{x}^1 = \mathbf{x}^0 - \frac{\mathbf{f}(\mathbf{x}^0)}{\mathbf{f}'(\mathbf{x}^0)}.$$

Geometrically, the affine linear function $\mathbf{g}$ describes the tangent line to the graph of the function $\mathbf{f}$ at the point $\mathbf{x}^0$. This consideration can be extended to the case of more than one variable. Given an approximation $\mathbf{x}^0$ to a zero of $\mathbf{f}$, by Taylor's formula we still have the approximation (2.3), where now, as in the previous section,

$$J(\mathbf{x}) = (\frac{\partial \mathbf{f}_j}{\partial x_k})_{j,k=1,\ldots,n}$$

denotes the Jacobian matrix of $\mathbf{f}$. Again we obtain a new approximation $\mathbf{x}^1$ for the solution of $\mathbf{f}(\mathbf{x}) = 0$ by solving the linearized equation (2.4), by

$$\mathbf{x}^1 = \mathbf{x}^0 - [J(\mathbf{x}^0)]^{-1}\mathbf{f}(\mathbf{x}^0).$$

Geometrically, the function $\mathbf{g}$ of (2.3) corresponds to the hyperplane tangent to $\mathbf{f}$ at the point $\mathbf{x}^0$.

Definition 2.4.1 Let $D \subset \mathbb{R}^n$ be open and let $\mathbf{f} : D \to \mathbb{R}^n$ be a continuously differentiable function such that the Jacobian matrix $\mathbf{f}'(\mathbf{x})$ is nonsingular for all $\mathbf{x} \in D$. Then Newton's method for the solution of the equation

$$\mathbf{f}(\mathbf{x}) = 0$$

is given by the iteration scheme

$$\mathbf{x}^{n+1} = \mathbf{x}^n - [J(\mathbf{x}^n)]^{-1}\mathbf{f}(\mathbf{x}^n), \;\; n = 0, 1, \ldots, \tag{2.5}$$

starting with some $\mathbf{x}^0 \in D$.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Theorem 2.4.2 Let $D \subset \mathbb{R}^n$ be open and convex and let $\mathbf{f} : D \to \mathbb{R}^n$ be a continuously differentiable. Assume that for some norm $\|\cdot\|$ on $\mathbb{R}^n$ and some $\mathbf{x}^0 \in D$ the following conditions hold:

(a) $\mathbf{f}$ satisfies

$$\|J(\mathbf{x}) - J(\mathbf{y})\| \le \gamma\|\mathbf{x} - \mathbf{y}\| \tag{2.6}$$

for all $\mathbf{x}, \mathbf{y} \in D$ and some constant $\gamma > 0$.

(b) The Jacobian matrix $J(\mathbf{x})$ is nonsingular for all $\mathbf{x} \in D$, and there exists a constant $\beta > 0$ such that

$$\|[J(\mathbf{x})]^{-1}\| \le \beta, \ \ \mathbf{x} \in D. \tag{2.7}$$

(c) For the constants

$$\alpha = \|[J(\mathbf{x}^0)]^{-1}\mathbf{f}(\mathbf{x}^0)\| \text{ and } \ q = \alpha\beta\gamma \tag{2.8}$$

the inequality

$$q < \frac{1}{2} \tag{2.9}$$

is satisfied.

(d) For $r = 2\alpha$ the closed ball $\overline{B(\mathbf{x}^0, r)} = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}^0\| \le r\}$ is contained in D.

Then $\mathbf{f}$ has a unique zero $\mathbf{x}^*$ in $\overline{B(\mathbf{x}^0, r)}$. Starting with $\mathbf{x}^0$ the Newton iteration

$$\mathbf{x}^{n+1} = \mathbf{x}^n - [J(\mathbf{x}^n)]^{-1}\mathbf{f}(\mathbf{x}^n), \ \ n = 0, 1, \ldots, \tag{2.10}$$

is well-defined. The sequence $(\mathbf{x}^n)$ converges to the zero $\mathbf{x}^*$ of $\mathbf{f}$, and we have the error estimate

$$\|\mathbf{x}^n - \mathbf{x}^*\| \le 2\alpha q^{2^n - 1}, \ \ n = 0, 1, \ldots. \tag{2.11}$$

Proof. See [5].

