Notes

2020년 7월 21일 화요일 오전 2:06

Algebraic Properties of $\mathbb{R}^n$

1. $u+v = v+u$
2. $(u+v)+w = u+(v+w)$
3. $u+0 = 0+u = u$
4. $u+(-u) = -u+u = 0$
5. $c(u+v) = cu+cv$
6. $(c+d)u = cu+du$
7. $c(du) = (cd)u$
8. $1u = u$

Chapter 1

1.2 row reduction and echelon forms

DEFINITION A **pivot position** in a matrix A is a location in A that corresponds to a leading 1 in the reduced echelon form of A. A **pivot column** is a column of A that contains a pivot position.

1.4

Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
c. The columns of A span $\mathbb{R}^m$.
d. A has a pivot position in every row.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 + .2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} .2x_3 \\ 0 \\ x_3 \end{bmatrix}$$

$$\boxed{= x_2 \underbrace{\begin{bmatrix} .3 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{u}} + x_3 \underbrace{\begin{bmatrix} .2 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}}} \quad \text{(with } x_2, x_3 \text{ free)} \qquad (2)$$

— parametric Vector Equation

This calculation shows that every solution of (1) is a linear combination of the vectors $\mathbf{u}$ and $\mathbf{v}$, shown in (2). That is, the solution set is Span $\{\mathbf{u}, \mathbf{v}\}$. Since neither $\mathbf{u}$ nor $\mathbf{v}$ is a scalar multiple of the other, the solution set is a plane through the origin. See Figure 2. ∎

1.7 _Linear Independence

DEFINITION An indexed set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution. The set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is said to be **linearly dependent** if there exist weights $c_1, \dots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \qquad (2)$$

DEFINITION A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **onto** $\mathbb{R}^m$ if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at least one* $\mathbf{x}$ in $\mathbb{R}^n$.

DEFINITION A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **one-to-one** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at most one* $\mathbf{x}$ in $\mathbb{R}^n$.

★ Theorem 12

THEOREM 12 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let A be the standard matrix for T. Then:

a. T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$;
b. T is one-to-one if and only if the columns of A are linearly independent.

if there are pivots for every row, it is onto.

chapter2 2.1 powers of matrix

THEOREM 2 Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

a. $A(BC) = (AB)C$ (associative law of multiplication)
b. $A(B + C) = AB + AC$ (left distributive law)
c. $(B + C)A = BA + CA$ (right distributive law)
d. $r(AB) = (rA)B = A(rB)$ for any scalar r
e. $I_mA = A = AI_n$ (identity for matrix multiplication)

WARNINGS:

1. In general, $AB \neq BA$.
2. The cancellation laws do *not* hold for matrix multiplication. That is, if $AB = AC$, then it is *not* true in general that $B = C$. (See Exercise 10.)
3. If a product AB is the zero matrix, you *cannot* conclude in general that either $A = 0$ or $B = 0$. (See Exercise 12.)

THEOREM 3 Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^T)^T = A$
b. $(A + B)^T = A^T + B^T$
c. For any scalar r, $(rA)^T = rA^T$
d. $(AB)^T = B^TA^T$

2.2 inverse of matrix

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

A matrix that is *not* invertible is sometimes called a **singular matrix**, and an invertible matrix is called a **nonsingular matrix**.

Theorem 4 : test to find out whether 2x2 matrix is invertible

THEOREM 4 Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

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

If $ad - bc = 0$, then A is not invertible.

THEOREM 5 If A is an invertible $n \times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

THEOREM 6

a. If A is an invertible matrix, then A^{-1} is invertible and
$$(A^{-1})^{-1} = A$$
b. If A and B are $n \times n$ invertible matrices, then so is AB, and the inverse of AB is the product of the inverses of A and B in the reverse order. That is,
$$(AB)^{-1} = B^{-1}A^{-1}$$
c. If A is an invertible matrix, then so is A^T, and the inverse of A^T is the transpose of A^{-1}. That is,
$$(A^T)^{-1} = (A^{-1})^T$$

THEOREM 7 An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.

ALGORITHM FOR FINDING A^{-1}

Row reduce the augmented matrix $[\,A \quad I\,]$. If A is row equivalent to I, then $[\,A \quad I\,]$ is row equivalent to $[\,I \quad A^{-1}\,]$. Otherwise, A does not have an inverse.

2.3 characterization of invertible matrices

THEOREM 8 The Invertible Matrix Theorem

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.
b. A is row equivalent to the $n \times n$ identity matrix.
c. A has n pivot positions.
d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
e. The columns of A form a linearly independent set.
f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
h. The columns of A span $\mathbb{R}^n$.
i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
j. There is an $n \times n$ matrix C such that $CA = I$.
k. There is an $n \times n$ matrix D such that $AD = I$.
l. A^T is an invertible matrix.

2.5 matrix factorization

ALGORITHM FOR AN LU FACTORIZATION

1. Reduce A to an echelon form U by a sequence of row replacement operations, if possible.
2. Place entries in L such that the *same sequence of row operations* reduces L to I.

THE LEONTIEF INPUT–OUTPUT MODEL, OR PRODUCTION EQUATION

$$\underset{\substack{\text{Amount}\\ \text{produced}}}{\mathbf{x}} = \underset{\substack{\text{Intermediate}\\ \text{demand}}}{C\mathbf{x}} + \underset{\substack{\text{Final}\\ \text{demand}}}{\mathbf{d}} \tag{4}$$

chapter 3
3.2 properties of determinants

THEOREM 3 Row Operations

Let A be a square matrix.

because A & B are row-equivalent

a. If a multiple of one row of A is added to another row to produce a matrix B, then $\det B = \det A$.
b. If two rows of A are interchanged to produce B, then $\det B = -\det A$.
c. If one row of A is multiplied by k to produce B, then $\det B = k \cdot \det A$.

$$\det A = \begin{cases} (-1)^r \cdot \begin{pmatrix}\text{product of}\\ \text{pivots in } U\end{pmatrix} & \text{when } A \text{ is invertible} \\ 0 & \text{when } A \text{ is not invertible} \end{cases}$$

THEOREM 5 If A is an $n \times n$ matrix, then $\det A^T = \det A$.

THEOREM 6 Multiplicative Property

If A and B are $n \times n$ matrices, then $\det AB = (\det A)(\det B)$.

Warning: A common misconception is that Theorem 6 has an analogue for *sums* of matrices. However, $\det(A + B)$ is *not* equal to $\det A + \det B$, in general.

3.3 cramer's rule, volume, linear transformation

THEOREM 7 Cramer's Rule

Let A be an invertible $n \times n$ matrix. For any $\mathbf{b}$ in $\mathbb{R}^n$, the unique solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ has entries given by

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \quad i = 1, 2, \ldots, n \tag{1}$$

A Formula for A^{-1}

Cramer's rule leads easily to a general formula for the inverse of an $n \times n$ matrix A. The jth column of A^{-1} is a vector $\mathbf{x}$ that satisfies

$$A\mathbf{x} = \mathbf{e}_j$$

where $\mathbf{e}_j$ is the jth column of the identity matrix, and the ith entry of $\mathbf{x}$ is the (i, j)-entry of A^{-1}. By Cramer's rule,

$$\{(i, j)\text{-entry of } A^{-1}\} = x_i = \frac{\det A_i(\mathbf{e}_j)}{\det A} \tag{2}$$

Recall that A_{ji} denotes the submatrix of A formed by deleting row j and column i. A cofactor expansion down column i of $A_i(\mathbf{e}_j)$ shows that

$$\det A_i(\mathbf{e}_j) = (-1)^{i+j} \det A_{ji} = C_{ji} \tag{3}$$

where C_{ji} is a cofactor of A. By (2), the (i, j)-entry of A^{-1} is the cofactor C_{ji} divided by $\det A$. [Note that the subscripts on C_{ji} are the reverse of (i, j).] Thus

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \tag{4}$$

The matrix of cofactors on the right side of (4) is called the **adjugate** (or **classical adjoint**) of A, denoted by adj A. (The term *adjoint* also has another meaning in advanced texts on linear transformations.) The next theorem simply restates (4).

R

An Inverse Formula

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

$$A^{-1} = \frac{1}{\det A}\,\text{adj}\,A$$

THEOREM 9 If A is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|\det A|$. If A is a 3×3 matrix, the volume of the parallelepiped determined by the columns of A is $|\det A|$.

chapter 4
4.1 vector spaces and subspaces

DEFINITION A **vector space** is a nonempty set V of objects, called *vectors*, on which are defined two operations, called *addition* and *multiplication by scalars* (real numbers), subject to the ten axioms (or rules) listed below. The axioms must hold for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and for all scalars c and d.

1. The sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} + \mathbf{v}$, is in V.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. There is a **zero** vector $\mathbf{0}$ in V such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. For each $\mathbf{u}$ in V, there is a vector $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. The scalar multiple of $\mathbf{u}$ by c, denoted by $c\mathbf{u}$, is in V.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

DEFINITION A **subspace** of a vector space V is a subset H of V that has three properties:

a. The zero vector of V is in H.
b. H is closed under vector addition. That is, for each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u} + \mathbf{v}$ is in H.
c. H is closed under multiplication by scalars. That is, for each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

4.2

DEFINITION The **null space** of an $m \times n$ matrix A, written as Nul A, is the set of all solutions of the homogeneous equation $A\mathbf{x} = \mathbf{0}$. In set notation,

$$\text{Nul } A = \{\mathbf{x} : \mathbf{x} \text{ is in } \mathbb{R}^n \text{ and } A\mathbf{x} = \mathbf{0}\}$$

THEOREM 2 The null space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^n$. Equivalently, the set of all solutions to a system $A\mathbf{x} = \mathbf{0}$ of m homogeneous linear equations in n unknowns is a subspace of $\mathbb{R}^n$.

Contrast Between Nul A and Col A for an $m \times n$ Matrix A

Nul A	Col A
1. Nul A is a subspace of $\mathbb{R}^n$.	1. Col A is a subspace of $\mathbb{R}^m$.
2. Nul A is implicitly defined; that is, you are given only a condition ($A\mathbf{x} = \mathbf{0}$) that vectors in Nul A must satisfy.	2. Col A is explicitly defined; that is, you are told how to build vectors in Col A.
3. It takes time to find vectors in Nul A. Row operations on $[\,A \quad \mathbf{0}\,]$ are required.	3. It is easy to find vectors in Col A. The columns of A are displayed; others are formed from them.
4. There is no obvious relation between Nul A and the entries in A.	4. There is an obvious relation between Col A and the entries in A, since each column of A is in Col A.
5. A typical vector $\mathbf{v}$ in Nul A has the property that $A\mathbf{v} = \mathbf{0}$.	5. A typical vector $\mathbf{v}$ in Col A has the property that the equation $A\mathbf{x} = \mathbf{v}$ is consistent.
6. Given a specific vector $\mathbf{v}$, it is easy to tell if $\mathbf{v}$ is in Nul A. Just compute $A\mathbf{v}$.	6. Given a specific vector $\mathbf{v}$, it may take time to tell if $\mathbf{v}$ is in Col A. Row operations on $[\,A \quad \mathbf{v}\,]$ are required.
7. Nul $A = \{\mathbf{0}\}$ if and only if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.	7. Col $A = \mathbb{R}^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^m$.
8. Nul $A = \{\mathbf{0}\}$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.	8. Col $A = \mathbb{R}^m$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ *onto* $\mathbb{R}^m$.

DEFINITION A **linear transformation** T from a vector space V into a vector space W is a rule that assigns to each vector $\mathbf{x}$ in V a unique vector $T(\mathbf{x})$ in W, such that

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in V, and
(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for all $\mathbf{u}$ in V and all scalars c.

DEFINITION Let H be a subspace of a vector space V. An indexed set of vectors $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ in V is a **basis** for H if

(i) $\mathcal{B}$ is a linearly independent set, and
(ii) the subspace spanned by $\mathcal{B}$ coincides with H; that is,

$$H = \text{Span}\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$$

THEOREM 6 The pivot columns of a matrix A form a basis for Col A.

is equivalent to

$$\boxed{\mathbf{x} = P_{\mathcal{B}}[\,\mathbf{x}\,]_{\mathcal{B}}} \qquad (4)$$

We call $P_{\mathcal{B}}$ the **change-of-coordinates matrix** from $\mathcal{B}$ to the standard basis in $\mathbb{R}^n$. Left-multiplication by $P_{\mathcal{B}}$ transforms the coordinate vector $[\,\mathbf{x}\,]_{\mathcal{B}}$ into $\mathbf{x}$. The change-of-coordinates equation (4) is important and will be needed at several points in Chapters 5 and 7.

Since the columns of $P_{\mathcal{B}}$ form a basis for $\mathbb{R}^n$, $P_{\mathcal{B}}$ is invertible (by the Invertible Matrix Theorem). Left-multiplication by $P_{\mathcal{B}}^{-1}$ converts $\mathbf{x}$ into its $\mathcal{B}$-coordinate vector:

$$P_{\mathcal{B}}^{-1}\mathbf{x} = [\,\mathbf{x}\,]_{\mathcal{B}}$$

DEFINITION If V is spanned by a finite set, then V is said to be **finite-dimensional**, and the **dimension** of V, written as dim V, is the number of vectors in a basis for V. The dimension of the zero vector space $\{\mathbf{0}\}$ is defined to be zero. If V is not spanned by a finite set, then V is said to be **infinite-dimensional**.

4.5 Rank

- The rank of A is the dimension of the column space of A.
- The nullity of A is the dimension of the null space of A.

"The Rank Theorem"
→ rank A + dim Nul A = n

THEOREM 15 Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be bases of a vector space V. Then there is a unique $n \times n$ matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ such that

$$[\,\mathbf{x}\,]_{\mathcal{C}} = \underset{\mathcal{C} \leftarrow \mathcal{B}}{P}[\,\mathbf{x}\,]_{\mathcal{B}} \qquad (4)$$

The columns of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ are the $\mathcal{C}$-coordinate vectors of the vectors in the basis $\mathcal{B}$. That is,

$$\underset{\mathcal{C} \leftarrow \mathcal{B}}{P} = [\,[\mathbf{b}_1]_{\mathcal{C}} \quad [\mathbf{b}_2]_{\mathcal{C}} \quad \cdots \quad [\mathbf{b}_n]_{\mathcal{C}}\,] \qquad (5)$$

Observe that the matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ in Example 2 already appeared in (7). This is not surprising because the first column of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ results from row reducing $[\,\mathbf{c}_1 \quad \mathbf{c}_2 \,\vdots\, \mathbf{b}_1\,]$ to $[\,I \,\vdots\, [\mathbf{b}_1]_{\mathcal{C}}\,]$, and similarly for the second column of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$. Thus

$$\boxed{[\,\mathbf{c}_1 \quad \mathbf{c}_2 \,\vdots\, \mathbf{b}_1 \quad \mathbf{b}_2\,] \sim [\,I \,\vdots\, \underset{\mathcal{C} \leftarrow \mathcal{B}}{P}\,]}$$

4.7

$[x]_C = \underset{C \leftarrow B}{P}[x]_B$ Change-of-coordinates matrix from B to C

$\left(\underset{C \leftarrow B}{P}\right)^{-1} = \underset{B \leftarrow C}{P}$

4.8

Casoratian : determinant of casorati matrix

★ If a casorati matrix is not invertible, the associated signals being tested may or may not be linearly dependent.

5.1 Eigenvectors and Eigenvalues

Eigenvalues : a nonzero vector x satisfying $Ax = \lambda x$, and λ (lambda) is the eigenvalue, a nontrivial solution for x
Eigen value 를 갖는다는 의미는 $(A - \lambda I)x = 0$ 를 만족하는 nontrivial solution 있다는 것으로, 이를 만족하는 모든 x-vector 의 모음을 eigenspace 라고 한다. (" consists of all the eigenvectors ~"

Theorem 1 : The eigenvalues of a triangular matrix are the entries on its main diagonal

The scalar λ is an eigenvalue of A if and only if the equation $(A - \lambda I)x = 0$ has a nontrivial solution, that is, if and only if the equation has a free variable.
↳ If the value 0 happens to be an eigenvalue, then the corresponding matrix A is not invertible.

Theorem 2 : If $V_1 \dots, V_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \dots \lambda_r$ of a square matrix A, the the set $\{V_1, \dots V_r\}$ is linearly independent.

Chapter 5.2 : The Characteristic Equation

Theorem : matrix A is invertible if and only if "0" is not an eigenvalue .... (etc..)

Theorem 3 : Properties of Determinants

a. A is invertible if and only if $\det A \neq 0$.
b. $\det AB = (\det A)(\det B)$

Theorem 3 : Properties of Determinants

a. A is invertible if and only if $\det A \neq 0$.
b. $\det AB = (\det A)(\det B)$
c. $\det A^T = \det A$
d. If A is triangular, then det A is the product of main diagonals
e. $\det A = (-1)^r$ (products of main diagonal)

A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if λ satisfies the characteristic equation $\det(A - \lambda I) = 0$

Multiplicity - check the example 3 and 4 on page 271

Theorem 4: If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

Similarity Transformation: $P^{-1}AP = B$

⇒ If $n \times n$ matrix A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues with the same multiplicities.

☆!! 1. Some matrices are not similar though they have the same eigenvalues.
2. Similarity is not the same as row equivalence.
3. Row operations usually change its eigenvalues.

Chapter 5.3 - Diagonalization

Diagonalization Theorem : - $n \times n$ matrix A is diagonalization if and only if A has n linearly independent eigenvectors.
- $n \times n$ matrix A with n distinct eigenvalues is diagonalizable

Diagonalizing Matrices : - Find the eigenvalues of A
- Find n linearly independent eigenvectors of A
- Construct P from the vectors in step 2.
- Construct D from the corresponding eigenvalues.

☆ it is essentia tha the order of the eigenvalues maches the order chosen for the columns of P

Theorem : An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Chapter 5.4 - Eigenvectors and linear Transformations

THEOREM 8 Diagonal Matrix Representation

Suppose $A = PDP^{-1}$, where D is a diagonal $n \times n$ matrix. If $\mathcal{B}$ is the basis for $\mathbb{R}^n$ formed from the columns of P, then D is the $\mathcal{B}$-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$.

THEOREM 9 Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ $(b \neq 0)$ and an associated eigenvector $\mathbf{v}$ in $\mathbb{C}^2$. Then

$$A = PCP^{-1}, \quad \text{where} \quad P = [\,\mathrm{Re}\,\mathbf{v} \quad \mathrm{Im}\,\mathbf{v}\,] \quad \text{and} \quad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

6.1 Inner product, length, and orthogonality

DEFINITION The **length** (or **norm**) of $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$$

DEFINITION For $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, the **distance between u and v**, written as $\mathrm{dist}(\mathbf{u}, \mathbf{v})$, is the length of the vector $\mathbf{u} - \mathbf{v}$. That is,

$$\mathrm{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

DEFINITION Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\mathbf{u}\cdot\mathbf{v} = 0$.

THEOREM 2 The Pythagorean Theorem

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

1. A vector $\mathbf{x}$ is in $W^\perp$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans W.
2. $W^\perp$ is a subspace of $\mathbb{R}^n$.

Theorem 3 : $(\mathrm{Row}\ A)^\perp = \mathrm{Nul}\ A$, $(\mathrm{Col}\ A)^\perp = \mathrm{Nul}\ A^T$

$u \cdot v = \|u\| \cdot \|v\| \cos\theta$

Definition : An orthogonal basis for a subspace W of R^n is a basis for W that is also an orthogonal set.

Theorem 5 : Let $\{u_1, \ldots, u_p\}$ be an orthogonal basis, $y = c_1 u_1 + \cdots + c_p u_p$
Then, $c_j = \dfrac{y \cdot u_j}{u_j \cdot u_j}$ for all j

Theorem 6 : Matrix U has orthonormal columns if and only if $U^T U = I$

Theorem 7: U : $m \times n$ matrix, x and y in $\mathbb{R}^n$

Theorem 6 : Matrix U has orthonormal columns if and only if $U^TU = I$

Theorem 7 : U : $m \times n$ matrix, x and y in $\mathbb{R}^n$

- $\|Ux\| = \|x\|$
- $(Ux)\cdot(Uy) = x\cdot y$
- $(Ux)\cdot(Uy) = 0$ if and only if $x\cdot y = 0$

Orthogonal Matrix : $U^{-1} = U^T$

Theorem 8 :

Theorem 9 : The Best Approximation Theorem

$\|y - \hat{y}\| < \|y - v\|$

Theorem 10 : Following equations . . .

1. $\hat{y} = \dfrac{y\cdot u_1}{u_1\cdot u_1}u_1 + \cdots + \dfrac{y\cdot u_p}{u_p\cdot u_p}u_p$

2. $\mathrm{proj}_W y = \; = (y\cdot u_1)u_1 + (y\cdot u_2)u_2 + \cdots + (y\cdot u_p)u_p$

3. $U = [u_1 \; \dots \; u_p]$,

$\mathrm{Proj}_W y = \hat{y} = UU^Ty$

Suppose U is an $n\times p$ matrix with orthonormal columns, and let W be the column space of U. Then,

$U^TUx = I_px = x$

$UU^Ty = \mathrm{proj}_W y$

6.4 The Gram-Schmidt Process

Theorem 11 : Gram-Schmidt Process

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2 v_1}{v_1 v_1}v_1$$

$$v_3 = x_3 - \frac{x_3 v_1}{v_1 v_1}v_1 - \frac{x_3 v_2}{v_2 v_2}v_2$$

$$\vdots$$

$\{v_1, \dots v_p\}$ is an orthogonal basis for W.

Also, $\mathrm{Span}\{v_1 \dots v_k\} = \mathrm{Span}\{x_1 \dots x_k\}$ for $1 \le k \le p$

Theorem 12 : QR Factorization

A : $m\times n$, linearly independent

$A = QR$, $Q = m\times n$, columns being orthonormal basis for Col A

$R = n\times n$ upper triangular invertible matrix with positive entries on its diagonal.

$Q = [v_1 \; \dots \; v_p]$

$R = Q^TA = Q^T(QR) = IR = R$

6.5 Least-Squares Problems

Theorem 14 : Following statements are logically equivalent.

1. $Ax = b$ has a unique least-squares solution for each b in $\mathbb{R}^m$
2. The columns of A are linearly independent.
3. A^TA is invertible.

When these statements are true, the least-squares solution $\hat{x}$ is given by

$\hat{x} = (A^TA)^{-1}A^Tb$

Theorem 15 : let $A = QR$, (QR-factorization)

then $Ax = b$ has solutions,

$\hat{x} = R^{-1}Q^Tb$

6.6 Applications to Linears Models

6.7 Inner Product Spaces

DEFINITION An **inner product** on a vector space V is a function that, to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V, associates a real number $\langle \mathbf{u}, \mathbf{v}\rangle$ and satisfies the following axioms, for all $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and all scalars c:

1. $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$
2. $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$
3. $\langle c\mathbf{u}, \mathbf{v}\rangle = c\langle \mathbf{u}, \mathbf{v}\rangle$
4. $\langle \mathbf{u}, \mathbf{u}\rangle \ge 0$ and $\langle \mathbf{u}, \mathbf{u}\rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$

A vector space with an inner product is called an **inner product space**.