Definition 2.4.3 A convergent sequence $(\mathbf{x}^n)$ from a normed space with limit $\mathbf{x}$ is said to be *convergent of order* $p \ge 1$ if there exists a constant $C > 0$ such that

$$\|\mathbf{x}^{n+1} - \mathbf{x}\| \le C\|\mathbf{x}^n - \mathbf{x}\|^p, \ \ \ n = 1, 2, \ldots.$$

Convergence of order one or two is also called *linear* or *quadratic convergence*, respectively.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Example 2.4.4 Solve the nonlinear system

$$f_1(x_1, x_2) = x_1^3 + 3x_2^2 - 21 = 0$$
$$f_2(x_1, x_2) = x_1^2 + 2x_2 + 2 = 0$$

by Newton's method starting, with the initial estimate $\mathbf{x}^0 = (x_1^0, x_2^0) = (1, -1)$. Iterate until $\|\mathbf{x}^k - \mathbf{x}^{k-1}\|_\infty \le 10^{-6}$.

The Jacobian matrix is

$$J(x_1, x_2) = \begin{bmatrix} 3x_1^2 & 6x_2 \\ 2x_1 & 2 \end{bmatrix}.$$

At the point $(1, -1)$ the function vector and the Jacobian matrix take on the values

$$F(1, -1) = \begin{bmatrix} -17 \\ 1 \end{bmatrix}, J(1, -1) = \begin{bmatrix} 3 & -6 \\ 2 & 2 \end{bmatrix}.$$

The differentials $\triangle x_1^0$ and $\triangle x_2^0$ are the solution of the system

$$\begin{bmatrix} 3 & -6 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} \triangle x_1^0 \\ \triangle x_2^0 \end{bmatrix} = - \begin{bmatrix} -17 \\ 1 \end{bmatrix}.$$

Its solution is

$$\triangle \mathbf{x}^0 = \begin{bmatrix} \triangle x_1^0 \\ \triangle x_2^0 \end{bmatrix} = \begin{bmatrix} 1.555556 \\ -2.055556 \end{bmatrix}.$$

Thus, the next point of iteration is

$$\mathbf{x}^1 = \begin{bmatrix} x_1^1 \\ x_2^1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1.555556 \\ -2.055556 \end{bmatrix} = \begin{bmatrix} 2.555556 \\ -3.055556 \end{bmatrix}.$$

Similarly, the next three points are

$$\mathbf{x}^2 = \begin{bmatrix} 1.865049 \\ -2.500801 \end{bmatrix}, \mathbf{x}^3 = \begin{bmatrix} 1.661337 \\ -2.359271 \end{bmatrix} \text{ and } \mathbf{x}^4 = \begin{bmatrix} 1.643173 \\ -2.349844 \end{bmatrix}.$$

The results are summarized in the table

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Iteration k	$\mathbf{x}^k$	$\lVert\mathbf{x}^k - \mathbf{x}^{k-1}\rVert_\infty$
0	[1,-1]	-
1	[2.555556,-3.055556]	2.055556
2	[1.865049,-2.500801]	0.690507
3	[1.661337,-2.359271]	0.203712
4	[1.643173,-2.349844]	0.018164
5	[1.643038,-2.349787]	0.000135
6	[1.643038,-2.349787]	7.3×10^{-9}

2.5 Generalized Inverses of Matrices

A matrix has an inverse only if it is *square*, and even then only if it is *nonsingular*, or, in other words, if its columns (or rows) are linearly independent. By a *generalized inverse* of a given matrix A, we shall mean a matrix X associated in some way with A that

(i) exists for a class of matrices larger than the class of nonsingular matrices,

(ii) has some of the properties of the usual inverse, and

(iii) reduces to the usual inverse when A is nonsingular.

In theory, there are many different generalized inverses that exist. We shall consider some of these.

Definition 2.5.1 Let A be an m by n matrix. The *Moore-Penrose generalized inverse* of A, denoted $A^\dagger$, is the unique n by m matrix X which satisfies the four Penrose conditions:

$$\begin{aligned} AXA &= A, \\ XAX &= X, \\ (AX)^T &= AX, \text{ and} \\ (XA)^T &= XA. \end{aligned}$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

Definition 2.5.2 A *{2}-inverse* (also *outer inverse*) of $A \in \mathbb{R}^{m\times n}$ is a matrix $X \in \mathbb{R}^{n\times m}$ satisfying $XAX = X$, in which case rank$X \leq$ rank$A$, with equality if $X = A^{\dagger}$. We say that X is a *low rank* [*high rank*]*{2}-inverse* of A if its rank is near 0 [near rank$A$], respectively.

2.6 Gram-Schmidt Orthogonalization

Definition 2.6.1 Let X be a complex (or real) linear space. Then a function $\langle\cdot,\cdot\rangle : X \times X \longrightarrow \mathbb{C}$(or $\mathbb{R}$) with the properties;

(H1) $\langle \mathbf{x},\mathbf{x}\rangle \geq 0$, (positivity)

(H2) $\langle \mathbf{x},\mathbf{x}\rangle = 0$ if and only if $\mathbf{x} = 0$, (definiteness)

(H3) $\langle \mathbf{x},\mathbf{y}\rangle = \overline{\langle \mathbf{y},\mathbf{x}\rangle}$, (symmetry)

(H4) $\langle \alpha\mathbf{x} + \beta\mathbf{y},\mathbf{z}\rangle = \alpha\langle \mathbf{x},\mathbf{z}\rangle + \beta\langle \mathbf{y},\mathbf{z}\rangle$, (linearity)

for all **x**,**y**,**z**$\in X$ and $\alpha, \beta \in \mathbb{C}$ (or $\mathbb{R}$) is called a *scalar product*, or an *inner product*, on X. (By the bar we denote the complex conjugate.) A linear space X equipped with a scalar product is called a *pre-Hilbert space*, or an *inner product space.*

The Gram-Schmidt orthogonalization procedure as described in the following theorem. For a subset U of a linear space X we denote the set *spanned* by all linear combinations of elements of U by span{U}.

Theorem 2.6.2 Let $\{u_0, u_1, \ldots\}$ be a finite or countable number of linearly independent elements of a pre-Hilbert space. Then there exists a uniquely determined orthogonal system $\{q_0, q_1, \ldots\}$ of the form

$$q_n = u_n + r_n,\ n = 0, 1, \ldots,$$

with $r_0 = 0$ and $r_n \in \mathrm{span}\{u_0, \ldots, u_{n-1}\}$, n = 1, 2, ..., satisfying

$$\mathrm{span}\{u_0, \ldots, u_n\} = \mathrm{span}\{q_0, \ldots, q_n\},\ n = 0, 1, \ldots.$$

Proof. See [5].

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

2.7 Singular Value Decomposition(SVD)

Theorem 2.7.1 (Singular Value Decomposition) Let A be an $n \times m$ matrix. Then there are orthogonal matrices U and V, of order m and n, respectively such that

$$V^T AU = D \tag{2.12}$$

is diagonal rectangular matrix of order $n \times m$

$$D = \begin{bmatrix} \mu_1 & & & 0 & \\ & \mu_2 & & & \vdots \\ & & \ddots & & \vdots \\ 0 & & & \mu_r & \\ & \cdots & \cdots & & \ddots \end{bmatrix}.$$

The number $\mu_1, \mu_2, \ldots, \mu_r$ are called the *singular value* of A. They are all real and positive, and they can be arranged so that

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_r > \mu_{r+1} = \cdots = \mu_n = 0$$

where r is the rank of the matrix A.

Proof. See [1].

Algorithm of Singular Value Decomposition

To find the singular value decomposition of the matrix $A \in \mathbb{R}^{m \times n}$ one has to:

(1) Find the eigenvalues of the matrix $A^T A$ and arrange them in descending order.

(2) Find the number of nonzero eigenvalues of matrix $A^T A$, say r.

(3) Find the orthonormal eigenvectors of the matrix $A^T A$ corresponding to the obtained eigenvalues, and arrange them in the same order of form the column-vectors of the matrix $V \in \mathbb{R}^{n \times n}$.

(4) Form a diagonal matrix $\sum \in \mathbb{R}^{m \times n}$ placing on the leading diagonal of it the square roots $\sigma_i = \sqrt{\lambda_i}$ of $p = \min\{m, n\}$ first eigenvalues of matrix $A^T A$ got in (1) in descending order.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

(5) Find the first column-vectors of the matrix $U \in \mathbb{R}^{m\times m}$:

$$\mathbf{u}_i = \sigma_i^{-1} A\mathbf{v}_i, \quad i = 1, ..., r. \tag{2.13}$$

(6) Add to the matrix U the rest of m-r vectors using the Gram-Schmidt orthogonalization process.

Example 2.7.2 Let us find the singular value decomposition of the matrix

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

1. Find the eigenvalues of the matrix

$$A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}:$$

$$det(A^T A - \lambda I) = 0 \Longleftrightarrow \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0$$

we have $\lambda_1 = 3, \lambda_2 = 1$.

2. Find the number of nonzero eigenvalues of the matrix $A^T A : r = 2$.

3. Find the orthonormal eigenvectors of the matrix $A^T A$ corresponding to the eigenvalues λ_1 and λ_2 :

$$\mathbf{v}_1 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$$

forming a matrix

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix}.$$

4. Find the singular value and the diagonal matrix $\Sigma \in \mathbb{R}^{3\times 2}$:

$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3}$ and $\sigma_2 = \sqrt{\lambda_2} = \sqrt{1} = 1$

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

on the leading diagonal of which are the square roots of the eigenvalues of the matrix A^TA(in descending order) and the rest of the entries of the matrix Σ are zeros.

5. Find the first two column-vectors of the matrix $U \in \mathbb{R}^{3\times 3}$using formula (2.13)

$$\mathbf{u}_1 = \sigma_1^{-1}A\mathbf{v}_1 = \frac{\sqrt{3}}{3}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} \sqrt{6}/3 \\ \sqrt{6}/6 \\ \sqrt{6}/6 \end{bmatrix}$$

and

$$\mathbf{u}_2 = \sigma_2^{-1}A\mathbf{v}_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}.$$

6. To find the vector $\mathbf{u}_3$ we shall first find, applying the Gram-Schmitd process, a vector $\acute{\mathbf{u}}_3$ perpendicular to $\mathbf{u}_1$ and $\mathbf{u}_2$:

$$\acute{\mathbf{u}}_3 = \mathbf{e}_1 - (\mathbf{u}_1^T\mathbf{e}_1)\mathbf{u}_1 - (\mathbf{u}_2^T\mathbf{e}_2)\mathbf{u}_2 = \begin{bmatrix} 1/3 & -1/3 & -1/3 \end{bmatrix}^T.$$

Norming the vector $\acute{\mathbf{u}}_3$, we get

$$\mathbf{u}_3 = \begin{bmatrix} \sqrt{3}/3 \\ -\sqrt{3}/3 \\ -\sqrt{3}/3 \end{bmatrix}.$$

Hence

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \sqrt{6}/3 & 0 & \sqrt{3}/3 \\ \sqrt{6}/6 & -\sqrt{2}/2 & -\sqrt{3}/3 \\ \sqrt{6}/6 & \sqrt{2}/2 & -\sqrt{3}/3 \end{bmatrix}$$

and the singular value decomposition of the matrix A is

$$A = \begin{bmatrix} \sqrt{6}/3 & 0 & \sqrt{3}/3 \\ \sqrt{6}/6 & -\sqrt{2}/2 & -\sqrt{3}/3 \\ \sqrt{6}/6 & \sqrt{2}/2 & -\sqrt{3}/3 \end{bmatrix}\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix}.$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved