

데이터과학을 위한 선형대수학

(Linear Algebra for Data Science)

강완모
한국과학기술원
wanmo.kang@kaist.ac.kr

조경현
뉴욕대학교
Genentech
kyunghyun.cho@nyu.edu

Version 1.7
February 25, 2024

Contents

Preface	v
1 Introduction	1
2 Matrices and Gaussian Elimination	5
2.1 Matrix Operations	5
2.2 Solving Simultaneous Linear Equations	7
2.3 An Example of Gaussian Elimination	9
2.4 Block Matrices	11
2.5 Inverse of a Matrix	13
2.5.1 Triangular Factors	15
2.5.2 LU-Factorization	18
2.6 Inverse of Block Matrices	19
2.7 Application to Data Science: Graphs and Matrices	20
3 Vector Spaces	23
3.1 Vector Spaces and Subspaces	24
3.1.1 Operations in a Vector Space	24
3.1.2 Two Fundamental Subspaces induced by Matrices	26
3.2 Solving $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x} = \mathbf{b}$	27
3.2.1 Row Echelon Form U	27
3.2.2 Pivot Variables and Free Variables	29
3.3 Linear Independence, Basis, and Dimension	32
3.4 Rank of Matrices	35
3.5 The Four Fundamental Subspaces	37
3.6 Existence of Inverses	39
3.7 Matrices of Rank-one	42
3.8 Linear Transformations	43
3.8.1 Matrix Representation of Linear Transformations	44

3.8.2	Interpretable Linear Transformations	47
3.9	Application to Data Science: Neural Networks	49
3.9.1	Flexibility of Neural Network Representations	51
3.10	Application to Graph Theory	51
4	Orthogonality and Projections	55
4.1	Inner Products	56
4.2	Orthogonal Vectors and Subspaces	61
4.3	Orthogonal Projection	65
4.3.1	Projection onto the Direction of a Vector	65
4.3.2	Projection onto a Subspace Spanned by Orthonormal Vectors	67
4.3.3	Projection onto a Subspace Spanned by Independent Vectors	68
4.4	How to Build an Orthonormal Basis: Gram-Schmidt Procedure	70
4.4.1	Gram-Schmidt Procedure for Given Linearly Independent Vectors	71
4.4.2	Projection as a Distance-Minimization	72
4.5	Decomposition into Orthogonal Complements	73
4.6	Orthogonality in Euclidean Spaces \mathbb{R}^n	73
4.6.1	Orthogonal Complements of Fundamental Subspaces	74
4.7	Orthogonal Matrices	75
4.7.1	<i>QR</i> Decomposition	76
4.7.2	Isometry induced by an Orthogonal Matrix	76
4.8	Matrix Norms	77
4.9	Application to Data Science: Least Square and Projection	79
4.9.1	Least Square as a Convex Quadratic Minimization	80
4.9.2	Equivalence of Least Square and Projection	81
5	Singular Value Decomposition (SVD)	83
5.1	Variational Formulation for the Best-fit Subspaces	84
5.1.1	Best-fit 1-dimensional subspace	84
5.1.2	Best-fit 2-dimensional subspace	85
5.1.3	Best-fit k -dimensional subspace	86
5.2	Orthogonality of Left Singular Vectors	88
5.3	Representation of SVD in Various Forms	89
5.4	Properties of $\sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top$	91
5.5	Spectral Decomposition of Symmetric Matrices	93
5.6	Relationship between Singular Values and Eigenvalues	96
5.6.1	Symmetrization	97
5.7	Low rank approximation	98

5.8	Pseudoinverse	100
5.8.1	Generalized Projection and Least Squares	103
5.9	How to obtain SVD	104
5.10	Application to Statistics: PCA from SVD	106
6	SVD in Real Life	109
6.1	Compression of Images via SVD	109
6.2	Left-singular Vectors as the Coordinates of Embedding Vectors in the Latent Space	112
6.3	Approximation of MNIST Images via SVD	114
6.3.1	Geometry of MNIST Images in the Latent Space of Variational Auto-Encoder	117
6.4	Word Embeddings via SVD	118
6.5	Decomposition of Financial Time-Series via SVD	120
7	Positive Definite Matrices	123
7.1	Positive (Semi-)Definite Matrices	123
7.2	The Square Root of Positive Semi-definite Matrices	125
7.3	Cholesky Factorization of Positive Definite Matrices	126
7.4	Variational Characterization of Symmetric Eigenvalues	127
7.4.1	Eigenvalues and Singular Values of Matrix Sums	129
7.5	Ellipsoidal Geometry of Positive Definite Matrices	133
7.6	Application to Data Science: Kernel Trick in Machine Learning	134
8	Determinants	139
8.1	Definition and Properties	140
8.2	Formulas for the Determinant	144
8.2.1	Determinant of Block Matrices	147
8.2.2	Matrix Determinant Lemma	149
8.3	Applications of Determinant	150
8.3.1	Volume of a Parallelograms in \mathbb{R}^n	150
8.3.2	Computation of A^{-1}	151
8.3.3	Cramer's Rule: Solution of $A\mathbf{x} = \mathbf{b}$	152
8.4	Sherman-Morrison and Woodbury Formulas	152
8.5	Application to Optimization: Rank-one update of Hessians	153
9	Eigenvalues and Eigenvectors	157
9.1	Examples of Eigen Analysis	158
9.2	Some Properties of Eigenpairs	159
9.3	Similarity and Change of Basis	162
9.3.1	Change of Basis	162

9.3.2 Change of Orthonormal Bases	163
9.3.3 Similarity	163
9.4 Diagonalization	164
9.4.1 An Example of Non-diagonalizable Matrix	166
9.5 Alternative Proof of the Spectral Decomposition Theorem	167
9.6 How to obtain Eigenvalues and Eigenvectors	168
9.7 Application to Data Science: Power Iteration and Google PageRank	169
10 Advanced Results in Linear Algebra	173
10.1 Dual Space	173
10.2 Transpose of Matrices and Adjoint of Linear Transformations	174
10.2.1 Adjoint and Projection	177
10.3 Further Theory of Positive Definite Matrices	178
10.3.1 Congruence Transformations	178
10.3.2 Positive Semi-definite Cone and Partial Order	180
10.4 Schur triangularization	184
10.5 Perron-Frobenius Theorem	185
10.6 Eigenvalue Adjustments and Applications	187
11 Big Theorems in Linear Algebra	189
11.1 The First Big Theorem: Cayley-Hamilton Theorem	189
11.2 Decomposition of Nilpotency into Cyclic Subspaces	190
11.3 Nilpotency of $A - \lambda I$	194
11.4 The Second Big Theorem: Jordan Normal Form Theorem	196
12 Homework Assignments	199
13 Problems	207
13.1 Problems for Chapter 1 ~ 4	207
13.2 Problems for Chapter 5 ~ 8	210
Bibliography	215
Appendices	217
A Permutation and its Matrix Representation	217
B Existence of Optimizers	219
C Complex Numbers and Matrices	223

D Covariance Matrices	225
D.1 Covariance Matrix of Random Vector	226
D.1.1 Positive Definiteness of Covariance Matrices	227
D.1.2 A Useful Quadratic Identity	227
D.2 Multivariate Gaussian Distribution	228
D.3 Conditional Multivariate Gaussian Distribution	229
D.4 Multivariate Gaussian Sampling using Cholesky Decomposition	229
D.5 Ill-conditioned Sample Covariance Matrices	230
Index	231

Kang & Cho (2024)

Preface

계량적인 방법론을 사용하는 모든 학문분야에서 가장 자주 활용되는 수학 분야 중의 하나가 선형대수학이다. 선형대수학은 공학자와 기술자, 경제학자들의 연구와 개발에서 일상적인 도구가 되었다. 더우기 최근 급속도로 발전하고 있는 data science와 artificial intelligence에서는 선형대수학이 해당 분야의 이해와 연구, 활용에 필수적이라고 인식되고 있다. 이러한 분야에서는 고차원 공간 속의 데이터를 저차원 공간 상으로 변환하여 다양한 분석을 통해 데이터에 대해 이해하게 된다. 여기서 저차원 공간 상으로의 변환이 다양한데 가장 손쉽게 사용가능하고 많이 활용되는 방법이 부분공간 상으로 projection을 통하여 근사하는 방법이다. 이때 고차원 데이터를 잘 근사하는 저차원 부분공간을 결정하는 문제가 singular value decomposition (SVD)이다. 그리고 데이터들을 기술하는 분포에 연관된 행렬들이 갖는 positive definiteness의 다양한 특성을 이해하면 데이터에 대한 이해를 더욱 깊게 할 수 있다.

훌륭한 많은 선형대수학 교과서들이 서점에 있다. 본 교과서가 기존 책들과 가장 다른 점은 가능한 빠르게 projection과 SVD, positive definiteness를 수학적으로 소개한다는 점이다. 기존 교과서 대부분은 determinant, eigenvalue/eigenvector를 공부한 뒤에 corollary로 SVD를 도입하였지만 본 교과서에서는 순서를 바꾸어 SVD와 positive definiteness를 먼저 공부한다. 특히 SVD를 행렬의 분해보다는 데이터의 근사 표현으로 도입하여 독자들이 자신있게 응용할 수 있도록 의도하였다. 이렇게 순서가 바뀌었어도 수학적 논리 전개에서 문제가 없도록 구성하였다. 그리고 후반부에 determinant를 도입하기 전까지 SVD와 positive definiteness를 복소수 도입없이 설명할 수 있다는 것도 기존 교과서와 큰 차이이다. 선형대수학의 후반부에 위치하여 독자들이 접근하기 어려웠던 Cayley-Hamilton정리와 Jordan normal form정리도 심화된 대수학 지식없이 이해할 수 있도록 서술하였고 행렬의 transpose와 선형사상의 adjoint, Perron-Frobenius정리 등 유용한 결과들도 유도하였다. 또한 행렬을 활용할 때 작은 행렬들로 구성된 block matrix가 유용한데 많은 결과들에 대해서 block matrix 표현도 병행하여 유도하였다.

이 책을 저술하는 데 있어서 [2]와 [5]를 많이 참조하였다. SVD에 관련해서는 [1]의 접근 방식을 따랐다. 그리고 Google행렬과 관련해서는 [4]를 참조하였다. 교과서로 사용되는 대부분의 수학책은 보통 많은 연습문제가 본문에 포함되어 있다. 이러한 구성은 장점도 많지만 단점도 있다. 본 교과서에서는 간단한 결과나 연습문제 중 일부를 Fact로 요약하고 증명하였다. 독자들은 모든 Fact들을 증명하려고 시도해보기 바란다. 책 말미에는 homework assignments가 제시되어 있는데 이를 통해서 독자들이 여러 개념이나 계산들에 익숙해 지기를 기대한다. 그리고 KAIST에 두 학기동안 사용된 중간 및 기말고사 문제들도 제공되므로 내용의 이해도를 테스트해보기 바란다.

독자들이 선형대수학과 행렬이론의 여러 유용한 결과를 수학적으로 올바르게 이해하고 세상의 여러 문

제에 자신있게 적용하는 데에 본 교과서가 도움이 되기를 기대한다.

Chapter 1

Introduction

수학은 다양한 공간과 공간 사이에 정의된 함수, 그리고 함수가 보존하는 공간의 성질에 대해 주로 고민한다. 수학의 여러 분야 중에서 선형대수학은 상수배가 가능하고 더할 수 있는 원소들로 구성된 공간과 그 공간의 원소들을 대응시키는 함수 중에 상수배와 덧셈을 잘 보존하는 함수들에 관심이 많다. 이러한 원소와 공간, 함수를 각각 벡터, 벡터공간, 선형사상이라고 부른다. 특히 상수배와 덧셈이 보존되는 것이 선형성인데, 이 선형성 때문에 벡터공간과 선형사상에 대해서 상당히 많은 것이 밝혀져 있다. 예를 들면 차원이 동일한 벡터공간들은 거의 동일한 구조를 가진다. 여러 숫자를 사각 형태로 배열한 것을 행렬이라고 부르는데, 선형사상과 행렬이 1-1로 대응하기 때문에 선형사상에 대한 분석은 대응되는 행렬에 대한 연구와 동일하다. 행렬을 분류하는 연구를 통해 동일한 구조의 선형사상들은 Jordan form이라고 불리는 간단한 형태의 행렬을 공유한다는 것이 밝혀졌다. 이러한 분류를 연구하는 과정에서 부산물로 유용한 결과들이 많이 얻어졌다. 이 책에서는 이러한 분류의 결과를 소개하면서 여러 유용한 결과를 논리의 비약없이 소개하고자 한다.

선형대수학(linear algebra)을 기술하고 전개하는 데 가장 중요한 수학적 개념이 벡터(vector)와 행렬(matrix)이다. 벡터는 더하기(addition)와 상수배(scalar multiplication)가 의미를 갖는 대상을 말한다. 덧셈의 항등원에 해당하는 대상을 영벡터(vector of zeros, zero vector)라고 부르고 **0**으로 표시한다. 더하기의 역원이 존재하는 경우에 실수에서와 같이 마이너스 기호(-)를 원래 대상의 앞에 붙여서 표시한다. 벡터 \mathbf{v} 에 대해서 $1\mathbf{v} = \mathbf{v}$ 로 표기하는 경우,

$$\mathbf{v} + \mathbf{v} = 1\mathbf{v} + 1\mathbf{v} = (1 + 1)\mathbf{v} = 2\mathbf{v}$$

가 자연스럽도록 벡터의 더하기와 상수배 사이의 분배 법칙(distributive laws)을 만족하는 벡터를 모아 놓은 공간을 벡터공간(vector space)이라고 부른다. 벡터공간에 대한 자세한 공부는 Chapter 3에서 시작된다. 가장 직관적이면서 중요한 벡터의 예는 유한개의 숫자를 순서있게 모아 놓은 배열들이다. 즉 m 개의 실수 v_1, v_2, \dots, v_m 를 세로로 쌓은 배열인 m 차원 벡터를 \mathbf{v} 로 놓으면

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

로 표시할 수 있다. v_i 들을 벡터의 원소(entry, element, component)라 부르는데 v_i 는 \mathbf{v} 의 i 번째 원소이다. m 차원 벡터는 $\mathbf{v} \in \mathbb{R}^m$ 로 표시하고 m -벡터(m -vector)나 \mathbb{R}^m -벡터(\mathbb{R}^m -vector)라고 부르기도 한다. 두 벡터의 덧셈과 스칼라에 대한 상수배는

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_m + w_m \end{pmatrix}, \quad c\mathbf{v} = \begin{pmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_m \end{pmatrix}$$

로 정의하는데 벡터의 원소 별로 연산을 수행하면 된다. 이러한 벡터들의 덧셈에 대한 항등원은 모든 원소가 0인 영벡터이다. 만약 영벡터의 크기가 중요하면 $\mathbf{0}_m$ 으로 첨자를 이용하여 m 차원 영벡터를 표시한다.

행렬은 다양한 방법으로 정의할 수 있는데, 이 강의록에서는 \mathbb{R}^m -벡터를 가로방향으로 붙여서(concatenating) 만들어진 숫자들의 이차원 배열로 정의한다. 즉 n 개의 \mathbb{R}^m -벡터

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, \mathbf{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

를 가로방향으로 붙여서 만들어진

$$A = [\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_n] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

를 $m \times n$ 행렬이라고 부른다. 이차원 배열로서 행렬의 가로줄은 행(row), 세로줄은 열(column) 또는 열벡터(column vector)라고 부른다. 행렬은 보통 알파벳 대문자로 표시하고 원소는 소문자를 사용한다. 예를 들면 행렬 A 의 (i, j) 원소, 즉 i 행, j 열의 원소는 a_{ij} 인데 행렬의 이름 A 만 주어진 경우에는 $(A)_{ij}$ 로 표시하면 편리하다. $A = (a_{ij})$ 로 표시하는 경우도 많은데 행렬 A 가 a_{ij} 를 원소로 갖는다는 의미이다. 두 행렬 A 와 B 의 합 $A + B$ 는 원소 별로 합친 $((A)_{ij} + (B)_{ij})$, 즉 다음과 같이 정의한다.

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

그리고 scalar c 의 상수배 cA 는 행렬의 원소 각각에 스칼라를 곱해준다. 즉

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

로 정의한다. 따라서 \mathbb{R}^n -벡터는 $n \times 1$ 행렬도 된다.

벡터공간에는 벡터의 종류에 따라서 다양한 구조나 연산이 추가되면서 의미가 부여되고 활용된다. 덧셈과 상수배가 정의되는 행렬로 구성된 집합도 벡터공간으로 생각할 수 있으며 행렬곱이 연산으로 추가될 수 있다. 행렬곱의 자세한 내용은 다음 장에서 다룬다. 본 장에서는 행렬곱의 간단한 경우로서 행렬과 \mathbb{R}^n -벡터의 곱(또는 $m \times n$ 행렬과 $n \times 1$ 행렬의 곱)을 살펴본다. $m \times n$ 행렬 $A = [\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n]$ 와 v_i 를 원소로 갖는 \mathbb{R}^n -벡터 \mathbf{v} 의 곱셈은

$$A\mathbf{v} = v_1\mathbf{a}_1 + \cdots + v_n\mathbf{a}_n = \sum_{i=1}^n v_i\mathbf{a}_i \quad (1.1)$$

로 정의한다. 곱셈의 결과물은 \mathbb{R}^m -벡터, 즉 $A\mathbf{v} \in \mathbb{R}^m$ 이 되고, 이러한 행렬과 벡터의 곱을 행렬의 열벡터 \mathbf{a}_i 들의 선형결합(linear combination)이라고 부르고 v_i 를 선형결합의 가중치/계수(weight/coefficient of linear combination)라고 부른다. 그리고 이 강의록 전체에 걸쳐서 반복적으로 사용될 특별한 벡터들로 $\mathbf{1}, \mathbf{e}_i$ 등이 있다.¹

행렬과 \mathbb{R}^m -벡터는 데이터를 모형화하는데 유용하다. 예를 들어 실수로 표현되는 개인별 데이터를 \mathbb{R}^m 벡터로 표시하면 n 명의 데이터는 이를 열로 붙여서 $m \times n$ 행렬로 생각할 수 있다. 이 행렬을 A 라 하자. 그러면 $(A)_{ij}$ 는 j 번째 사람의 데이터의 i 번째 항목 값이 된다. 각 항목의 평균을 벡터로 생각하여 한번에 계산하면 벡터 \mathbf{a}_i 들의 평균으로 $\frac{1}{n}\mathbf{a}_1 + \cdots + \frac{1}{n}\mathbf{a}_n = \frac{1}{n}(\mathbf{a}_1 + \cdots + \mathbf{a}_n) = \frac{1}{n}(A\mathbf{1}) = A(\frac{1}{n}\mathbf{1})$ 와 같이 행렬과 벡터의 곱으로 표시된다.

다른 예로 다음과 같은 경제모형을 생각하자.

- a_{ij} : contribution of material j to i -th product, \mathbf{a}_j : \mathbb{R}^m -vector whose i -th entry is a_{ij} ;
- x_j : amount of material j available, \mathbf{x} : \mathbb{R}^n -vector whose j -th entry is x_j .

이러한 모형에서 x_j 가 가중치인 \mathbf{a}_j 들의 선형결합인 $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \sum_{j=1}^n x_j\mathbf{a}_j$ 은 원자재 j 를 x_j 만큼 구매할 때 m 개 제품의 생산량을 나타내는 것으로 해석할 수 있다. 이 선형결합은 행렬곱 $A\mathbf{x}$ 로 표현할 수 있다. 따라서 목표 생산량을 원자재의 낭비없이 생산할 수 있는 원자재의 조합을 찾는 문제는 i -번째 제품의 필요량이 b_i 인 경우 생산량 벡터 $\mathbf{b} = (b_i)$ 에 대해서

$$A\mathbf{x} = \mathbf{b}$$

와 같이 행렬과 벡터로 간단하게 표현된다. 방금 살펴본 것처럼 데이터를 다룰 때 행렬과 벡터를 이용하여 사고하면 여러모로 편리하다.

$${}^1\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{where only the } i\text{-th component is 1. Data science에서 } \mathbf{e}_i \text{를 one-hot vector라고 부르기도 한다.}$$

선형대수학은 $Ax = \mathbf{b}$ 로 표현되는 연립 일차방정식을 푸는 과정에서 공부를 시작할 수 있다. 그리고 관련된 여러 개념들이 자연스럽게 도입되고 새로운 결과들을 얻게된다. 본 강의록에서 다룰 내용들은 다음과 같다.

- How to solve linear systems: Gaussian elimination;
- Abstraction of data and manipulations: a vector space and linear transformation;
- Approximation of data: orthogonality, projection, and least squares;
- Factorization of data: SVD (singular value decomposition) and PCA (principal components analysis), Pseudoinverse;
- Shapes of data: covariance, positive definiteness, and convexity.
- Key features of matrices: determinant, eigenvalue and eigenvector;
- Advanced results for matrices: adjoint, positive definite cone, and Perron-Frobenius theorem;
- Theorems by Cayley-Hamilton and Jordan.

Data science 분야 연구에서는 종종 고차원 데이터 공간 속의 벡터(데이터)들을 저차원 부분공간에 정사영하여 근사된 벡터들을 탐색/분석한다. 이때 어떤 저차원 부분공간이 주어진 고차원 데이터들을 잘 근사하는지 결정하는 문제가 SVD이다. 그리고 데이터간의 관계를 기술하는 과정에서 positive definite 행렬이 자주 이용되는 데 이러한 행렬들의 여러 성질들은 data science나 공학에서 자주 활용된다. 이 책에서는 determinant나 eigenvalue/eigenvector 개념의 도움없이 projection, SVD와 positive definiteness를 수학적으로 소개한다.

Chapter 2

Matrices and Gaussian Elimination

선형연립방정식을 푼다는 것은 각각의 일차방정식을 동시에 만족하는 변수들의 값을 정하는 것이다. 이 풀이 과정을 체계화 한 것이 Gaussian elimination으로 두 번째 일차식에서 변수 한개를 소거하고 세 번째 일차식에서는 두 번째 식에서 소거된 변수 외에 또 다른 변수까지 두 변수를 소거하는 방식으로 진행하여 마지막 식에는 변수가 한 개만 남도록 만든다. 그러면 마지막 식에서 식에 포함된 변수의 값을 읽어 낼 수 있고, 직전 식에 해당 변수 값을 대입하면 또 다른 변수의 값이 정해지는 방식으로 모든 변수의 값을 정할 수 있다. 그런데 변수나 등식 기호 없이 방정식의 계수들과 우측 상수항들로 행렬을 구성하고 Gaussian elimination 을 이 행렬에 대한 연산으로 나타내고자 하면 상수배와 덧셈, 그리고 행렬과 벡터의 곱셈 만으로는 부족하고 행렬과 행렬의 곱이 필요해진다. 즉 Gaussian elimination의 변수 소거가 특별한 형태의 행렬을 연립방정식 을 나타내는 행렬의 왼쪽에 곱하는 것과 동일해진다. 이렇게 선형연립방정식을 푸는 과정을 행렬 연산으로 살펴보는 과정에서 행렬에 대해 놀라울 정도로 많은 발견을하게 된다.

2.1 Matrix Operations

행렬과 벡터의 곱을 (1.1)로 정의하였다. 이를 행렬 간의 곱셈으로 확장해보자. 행렬 간의 곱셈을 생각하기 위해 두 번째 행렬로 크기가 $n \times \ell$ 인 $B = [\mathbf{b}_1 | \mathbf{b}_2 | \cdots | \mathbf{b}_\ell] = (b_{jk})$ 가 주어졌다고 하자. k 번째 열 \mathbf{b}_k 와 A 의 곱은 $A\mathbf{b}_k = b_{1k}\mathbf{a}_1 + b_{2k}\mathbf{a}_2 + \cdots + b_{nk}\mathbf{a}_n$ 이다. 이를 이용하여 두 행렬 A 와 B 의 곱 AB 를

$$AB = [A\mathbf{b}_1 | A\mathbf{b}_2 | \cdots | A\mathbf{b}_\ell]$$

로 정의한다. 이 곱셈이 잘 정의되기 위해서는 첫 번째 행렬과 두 번째 행렬의 열이 곱해지므로 첫 번째 행렬의 열의 개수와 두 번째 행렬의 행의 개수가 일치해야 한다. 즉 $m \times n$ 행렬과 $n \times \ell$ 행렬을 곱하면 $m \times \ell$ 행렬이 생겨난다. 곱 AB 의 (i, j) 원소는

$$\begin{aligned} (AB)_{ij} &= b_{1j}a_{i1} + \cdots + b_{nj}a_{in} \\ &= \sum_{k=1}^n a_{ik}b_{kj} \end{aligned}$$

$$= \begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix}$$

와 같이 다양하게 표현된다. 행렬 간의 곱은 결합법칙 $(AB)C = A(BC)$ 가 성립하고 행렬 간의 덧셈과 곱셈은 분배 법칙 $A(B + C) = AB + AC$, $(B + C)D = BD + CD$ 가 성립하는데 각자 확인해 보기 바란다. 행렬 간의 곱셈이 실수의 곱셈과 다른 점은 교환법칙이 성립하지 않는 것이다. $EF \neq FE$ 인 두 행렬 E 와 F 의 예를 쉽게 찾을 수 있다. 행렬 덧셈의 항등원은 모든 원소가 0인 영행렬(matrix of zeros)이고 $\mathbf{0}$ 으로 표시한다. 다양한 크기의 영행렬을 $\mathbf{0}$ 으로 표시하지만 대부분의 경우에 혼동되지 않는다. 영행렬의 크기를 특별히 표시할 필요가 있는 경우에는 $m \times n$ 영행렬을 $\mathbf{0}_{m,n}$ 로 표시하기도 한다. 행렬 곱셈의 항등원은 행 번호와 열번호가 같은 대각원소(diagonal entry)들이 1이고 나머지 비대각원소(off-diagonal entry)들은 0인 정사각행렬(square matrix)로서 항등행렬(identity matrix)이라고 부르고

$$I = I_n = \begin{bmatrix} 1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1 \end{bmatrix}$$

로 표시한다. 단위행렬은 종종 크기를 첨자로 알려주는 것이 필요하다. $m \times n$ 행렬 A 에 대해서

$$AI_n = A = I_m A$$

가 성립함을 확인해 보기 바란다. 행렬의 곱셈에서 유용한 행렬 중 한 가지가 대각행렬(diagonal matrix)로 비대각원소가 모두 0인 행렬이다. i 번째 대각원소가 d_i 이면

$$D = \begin{bmatrix} d_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & d_n \end{bmatrix} = \text{diag}(d_i) = \text{diag}(d_1, \dots, d_n)$$

으로 표시한다. 물론, $d_i = 0$ 일 수도 있다. 대각행렬은 다른 행렬의 왼쪽 또는 오른쪽에 곱해지면서 scaling이라 부르는 행 또는 열에 대한 상수배 작용을 한다. 즉 어떤 행렬 A 에 대해서 AD 는 A 의 j 번째 열벡터에 d_j 를 상수배(scaling)한 것이다. 즉 AD 는 A 를 각 열 별로 scaling한 행렬이 된다. DA 는 i 번째 행을 d_i 값으로 scaling한 행렬이 된다.

행렬을 다룰 때 유용한 행렬의 변형으로 행렬의 전치(transpose)가 있다.

Definition 2.1 An $n \times m$ matrix A^\top is the transpose of an $m \times n$ matrix A if $(A^\top)_{ij} = (A)_{ji}$.

간단한 예를 통해 행렬의 transpose를 쉽게 이해할 수 있다:

$$\begin{bmatrix} 2 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix}^\top = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 4 & 3 \end{bmatrix}.$$

행렬의 transpose는 행렬의 덧셈, 곱셈과 자연스럽게 어우러진다.

- $(A + B)^\top = A^\top + B^\top$

since $((A + B)^\top)_{ij} = (A + B)_{ji} = (A)_{ji} + (B)_{ji} = (A^\top)_{ij} + (B^\top)_{ij} = (A^\top + B^\top)_{ij}$;

- $(AB)^\top = B^\top A^\top$

since $((AB)^\top)_{ij} = (AB)_{ji} = \sum_{k=1}^{\ell} (A)_{jk} (B)_{ki} = \sum_{k=1}^{\ell} (B^\top)_{ik} (A^\top)_{kj} = (B^\top A^\top)_{ij}$.

행렬의 transpose를 벡터의 transpose로 확장해보자. a_i 가 i 번째 원소인 \mathbb{R}^n -벡터 \mathbf{a} 는 $n \times 1$ 행렬로 생각할 수 있으므로 \mathbf{a}^\top 는 $[a_1 \ a_2 \ \dots \ a_n]$ 과 같이 $1 \times n$ 행렬로 생각할 수 있다. 이 개념을 $m \times n$ 행렬 A 와 \mathbb{R}^n -벡터 \mathbf{v} 의 곱셈을 정의한 (1.1)에서 $m = 1$, 즉 행렬 A 가 $1 \times n$ 행렬 $A = \mathbf{a}^\top$ 인 경우에 적용해보자. 그러면 $A\mathbf{v} = \mathbf{a}^\top \mathbf{v}$ 인데 곱이 1×1 행렬, 즉 실수이므로 행렬표시를 하지 않고

$$\mathbf{a}^\top \mathbf{v} = \sum_{i=1}^n a_i v_i \quad (2.1)$$

로 쓸 수도 있다. 이를 \mathbb{R}^n -벡터 \mathbf{a} 와 \mathbf{v} 의 (표준)내적(inner product)이라고 부르는데 나중에 벡터공간에서 내적을 도입할 때 자세하게 공부한다(Definition 4.3). 당분간은 \mathbb{R}^n -벡터 간의 내적을 (2.1)로 생각해도 된다. 벡터 간의 내적을 도입하면 여러 행을 가진 행렬과 벡터가 곱해서 얻어지는 벡터는 각각의 행과 벡터의 곱을 내적으로 생각하여 성분이 계산되는 것으로 해석할 수 있다. 따라서 행렬과 벡터의 곱은 벡터 내적의 반복으로 생각할 수도 있다.

행렬의 transpose를 알고나면 transpose에 의해 변하지 않는 행렬은 어떤 성질을 가지고 있을까라는 궁금함이 생겨날 것이다. 이러한 행렬을 대칭행렬(symmetric matrix)이라고 부르는데, 대칭행렬들은 좋은 성질들을 풍부하게 가지고 있어서 선형대수학에서 아주 중요한 대상이다.

Definition 2.2 A is symmetric if $A^\top = A$.

대칭행렬의 간단한 성질 몇 가지는 다음과 같다.

- Symmetric matrix is square.
- Every diagonal matrix is symmetric.
- For any matrix A , $A^\top A$ and AA^\top are symmetric.
- ADA^\top and $A^\top DA$ are symmetric where D is a diagonal matrix.

이유는 각자 생각해 보기 바란다. 대칭행렬의 더 깊고 풍부한 성질들은 이후에 소개된다.

2.2 Solving Simultaneous Linear Equations

선형대수학과 행렬 이론을 전개하는 것은 연립선형방정식을 푸는 것과 밀접한 연관이 있다. 다음과 같은 2 원 일차 연립방정식을 생각해 보자. 이 방정식을 행렬을 사용하면 다음과 같다.

$$\left\{ \begin{array}{rcl} \text{(equation 1)} & 1x + 1y & = 5 \\ \text{(equation 2)} & 2x - 1y & = 1 \end{array} \right. \iff \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} \iff \begin{bmatrix} 1 & 1 & | & 5 \\ 2 & -1 & | & 1 \end{bmatrix}$$

특히 마지막에는 우변을 포함한 2×3 행렬로 연립방정식을 표현하였다. 이 연립방정식을 푸는 과정을 살펴보자. 첫 번째 방정식의 양변에 2배를 한 뒤에 두 번째 방정식에 뺄 것이다. Gaussian elimination이라고 부르는 이 과정은 elementary matrix라고 부르는 특별한 형태의 행렬을 연립방정식을 나타내는 2×3 행렬의 왼쪽에 곱하는 것과 동일하다. 현재 예에서는 elementary matrix가 $\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$ 인데 이 행렬의 두 번째 행이 소거과정을 담고 있다:

$$\text{(equation 2)} - 2 \times \text{(equation 1)} \Rightarrow \text{(equation 2)}: \begin{cases} 1x + 1y = 5 \\ - - 3y = -9 \end{cases}$$

$$\iff \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \left[\begin{array}{cc|c} 1 & 1 & 5 \\ 2 & -1 & 1 \end{array} \right] = \left[\begin{array}{cc|c} 1 & 1 & 5 \\ 0 & -3 & -9 \end{array} \right]$$

두 번째 방정식에서 첫 번째 변수를 소거하였으므로 두 번째 변수값을 결정할 수 있다.

$$-3y = -9 \implies y = 3$$

그리고 마지막으로 첫 번째 변수값을 첫 번째 방정식에서 결정한다.

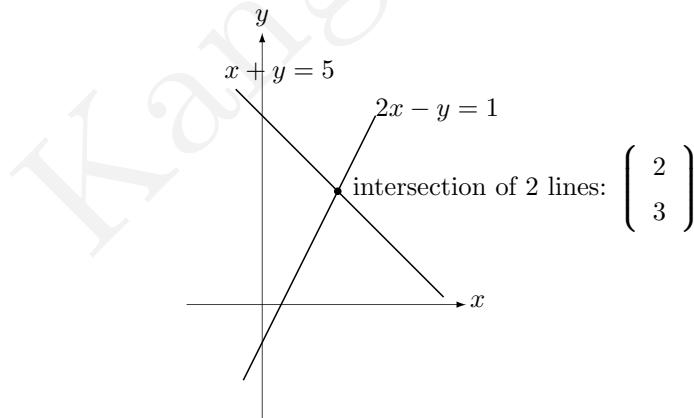
$$x + y = 5, y = 3 \implies x = 5 - 3 = 2$$

The Geometry of Linear Equations

변수가 3개 이상인 경우는 어렵지만 미지수가 2개인 연립방정식은 2차원 평면에서 방정식의 의미를 살펴볼 수 있다. 다음과 같이 두 가지 관점이 있다.

- **Row-wise:** 각 방정식의 그래프를 도시하여 그래프가 만나는 점이 두 방정식을 모두 만족하는 해로 이해하는 것이다. 이는 행렬로 표시된 방정식에서 계수행렬의 각 행과 변수 벡터가 곱해진 벡터의 각 원소별로 우변과 비교하는 것으로 이해할 수 있다. 선형연립방정식의 경우 아래와 같이 두 직선의 교점과 방정식의 해가 대응된다.

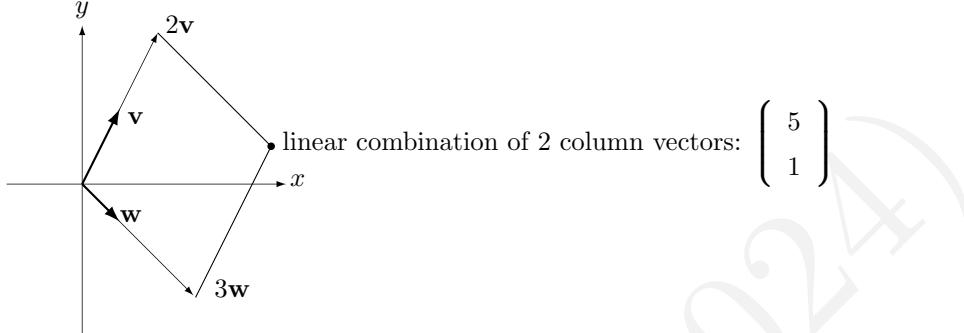
$$\begin{array}{rcl} x + y = 5 \\ 2x - y = 1 \end{array} \iff \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$



- **Column-wise:** 행렬로 표시된 방정식을 (1.1)로 이해하면 방정식의 우변 벡터를 계수행렬의 열벡터들의 선형조합(linear combination)으로 표시할 때 조합 가중치가 방정식의 해가 되는 것으로 이해할

수 있다.

$$\begin{array}{rcl} x + y = 5 & \Leftrightarrow & x \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{\mathbf{v}} + y \underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{\mathbf{w}} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} \\ 2x - y = 1 & & \end{array}$$



2차원 예를 변수가 n 개인 m 개의 선형방정식으로 확장해보자. 연립방정식은 $m \times n$ 행렬 A 와 m 차원 벡터 \mathbf{b} 에 대해서 $A\mathbf{x} = \mathbf{b}$ 로 표현할 수 있는데, 2차원 경우와 비슷하게 방정식의 해 \mathbf{x} 는 intersection of m hyperplanes in \mathbb{R}^n , 또는 combination of n column vectors in \mathbb{R}^m 으로 이해할 수 있다. 어느 차원에서나 직선 또는 hyperplane들의 상대적인 위치에 따라서 해가 유일하거나, 존재하지 않거나, 또는 무한히 많을 수 있다. 각 경우에 대한 예를 3차원에서 생각해 보기 바란다. 해의 존재에 대한 해석을 열벡터 관점에서 하려면 벡터들 간의 독립성이라는 개념이 필요하며 이후에 Definition 3.4에서 정의된다. 위의 예에서 두 2차원 벡터 \mathbf{v} 와 \mathbf{w} 는 선형독립(linearly independent)이고 주어진 $(5, 1)^\top$ 뿐만 아니라 임의의 2차원 우변 벡터에 대해서 해가 존재한다. 일반적으로 m 개의 선형방정식이 연립된 연립방정식의 계수행렬의 열벡터 중 m 개 이상이 linearly independent하면 모든 \mathbf{b} 에 대해서 해가 존재한다는 것을 나중에 살펴본다.

2.3 An Example of Gaussian Elimination

다음과 같은 3개 미지수에 대한 3개의 선형방정식으로 이루어진 연립방정식을 생각해보자:

$$\begin{array}{rcl} 2u + v + w & = & 5 \\ 4u - 6v & = & -2 \\ -2u + 7v + 2w & = & 9 \end{array}$$

이 연립방정식을 행렬로 간결하게 표현하면 다음과 같이 나타낼 수 있다: $\left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array} \right]$.

이 연립방정식을 푸는 과정을 살펴 보자. 먼저 두 번째 식에서 첫 번째 변수를 소거하고 세 번째 식에서는 첫 번째와 두 번째 변수를 소거한다. 이렇게 하고 나면 세 번째 식에서 세 번째 미지수를 결정하고, 이를 첫 번째와 두 번째 식에 대입하여 두 식의 미지수 개수를 하나씩 줄인 뒤에 나머지 미지수를 정해나가면 된다. 여기서 우리가 집중해야하는 것은 이 소거 과정을 elementary matrix라고 불리는 간단한 행렬들의 연속적인 왼쪽 곱으로 표현하는 것이다. Elementary matrix는 단위행렬에서 단 한개의 비대각원소만을 0이 아닌 숫자로 바꾼 행렬이다. 예를 들어 $i > j$ 일 때 단위행렬에서 (i, j) 원소만 0이 아닌 b 로 바뀐 행렬 E 를 행렬 A 의

원쪽에 곱하면

- i 행을 제외한 EA 의 행들은 A 의 행들과 동일하고;
- EA 의 i 행은 A 의 j 행에 b 를 곱하여 A 의 i 행에 더한 값이 된다.

1. (equation 2) - 2 (equation 1) \Rightarrow (equation 2): \Leftrightarrow left multiplication of $\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$\begin{array}{rcl} 2u + v + w = 5 \\ -8v - 2w = -12 \\ -2u + 7v + 2w = 9 \end{array} \Leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array} \right] = \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ -2 & 7 & 2 & 9 \end{array} \right]$$

2. (equation 3) + (equation 1) \Rightarrow (equation 3): \Leftrightarrow left multiplication of $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

$$\begin{array}{rcl} 2u + v + w = 5 \\ -8v - 2w = -12 \\ 8v + 3w = 14 \end{array} \Leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ -2 & 7 & 2 & 9 \end{array} \right] = \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{array} \right]$$

3. (equation 3) + (equation 2) \Rightarrow (equation 3): \Leftrightarrow left multiplication of $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

$$\begin{array}{rcl} 2u + v + w = 5 \\ -8v - 2w = -12 \\ w = 2 \end{array} \Leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{array} \right] = \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{array} \right]$$

위의 과정은 원래 방정식에 세 개의 elementary matrix를 차례로 (순서를 지키면서) 좌측에 곱한 것이다:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array} \right] = \left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{array} \right].$$

이러한 과정을 Gaussian elimination이라고 부른다. 이 과정을 거치면 i 번째 방정식에는 $1, \dots, (i-1)$ 번째 변수가 모두 소거되어 있고 변경된 계수행렬을 $C = (c_{ij})$ 라하면 계수행렬 관점에서는 $c_{i1} = \dots = c_{i(i-1)} = 0$ 이다. 어떤 행렬이 대각원소 아래쪽 원소들이 모두 0이면, 즉 $A = (a_{ij})$ 에서 $a_{ij} = 0$ if $i > j$ 인 경우 이런 행렬들을 상삼각행렬(upper triangular matrix, 위쪽 삼각형 영역에만 0이 아닌 원소가 존재하므로)이라고 부른다. 하삼각행렬(lower triangular matrix)도 유사하게 정의된다.

이렇게 연립방정식의 계수행렬이 상삼각행렬로 바뀌면 back-substitution이라고 불리는 절차에 의해 간단하게 해를 정할 수 있다. 위의 예에서는 세 번째 방정식(행)부터 두 번째 방정식(행), 그리고 첫 번째 방정식(행)의 순서로 뒤쪽 변수부터 차례로 변수값을 정할 수 있다.

Row 3 : $w = 2$

$$\Rightarrow \text{Row 2 : } -8v = -12 + 2w = -8, v = 1$$

$$\Rightarrow \text{Row 1 : } 2u = 5 - v - w = 5 - 1 - 2 = 2, u = 1.$$

연립방정식에 따라서는 이러한 절차가 적용되는 과정에서 다음과 같은 두 가지 이유로 중단될 수 있다.

- Non-singular (curable by row exchange): 두 번째 행에 대한 소거과정에서 너무 많은 변수가 소거되면 세 번째 행의 변수 소거가 진행되지 않을 수 있다. 이러한 경우는 Gaussian elimination 전에 두 번째 행과 세 번째 행을 교환(exchange)한 후에 Gaussian elimination을 수행하면 된다. 연립방정식에서 방정식의 나열순서는 문제가 되지 않는 것과 같은 원리이다.

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 5 \\ 4 & 6 & 8 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 3 \\ 0 & 2 & 4 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 1 & 1 \\ 4 & 6 & 8 \\ 2 & 2 & 5 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

- Singular (incurable): 어느 두 행이 서로 상수배 관계가 있으면 소거과정에서 한 행이 모두 0으로 채워진다. 이 경우도 (필요시 적절한 행 교환을 수행하면) Gaussian elimination을 통해 계수행렬을 상삼각행렬로 변경할 수 있는데¹ 방정식의 우변에 따라서 back-substitution이 적용될 수도 있고² 적용되지 않을 수도 있다(이 경우 해가 존재하지 않는다). 이런 의미에서 incurable이다.

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 5 \\ 4 & 4 & 8 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

이 예에서는 Gaussian elimination에 의해 변경된 우변의 세 번째 원소가 0인지 아닌지에 따라서 해가 무한개인지 존재하지 않을지 결정된다.

2.4 Block Matrices

\mathbb{R}^{n_1} -벡터 $\mathbf{u}_1, \mathbf{v}_1$ 과 \mathbb{R}^{n_2} -벡터 $\mathbf{u}_2, \mathbf{v}_2$ 에 대해서 $\mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}$ 와 $\mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix}$ 로 벡터를 확장하고 이렇게 표시된 두 $\mathbb{R}^{n_1+n_2}$ -벡터 간의 내적

$$\mathbf{u}^\top \mathbf{v} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}^\top \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} = \mathbf{u}_1^\top \mathbf{v}_1 + \mathbf{u}_2^\top \mathbf{v}_2$$

을 살펴보자. 특히 \mathbf{u}^\top 를 $1 \times (n_1 + n_2)$ 행렬로 생각하고 \mathbf{v} 를 $(n_1 + n_2) \times 1$ 행렬로 생각하여 행렬곱으로 내적을 살펴보면

$$\mathbf{u}^\top \mathbf{v} = \begin{bmatrix} \mathbf{u}_1^\top & \mathbf{u}_2^\top \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^\top \mathbf{v}_1 + \mathbf{u}_2^\top \mathbf{v}_2 \end{bmatrix}$$

¹상삼각행렬 여부는 대각원소와 그 위쪽 원소가 0인지 여부와 관계없이 대각원소 아래쪽 원소가 0인 것만 중요하다.

²이 경우 무한개의 해가 존재한다

으로 표현할 수 있다. 행렬과 벡터의 곱은 내적의 반복이므로 $m \times n_1$ 행렬 A_{11} 과 $m \times n_2$ 행렬 A_{12} 로 구성된 행렬 $A = [A_{11} \ A_{12}]$ 와 벡터 \mathbf{v} 의 곱은 A 의 각 행이 n_1 과 n_2 개의 원소를 가진 subvector가 합쳐진 결로 생각해서 위에서와 같은 곱을 m 번 반복해서 곱의 결과인 \mathbb{R}^m 벡터의 m 개의 원소를 계산할 수 있다. 즉

$$A\mathbf{v} = \begin{bmatrix} A_{11} & A_{12} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = A_{11}\mathbf{v}_1 + A_{12}\mathbf{v}_2$$

이다. \mathbf{v} 대신에 $(n_1 + n_2) \times \ell$ 행렬 $B = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$ 를 곱하는 경우는 B 의 각 열이 n_1, n_2 개의 원소를 갖는 subvector의 합으로 생각하면 각 열과 A 의 곱이 AB 의 열을 만들어 낸다. 이를 행렬로 표현하면

$$AB = \begin{bmatrix} A_{11} & A_{12} \end{bmatrix} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = A_{11}B_{11} + A_{12}B_{21} \quad (2.2)$$

가 됨을 알 수 있다. 가장 우측의 행렬 덧셈은 $A_{11}B_{11}$ 과 $A_{12}B_{21}$ 가 모두 $m \times \ell$ 행렬이어서 잘 정의된다. 그리고 A 와 B 의 역할이 바뀐 다음과 같은 형태의 행렬에도

$$\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} & A_{11}B_{12} \\ A_{21}B_{11} & A_{21}B_{12} \end{bmatrix} \quad (2.3)$$

이 성립하는 것을 확인할 수 있다. 더 많은 submatrix로 구성된 block matrix간의 곱은 여기에 기술된 두 가지 곱 표현 (2.2)과 (2.3)을 반복적(recursive)으로 적용하면 계산할 수 있다. 행렬 A 가 작은 크기의 행렬 A_{ij} 들에 대해 다음과 같이 block matrix로 표현된다고 하자:

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

A 를 구성하는 block들인 부분행렬(submatrix)들의 크기는 자연수 m_i, n_i 에 대해서 $A_{11} : m_1 \times n_1, A_{12} : m_1 \times n_2, A_{21} : m_2 \times n_1, A_{22} : m_2 \times n_2$ 라 하면 $A : (m_1 + m_2) \times (n_1 + n_2)$ 가 된다. 또 다른 행렬 B 도 적절한 크기의 B_{ij} 들에 대해 block matrix로 $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ 표현되었다면

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix} \quad (2.4)$$

로 표현됨을 (2.2)과 (2.3)을 이용하여 보일 수 있다.³ 이런 행렬곱이 성립하는 조건은 block matrix 곱에서 나타나는 $A_{ij}B_{jk}$ 들이 잘 정의될 조건인 A_{ij} 의 열 개수와 B_{jk} 의 행 개수가 일치하는 것이다.

비교를 위해서 2×2 행렬 두 개의 곱을 나타내보자.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}.$$

³ 이 표현이 성립하는 것은 각자 확인하기 바란다.

Block matrix간의 곱과 비교하면 block index와 entry index가 정확하게 일치함을 알 수 있다. Block matrix 간의 곱에서는 곱행렬의 block을 구성하는 submatrix곱들의 순서를 반드시 지켜야한다. 예를 들면, block matrix 곱의 첫 번째 block $A_{11}B_{11} + A_{12}B_{21}$ 이 $B_{11}A_{11} + A_{12}B_{21}$ 또는 $A_{11}B_{11} + B_{21}A_{12}$ 와 다를 수 있다는 것이 $a_{11}b_{11} + a_{12}b_{21} = b_{11}a_{11} + a_{12}b_{21} = a_{11}b_{11} + b_{21}a_{12}$ 인 숫자 간의 곱과 다른 점이다.

어떤 Block matrix의 diagonal block들이 square이고 diagonal block의 아래쪽 block들이 영행렬이면 행렬의 거듭제곱이 Fact 2.1처럼 간단하게 표현된다.

Fact 2.1 Let a square matrix A be $\begin{bmatrix} B & C_1 \\ 0 & D \end{bmatrix}$ where B and D are square matrices. Then, $A^k = \begin{bmatrix} B^k & C_k \\ 0 & D^k \end{bmatrix}$ for some C_k 's.

Proof: $A^2 = \begin{bmatrix} B^2 & BC_1 + C_1D \\ 0 & D^2 \end{bmatrix}$ and $C_2 = BC_1 + C_1D$. If we assume $A^{k-1} = \begin{bmatrix} B^{k-1} & C_{k-1} \\ 0 & D^{k-1} \end{bmatrix}$,
 $A^k = \begin{bmatrix} B^k & BC_{k-1} + C_1D^{k-1} \\ 0 & D^k \end{bmatrix}$ and $C_k = BC_{k-1} + C_1D^{k-1}$. ■

2.5 Inverse of a Matrix

행렬 간의 덧셈에 대해서는 항등원이 영행렬이고 자연스럽게 행렬의 모든 원소에 minus를 취한 행렬이 역원이다. 그리고 행렬 간의 곱셈에 대해서는 항등원으로 identity matrix가 정의되는 것을 공부했는데 이 절에서는 곱셈에 대한 역원의 역할을 하는 행렬을 찾아보자. 먼저 곱셈에 대한 역원의 정의는 다음과 같다.

Definition 2.3 Let A be an $m \times n$ matrix. A matrix B is a left-inverse of A if $BA = I_n$ and C is a right-inverse of A if $AC = I_m$. If B is both left-inverse and right-inverse of A , then we say that A is invertible and has an inverse.

이 정의와 관련되어 관례나 정의로부터 추가적으로 생각해 낼 수 있는 간단한 사실들은 다음과 같다.

- If B is a $k \times \ell$ matrix, then $n = k$ and $\ell = m$ for AB and BA to be well-defined. That is, B has to be $n \times m$. In fact, $m = n$ if B is an inverse of A , which will be shown later.
- Inverse is unique if it exists: if both B and C are inverses of A , then

$$B = BI_m = B(AC) = (BA)C = I_nC = C.$$

We usually denote the inverse of A as A^{-1} .

- **Useful fact:** 간단한 다음 사실은 이후에 특별한 언급없이 자주 사용되니 잘 이해해두면 유용하다.

If there exists an $\mathbf{x} \neq \mathbf{0}$ satisfying $A\mathbf{x} = \mathbf{0}$, then A has no left-inverse and is not invertible.

만약 left-inverse B 가 존재한다면 $\mathbf{0} = B(A\mathbf{x}) = (BA)\mathbf{x} = I_n\mathbf{x} = \mathbf{x}$ 어서 모순이다. 따라서 A 는 left-inverse를 갖지 못하고 A 는 invertible하지 않다.

- **Caution for the use of a left-inverse while solving $Ax = b$:** Assume that a left-inverse B of A exists. Then, for $Ax = b$, left multiplication of B to both sides gives $B(Ax) = Bb$ and hence $x = I_n x = (BA)x = B(Ax) = Bb$. However, for $x = Bb$, $Ax = A(Bb) = (AB)b$ may not reduce to b unless B is a right-inverse of A . Once A has its inverse B , then Bb is a solution of $Ax = b$.

Example 2.1 예를 들면, $A = \begin{bmatrix} 1 & \\ 1 & \end{bmatrix}$ 이고 $\mathbf{b} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ 일 때 $Ax = b$ 를 생각하자. 1×2 행렬 $B = [1 \ 0]$

는 $BA = [1]$ 이어서 A 의 left-inverse이다. $Bb = (2)$ 이지만 $x = (2)$ 는 $Ax = b$ 의 해가 아니다.⁴ 만약 B 가 right-inverse가 아닌 left-inverse이라면, B 를 연립방정식의 양변에 곱하는 것은 마치 등식이 아닌 식의 양변에 0을 곱하는 것과 비슷한 일을 하는 것이다. ■

- If $ad - bc \neq 0$, then $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Please check that this is the inverse!
- Inverse of diagonal matrix is simple: $A = \begin{bmatrix} d_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & d_n \end{bmatrix} \Rightarrow A^{-1} = \begin{bmatrix} d_1^{-1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & d_n^{-1} \end{bmatrix}$ if all $d_i \neq 0$. If some $d_i = 0$, then, for any B , $(AB)_{ij} = 0$ for $j = 1, \dots, n$ and $AB \neq I$. That is, A is not invertible.
- If both A and B are invertible, then $(AB)^{-1} = B^{-1}A^{-1}$. Check that $(AB)(B^{-1}A^{-1}) = I$ and $(B^{-1}A^{-1})(AB) = I$.
- If A is invertible, then $(A^{-1})^\top = (A^\top)^{-1}$ since $A^\top(A^{-1})^\top = (A^{-1}A)^\top = I^\top = I$.
- If A is symmetric and invertible, then A^{-1} is symmetric since $(A^{-1})^\top = (A^\top)^{-1} = A^{-1}$.

상대적으로 간단하지만 이후 여러번 활용되는 삼각행렬의 역행렬에 대한 결과 한 가지를 소개한다. 이 결과에 대한 증명 방법을 잘 익혀두는 것은 결과 자체의 활용 이상으로 유용하다.

Theorem 2.1 Assume A is an upper triangular matrix. Then, A is invertible if and only if every diagonal element of A is non-zero. A^{-1} is also upper triangular if it exists.

Proof: We use mathematical induction on n . As the induction hypothesis, we assume that the statement holds for matrices of size less than n . Note that the fact holds for all 1×1 matrices. Denote an $n \times n$ upper triangular matrix A as $\begin{bmatrix} A_{n-1} & \mathbf{u} \\ \mathbf{0}_{n-1}^\top & a \end{bmatrix}$ where A_{n-1} is an $(n-1) \times (n-1)$ upper triangular matrix, $\mathbf{u} \in \mathbb{R}^{n-1}$, and $a \in \mathbb{R}$. We may use $\mathbf{0}$ instead of $\mathbf{0}_{n-1}$ for simplicity of presentation.

⁴사실 $Ax = b$ 를 만족하는 해가 없다.

- “only if” part: Assume that the $n \times n$ upper triangular matrix A is invertible. If $a = 0$, then the last row of A vanishes, which makes the last row of AB vanish as well whatever B is. This contradicts to the invertibility of A . So, $a \neq 0$. Let $B = \begin{bmatrix} B_{n-1} & \mathbf{v} \\ \mathbf{w}^\top & b \end{bmatrix}$ be an inverse of A . From

$$AB = \begin{bmatrix} A_{n-1} & \mathbf{u} \\ \mathbf{0}^\top & a \end{bmatrix} \begin{bmatrix} B_{n-1} & \mathbf{v} \\ \mathbf{w}^\top & b \end{bmatrix} = \begin{bmatrix} A_{n-1}B_{n-1} + \mathbf{u}\mathbf{w}^\top & A_{n-1}\mathbf{v} + \mathbf{u}b \\ a\mathbf{w}^\top & ab \end{bmatrix} = I_n,$$

we need $a\mathbf{w}^\top = \mathbf{0}^\top$ which implies $\mathbf{w} = \mathbf{0}$ from $a \neq 0$. Then $\mathbf{u}\mathbf{w}^\top = \mathbf{0}_{n-1,n-1}$ and the first block of AB has to satisfy $A_{n-1}B_{n-1} = I_{n-1}$. On the other hand,

$$BA = \begin{bmatrix} B_{n-1} & \mathbf{v} \\ \mathbf{0}^\top & b \end{bmatrix} \begin{bmatrix} A_{n-1} & \mathbf{u} \\ \mathbf{0}^\top & a \end{bmatrix} = \begin{bmatrix} B_{n-1}A_{n-1} & B_{n-1}\mathbf{u} + \mathbf{v}a \\ \mathbf{0}^\top & ba \end{bmatrix} = I_n$$

also implies $B_{n-1}A_{n-1} = I_{n-1}$. Therefore, A_{n-1} is an invertible upper triangular matrix of size $n - 1$, and its diagonal should be non-zero by the induction hypothesis. Combining with $a \neq 0$, all diagonals of A are non-zero, and the “only if” statement holds for matrices of size n .

- “if” part: Assume that A has non-zero diagonals. Then, A_{n-1} is invertible by the induction hypothesis since A_{n-1} is an $(n - 1) \times (n - 1)$ upper triangular matrix with non-zero diagonals. If we set $B = \begin{bmatrix} A_{n-1}^{-1} & -a^{-1}A_{n-1}^{-1}\mathbf{u} \\ \mathbf{0}^\top & a^{-1} \end{bmatrix}$, then it is easy to check $BA = AB = I_n$. So $A^{-1} = B$ and A is invertible.

If A is invertible, $A^{-1} = \begin{bmatrix} A_{n-1}^{-1} & -a^{-1}A_{n-1}^{-1}\mathbf{u} \\ \mathbf{0}^\top & a^{-1} \end{bmatrix}$. Since A_{n-1}^{-1} is upper triangular by the induction hypothesis, A^{-1} is also upper triangular. ■

행렬의 inverse와 transpose 간의 관계를 떠올리면 lower triangular matrix에 대해서도 Theorem 2.1이 동일하게 성립하는 것을 알 수 있다.

2.5.1 Triangular Factors

Gaussian elimination 과정에서 계수행렬의 왼쪽에 곱해지던 행렬들은 하삼각행렬이었는데 이는 우연이 아니다. 다음과 같이 하삼각행렬을 만드는 세번의 과정을 살펴보자.

$$\begin{bmatrix} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix} \xrightarrow{\textcircled{1}} \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ -2 & 7 & 2 & 9 \end{bmatrix} \xrightarrow{\textcircled{2}} \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{bmatrix} \xrightarrow{\textcircled{3}} \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

여기에서 사용된 세 하삼각행렬과 그 역행렬은 다음과 같다.

$$\textcircled{1}: (\text{equation 2}) - 2(\text{equation 1}), \quad \tilde{L}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad L_1 = \tilde{L}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\textcircled{2}: \text{(equation 3)} + \text{(equation 1)}, \quad \tilde{L}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad L_2 = \tilde{L}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\textcircled{3}: \text{(equation 3)} + \text{(equation 2)}, \quad \tilde{L}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad L_3 = \tilde{L}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

세 하삼각행렬을 순서대로 곱하면 여러 번의 Gaussian elimination을 표현하는 하삼각행렬을 얻을 수 있다. 일반적으로 여러 개의 하삼각행렬을 곱하더라도 그 곱이 하삼각행렬임을 확인해보기 바란다. 얻어진 하삼각행렬 \tilde{L} 과 그 역행렬 L 이 아래에 기술되어 있다.

$$\tilde{L} = \tilde{L}_3 \tilde{L}_2 \tilde{L}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}, \quad L = \tilde{L}^{-1} = \tilde{L}_1^{-1} \tilde{L}_2^{-1} \tilde{L}_3^{-1} = L_1 L_2 L_3 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}.$$

Gaussian elimination에 의해 변경된 계수행렬은 상삼각행렬이 되고 현재 살펴보는 예에서는 다음과 같다.

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}.$$

왼쪽에 곱해진 하삼각행렬을 우변으로 이항하고 변경된 계수행렬을 U 라 표시하면

$$\begin{aligned} \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/2 \\ 0 & 1 & 1/4 \\ 0 & 0 & 1 \end{bmatrix} \\ &= LDU \end{aligned}$$

와 같이 (하삼각행렬) \times (대각행렬) \times (상삼각행렬)을 얻는다. L, D, U 를 구해서 표현한 LDU 를 행렬 A 의 LDU -decomposition 또는 LU -decomposition이라고 부른다. LU -분해는 A 의 역행렬을 구하는 경우에도 활용 가능하다. U 도 invertible하면 $A^{-1} = U^{-1}L^{-1}$ 이므로 $[A|I]$ 로 확장시킨 $n \times (2n)$ 행렬의 앞쪽 절반이 단위행렬이 되도록 Gaussian elimination을 행렬 전체에 수행하면

$$U^{-1}\tilde{L} [A|I] = U^{-1} [U|\tilde{L}] = [I|U^{-1}\tilde{L}] = [I|A^{-1}]$$

에 의해 단위행렬로 확장된 부분이 역행렬로 바뀌게 된다. 이 방법으로

$$A^{-1} = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}^{-1}$$

를 계산해보자. 역행렬을 구하기 위해 원래 계수행렬 옆에 identity matrix를 덧붙여서(augment)

$$\left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & -6 & 0 & 0 & 1 & 0 \\ -2 & 7 & 2 & 0 & 0 & 1 \end{array} \right]$$

와 같이 큰 행렬을 만든다. 그리고 다음과 같이 왼쪽에 Gaussian elimination에 사용된 하삼각행렬 \tilde{L} 을 곱하여 계수행렬을 상삼각행렬 U 로 만든 뒤에 다시 적절한 상삼각행렬 U^{-1} 를 왼쪽에 곱하여 계수행렬을 대각행렬로 바꾼다.

$$\begin{aligned} \left[\begin{array}{ccc} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{array} \right] \left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & -6 & 0 & 0 & 1 & 0 \\ -2 & 7 & 2 & 0 & 0 & 1 \end{array} \right] &= \left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array} \right] \\ \Rightarrow \left[\begin{array}{ccc} 1 & 1/8 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right] \left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array} \right] &= \left[\begin{array}{ccc|ccc} 2 & 0 & 3/4 & 3/4 & 1/8 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array} \right] \\ \Rightarrow \left[\begin{array}{ccc} 1 & 0 & -3/4 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{array} \right] \left[\begin{array}{ccc|ccc} 2 & 0 & 3/4 & 3/4 & 1/8 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array} \right] &= \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 3/2 & -5/8 & -3/4 \\ 0 & -8 & 0 & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array} \right] \\ \Rightarrow \left[\begin{array}{ccc} 1/2 & 0 & 0 \\ 0 & -1/8 & 0 \\ 0 & 0 & 1 \end{array} \right] \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 3/2 & -5/8 & -3/4 \\ 0 & -8 & 0 & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array} \right] &= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3/4 & -5/16 & -3/8 \\ 0 & 1 & 0 & 1/2 & -3/8 & -1/4 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array} \right] \end{aligned}$$

마지막에는 적절한 대각행렬을 왼쪽에 곱하여 계수행렬이 identity matrix가 되도록 만든다. 원래 계수행렬 오른쪽에 augment된 identity matrix는 왼쪽에 곱해지는 모든 행렬의 곱을 나타내는데, 계수행렬이 identity matrix로 변경되었으므로 계수행렬 오른쪽의 행렬이 역행렬이 됨을 알 수 있다.

$$A^{-1} = \begin{bmatrix} 3/4 & -5/16 & -3/8 \\ 1/2 & -3/8 & -1/4 \\ -1 & 1 & 1 \end{bmatrix}.$$

이를 알파벳기호로 표현하면 다음과 같이 (보기)e) 간단하게

$$A^{-1} = (LDU)^{-1} = U^{-1}D^{-1}L^{-1}$$

으로 나타낼 수 있지만 L, D, U 는 결국 Gaussian elimination을 통해 알 수 있다.

2.5.2 LU-Factorization

우리가 위에서 수행한 Gaussian elimination은 i 번째 행을 변경할 때, i 번째 행보다 위쪽 행들(즉 $1, \dots, i-1$ 번째 행들)의 선형조합을 i 번째 행에 더하는 방법으로 이루어졌다. 따라서 이를 나타내는 행렬이 하삼각행렬 일 뿐만 아니라 대각원소가 반드시 1이 됨을 알 수 있다. 이 과정을 나타내는 대각원소가 1인 하삼각행렬의 역행렬을 L 로 나타내고 변경된 계수행렬을 상삼각행렬 U 라 하면

$$A = LU$$

로 분해되는데 이를 ***LU-factorization***이라고 부른다. *LU-factorization*을 얻으면

$$\begin{aligned} Ax = b &\Rightarrow LUx = b \quad \text{or} \quad Ly = b \\ &\Rightarrow y = L^{-1}b \\ &\Rightarrow Ux = y = L^{-1}b \\ &\Rightarrow x = U^{-1}L^{-1}b \end{aligned}$$

를 통해서 선형연립방정식을 풀 수 있다. 이전에 언급한 바와 같이 계수행렬이 삼각행렬이면 back-substitution 에 의해 방정식을 간단하게 풀 수 있다. 적절한 대각행렬 D 를 이용하면 U 도 대각원소가 1이 되도록 변경하여 $A = LDU$ 분해가 가능해진다.

대각행렬 D 의 원소가 모두 0이 아닌 경우 이러한 분해가 유일함을 다음과 같이 살펴볼 수 있다. 한 행렬의 두 가지 분해 $L_1 D_1 U_1 = L_2 D_2 U_2$ 를 고려하자. 이 등식에서 L_2 와 U_1 을 반대편으로 이항시키면 $L_2^{-1} L_1 D_1 = D_2 U_2 U_1^{-1}$ 를 얻는다. 한번 더 D_1 을 오른쪽으로 이항시키면 $L_2^{-1} L_1 = D_2 U_2 U_1^{-1} D_1^{-1}$ 를 얻는다. Theorem 2.1에 의해 $L_2^{-1} L_1$ 과 $U_2 U_1^{-1}$ 은 각각 하삼각행렬과 상삼각행렬이다. 그리고 $L_2^{-1} L_1$ 의 대각원소들이 모두 1임도 쉽게 알 수 있다. 또한, 대각행렬 D_2 와 D_1^{-1} 을 곱하는 것은 상삼각행렬의 형태에 영향을 끼치지 못하므로 $L_2^{-1} L_1 = D_2 U_2 U_1^{-1} D_1^{-1}$ 의 좌변은 대각원소가 1인 하삼각행렬이고 우변은 상삼각행렬이다. 따라서 양변 모두 identity matrix이어야 한다. $L_2^{-1} L_1 = I$ 로부터 $L_1 = L_2$ 를 얻게되고 $D_1 = D_2$ 과 $U_1 = U_2$ 를 차례로 얻을 수 있다.

*LU-factorization*과 관련하여 대칭행렬의 경우에 관찰되는 특별한 성질을 살펴보자. 대칭행렬 A 가 행교환 없이 $A = LDU$ 로 분해되었다고 하자. 대칭성 때문에 $LDU = U^\top D L^\top$ 인데 *LU-factorization*의 uniqueness 에 의해 $U = L^\top$ 임을 알 수 있어 $A = LDL^\top$ 로 분해된다.

LU-Factorization with Row Exchange

Permutation matrix는⁵ identity matrix의 두 열을 (여러번) 교환한 행렬로 생각할 수 있다. 여러번 교환했지만 permutation matrix의 각 행과 열들은 한 원소만 1을 값으로 갖고 나머지 원소들은 모두 0이다. 다음과 같은 행렬

$$A = \begin{bmatrix} 0 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

⁵Permutation matrix에 대한 자세한 내용은 Appendix A를 참조하기 바란다.

에 대해서는 Gaussian elimination을 시작할 수 없다. 하지만, 두 방정식의 순서를 바꾸듯이 첫 번째 행과 두 번째 행을 교환하면 Gaussian elimination을 수행할 수 있다. 이 교환에 대응되는 permutation matrix는

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

이고, 이렇게 행을 교환한⁶ 행렬

$$PA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 6 \\ 0 & 2 & 3 \end{bmatrix}$$

에 대해서는 앞 절에서처럼 LU-factorization을 수행하여 $PA = LU$ 를 만족하는 하삼각행렬 L 과 상삼각행렬 U 를 얻을 수 있다.

2.6 Inverse of Block Matrices

A 가 정사각행렬로서 작은 크기의 정사각행렬 A_{11} 과 A_{22} 에 대해 다음과 같이 block matrix로 표현된다고 하자:

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

만약 A_{11} 이 invertible하다면 Gaussian elimination에 의해 A_{21} 을 소거할 수 있다. 이 소거과정을 행렬을 이용하여 표현하면 다음과 같다.

$$\begin{bmatrix} A_{11}^{-1} & \mathbf{0} \\ -A_{21}A_{11}^{-1} & I_{22} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_{11} & A_{11}^{-1}A_{12} \\ \mathbf{0} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}. \quad (2.5)$$

마찬가지로 A_{22} 이 invertible하다면 A_{12} 를 소거할 수 있으며 이 Gaussian elimination은

$$\begin{bmatrix} I_{11} & -A_{12}A_{22}^{-1} \\ \mathbf{0} & A_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & \mathbf{0} \\ A_{22}^{-1}A_{21} & I_{22} \end{bmatrix} \quad (2.6)$$

로 표현된다. (2.5)의 우변의 행렬을 간단하게 나타내기 위해

$$S_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12}$$

라는 행렬을 도입하자. 이 행렬 S_{22} 를 **Schur complement of A_{11} with respect to A** 라고 부른다.⁷ 그러면 (2.5)의 우변 행렬은 $\begin{bmatrix} I_{11} & A_{11}^{-1}A_{12} \\ \mathbf{0} & S_{22} \end{bmatrix}$ 로 나타낼 수 있다. 추가적으로 S_{22} 가 invertible하다면 (2.6)을 활용하여⁸ 다음과 같은 Gaussian elimination을 수행할 수 있다:

$$\begin{bmatrix} I_{11} & -A_{11}^{-1}A_{12}S_{22}^{-1} \\ \mathbf{0} & S_{22}^{-1} \end{bmatrix} \begin{bmatrix} I_{11} & A_{11}^{-1}A_{12} \\ \mathbf{0} & S_{22} \end{bmatrix} = \begin{bmatrix} I_{11} & \mathbf{0} \\ \mathbf{0} & I_{22} \end{bmatrix}.$$

⁶Permutation matrix을 다른 행렬의 앞에 곱하면 행을 교환하고 뒤에 곱하면 열을 교환한다.

⁷만약 A_{22} 가 invertible하다면 $S_{11} = A_{11} - A_{12}A_{22}^{-1}A_{21}$ 가 **Schur complement of A_{22} with respect to A** 이다.

⁸ $A_{11} \Leftarrow I_{11}$, $A_{12} \Leftarrow A_{11}^{-1}A_{12}$, $A_{21} \Leftarrow \mathbf{0}$, $A_{22} \Leftarrow S_{22}$ 로 놓고 (2.6)를 적용하란 의미이다.

행렬 A 의 왼쪽에 곱해진 두 행렬을 곱하여 표현하면

$$\begin{bmatrix} I_{11} & -A_{11}^{-1}A_{12}S_{22}^{-1} \\ \mathbf{0} & S_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & \mathbf{0} \\ -A_{21}A_{11}^{-1} & I_{22} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_{11} & \mathbf{0} \\ \mathbf{0} & I_{22} \end{bmatrix}$$

을 얻게 되어 원하던 inverse는 왼쪽에 곱해진 두 행렬의 곱인

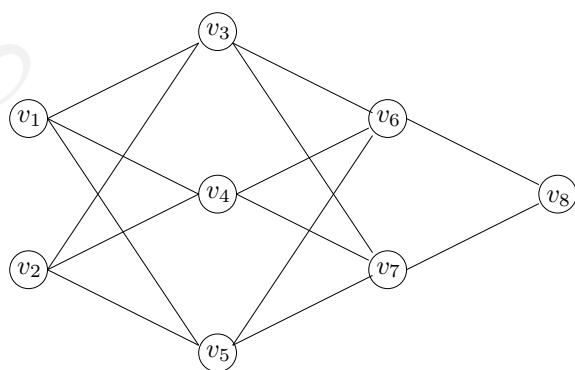
$$\begin{aligned} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} &= \begin{bmatrix} I_{11} & -A_{11}^{-1}A_{12}S_{22}^{-1} \\ \mathbf{0} & S_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & \mathbf{0} \\ -A_{21}A_{11}^{-1} & I_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}S_{22}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}S_{22}^{-1} \\ -S_{22}^{-1}A_{21}A_{11}^{-1} & S_{22}^{-1} \end{bmatrix} \end{aligned} \quad (2.7)$$

이다.

Gaussian elimination을 표현하는 식 (2.5)과 (2.6)은 행렬과 관련된 다양한 사고실험에 유용하다. 특히 두 표현은 block matrix의 determinant를 구하는데도 활용된다.

2.7 Application to Data Science: Graphs and Matrices

사회적 관계나 engineering system 등의 표현에서 선분으로 여러 객체 간의 관계를 나타낸 수학적 대상을 그래프 또는 네트워크(graph/network)라고 부른다. 아래 그림에서 v_i 를 둘러싼 원들은 node라고 부르는데 사람, 조직, 기계, 컴퓨터 등을 나타내며 다른 node와 관계를 맺을 수 있는 모든 대상을 표현한다. 그리고 두 node 간에 우리가 관심있는 관계가 있으면 선분으로 연결하고 이 선분을 edge라고 부른다. 이러한 시각적 표현은 직관적이어서 좋으나 node의 상대적 위치 등에 따라서 동일한 형태의 그래프가 다르게 보일 수도 있는 단점이 있다. 따라서 그래프의 특성을 파악하거나 그래프와 연관된 여러 계산을 체계적으로 수행하기 위해 이렇게 시각적으로 표현된 그래프를 행렬로 표현한다.



그래프의 node의 개수가 n 개 이면 $n \times n$ 행렬을 만드는데 node에 일련번호를 붙인 뒤에 그 일련번호와 행과 열을 대응시킨다. 그리고 i 번째 node와 j 번째 node가 edge로 연결되었으면 행렬의 (i, j) 원소의 값을 1로 정하고 연결되지 않았으면 0으로 정한 행렬이 adjacency matrix이다. 관계를 나타내는 edge가 방향성이 없으므로 1을 할당하는 규칙이 대칭적이어서 adjacency matrix는 대칭행렬이다.

예로 위에 제시한 그래프의 경우 node v_i 를 i 번째 행과 i 번째 열에 대응시키면 다음과 같은 adjacency matrix를 얻는다.

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

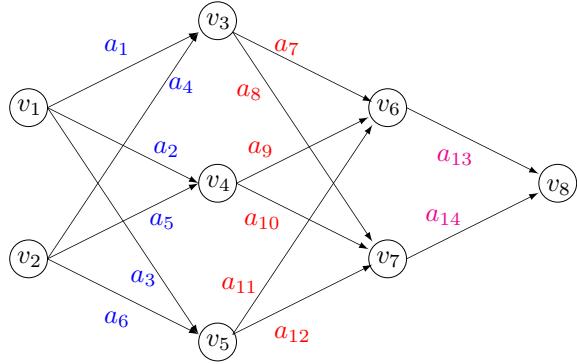
동일한 그래프에 대해 node를 $v_3, v_4, v_5, v_8, v_1, v_2, v_6, v_7$ 의 순서로 adjacency matrix를 작성해보자. 0은 편의상 공백으로 처리했다. 그러면 아래와 같은 block matrix를 얻는다.

$$\left[\begin{array}{c|ccccc} & 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 & 1 \\ & 1 & 1 & & & \\ \hline 1 & 1 & 1 & & & \\ 1 & 1 & 1 & & & \\ 1 & 1 & 1 & 1 & & \\ 1 & 1 & 1 & 1 & & \end{array} \right] = \begin{bmatrix} \mathbf{0} & B \\ B^\top & \mathbf{0} \end{bmatrix}$$

이렇게 표현된 행렬에서 block B 의 행에 대응되는 node $\{v_3, v_4, v_5, v_8\}$ 를 그룹 1, 열에 대응되는 node $\{v_1, v_2, v_6, v_7\}$ 를 그룹 2라 하면 그래프의 모든 edge는 그룹 1과 그룹 2의 node를 연결하고 각 그룹 내의 node를 연결하는 edge는 없다는 것을 알 수 있다.⁹ 그래프의 adjacency matrix를 분석하고 가공하여 그래프의 여러 특성을 알아내고 그래프로 표현된 시스템에 관한 다양한 분석이 가능하다. Adjacency matrix는 data science 뿐만 아니라 응용 수학과 공학, 사회 과학 분야에서 널리 활용된다. 그리고 이런 분석에서 block matrix를 적절하게 활용하면 유용하다. 예를 들면 $\begin{bmatrix} B & C \\ \mathbf{0} & D \end{bmatrix}$ 행렬의 node들의 연결 관계를 생각하거나, block 구조가 자연스러운 여러 행렬의 역행렬을 원래 block에 대해서 해석하는 등 여러 방면에서 block matrix가 중요하다.

만약 관계를 나타내는 선분이 방향이 있는 화살표인 경우 이 화살표를 arc라고 부르고 edge와는 다르게 다음 그림과 같이 고유한 이름을 붙이는 경우가 많다

⁹이런 그래프를 bipartite graph라고 부른다.



이러한 유방향 그래프(directed graph)는 네트워크(network)라고도 부르며 행을 node에 대응시키고 열은 arc에 대응시켜서 arc j 가 시작되는 node i 에 대응되는 행렬의 원소 $a_{ij} = 1$ 로, arc j 가 끝나는 node i 대해서는 $a_{ij} = -1$ 로 놓고 나머지 원소를 0으로 놓은 행렬을 incidence matrix라고 부른다. 위 네트워크의 incidence matrix A 는

$$A = \begin{bmatrix} 1 & 1 & 1 & & & & & \\ & 1 & 1 & 1 & & & & \\ -1 & & -1 & & 1 & 1 & & \\ & -1 & & -1 & & 1 & 1 & \\ & & -1 & & -1 & & 1 & 1 \\ & & & -1 & & -1 & -1 & 1 \\ & & & & -1 & & -1 & -1 \\ & & & & & -1 & -1 & -1 \\ & & & & & & -1 & -1 \end{bmatrix}$$

이다. 이 행렬의 특징은 각 arc의 시작과 끝 node가 한 개씩이므로 각 열에 1과 -1이 한 개씩만 존재한다는 것이다. 이 network이 원유를 수송하는 파이프 라인이라고 생각하고 원유가 흐르는 arc j 의 시간당 유량을 x_j 라 표시하면 $\mathbf{x} = (x_j)$ 는 모든 arc의 시간당 유량을 나타내는 벡터이다. 이때 $A\mathbf{x}$ 는 무엇을 나타낼까? $A\mathbf{x}$ 는 파이프 라인의 각 node에서 유입된 원유와 유출된 원유의 차이를 나타낸다. 따라서 $b_1 > 0, b_2 > 0, b_8 = -(b_1 + b_2) < 0, b_3 = \dots = b_7 = 0$ 인 \mathbf{b} 에 대해 $A\mathbf{x} = \mathbf{b}$ 를 만족하는 \mathbf{x} 는 v_1, v_2 에서 각각 b_1, b_2 만큼 공급된 원유가 v_8 로 전량 흐르고 나머지 node에서는 원유의 소비나 공급 없이 흐름이 보존되는 network의 흐름을 나타낸다. 물론 현실에 해석이 어려운 음수인 흐름을 방지하고 싶으면 $x_j \geq 0$ 이라는 조건이 추가되어야 한다. $A\mathbf{x} = \mathbf{b}$ 는 흐름의 보존을 표현하는 incidence matrix의 좋은 활용 예이다.

Chapter 3

Vector Spaces

물리적 힘이 가해지고 있는 물체를 생각해 보자. 이 물체의 운동을 변화시키기 위해서는 기존에 가해지는 힘의 크기를 변화시키거나 새로운 힘을 가해야 한다. 이 경우 물체에는 원래 가해지던 힘에 새로 추가된 힘이 동시에 작용하거나, 크기만 바뀐 힘이 작용한다. 두 힘이 동시에 작용하는 효과를 하나의 힘으로 나타내는 것이 힘의 덧셈이고, 힘의 크기만을 바꾸는 것은 힘을 상수배하는 것이다. 물체의 운동을 변화시키는 방법은 결국 이 두 가지 방법을 잘 조합하여 반복하는 것이다. 이 간단한 예에 상수배가 가능한 대상들의 덧셈이 주어지는 벡터공간의 구조가 담겨있다. 사실 벡터공간에서 연산의 핵심은 상수배가 가능한 벡터로 불리는 대상들의 덧셈이 상수배와 조화를 이루게 하는 것이다. 이러한 벡터공간의 예는 2차원 점들로 이루어진 평면, 같은 형태의 행렬을 모아 놓은 집합, 이차 다항식을 모두 모은 집합, 동일한 확률공간 상의 확률변수 등을 생각할 수 있다. 이렇게 간단한 연산 구조를 갖는 벡터공간에서도 흥미롭고 유용한 결과를 많이 관찰할 수 있다. 상수배한 여러 벡터들을 더하는 것을 선형조합이라고 부르는데 유한 개의 벡터들이 중복이 없이 최소로 모은 것인지 여부를 어떤 벡터도 다른 벡터들의 선형조합으로 표시할 수 없는지로 판단할 수 있는데 중복이 없는 경우 선형독립이라고 한다. 또한 벡터공간의 모든 벡터들이 고정된 유한 개의 벡터들의 선형조합으로 표시할 수 있으면 벡터공간을 효율적으로 표현할 수 있다. 특히 선형독립인 유한개의 벡터가 선형조합으로 벡터공간의 모든 벡터를 표현하는 경우 이 벡터들을 기저라고 부른다. 그리고 기저를 구성하는 벡터들이 늘 동일하지 않을 수 있지만 벡터들의 개수는 벡터공간마다 고유하고 이를 그 벡터공간의 차원이라고 부른다. 그리고 상수배되는 수의 집합이 동일한 경우, 차원이 같은 벡터공간은 거의 동일한 구조를 갖는다.

두 벡터공간 사이에 벡터연산을 보존하는 선형사상이라 불리는 함수를 도입하면 더 풍부한 결과를 관찰할 수 있다. 선형사상은 정의된 두 벡터공간의 기저가 정해지면 행렬로 표현되는데 이 대응이 일대일 관계이어서 행렬과 선형사상의 구조에 대해 한쪽에서 얻은 결과를 다른 쪽에 적용할 수 있다. 예를 들면, 행렬의 열공간과 영공간에 대한 결과는 선형사상의 치역과 커널에 대한 결과와 동일하다. 행렬의 열공간과 영공간은 Gaussian elimination을 통해 알아낼 수 있다. 이런 작업들을 좀 더 체계화하기 위해서 row echelon form을 도입하고 이 행렬의 각 행의 영이 아닌 첫번째 원소들의 개수와 위치가 열공간과 영공간을 찾게 해주고 그 개수가 행렬의 rank인데 행렬이 작용하는 자유도에 해당하여 행렬과 벡터공간이 차원에 관해 많은 정보를 제공한다. 선형사상이 기하학적인 대응관계를 나타내는 회전, 대칭, 정사영 등에 대응되는 행렬이 갖는 특별한 형태에 대해서도 알아본다.

3.1 Vector Spaces and Subspaces

선형연립방정식 $A\mathbf{x} = \mathbf{b}$ 의 해를 찾는 과정에서 필요한 여러 연산과 방정식에 연관된 homogeneous equation인 $A\mathbf{x} = \mathbf{0}$ 의 해집합의 특성을 추상화하여 vector space와 vector operation을 정의한다. vector space를 채우는 vector로서 우리가 익숙한 \mathbb{R}^n 의 점을 떠올리겠지만, 같은 크기의 행렬들이나 동일한 정의구역 상에서 정의된 실수값을 갖는 함수들을 모아도 vector space가 되므로 vector space는 매우 일반적인 개념이다.

3.1.1 Operations in a Vector Space

Vector space의 기본적인 연산(operation)은 vector addition과 scalar multiplication 두 가지이다. 나머지 연산은 모두 이 두 가지 기본연산에서 파생된다.

1. Scalar multiplication: 상수배를 생각하려면 먼저 scalar를 알아야 한다. 본 강의록에서는 \mathbb{R} 과 \mathbb{C} 가 가능한데 주로 \mathbb{R} 을 scalar의 집합으로 생각하면 된다. Scalar c 와 vector \mathbf{v} 에 대해서 상수배 $c\mathbf{v}$ 역시 vector가 된다. Scalar 집합의 곱셈에 대한 항등원(multiplicative identity)은 1로 표시하는데 임의의 vector \mathbf{v} 에 대해 $1\mathbf{v} = \mathbf{v}$ 가 성립하는 것으로 약속한다. 그리고 두 scalar를 상수배할 때 일종의 결합법칙(the association of scalar multiples and a vector)인 $(c_1 c_2)\mathbf{v} = c_1(c_2\mathbf{v})$ 이 성립한다고 약속한다.
2. Vector addition: 두 vector \mathbf{v}_1 과 \mathbf{v}_2 에 대해서 $\mathbf{v}_1 + \mathbf{v}_2$ 는 vector이다. 벡터의 덧셈은 교환법칙과 결합법칙이 성립한다(commutative and associative): $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1$, $\mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3) = (\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3$. 덧셈의 항등원은 $\mathbf{0}$ 으로 표시하는데 대부분의 vector space에서 직관적으로 파악할 수 있다. 예를 들면, zero vector, 영행렬, 0값만을 취하는 상수함수 등이다. Scalar집합의 덧셈에 대한 항등원을 0으로 표시하는데 임의의 vector \mathbf{v} 에 대해서 $0\mathbf{v} = \mathbf{0}$ 으로 약속하는 것이 자연스럽다. Scalar 1의 덧셈에 대한 역원 -1 에 대해서 $(-1)\mathbf{v} = -\mathbf{v}$ 로 표시한다.
3. Two distributive interactions between vector addition and scalar multiplication: $c(\mathbf{v}_1 + \mathbf{v}_2) = c\mathbf{v}_1 + c\mathbf{v}_2$, $(c_1 + c_2)\mathbf{v} = c_1\mathbf{v} + c_2\mathbf{v}$. 상수배와 vector addition의 분배 법칙이다. Vector 관련된 계산을 할 때, 우리가 매번 인지하지 않더라도 거의 모든 계산 단계에서 사용되고 있다. 예를 들면, 임의의 \mathbf{v} 에 대해 $\mathbf{v} - \mathbf{v} = 1\mathbf{v} + (-1)\mathbf{v} = (1 - 1)\mathbf{v} = \mathbf{0}$ 이므로 $-\mathbf{v}$ 가 \mathbf{v} 의 vector addition에 대한 역원이 된다.

Definition 3.1 A set \mathbb{V} is a vector space (with its scalar \mathbb{R}) if all vectors in \mathbb{V} and scalars in \mathbb{R} satisfy the above operational rules.

Vector space의 예로는 \mathbb{R}^n , \mathbb{R}^∞ , 3×2 matrices $\approx \mathbb{R}^6$, space of vector-valued functions들이 있다. 이 중에서 \mathbb{R}^n 이 표준적인 유한차원 vector space이다. 차원에 관한 이야기는 나중에 이루어진다. Vector space을 도입하면 바로 연결되는 이야기가 부분공간(subspace)에 관한 것이다. Vector space의 일부 벡터들이 자기들만으로도 vector space의 규칙을 충분히 지킬 수 있는 경우, 이 벡터들의 집합을 원래 vector space의 subspace라 부른다. 기존 벡터들로부터 새로운 벡터를 만들어 내는 vector addition과 scalar multiplication를 합하여 벡터들의 선형조합(linear combination of vectors)을 정의한다: vector $\mathbf{v}_1, \dots, \mathbf{v}_k$ 와 scalar c_1, \dots, c_k 에

대해서

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \sum_{i=1}^k c_i\mathbf{v}_i$$

를 선형조합(linear combination)이라고 부른다.

Definition 3.2 A subspace of a vector space is a non-empty subset of the vector space such that all linear combinations stay in the subset.

Vector space \mathbb{V} 의 공집합이 아닌 부분집합 \mathbb{W} 가 subspace임을 보일 때 vector space의 모든 규칙을 확인하지 않고 vector addition과 scalar multiplication에 대해서 \mathbb{W} 가 닫혀있는지만 확인하면 된다. 즉

- $\mathbb{W} \subset \mathbb{V}$;
- $\mathbf{x}, \mathbf{y} \in \mathbb{W} \Rightarrow \mathbf{x} + \mathbf{y} \in \mathbb{W}$;
- $c \in \mathbb{R}, \mathbf{x} \in \mathbb{W} \Rightarrow c\mathbf{x} \in \mathbb{W}$

만을 확인하면 된다.

Vector space \mathbb{V} 의 subspace로는 가장 간단한 $\{\mathbf{0}\}$ 으로부터 $\{c\mathbf{x} : c \in \mathbb{R}\}$ for $\mathbf{x} \in \mathbb{V}$ 와 같은 1차원 subspace, 그리고 종종 필요한 $\{c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n : c_1, \dots, c_n \in \mathbb{R}\}$ for $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{V}$ 등이 있다. $\{(x, y) : x \geq 0, y \geq 0\}$ 는 \mathbb{R}^2 의 subspace가 아니다. 행렬을 모아 놓은 vector space에서는 하삼각행렬을 모은 집합이나 대칭행렬로 구성된 집합들이 subspace이다.

Vector의 상수배에 관해서 유용한 사실 한 가지를 추가하자. Scalar c 와 vector $\mathbf{v} \in \mathbb{R}^n$ 에 대해서 $c\mathbf{v}$ 와 $\mathbf{v}c$ 가 모두 의미있는 표현이다. $c\mathbf{v}$ 는 vector space에서 scalar multiplication의 표준적인 표현방식이다. $\mathbf{v}c$ 는 \mathbf{v} 를 $n \times 1$ 행렬로 보고 c 를 1×1 행렬로 보아 $\mathbf{v}[c]$ 로 해석할 수 있다. 종종 $c\mathbf{v}$ 표현에서는 파악하기 어렵지만 $\mathbf{v}c$ 방식의 표현에서는 상대적으로 쉽게 새로운 표현을 인지하는 경우가 있다. 그리고 다음과 같이 집합과 벡터를 더하는 표기를 도입하면 편리하다.

Definition 3.3 For any subsets A and B of a vector space \mathbb{V} , we define the sum of A and B as

$$A + B = \{\mathbf{u} + \mathbf{v} : \mathbf{u} \in A, \mathbf{v} \in B\}.$$

If both \mathbb{U} and \mathbb{W} are subspaces of \mathbb{V} and $\mathbb{U} \cap \mathbb{W} = \{\mathbf{0}\}$, we denote $\mathbb{U} + \mathbb{W}$ as $\mathbb{U} \oplus \mathbb{W}$ and call it the **direct sum**. For simplicity, we also denote $\{\mathbf{v}\} + A = \mathbf{v} + A$ for $\mathbf{v} \in \mathbb{V}$ and $A \subset \mathbb{V}$.

이 정의에서 direct sum의 중요한 성질을 한가지 유추할 수 있다. $\mathbb{U} \oplus \mathbb{W}$ 의 어느 벡터가 $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}$ 와 $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{W}$ 에 대해서 $\mathbf{u}_1 + \mathbf{v}_1 = \mathbf{u}_2 + \mathbf{v}_2$ 로 표현된다면 $\mathbf{u}_1 - \mathbf{u}_2 = \mathbf{v}_2 - \mathbf{v}_1 \in \mathbb{U} \cap \mathbb{W}$ 인데 Definition 3.3에 따라서 $\mathbb{U} \cap \mathbb{W} = \{\mathbf{0}\}$ 이므로 $\mathbf{u}_1 = \mathbf{u}_2$ 이고 $\mathbf{v}_1 = \mathbf{v}_2$ 이다. 즉 $\mathbb{U} \oplus \mathbb{W}$ 의 벡터들은 \mathbb{U} 와 \mathbb{W} 의 성분들로 유일하게 표현된다는 것을 알 수 있다.

Fact 3.1 A vector in a direct sum has a unique representation: For $\mathbf{v} \in \mathbb{U} \oplus \mathbb{W}$, there exists unique $\mathbf{u} \in \mathbb{U}$ and $\mathbf{w} \in \mathbb{W}$ such that $\mathbf{v} = \mathbf{u} + \mathbf{w}$.

Fact 3.1에 의하면 부분공간 A, B 와 $A \oplus B$ 의 관계가 2차원 유클리드 공간과 그 공간을 구성하는 두 축의 관계와 비슷하다. 즉 $\mathbb{R}^2 = (\mathbb{R} \times \{\mathbf{0}\}) \oplus (\{\mathbf{0}\} \times \mathbb{R})$ 이다. 이 유사성을 염두에 두면 이후에 direct sum으로 표현되는 벡터공간을 이해하기 용이하다.

3.1.2 Two Fundamental Subspaces induced by Matrices

$m \times n$ 행렬 A 와 관련하여 쉽게 생각할 수 있는 \mathbb{R}^m 과 \mathbb{R}^n 의 subspace는 column space와 null space이다. 행렬 $A = [\mathbf{a}_1 | \cdots | \mathbf{a}_n]$, $\mathbf{a}_i \in \mathbb{R}^m$ 를 생각하자.

- The column space of A , $\text{Col}(A)$: the collection of linear combinations of columns of A .

$$\text{Col}(A) = \{v_1\mathbf{a}_1 + \cdots + v_n\mathbf{a}_n : v_1, \dots, v_n \in \mathbb{R}\} = \{A\mathbf{v} : \mathbf{v} \in \mathbb{R}^n\} \subset \mathbb{R}^m.$$

다음은 column space의 (간단한) 성질들로서 이유를 생각해보는 것이 좋은 공부이다.

1. $A\mathbf{x} = \mathbf{b}$ is solvable if and only if $\mathbf{b} \in \text{Col}(A)$;
2. $\text{Col}(I_n) = \mathbb{R}^n$;
3. If A is an $n \times n$ invertible matrix, then $\text{Col}(A) = \mathbb{R}^n$.
4. $\text{Col}(A)$ is a subspace of \mathbb{R}^m .

- The null space of A , $\text{Null}(A)$: the collection of vectors being mapped to $\mathbf{0}$ via the matrix A .

$$\text{Null}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

$\text{Null}(A)$ 는 kernel of A 라고도 불린다. 다음은 null space의 (간단한) 성질들이다. 각자 이유를 생각해보자.

1. $\text{Null}(A) = \mathbb{R}^n$ if and only if $A = \mathbf{0}$;
2. $\text{Null}(I_n) = \{\mathbf{0}\}$;
3. If A is an $n \times n$ invertible matrix, then $\text{Null}(A) = \{\mathbf{0}\}$.
4. $\text{Null}(A)$ is a subspace of \mathbb{R}^n .

한 행렬이 다른 행렬과 곱해지면 보통 column space가 작아지며 오른쪽에 곱해지는 행렬이 invertible한 경우에는 column space에 변화가 없다.

Lemma 3.1 For any matrices A and B where the number of columns of A and the number of rows of B coincide, $\text{Col}(AB) \subset \text{Col}(A)$. If B is invertible, then $\text{Col}(AB) = \text{Col}(A)$.

Proof: For any \mathbf{v} , $AB\mathbf{v} = A(B\mathbf{v}) \in \text{Col}(A)$. Therefore, $\text{Col}(AB) \subset \text{Col}(A)$. Assume B is invertible. Set $C = AB$. Then $A = CB^{-1}$ and $\text{Col}(A) = \text{Col}(CB^{-1}) \subset \text{Col}(C) = \text{Col}(AB)$ by the first part of the lemma. ■

다음 결과에서 null space가 선형연립방정식의 해를 모아 놓은 집합의 구조를 결정하는 중요한 역할을 한다는 것을 알 수 있다. 즉 선형연립방정식의 해집합은 null space를 평행이동하면 얻을 수 있다.

Fact 3.2 Assume $Ax^* = \mathbf{b}$. Then, $\{\mathbf{x} : Ax = \mathbf{b}\} = \{\mathbf{x}^* + \mathbf{y} : \mathbf{y} \in \text{Null}(A)\} = \mathbf{x}^* + \text{Null}(A)$.

Proof: Assume $Ax = \mathbf{b}$. Then, $A(\mathbf{x} - \mathbf{x}^*) = Ax - Ax^* = \mathbf{b} - \mathbf{b} = \mathbf{0}$ and $\mathbf{x} - \mathbf{x}^* = \mathbf{y} \in \text{Null}(A)$, which proves one direction of inclusion. For the other direction, $A(\mathbf{x}^* + \mathbf{y}) = Ax^* + Ay = \mathbf{b} + \mathbf{0} = \mathbf{b}$ for $\mathbf{y} \in \text{Null}(A)$. ■

3.2 Solving $Ax = \mathbf{0}$ and $Ax = \mathbf{b}$

선형연립방정식 $Ax = \mathbf{b}$ 는 계수행렬 A 의 열벡터들이 독립적일수록¹ 더 많은 우변상수 \mathbf{b} 에 대해 해를 갖는다. 열벡터들의 독립성이 최고인 경우로 생각할 수 있는 invertible한 A 의 경우에는 어떠한 \mathbf{b} 에 대해서도 $\mathbf{x} = A^{-1}\mathbf{b}$ 가 $Ax = \mathbf{b}$ 의 해가 된다. 이 경우에 비해 열벡터들의 독립성이 낮은 경우에 해당하는 $\mathbf{0} \neq \mathbf{y} \in \text{Null}(A)$ 가 존재하여 열벡터 간에 $Ay = \mathbf{0}$ 이라는 상호관계가 존재하는 경우에는 \mathbf{b} 에 따라서 $Ax = \mathbf{b}$ 의 해가 존재하지 않거나 존재하더라도 Fact 3.2에 의해 무한히 많은 해가 존재한다. 일반적으로 column space 가 우변상수 벡터들의 공간인 \mathbb{R}^m 보다 작으면(즉 $\text{Col}(A) \subsetneq \mathbb{R}^m$ 이면) $\mathbf{b} \in \text{Col}(A)$ 인 \mathbf{b} 에 대해서만 해가 존재한다.

연립방정식 $Ax = \mathbf{b}$ 을 푸는 과정은 방정식들을 상수배하여 다른 방정식에 더하는 과정을 반복하되 0으로 변하는 미지수의 계수들이 많아지도록 만드는 소거작업을 수행하는 것이다. 이 때 0이 되는 계수의 위치를 가능한 왼쪽 위부터 정하는 방법이 Gaussian elimination이다. 일단 어떤 계수가 0이되면 그 계수 아래쪽의 모든 계수를 0으로 만드는데 이런 과정을 마치면 0인 계수의 모습이 좌측에서 우측으로 내려가는 계단의 모습을 연상시켜 이렇게 소거된 계수행렬을 row echelon form이라고 부른다.

3.2.1 Row Echelon Form U

행렬 $A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix}$ 의 Gaussian elimination을 수행해 보자.

- First pivoting: A 의 첫 번째 열에서 a_{11} 을 제외한 나머지 원소를 0으로 소거하는 행렬이 \tilde{L}_1 이다.

$$\tilde{L}_1 A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 6 & 6 \end{bmatrix}$$

- Second pivoting: $\tilde{L}_1 A$ 의 세 번째 열에서 a_{33} 밑에 위치한 원소를 0으로 소거하는 행렬이 \tilde{L}_2 이다.

$$\tilde{L}_2 \tilde{L}_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 6 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

왼쪽에 곱해지는 행렬들을 elementary matrix라고 부르는데 대각원소가 모두 1이고 대각원소 위쪽의 원소들은 모두 0인 특징을 갖는 하삼각행렬이다. 위의 소거 과정에서 열의 다른 원소를 0으로 만드는데 쓰이고

¹나중에 Definition 3.4에서 벡터들의 독립성에 대해서 정확하게 논의한다.

남은 원소인 $(A)_{11}$ 과 $(\tilde{L}_1 A)_{23}$ 를 **pivot**이라고 부른다. Gaussian elimination 과정을 거치면 아래에 기술하는 계단형태의 상삼각행렬을 얻는데, 0만 있는 행은 아래쪽 행에 위치하는데 이러한 형태를 (**row**) **echelon form**이라고 부른다. Row echelon form은 자동적으로 upper triangular matrix이다.

1. Pivots are the first non-zero entries in their rows;
2. Below each pivot is a column of zeros, obtained by elimination;
3. Each pivot lies to the right of the pivot in the row above.

왼쪽에 곱해져서 소거를 담당하는 하삼각행렬들을 모두 (순서대로) 곱하면

$$\tilde{L} = \tilde{L}_2 \tilde{L}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -2 & 1 \end{bmatrix}$$

이 되는데 Theorem 2.1에 의해 invertible하므로 역행렬을 구하면

$$L = \tilde{L}^{-1} = \tilde{L}_1^{-1} \tilde{L}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$$

이 된다. $\tilde{L}A = U$ 는 $L^{-1}A = U$ 로 표시되고 A 에 대해서 풀면

$$A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \text{lower triangular} \times \text{row echelon form}$$

를 최종적으로 얻는다. Gaussian elimination을 수행하는 과정에서 행을 교환할 필요가 있는 경우 가능한 위쪽의 행과 교환하는 규칙을 적용하면 row echelon form은 유일하다. 별도로 엄밀하게 증명하지 않더라도 반복된 예제로부터 다음 결과를 받아들일 수 있을것이다.

Fact 3.3 For any $m \times n$ matrix A , there is a permutation matrix P , a lower triangular square matrix L with unit diagonals, and an $m \times n$ echelon matrix U which is upper triangular such that $PA = LU$.

Row echelon form에서는 pivot원소가 반드시 1일 필요는 없다는 사실에 유의하자.

Reduced Row Echelon Form R

Row echelon form에서는 pivot이 1이 아닐 수도 있고 pivot원소가 속한 열에서 pivot원소보다 위쪽 행에 있는 원소들이 0이 아닐 수도 있다. 왼쪽에 적당한 대각행렬을 곱하여 pivot을 모두 1로 scale하고 다시 왼쪽에 적절한 상삼각행렬을 곱하여 pivot 위쪽의 0이 아닌 원소들을 소거한 행렬을 **reduced row echelon form**이라고 부른다.² 행렬 A 의 reduced row echelon form이 R 이면 $R = \text{rref}(A)$ 로 표시하기도 한다.

² 행의 교환없이 Gaussian elimination이 수행된 경우 reduced row echelon form은 유일하게 결정된다.

이전 절에서 얻어진 row echelon form $L^{-1}PA = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ 에 추가적인 scaling과 소거를 수행해 보자.

- Scaling of the second row: pivot 3을 1로 바꾸기 위해서 원쪽에 대각행렬을 곱한다.

$$\tilde{D}L^{-1}PA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Reduction above pivots: pivot 위쪽의 원소를 소거하기 위해서 행렬 원쪽에 상삼각행렬을 곱한다.

$$\tilde{U}\tilde{D}L^{-1}PA = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R = \text{rref}(A)$$

지금까지 수행한 과정을 모두 행렬곱으로 표현하면 $\tilde{U}\tilde{D}L^{-1}PA = R$ 이므로 정리하면

$$PA = L\tilde{D}^{-1}\tilde{U}^{-1}R = LDUR \quad (3.1)$$

을 얻는다. 이 표현에서 중요한 점 한 가지는 A 를 변형시킨 행렬 P, L, D, U 가 모두 가역행렬이라는 것이다. 또한 두 상삼각행렬의 곱인 UR 은 상삼각행렬이다.

Row echelon form U 와 reduced row echelon form R 을 도식화하면 다음과 같다.

$$U = \begin{bmatrix} * & * & * & * & * & * & * & * \\ 0 & * & * & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{scaling and reduction}} R = \begin{bmatrix} 1 & 0 & *' & 0 & *' & *' & *' & 0 \\ 0 & 1 & *' & 0 & *' & *' & *' & 0 \\ 0 & 0 & 0 & 1 & *' & *' & *' & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3.2.2 Pivot Variables and Free Variables

Reduced row echelon form R 에 대해 $R\mathbf{x}$ 에서 pivot이 들어있는 열에 대응되는 변수는 pivot variable이라고 부르고, 나머지 변수를 free variable이라고 부른다. $R\mathbf{x} = \mathbf{0}$ 의 해를 찾는 방법은 다양하지만, 체계적인 한 가지 방법이 free variable에 (이름 그대로) 임의의 값을 할당한 후에 pivot variable의 값을 정하는 방법이다. 이전 예를 살펴보면

$$\begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

에서 pivot variable을 free variable에 종속적인 형태로 표현하면 $\begin{cases} u = -3v + y \\ w = -y \end{cases}$ 를 얻는다. 따라서 free variable v, y 에 임의의 값을 할당하면 pivot variable u, w 가 바로 결정된다. 아직 차원을 정확하게 정의

하지는 않았지만³ free variable의 개수가 $R\mathbf{x} = \mathbf{0}$ 의 해집합의 차원일 것이라는 것을 추측할 수 있다. 몇 가지 쉽게 유추할 수 있는 결과들은 다음과 같다.

- $A\mathbf{x} = \mathbf{0} \Leftrightarrow R\mathbf{x} = \mathbf{0}$: (3.1)에서 P, L, D, U 가 invertible하므로 두 연립방정식은 동치이다.
- $\text{Null}(A)$ 에 속하는 vector를 일반적으로 표현하기 위해서 pivot variable을 free variable에 대해서 표현하여 대입하면 두 free variable에 대해 $\begin{bmatrix} -3v + y \\ v \\ -y \\ y \end{bmatrix} = v \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$ 로 표현된다. 이는 기하학적으로 \mathbb{R}^4 공간의 2차원 평면으로 생각할 수도 있고 column space 표현으로는 $\text{Col} \left(\begin{bmatrix} -3 & 1 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} \right)$ 로 표현할 수도 있다. 마지막 표현 속의 행렬에서 첫 번째 열은 $v = 1, y = 0$ 에 대응하는 해이고 두 번째 열은 $v = 0, y = 1$ 에 대응하는 $A\mathbf{x} = \mathbf{0}$ 의 해이다.

일반적으로 A 가 $m \times n$ 행렬이라고 하자. \dim 은 vector space의 차원을 나타내는 기호이다. 그러면,

- the number of pivots $\leq \min(m, n)$;
- $\dim \text{Null}(A) = \text{the number of free variables} = n - \text{the number of pivots}$;
- if $n > m$, there exists at least one free variable, and $A\mathbf{x} = \mathbf{0}$ has a non-zero solution.

마지막 결과는 간단하지만 이후에 유용하게 활용되므로 따로 정리한다.

Lemma 3.2 If a matrix A has more columns than rows, $A\mathbf{x} = \mathbf{0}$ has a non-zero solution.

Solving $A\mathbf{x} = \mathbf{b}$, $U\mathbf{x} = \mathbf{c}$, $R\mathbf{x} = \mathbf{d}$

이제 선형연립방정식의 풀이의 마지막 단계이다. 방정식의 순서를 바꾸지 않아도 되는 경우, 즉 row echelon form을 permutation matrix를 곱하지 않고 얻는 경우를 가정하자. 그러면 첫 단계로 $A\mathbf{x} = \mathbf{b}$ 의 양변에 하삼각행렬을 곱하여 $L^{-1}(A\mathbf{x}) = L^{-1}(\mathbf{b})$, 즉 row echelon form으로 기술되는 $U\mathbf{x} = \mathbf{c}$ 를 얻는다. 앞의 계수행렬 예제에서 일반적인 우변상수를 고려하면 다음과 같다:

$$A\mathbf{x} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \mathbf{b}.$$

³나중에 Definition 3.7에서 벡터공간의 차원에 대해서 수학적으로 정의한다.

원편에 하삼각행렬을 다음과 같이 곱하면

$$L^{-1}A\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = L^{-1}\mathbf{b}$$

row echelon form으로 기술된

$$U\mathbf{x} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b_1 \end{bmatrix} = \mathbf{c}$$

을 얻는다. 따라서 $U\mathbf{x} = \mathbf{c}$ ($\Leftrightarrow A\mathbf{x} = \mathbf{b}$)가 해를 갖을 필요충분조건은 $b_3 - 2b_2 + 5b_1 = 0$ 이다. $b_3 - 2b_2 + 5b_1 = 0$ 조건 하에서 연립방정식은

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ 0 \end{bmatrix}$$

이다. Free variable을 모두 0으로 놓고($v = 0, y = 0$) 풀면 $3w = b_2 - 2b_1$, $w = \frac{1}{3}b_2 - \frac{2}{3}b_1$, $u + 3w = b_1$, $u = 3b_1 - b_2$ 이므로 particular solution으로 $\mathbf{x}_p = (3b_1 - b_2, 0, \frac{1}{3}b_2 - \frac{2}{3}b_1, 0)^\top$ 을 얻을 수 있다. 원 방정식에 대응되는 homogeneous equation인 $A\mathbf{x} = \mathbf{0}$ 와 동치인

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \mathbf{0}$$

를 풀어보자. 앞에서 살펴본 바와 같이

- $v = 1, y = 0$: $\mathbf{x}_1 = (-3, 1, 0, 0)^\top$
- $v = 0, y = 1$: $\mathbf{x}_2 = (1, 0, -1, 1)^\top$

과 같은 두 개의 독립적인 해를 찾을 수 있다. 세 개의 해를 다음과 같이

$$\mathbf{x} = \mathbf{x}_p + \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 = \begin{bmatrix} 3b_1 - b_2 \\ 0 \\ \frac{1}{3}b_2 - \frac{2}{3}b_1 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

한꺼번에 표현할 수 있다. 이 vector가 정말로 $A\mathbf{x} = \mathbf{b}$ 를 만족함을 확인해보기 바란다. 정리하면 $A\mathbf{x} = \mathbf{b}$ 가 row echelon form으로 $U\mathbf{x} = \mathbf{c}$ 로 표현되었는데 U 가 r 개의 pivot을 가지고 있다고 가정하자. 그러면, U 의 마지막 $m - r$ 행들은 원소가 모두 0이므로 \mathbf{c} 의 마지막 $m - r$ 원소도 0이어야지 연립방정식이 해를 갖는다. 그리고 해를 갖는 경우, 모든 해는 $n - r$ 개의 free variable로 표현할 수 있다.

Gaussian elimination 과정에서 행의 교환이 필요한 경우 어떤 행을 교환하는지에 따라서 U 가 달라질 수는 있지만 한 행렬의 pivot의 개수는 일정하다는 것을 보일 수 있어서 이 개수를 행렬의 rank라고 부른다 (Section 3.4).

3.3 Linear Independence, Basis, and Dimension

이 절에서는 linear independence, spanning a subspace, a basis for a subspace, dimension of a subspace 등의 개념을 배운다.

Linear Independence

Vector들이 linearly independent하다는 것은 말 그대로 중복되지 않는 최소한의 벡터들을 모아 놓은 것이라는 의미이다. 이를 정식화한 것이 다음 정의이다.

Definition 3.4 For vectors $\mathbf{v}_i \in \mathbb{V}$ and scalars c_i , suppose $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$ only happens when $c_1 = \cdots = c_n = 0$. Then, $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is **linearly independent**.

If the linear combination vanishes for some non-zero c_i 's, the $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is called **linearly dependent**, and some \mathbf{v}_i can be represented as a linear combination of others.

아래 내용은 linear dependency에 관한 (유용한) 결과들이다. 결과들이 성립하는 이유를 잘 이해하여야 한다.

- $\{\mathbf{0}\}$ is linear dependent.

- $A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$: Columns of A are linearly dependent because $(-3)\mathbf{a}_1 + 1\mathbf{a}_2 + 0\mathbf{a}_3 + 0\mathbf{a}_4 = \mathbf{0}$.

Rows of A are linearly dependent. For $B = A^\top$, $5\mathbf{b}_1 + (-2)\mathbf{b}_2 + 1\mathbf{b}_3 = \mathbf{0}$.

- $A = \begin{bmatrix} 3 & 4 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 2 \end{bmatrix}$: Columns of A are linearly independent.

- $A = [\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n]$ and $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top \in \mathbb{R}^n$:

$\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ is linearly independent

$$\Leftrightarrow x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0} \text{ implies } \mathbf{x} = \mathbf{0}$$

$$\Leftrightarrow A\mathbf{x} = \mathbf{0} \text{ implies } \mathbf{x} = \mathbf{0}$$

$$\Leftrightarrow \text{Null}(A) = \{\mathbf{0}\}$$

- Non-zero rows of row echelon form U are linearly independent. Columns containing pivots are linearly independent.
- A set of n vectors in \mathbb{R}^m is always linearly dependent if $n > m$ because of Lemma 3.2.

Spanning Subspaces

어떤 벡터들이 선형조합으로 벡터공간의 모든 벡터들을 표현하는 것을 span한다고 말한다.

Definition 3.5 For vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, their **span** is a minimal subspace containing $\mathbf{v}_1, \dots, \mathbf{v}_n$, and is described formally as

$$\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\} = \{c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n : c_i \in \mathbb{R}\}.$$

If $\mathbb{V} = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, then we say that $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ spans \mathbb{V} .

- For $A = [\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_n]$, $\text{Col}(A) = \text{span}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$.
- For $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^\top$ where the i -th entry is 1, $\mathbb{R}^n = \text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$.

Fact 3.4 Assume that $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly independent. Then, the representation with respect to these vectors is unique. That is, if a vector \mathbf{v} has a representation, $\mathbf{v} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n$, then the coefficients x_i 's are unique.

Proof: Suppose that $\mathbf{v} = y_1\mathbf{v}_1 + \dots + y_n\mathbf{v}_n$ holds for some scalars y_i 's. If we equate these two representations, we get

$$(x_1 - y_1)\mathbf{v}_1 + \dots + (x_n - y_n)\mathbf{v}_n = \mathbf{0}.$$

Then, the linear independence implies $x_i - y_i = 0$ or $x_i = y_i$ for all $i = 1, \dots, n$. ■

이 결과에 의하면 linearly independent한 벡터들로 span하면 span된 벡터들의 linear combination 표현이 유일하다는 것을 알 수 있다. 또한 linearly independent한 벡터들로 span한 subspace 밖에 있는 벡터는 independent하다는 사실을 다음 결과가 말한다.

Fact 3.5 If $\mathbf{v} \notin \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ whereas $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly independent, then $\{\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{v}\}$ is linearly independent.

Proof: Notice that $\mathbf{v} \neq \mathbf{0}$. Suppose that $c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n + c_{n+1}\mathbf{v} = \mathbf{0}$ holds for some scalars c_i 's. If $c_{n+1} \neq 0$, then we have

$$\mathbf{v} = -\frac{c_1}{c_{n+1}}\mathbf{v}_1 - \dots - \frac{c_n}{c_{n+1}}\mathbf{v}_n \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\},$$

which contradicts the assumption. Hence $c_{n+1} = 0$ and the linear independence of spanning vectors implies $c_1 = \dots = c_n = 0$. ■

Basis for a Vector Space

Definition 3.6 A **basis** for a vector space \mathbb{V} is a set of vectors satisfying both of

1. (**independence**) The vectors in the set are linearly independent;
2. (**spanning property**) The vectors in the set span the space \mathbb{V} .

A vector in a basis is called a **basic vector**.

위 정의에서 independence는 minimality에 관한 것이고 spanning property는 sufficiency에 관한 것이다. Linearly dependent한 벡터들이 vector space를 span한다면 vector space를 더 적은 개수의 vector로 span 할 수 있다. Linearly independent한 벡터들이 basis가 아니라면 vector space를 span하지 못하는 것이므로 basis가 될 때까지 linearly independent한 vector를 추가할 수 있다(Fact 3.5).

- Basis를 통해서 vector space의 임의의 벡터들을 일의적으로 표현할 수 있어서 이 표현을 좌표로 생각할 수 있다. 즉 basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 이 주어졌을 때 임의의 vector $\mathbf{v} \in \mathbb{V}$ 에 대해서

$$\mathbf{v} = x_1 \mathbf{v}_1 + \cdots + x_n \mathbf{v}_n \quad (3.2)$$

으로 표현되면 \mathbf{v} 를 $(x_1, \dots, x_n)^\top \in \mathbb{R}^n$ 으로 생각할 수 있어 여러모로 편리하다. 이렇게 임의의 벡터를 \mathbb{R}^n 으로 바꾸어 생각하기 위해서는 이러한 대응관계가 일대일 관계여야 한다. 그런데 basis의 spanning property에 의해 $\mathbf{v} \in \mathbb{V}$ 에 대해서는 (3.2)와 같은 표현이 가능하고, basis의 independence와 Fact 3.4에 의해서 그 표현이 unique함을 알 수 있다. 즉 basic vector들에 대해 linear combination으로 표현되는 weighting coefficient는 unique하다. 요약하면 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 이 linearly independent일 때,

$$\mathbf{v} \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \stackrel{1 \leftarrow \text{to} \rightarrow 1}{\iff} (x_1, \dots, x_n) \in \mathbb{R}^n. \quad (3.3)$$

- 동일한 vector space에 대해서 basis는 여러 개가 존재할 수 있다. 예를 들면, $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 가 어떤 공간의 기저일 때, $B = [\mathbf{v}_1 | \cdots | \mathbf{v}_n]$ 로 놓으면 임의의 invertible matrix P 에 대해서 BP 의 열벡터들은 새로운 기저가 된다. 이유는 각자 생각해 보자. 예로서, $\mathbb{R}^2 = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ 에 대해서 $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 에 invertible matrix $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ 를 곱하면 $BP = P$ 여서 $\mathbb{R}^2 = \text{span} \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$ 이기도 하다.

Dimension of a Vector Space

Theorem 3.1 *If both $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ are bases for a vector space \mathbb{V} , then $n = m$.*

Proof: Since both sets of vectors span \mathbb{V} , there exist a_{ij} 's such that $\mathbf{v}_j = a_{1j} \mathbf{w}_1 + \cdots + a_{mj} \mathbf{w}_m$ for all $j = 1, \dots, n$. Let us set an $m \times n$ matrix $A = (a_{ij})$, $B = [\mathbf{v}_1 | \cdots | \mathbf{v}_n]$, and $C = [\mathbf{w}_1 | \cdots | \mathbf{w}_m]$. Observe the relationship $B = CA$ among three matrices. If $A\mathbf{x} = \mathbf{0}$ for some $\mathbf{x} \in \mathbb{R}^n$, then $B\mathbf{x} = CA\mathbf{x} = \mathbf{0}$. By the independence of columns of B , $\mathbf{x} = \mathbf{0}$ and hence $\text{Null}(A) = \{\mathbf{0}\}$. Hence, A should have more rows than columns, that is, $m \geq n$. If we change the roles of \mathbf{v}_i and \mathbf{w}_i , then we have $n \geq m$. ■

어떤 vector space의 basis가 유일할 필요는 없지만, Theorem 3.1에 의해 basis를 구성하는 basic vector들의 개수는 일정하다. 따라서 이 일정한 숫자로 vector space의 크기를 나타낸다.

Definition 3.7 *The dimension of a vector space \mathbb{V} is the number of basic vectors in a basis.*

다음 사실들의 이유를 잘 생각해 보자.

- $\dim(\mathbb{R}^n) = n$.

- $k + 1$ vectors in k -dimensional vector space are linearly dependent.
- Any spanning set can be reduced to a basis, i.e., a minimal spanning set.

Lemma 3.3 In a finite dimensional vector space, any linearly independent set of vectors can be extended to a basis.

Proof: 유한차원 벡터공간 \mathbb{V} 속의 선형독립인 k 개의 벡터들 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subset \mathbb{V}$ 를 생각하자. $k < \dim \mathbb{V} < \infty$ 라고 가정하자. 만약 $\mathbb{V} = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 라면 차원 정의에 따라서 $\dim \mathbb{V} = k$ 이어야 하므로 모순이 생긴다. 따라서 $\mathbf{v} \in \mathbb{V}$ 이고 $\mathbf{v} \notin \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 인 벡터 \mathbf{v} 가 존재한다. 그러면 Fact 3.5에 의해서 $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}\}$ 가 선형독립이어서 주어진 k 개의 선형독립인 벡터들 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 를 포함하는 $k + 1$ 개의 선형독립인 벡터를 얻는다. 선형독립인 벡터의 개수가 $\dim \mathbb{V}$ 보다 적은 경우 위의 과정을 반복하여 선형독립인 벡터를 한 개씩 추가할 수 있다. 따라서 최초의 벡터를 포함하는 $\dim \mathbb{V}$ 개의 선형독립인 벡터로 이루어진 기저를 얻을 수 있다. ■

Fact 3.6 Let \mathbb{V} be a vector space of dimension n . \mathbb{W}_1 and \mathbb{W}_2 are two subspaces of \mathbb{V} . Suppose that $\dim \mathbb{W}_1 = n_1$, $\dim \mathbb{W}_2 = n_2$, and $n_1 + n_2 > n$. Then $\dim(\mathbb{W}_1 \cap \mathbb{W}_2) \geq n_1 + n_2 - n$.

Proof: $\dim(\mathbb{W}_1 \cap \mathbb{W}_2) = k < n_1 + n_2 - n$ 이라 하고 $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 가 $\mathbb{W}_1 \cap \mathbb{W}_2$ 의 기저라 하자. \mathcal{B} 는 선형독립이므로 Lemma 3.3에 의해서 \mathbb{W}_1 의 기저 $\mathcal{B} \cup \{\mathbf{w}_{k+1}, \dots, \mathbf{w}_{n_1}\}$ 과 \mathbb{W}_2 의 기저 $\mathcal{B} \cup \{\mathbf{u}_{k+1}, \dots, \mathbf{u}_{n_2}\}$ 를 찾을 수 있다.

- $\mathbf{u} = z_{k+1}\mathbf{u}_{k+1} + \dots + z_{n_2}\mathbf{u}_{n_2}$ 가 $\mathbf{u} \in \mathbb{W}_1 \cap \mathbb{W}_2$ 라 가정하자. \mathcal{B} 가 $\mathbb{W}_1 \cap \mathbb{W}_2$ 의 기저이므로

$$\mathbf{u} = x_1\mathbf{v}_1 + \dots + x_k\mathbf{v}_k \quad \text{or} \quad z_{k+1}\mathbf{u}_{k+1} + \dots + z_{n_2}\mathbf{u}_{n_2} - x_1\mathbf{v}_1 - \dots - x_k\mathbf{v}_k = \mathbf{0}$$

인데 $\mathcal{B} \cup \{\mathbf{u}_{k+1}, \dots, \mathbf{u}_{n_2}\}$ 가 기저이므로 $x_1 = \dots = x_k = z_{k+1} = \dots = z_{n_2} = 0$ 이고 $\mathbf{u} = \mathbf{0}$ 된다.

- 가중치 $x_1, \dots, x_k, y_{k+1}, \dots, y_{n_1}, z_{k+1}, \dots, z_{n_2}$ 에 의한 선형조합이 영벡터인

$$\underbrace{x_1\mathbf{v}_1 + \dots + x_k\mathbf{v}_k}_{= \mathbf{v}} + \underbrace{y_{k+1}\mathbf{w}_{k+1} + \dots + y_{n_1}\mathbf{w}_{n_1}}_{= \mathbf{w}} + \underbrace{z_{k+1}\mathbf{u}_{k+1} + \dots + z_{n_2}\mathbf{u}_{n_2}}_{= \mathbf{u}} = \mathbf{0}$$

를 생각해 보자. $\mathbf{u} = -\mathbf{v} - \mathbf{w} \in \mathbb{W}_1$ 이므로 $\mathbf{u} \in \mathbb{W}_1 \cap \mathbb{W}_2$ 가 되고 직전에 얻어진 결과에 의해 $\mathbf{u} = \mathbf{0}$ 가 되어 $z_{k+1} = \dots = z_{n_2} = 0$ 이다. 또한 $\mathbf{v} + \mathbf{w} = \mathbf{0}$ 이어서 $\mathcal{B} \cup \{\mathbf{w}_{k+1}, \dots, \mathbf{w}_{n_1}\}$ 가 기저라는 사실과 합치면 $x_1 = \dots = x_k = y_{k+1} = \dots = y_{n_1} = 0$ 을 얻는다. 따라서 $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w}_{k+1}, \dots, \mathbf{w}_{n_1}, \mathbf{u}_{k+1}, \dots, \mathbf{u}_{n_2}\}$ 은 선형독립인데 벡터의 개수가 $n_1 + n_2 - k$ 로 벡터공간 \mathbb{V} 의 차원보다 커서 모순이 발생한다.

따라서 $\dim(\mathbb{W}_1 \cap \mathbb{W}_2) \geq n_1 + n_2 - n$ 이 성립한다. ■

3.4 Rank of Matrices

Pivot의 개수가 r 인 행렬 A 의 선형독립인 열의 최대 개수는 얼마일까? 먼저 A 의 row echelon form U 에 대해 생각해보자. U 의 pivot이 r 개이고 U 의 선형독립인 열 개수의 최댓값을 M 이라 하자.

- U 의 pivot을 포함하는 r 개의 열로 이루어진 집합을 $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ 이라 하고, 편의상 \mathbf{u}_i 의 pivot은 i 번째 원소인 p_i 라고 가정하자. 그러면, $x_1\mathbf{u}_1 + \dots + x_r\mathbf{u}_r = \mathbf{0}$ 을 연립방정식으로 생각할 때 r 번째 식은 $0x_1 + \dots + 0x_{r-1} + p_r x_r = 0$ 이고 \mathbf{u}_r 의 pivot인 $p_r \neq 0$ 이므로 $x_r = 0$ 이어야 한다. 따라서 식은 $x_1\mathbf{u}_1 + \dots + x_{r-1}\mathbf{u}_{r-1} = \mathbf{0}$ 으로 줄어들고 같은 논리를 반복하면 $x_1 = \dots = x_r = 0$ 을 얻게되어 $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ 이 선형독립임을 알 수 있다. 즉 $M \geq r$ 이다.
- $k > r$ 일 때 임의로 선택한 k 개의 U 의 열을 $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ 로 표시하자. U 의 pivot이 r 개이므로 \mathbf{u}_i 의 아래쪽 $m-r$ 개 원소는 모두 0이다. 따라서 m 개의 연립방정식인 $x_1\mathbf{u}_1 + \dots + x_k\mathbf{u}_k = \mathbf{0}$ 실제로는 k 개의 미지수에 대해 r 개의 선형방정식으로 구성된 것으로 생각할 수 있다. 따라서 Lemma 3.2에 의해 $\mathbf{0}$ 이 아닌 해가 존재하고 $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ 은 선형독립이 아니므로 $M \leq r$ 임을 알 수 있다.

따라서 $M = r$, 즉 U 의 선형독립인 열의 최대 개수는 r 이다. 이제 A 에 대해서 생각해 보자. $\tilde{L}PA = U$ 를 구성하는 P 와 \tilde{L} 이 invertible하므로

$$A\mathbf{x} = \mathbf{0} \Leftrightarrow U\mathbf{x} = \mathbf{0}$$

이다. 이 등식은 A 의 열들의 종속 관계와 U 의 열들의 종속 관계가 동일함을 나타낸다. 따라서 A 의 선형독립인 열의 최대 개수는 pivot의 개수인 r 임을 알 수 있다. 따라서 pivot의 개수는 Gaussian elimination 과정에 무관하게 각 행렬마다 일정하다는 것을 알 수 있고 이 개수를 행렬의 rank라고 부른다.

Definition 3.8 (Rank of Matrices) Let U be the row echelon form of a matrix A . If U has r pivots, then the **rank** of A is r . We denote it as $\text{rank } A$.

A 의 row echelon form U 와 reduced row echelon form R 은 자기 자신이 row echelon form이므로 A 와 U , R 의 rank는 모두 같다.

다음으로 행렬 A 의 rank와 선형독립인 행의 최대 개수에 대해 생각해 보자. A 와 PA 의 행들은 순서만 다르지 동일하므로 $P = I$ 인 경우, 즉 $A = LU$ 인 경우를 생각하자. $\text{rank } A = r$, 즉 pivot의 개수가 r 이라 하자. Row echelon form U 의 윗쪽 r 개의 행들이 선형독립이라는 것은 자명하다. 그리고 나머지 $m-r$ 개의 행은 모든 원소가 0이므로 U 의 선형독립인 행의 최대 개수는 r 이다. A 와 U 의 첫 r 개의 행으로 이루어진 행렬을 각각 \bar{A} 와 \bar{U} 라 하고 L 의 첫 r 개의 행과 열로 이루어진 가역인 하삼각행렬을 L_r 로 놓으면 L 이 하삼각행렬이므로 $\bar{A} = L_r \bar{U}$ 가 성립하며⁴ \bar{U} 의 행들은 선형독립이다. \bar{U} 의 행들이 선형독립이므로 만약 $\mathbf{y}^\top \bar{A} = \mathbf{y}^\top L_r \bar{U} = \mathbf{0}$ 이라면 $\mathbf{y}^\top L_r = \mathbf{0}$ 인데 L_r 이 invertible이어서 $\mathbf{y} = \mathbf{0}$ 이다. 즉 A 의 첫 r 개의 행이 선형독립이므로 최소 r 개의 행이 선형독립이다.

반대 방향을 보이자. $A = LU$ 에서 U 의 마지막 $m-r$ 개 행의 원소들이 모두 0이라는 사실로부터 L 의 첫 r 개의 열로 이루어진 행렬을 \tilde{L} 이라 하면 $A = \tilde{L} \bar{U}$ 가 성립한다. 위에서 얻은 $\bar{A} = L_r \bar{U}$ 로부터 $B = \tilde{L} L_r^{-1}$ 로 놓으면

$$A = \tilde{L} \bar{U} = \tilde{L} L_r^{-1} \bar{A} = B \bar{A}$$

를 얻는다. $k > r$ 일 때, A 의 임의의 k 개 행으로 이루어진 행렬 A' 과 선택된 k 개 행에 대응되는 B 의 행으로 이루어진 행렬 B' 에 대해서 $A' = B' \bar{A}$ 가 성립한다. 그런데 B' 은 $k \times r$ 행렬이므로 Lemma 3.2에 의해

⁴ L 이 대각원소가 1인 하삼각행렬이면 LU 의 i 번째 행은 $1, \dots, i-1$ 번째 행들의 선형조합으로 표현되므로 $\bar{A} = L_r \bar{U}$ 가 성립한다.

$\mathbf{y}^\top B' = \mathbf{0}^\top$ 가 되는 $\mathbf{y} \neq \mathbf{0}$ 가 존재하여 $\mathbf{y}^\top A' = \mathbf{y}^\top B' \bar{A} = \mathbf{0}^\top \bar{A} = \mathbf{0}^\top$ 가 되어 A' 의 행들이 선형종속이다. 따라서 선형독립인 A 의 행의 최대 개수는 $r = \text{rank } A$ 이다. 전반부 결과와 합치면

rank A 는 행렬 A 의 선형독립인 행 또는 열의 최대 개수이다.

그리고 자연스럽게

$$\text{rank } A = \text{rank } A^\top$$

임도 알 수 있다.

rank $A = r$ 인 $m \times n$ 행렬 $A = [\mathbf{a}_1 | \cdots | \mathbf{a}_n]$ 의 열들로 span되는 부분공간 $\mathbb{W} = \text{Col}(A)$ 를 생각하자. $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ 이 선형독립이라 하자. $k > r$ 이면 $\{\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{a}_k\}$ 가 종속이므로 x_i 가 존재하여 $x_1\mathbf{a}_1 + \cdots + x_r\mathbf{a}_r + x_k\mathbf{a}_k = \mathbf{0}$ 가 성립하며 $x_k \neq 0$ 이다. 따라서

$$\mathbf{a}_k = -\frac{x_1}{x_k}\mathbf{a}_1 - \cdots - \frac{x_r}{x_k}\mathbf{a}_r \in \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$$

이므로 $\text{Col}(A) \subset \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ 이다. 즉 $\mathbb{W} = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ 이다. 따라서 $\dim \mathbb{W} = \text{rank } A$ 이다. 정리하면,

Lemma 3.4 For $\mathbf{a}_i \in \mathbb{R}^m$,

$$\dim \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} = \text{rank } [\mathbf{a}_1 | \cdots | \mathbf{a}_n]$$

or, in matrix form,

$$\dim \text{Col}(A) = \text{rank } A$$

for any matrix A .

이다. A 의 행으로 span되는 부분공간에 대해서도 $\dim \text{Col}(A^\top) = \text{rank } A^\top = \text{rank } A$ 이므로 동일한 결과를 얻는다.

Fact 3.7 Suppose A and B are $n \times m$ and $m \times n$ matrices, respectively. Then $\text{rank}(AB) \leq \text{rank}(A)$ and $\text{rank}(AB) \leq \text{rank}(B)$.

Proof: Lemma 3.1에 의해 $\text{Col}(AB) \subset \text{Col}(A)$ 이다. 두 벡터공간 \mathbb{W}_1 과 \mathbb{W}_2 간에 $\mathbb{W}_1 \subset \mathbb{W}_2$ 인 포함관계가 성립하면 두 공간의 차원 간에 $\dim \mathbb{W}_1 \leq \dim \mathbb{W}_2$ 가 성립한다. 따라서 Lemma 3.4에 의해 $\text{rank}(AB) \leq \text{rank } A$ 이다. 그리고 첫번째 결과에 의해 $\text{rank}(AB) = \text{rank}((AB)^\top) = \text{rank}(B^\top A^\top) \leq \text{rank}(B^\top) = \text{rank}(B)$ 이다. ■

3.5 The Four Fundamental Subspaces

중학교 때 배운 집합의 두 가지 표현 방법(원소나열법, 조건제시법)과 유사하게 vector space도 다음과 같이 두 가지 방법으로 표현할 수 있다.

- a set of vectors that spans the subspace. e.g. $\text{Col}(A)$;

- a set of conditions that vectors in the space must satisfy. e.g. Null (A).

$\text{rank}(A) = r \in m \times n$ 행렬 A 에 대해서 다음과 같은 4가지 subspace를 생각하자:

1. Column space $\text{Col} (A) \subset \mathbb{R}^m$. $\dim \text{Col} (A) = r$;
2. Null space $\text{Null} (A) \subset \mathbb{R}^n$. $\dim \text{Null} (A) = n - r$;
3. Row space $\text{Row} (A) = \text{Col} (A^\top) \subset \mathbb{R}^n$. $\dim \text{Row} (A) = r$;
4. Left null space $\text{LeftNull} (A) = \text{Null} (A^\top) \subset \mathbb{R}^m$. $\dim \text{LeftNull} (A) = m - r$.

첫번째와 세번째 공간의 차원은 Lemma 3.4에 의해 알 수 있다. A 의 row rank와 column rank를 각각 $\text{Col} (A)$ 와 $\text{Row} (A)$ 의 dimension이라 하면

$$\boxed{\text{row rank} = \text{column rank}}$$

이다. Null space의 차원은 Gaussian elimination 과정에서 free variable의 개수임을 직관적으로 살펴보았다. 이를 좀 더 자세히 살펴보자. Gaussian elimination에서 얻은 동치 관계 $A\mathbf{x} = \mathbf{0} \Leftrightarrow U\mathbf{x} = \mathbf{0}$ 때문에 $\text{Null} (A) = \text{Null} (U)$ 이다. 그리고 $U\mathbf{x} = \mathbf{0}$ 에서 한 개의 free variable에 1을 부여하고 나머지 free variable은 0 으로 놓으면 선형 독립인 null space의 벡터를 free variable 개수만큼 얻는다. 이 벡터들이 null space의 기저를 구성하는 것을 보일 수 있다. 따라서 $\dim \text{Null} (U)$ 는 U 의 free variable의 개수 $n - r$ 이므로 $\text{Null} (A) = n - r$ 도 성립한다. 따라서 **rank-nullity theorem**이라고 불리는

$$\boxed{\text{rank}(A) + \dim \text{Null} (A) = \dim \text{Col} (A) + \dim \text{Null} (A) = \text{the number of columns of } A} \quad (3.4)$$

을 얻는다.

Example 3.1 Let us find the four fundamental subspaces of $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = U = R$. The four subspaces can be written down straightforwardly for this simple matrix.

$$1. \text{ Column space: } \text{Col} (A) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}.$$

$$2. \text{ Null space: } \text{Null} (A) = \text{span} \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

$$3. \text{ Row space: } \text{Row} (A) = \text{Col} (A^\top) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

$$4. \text{ Left null space: } \text{LeftNull} (A) = \text{Null} (A^\top) = \text{span} \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

■

보통 row echelon form U 에 대해서 4가지 fundamental subspace를 찾는 것이 A 에 대해서 찾는 것보다 상대적으로 쉽다. 이렇게 얻어진 U 에 대한 fundamental subspace로부터 어떻게 A 에 대한 fundamental space를 찾는지 다음 예를 통해 생각해보자.

An Example of Constructing Fundamental Spaces of A from Those of U

$$U = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ 는 } A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} \text{ 의 row echelon form이다.}$$

1. Finding the Row space, $\text{Col}(A^\top)$: Basis of $\text{Col}(A^\top) = \text{non-zero rows of } U$, 즉 $\text{Col}(A^\top) = \text{Col}(U^\top)$ 이다. 왜 그럴까? L 이 invertible하고 $A^\top = U^\top L^\top$ 이므로 Lemma 3.1에 의해 성립한다.
2. Finding the Column space, $\text{Col}(A)$: U 의 pivot을 포함한 열들이 $\text{Col}(U)$ 의 basis가 된다. 그런데 $A\mathbf{x} = \mathbf{0} \Leftrightarrow U\mathbf{x} = \mathbf{0}$ 으로 linear independence of some columns of $A \Leftrightarrow$ linear independence of columns of U whose indices are the same as those of A 이다. 이를 통해 $\dim \text{Col}(A) = \dim \text{Col}(U) = \text{rank } A = r$ 임을 알 수 있다. 즉 U 의 pivot 개수 만큼의 linearly independent한 A 의 열들을 찾으면 $\text{Col}(A)$ 의 basis가 된다. 그런데 U 의 pivot을 포함한 열들은 linearly independent하므로 동일한 A 의 열들도 linearly independent하다. 따라서 이렇게 찾은 A 의 열들이 $\text{Col}(A)$ 의 basis가 된다. 위의 예시에서는 행렬 A 의 1, 3번째 column이 $\text{Col}(A)$ 의 basis가 된다.
3. Finding the Null space, $\text{Null}(A)$: Gaussian elimination에서 얻은 동치 관계 $A\mathbf{x} = \mathbf{0} \Leftrightarrow U\mathbf{x} = \mathbf{0}$ 때문에 $\text{Null}(A) = \text{Null}(U)$ 이다. $\dim \text{Null}(U)$ 는 U 의 free variable의 개수 $n - r$ 이다. $U\mathbf{x} = \mathbf{0}$ 에서 한 개의 free variable에 1을 부여하고 나머지 free variable은 0으로 놓으면 선형 독립인 null space의 벡터를 free variable 개수만큼 얻고 이 벡터들이 null space의 기저를 구성한다.

How to Find a Basis of a Spanned Subspace

vector들이 주어졌을 때 이 벡터들이 span하는 공간의 basis를 구하는 방법을 생각해보자. 한 가지 방법은 벡터들을 행으로 적층한 행렬에 Gaussian elimination을 하여 행이 모두 0으로 바뀌지 않은 행들이 basis가 된다. 또 다른 방법으로 벡터들을 열로 적층한 행렬의 row echelon form의 pivot column에 해당하는 vector를 모으면 basis가 된다. 첫 번째 방법에서는 row echelon form의 행을 그대로 basic vector로 사용할 수 있지만 두 번째 방법에서는 pivot column에 해당하는 원래 vector로 basis를 구성해야 한다.

3.6 Existence of Inverses

$\text{rank}(A) = m < n$ 인 행렬 A 가 $A = LU$ 로 표현된다고 하자. 이 경우 총 m 개의 pivot이 존재한다. $n \times n$ permutation matrix P 를 A 의 오른쪽에 곱하여 열의 순서를 바꾸자. 이때 P 를 잘 선택하여 UP 의 앞쪽 m

개의 열로 구성된 행렬 \hat{U} 가 invertible하도록 만든다. 원래 U 가 열의 교환이 필요없는 경우 $P = I_n$ 으로 놓으면 된다. 예를 들면

$$L^{-1}A = U = \begin{bmatrix} * & * & * & * & * & * & \dots \\ 0 & 0 & * & * & * & * & \dots \\ 0 & 0 & 0 & * & * & * & \dots \\ 0 & 0 & 0 & 0 & \ddots & * & \vdots \\ 0 & 0 & 0 & 0 & 0 & * & \dots \end{bmatrix}$$

이면 적절한 permutation matrix P 를 오른쪽에 곱하여 두 번째 열을 뒤쪽으로 보내면

$$L^{-1}AP = UP = \left[\begin{array}{cccccc|cc} * & * & * & \dots & * & * & * & \dots \\ 0 & * & * & \dots & * & * & * & \dots \\ 0 & 0 & * & \dots & * & * & * & \dots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & * & * & * & \dots \\ 0 & 0 & \dots & 0 & 0 & * & * & \dots \end{array} \right] = \left[\begin{array}{ccc|cc} * & \dots & * & * & \dots \\ \ddots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0} & * & * & * & \dots \end{array} \right] = [\hat{U} | G]$$

를 얻는다. $GH = \mathbf{0}$ 인 $(n-m) \times m$ 행렬 H 를 생각하면

$$[\hat{U} | G] \begin{bmatrix} \hat{U}^{-1} \\ H \end{bmatrix} = \hat{U}\hat{U}^{-1} + GH = I_m$$

이 성립하므로 $C = P \begin{bmatrix} \hat{U}^{-1} \\ H \end{bmatrix} L^{-1}$ 로 놓으면

$$AC = (AP)(P^{-1}C) = (LUP)(P^{-1}C) = L \begin{bmatrix} \hat{U}^{-1} \\ H \end{bmatrix} L^{-1} = LI_m L^{-1} = I_m$$

가 되어 C 가 A 의 right-inverse가 된다. 이 경우 $H = \mathbf{0}$ 이면 $GH = \mathbf{0}$ 을 항상 만족하며, $\mathbf{0}$ 이외에도 다양한 H 가 존재하여 right-inverse가 여러 개일 수 있다.

만약 Gaussian elimination에서 행의 교환이 필요하고 이를 나타내는 permutation matrix가 Q 라면 위 유도과정에서 A 대신에 QA 를 생각하여 얻어진 right-inverse C 에 대해서 $QAC = I_m$ 이 되어 $AC = Q^{-1}$ 이므로 CQ 가 A 에 대한 right-inverse가 된다.

A 의 left-inverse B 가 존재한다면 $BA = I_n$ 을 만족해야 한다. 그런데 B 의 열이 m 개 이므로 $\text{rank}(B) \leq m$ 이므로 $\text{rank}(BA) \leq \text{rank}(B) \leq m$ 인데 우변으로부터 $\text{rank}(I_n) = n$ 이 되어 모순이다. 따라서 left-inverse는 존재하지 않는다.

- $\text{rank}(A) = n < m$ 인 경우, 위와 유사하게 left-inverse는 존재하지만 right-inverse가 존재하지 않는다.
- $\text{rank}(A) = m = n$ 인 경우를 생각하자. 앞에서 살펴본 $\text{rank } A = m < n$ 인 경우에서 G 와 H 가 없는 경우로 right-inverse가 $C = P\hat{U}^{-1}L^{-1}$ 인데 $L^{-1}AP = \hat{U}$ 이므로

$$CA = P\hat{U}^{-1}L^{-1}A = P\hat{U}^{-1}(L^{-1}AP)P^{-1} = P\hat{U}^{-1}\hat{U}P^{-1} = I_m$$

이어서 C 가 left-inverse도 된다. 즉 C 가 A 의 inverse이다.

정리하면 다음과 같이 요약할 수 있다.

Theorem 3.2 For an $m \times n$ matrix A , the inverse of A exists only when $\text{rank}(A) = m = n$. If $\text{rank}(A) = m \leq n$, the right-inverse of A exists. If $\text{rank}(A) = n \leq m$, the left-inverse of A exists.

또한 left-와 right-inverse는 $Ax = b$ 의 해의 존재 및 유일성에 다음과 같이 관련되어 있다.

For any $m \times n$ matrix A ,

- $\text{rank}(A) = m$ implies the **existence** of the solution for $Ax = b$: Row rank와 column rank가 일치하므로 $\text{Col}(A) = \mathbb{R}^m$ 이 되어서 방정식 $Ax = b$ 가 b 에 무관하게 적어도 하나의 해를 갖는다;
- $\text{rank}(A) = n$ implies the **uniqueness** of the solution for $Ax = b$: $\dim \text{Col}(A) = n$ 이므로 A 의 열들이 linearly independent하다. 따라서 $Ax = b$ 의 해가 존재한다면 유일해야 한다.

또한 행렬의 rank가 가질 수 있는 최댓값은 행의 개수와 열의 개수 중 작은 값이다. 어떤 행렬이 가능한 최대한의 rank를 갖는 경우 다음과 같이 rank에 대한 추가적인 결과와 특별한 left- 또는 right-inverse를 갖는다.

Fact 3.8 Let A be an $m \times n$ matrix.

1. $\text{rank } A = n$ case: Then the rank of $A^\top A$ is also n , and $A^\top A$ is invertible. $(A^\top A)^{-1} A^\top$ is a left-inverse of A .
2. $\text{rank } A = m$ case: Then the rank of AA^\top is also m , and AA^\top is invertible. $A^\top (AA^\top)^{-1}$ is a right-inverse of A .

Proof: For $\text{rank } A = n$ case, it is enough to show that the $n \times n$ matrix $A^\top A$ has a trivial nullspace $\{\mathbf{0}\}$. Assume $A^\top Ax = \mathbf{0}$. By multiplying x on the both sides of the equation, we get $x^\top A^\top Ax = 0$, which is equivalent to $|Ax| = 0$. For the norm of Ax to vanish, Ax itself has to be null, that is, $Ax = \mathbf{0}$. Hence $x \in \text{Null}(A)$. However the dimension balance (3.4) implies $\dim \text{Null}(A) = 0$ and $\text{Null}(A) = \{\mathbf{0}\}$. Therefore $\text{rank } A^\top A = n$ and $A^\top A$ is invertible. It is clear that $(A^\top A)^{-1} A^\top$ is a left-inverse.

The argument for the other case is almost identical. ■

간단한 예에서 right-inverse를 찾아보고 left-inverse와의 관계 등을 살펴보자.

Example 3.2 $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}$ 를 생각해 보자. 임의의 c_{31}, c_{32} 에 대해서 $C = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/5 \\ c_{31} & c_{32} \end{bmatrix}$ 는 $AC = I_2$

를 만족한다. 따라서 C 는 A 의 right-inverse이다. 하지만 A 의 left-inverse는 존재하지 않는다는 것을 앞에서 보였다. Right-inverse C 를 A 의 왼쪽에 곱해보면

$$CA = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/5 \\ c_{31} & c_{32} \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4c_{31} & 5c_{32} & 0 \end{bmatrix}$$

를 얻고 c_{31}, c_{32} 와 무관하게 left-inverse가 아님을 알 수 있다.

$$\text{rank } A = 2 \circ] \text{므로 Fact 3.8에 의해 } C^* = A^\top (AA^\top)^{-1} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}^\top \begin{bmatrix} 1/16 & 0 \\ 0 & 1/25 \end{bmatrix} = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/5 \\ 0 & 0 \end{bmatrix} \text{ 가}$$

right-inverse이다. 특히 이 right-inverse C^* 는 C 에서 $c_{31} = c_{32} = 0$ 인 경우로 나중에 Section 5.8에서 배우는 A 의 pseudo-inverse이다. ■

3.7 Matrices of Rank-one

어떤 행렬의 rank가 1이면 모든 열들이 서로 상수배 관계가 있다는 것이다. 따라서 rank-one 행렬은 적절한 vector \mathbf{u} 와 \mathbf{v} 에 대해 \mathbf{uv}^\top 으로 표현된다. 어떤 행렬의 rank가 1인 것을 아는 것이 어렵지 rank-one 행렬을

나타내는 두 vector를 찾는 것은 어렵지 않다. 예로서 $\begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix}$ 의 rank가 1인 것을 아는 것은 약간의

고민이 필요하지만, 일단 rank가 1이라는 것을 알고 이 행렬을 $\begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \end{bmatrix}$ 와 같이 표현하는 두 vector

\mathbf{u} 와 \mathbf{v} 를 찾는 것이 어렵진 않다. 물론 두 vector가 서로 임의의 상수와 역수를 곱해도 곱이 원래 행렬을 나타내므로 일의적으로 결정되지는 않는다. 하지만, rank-one 행렬로 일반적인 행렬이나 행렬곱을 표현하는 것은 유용하다. $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^\top \in \mathbb{R}^n$, 즉 i 번째 원소만 1이고 나머지는 0인 n 차원 벡터라 하자. 그리고 vector \mathbf{a}_i 를 i 번째 열로 갖는 행렬 $A = [\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n]$ 를 생각하자. 그러면, 행렬 A 는

$$A = \mathbf{a}_1 \mathbf{e}_1^\top + \cdots + \mathbf{a}_n \mathbf{e}_n^\top = \sum_{i=1}^n \mathbf{a}_i \mathbf{e}_i^\top$$

처럼 rank-one 행렬 $\mathbf{a}_i \mathbf{e}_i^\top$ 의 합으로 표현된다. 그리고 이 결과를 이용하면 두 행렬의 곱도 rank-one 행렬의 합으로 표현된다는 것을 쉽게 알 수 있다.

Lemma 3.5 Assume an $m \times n$ matrix $A = [\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n]$ for $\mathbf{a}_i \in \mathbb{R}^m$ and an $\ell \times n$ matrix $B = [\mathbf{b}_1 | \mathbf{b}_2 | \cdots | \mathbf{b}_n]$ for $\mathbf{b}_i \in \mathbb{R}^\ell$. Then,

$$AB^\top = \sum_{i=1}^n \mathbf{a}_i \mathbf{b}_i^\top. \quad (3.5)$$

Proof: 먼저 $A = \sum_{i=1}^n \mathbf{a}_i \mathbf{e}_i^\top$ 이고 $B = \sum_{i=1}^n \mathbf{b}_i \mathbf{e}_i^\top$ 이다. 그리고 $\mathbf{e}_i^\top \mathbf{e}_j$ 는 $i = j$ 일 때만 1이고 나머지 경우는 0이므로

$$AB^\top = \left(\sum_{i=1}^n \mathbf{a}_i \mathbf{e}_i^\top \right) \left(\sum_{j=1}^n \mathbf{b}_j \mathbf{e}_j^\top \right)^\top$$

$$\begin{aligned}
&= \left(\sum_{i=1}^n \mathbf{a}_i \mathbf{e}_i^\top \right) \left(\sum_{j=1}^n \mathbf{e}_j \mathbf{b}_j^\top \right) \\
&= \sum_{i=1}^n \sum_{j=1}^n \mathbf{a}_i \mathbf{e}_i^\top \mathbf{e}_j \mathbf{b}_j^\top \\
&= \sum_{i=1}^n \mathbf{a}_i \mathbf{b}_i^\top.
\end{aligned}$$

■

다음은 두 행렬 사이에 대각행렬이 추가로 곱해진 일반적인 경우에 대한 결과이다.

Corollary 3.1 Assume an $m \times n$ matrix $A = [\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n]$ for $\mathbf{a}_i \in \mathbb{R}^m$, an $\ell \times n$ matrix $B = [\mathbf{b}_1 | \mathbf{b}_2 | \cdots | \mathbf{b}_n]$ for $\mathbf{b}_i \in \mathbb{R}^\ell$, and an $n \times n$ diagonal matrix $D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Show that

$$ADB^\top = \sum_{i=1}^n \lambda_i \mathbf{a}_i \mathbf{b}_i^\top. \quad (3.6)$$

Proof: $AD = [d_1 \mathbf{a}_1 | d_2 \mathbf{a}_2 | \cdots | d_n \mathbf{a}_n]$ 이므로 $(AD)B^\top$ 에 Lemma 3.5를 적용한다. ■

이러한 rank-one 행렬을 이용한 표현은 이후에 유용하게 활용된다.

3.8 Linear Transformations

벡터가 변환되어 다른 벡터가 되는 대응 관계를 생각해 보자. 이 절에서는 벡터의 기본연산인 벡터 합 및 상수배와 잘 어울리는 벡터공간 간의 변환을 살펴본다. 다시 말하면 벡터 합 및 상수배와 호환되는 변환, 즉 벡터 합을 변환시키거나 변환된 벡터를 합하나 동일한 변환들을 다룬다. 그리고 상수배에 대해서도 상수를 벡터 변환 전과 후 언제 곱해도 상관없는 경우, 이러한 변환을 선형 변환(linear transformation)이라 부른다. 변환의 그래프가 1차원 공간에서는 원점을 지나는 직선, 2차원에서는 원점을 지나는 평면인 경우를 생각하면 된다. 이러한 선형 사상의 특징을 정의로 기술하면 다음과 같다.

Definition 3.9 두 벡터공간 \mathbb{V}, \mathbb{W} 사이에 정의된 함수 $T : \mathbb{V} \rightarrow \mathbb{W}$ 가 임의의 실수 $\alpha \in \mathbb{R}$ 와 임의의 벡터 $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{V}$ 에 대해서

- $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2);$
- $T(\alpha \mathbf{v}_1) = \alpha T(\mathbf{v}_1)$

이 성립하면 T 를 선형사상(linear map)이라 부른다.^a 선형대수학에서는 선형사상을 선형변환(linear transformation)이라고도 부른다.

^a선형사상이 되기 위한 두 가지 조건을 ”임의의 실수 $\alpha \in \mathbb{R}$ 와 임의의 벡터 $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{V}$ 에 대해서

$$T(\alpha \mathbf{v}_1 + \mathbf{v}_2) = \alpha T(\mathbf{v}_1) + T(\mathbf{v}_2)$$

이 성립한다”는 한 가지 조건으로 표현할 수도 있다.

이 정의에서 주의해서 살펴봐야 할 점은 선형사상/변환 T 를 기술하는 과정에서 정의역과 공변역인 두 벡터공간 \mathbb{V} 와 \mathbb{W} 의 기저들이 필요하지 않다. 그렇다면 ”선형사상은 정의구역에 속한 각각의 벡터들에 대해서 일일이 대응관계를 설명해야 하는가?”, 또는, ”선형사상/변환을 효율적으로 기술하는 어떤 방법이 있을까?”라는 질문이 자연스럽게 떠 오른다. 이에 대한 답을 찾아가는데 있어서 유한차원 벡터공간들을 고려하자. 즉 $\dim(\mathbb{V}) = n$, $\dim(\mathbb{W}) = m$ 인 경우를 고려하자.

3.8.1 Matrix Representation of Linear Transformations

벡터공간을 기술할 때 가장 중요한 역할을 하는 기저(basis)를 살펴보자. $\mathcal{B}_{\mathbb{V}} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 을 \mathbb{V} 의 기저라 하면 임의의 \mathbb{V} 의 벡터 $\mathbf{v} \in \mathbb{V}$ 는 유일한⁵ n 개의 실수 x_1, \dots, x_n 이 존재⁶하여 $\mathbf{v} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n$ 으로 표현된다. 그런데 T 가 선형사상이므로

$$T(\mathbf{v}) = T(x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n) = x_1T(\mathbf{v}_1) + \dots + x_nT(\mathbf{v}_n)$$

이 된다. 이로부터, T 를 기저 $\mathcal{B}_{\mathbb{V}}$ 에 대해 기술하는 것은 기저의 basic vector \mathbf{v}_j 에 대응되는 \mathbb{W} 의 벡터인 $T(\mathbf{v}_j)$ 를 정해주면 된다는 것을 알 수 있다. 즉 $\{T(\mathbf{v}_j) : j = 1, \dots, n\}$ 만 정해지면 임의의 $\mathbf{v} \in \mathbb{V}$ 에 대해서 $\mathcal{B}_{\mathbb{V}}$ 에 대한 좌표에 해당하는 $\{x_j : j = 1, \dots, n\}$ 을 찾아서 $\sum_{j=1}^n x_j T(\mathbf{v}_j)$ 를 계산하면 $T(\mathbf{v})$ 값이 된다.

$\mathcal{B}_{\mathbb{W}} = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ 가 벡터공간 \mathbb{W} 의 여러 기저 중 하나라 하자. $T(\mathbf{v}_j) \in \mathbb{W}$ 이므로 벡터 $T(\mathbf{v}_j)$ 의 기저 $\mathcal{B}_{\mathbb{W}}$ 에 대한 실수 좌표를 찾을 수 있는데 이를 $\{a_{1j}, \dots, a_{mj}\} \subset \mathbb{R}$ 라 하자. 즉

$$T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + \dots + a_{mj}\mathbf{w}_m = \sum_{i=1}^m a_{ij}\mathbf{w}_i$$

이다. 따라서 $T(\mathbf{v})$ 를 기저 $\mathcal{B}_{\mathbb{W}}$ 에 대해서 표현하면

$$\begin{aligned} T(\mathbf{v}) &= x_1T(\mathbf{v}_1) + \dots + x_nT(\mathbf{v}_n) = \sum_{j=1}^n x_j T(\mathbf{v}_j) \\ &= \sum_{j=1}^n x_j \left(\sum_{i=1}^m a_{ij}\mathbf{w}_i \right) \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}x_j \right) \mathbf{w}_i \end{aligned}$$

를 얻을 수 있다. 만약 $T(\mathbf{v})$ 의 기저 $\mathcal{B}_{\mathbb{W}}$ 에 대한 좌표를 $\{y_1, \dots, y_m\}$ 이라 하면 $T(\mathbf{v}) = \sum_{i=1}^m y_i \mathbf{w}_i$ 이므로 위 식과 비교하여

$$y_i = \sum_{j=1}^n a_{ij}x_j, \quad i = 1, \dots, m$$

을 얻을 수 있다. 따라서 n 차원 실수 벡터 $\mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$, m 차원 실수 벡터 $\mathbf{y} = (y_1, \dots, y_m)^\top \in \mathbb{R}^m$, 그리고 $m \times n$ 실수 행렬 $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ 에 대해

$$\mathbf{y} = A\mathbf{x}$$

가 성립함을 알 수 있다. 이러한 표현과 관련하여 몇 가지 생각할 내용은 다음과 같다.

⁵basic vector들이 선형독립이기 때문에

⁶basic vector들의 spanning 성질에 의해

- 우리가 고려한 벡터공간 \mathbb{V} 와 \mathbb{W} 는 \mathbb{R}^n , \mathbb{R}^m 과 상관없이 n 차 다항식과 같은 일반적인 벡터들로 구성될 수도 있는 공간임에 주의해야 한다.
- 기저를 구성하는 basic vector들 간의 관계는 linearly independent하다는 사실 외에는 다른 조건이 없다. 즉 우리가 벡터의 크기를 잴 줄 안다고 할 때, \mathbf{v}_j 의 크기는 1이 아닐 수도 있다. 또한, 서로 다른 두 basic vector간의 내적이 0이 아닐 수도 있다. 일반적인 내적은 Chapter 4에서 공부한다.
- x_1, \dots, x_n 을 찾는 것은 벡터공간과 기저 간의 관계이므로 T 의 역할과 무관하다. 벡터공간 \mathbb{V} , \mathbb{W} 와 선형사상 T 가 고정되어 있더라도 벡터공간 \mathbb{V} 의 기저가 바뀌면 동일한 벡터 $\mathbf{v} \in \mathbb{V}$ 의 좌표에 해당하는 $\mathbf{x} \in \mathbb{R}^n$ 이 바뀌고 따라서 변환관계를 나타내는 행렬 A 도 바뀐다. 또한, \mathbb{V} 의 기저가 고정되어 있어도 공변역인 \mathbb{W} 의 기저가 바뀌면 행렬 A 가 바뀌어야 한다. 즉 어떤 선형사상 T 를 표현하는 행렬은 정의구역과 공변역인 벡터공간의 기저 변화에 따라서 매번 달라진다.
- 우리가 익숙한 경우로 $\mathbb{V} = \mathbb{R}^n$, $\mathbb{W} = \mathbb{R}^m$ 일 때 표준기저 $\mathcal{B}_{\mathbb{V}} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, $\mathcal{B}_{\mathbb{W}} = \{\mathbf{e}'_1, \dots, \mathbf{e}'_m\}$ 가 있다. 선형사상 T 를 나타내는 행렬의 원소는 $a_{ij} = (T(\mathbf{e}_j))_i = T(\mathbf{e}_j)^\top \mathbf{e}'_i$ 인데 이 결과는 일반적이지 않은 특수한 경우라는 것에 유의해야 한다.

Example 3.3 A set \mathbb{V} includes all polynomials of degree n or less, that is, $\mathbb{V} = \{a_0 + a_1t + \dots + a_nt^n : a_0, a_1, \dots, a_n \in \mathbb{R}\}$. Let T map a polynomial $f(t)$ to its derivative $f'(t)$. n is a fixed integer in this question.

1. It is easy to check that \mathbb{V} is a vector space over the scalar set \mathbb{R} . Please do it by yourself.
2. Set $\mathcal{B}_{\mathbb{V}} = \{1, t, t^2, \dots, t^n\}$. It is clear that $\mathbb{V} = \text{span } \mathcal{B}_{\mathbb{V}}$. To see the linear independence, assume $a_0 + a_1t + \dots + a_nt^n = 0$ for all t . By plugging in various values into t , we get $a_0 = a_1 = \dots = a_n = 0$, and $\mathcal{B}_{\mathbb{V}}$ is linearly independent. Therefore, $\mathcal{B}_{\mathbb{V}}$ is a basis for \mathbb{V} and $\dim \mathbb{V} = |\mathcal{B}_{\mathbb{V}}| = n + 1$.
3. T can be regarded as a map from \mathbb{V} into \mathbb{V} . If we set $\mathbb{W} = \{a_0 + a_1t + \dots + a_{n-1}t^{n-1} : a_0, a_1, \dots, a_{n-1} \in \mathbb{R}\}$, T is also a map from \mathbb{V} onto \mathbb{W} . It can be also shown that T is a linear map in both cases. $\mathcal{B}_{\mathbb{W}} = \{1, t, t^2, \dots, t^{n-1}\}$ is a basis for \mathbb{W} .
4. For each $k = 0, 1, \dots, n$,

$$T(t^k) = kt^{k-1} = 0 + 0 \times t + \dots + 0 \times t^{k-2} + k \times t^{k-1} + 0 \times t^k + \dots + 0 \times t^{n-1}.$$

$n \times (n+1)$ matrix A representing T from $\mathcal{B}_{\mathbb{V}}$ onto $\mathcal{B}_{\mathbb{W}}$ is $A = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n-1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & n \end{bmatrix}$ and

$(n+1) \times (n+1)$ matrix A' representing T from \mathcal{B}_V into \mathcal{B}_V is $A' =$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n-1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & n \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Recall that matrix representations under different bases may be different.

■

Composition of Linear Transformations

두 개의 선형사상/변환 S 와 T 가 세 벡터공간 U, V, W 사이에 $S : U \rightarrow V, T : V \rightarrow W$ 로 정의되어 있고 $\dim(U) = n, \dim(V) = m, \dim(W) = \ell$ 이라 하자. 앞의 결과처럼 각 공간에서 선택한 세 기저 $\mathcal{B}_U, \mathcal{B}_V, \mathcal{B}_W$ 에 대해서 S 와 T 를 표현하는 행렬을 각각 $A = (a_{kj})$ 와 $B = (b_{ik})$ 라 하자. 그러면, 선형사상 $T \circ S : U \rightarrow W$ 를 기저 \mathcal{B}_U 와 \mathcal{B}_W 에 대해서 표현하는 $\ell \times n$ 행렬 C 는 어떻게 구할 수 있을까? 앞에서 계산했던 방식과 동일하게 할 수 있다. 기저 \mathcal{B}_U 의 임의의 basic vector \mathbf{u}_j 에 대해

$$S(\mathbf{u}_j) = \sum_{k=1}^m a_{kj} \mathbf{v}_k$$

이고, 기저 \mathcal{B}_V 의 임의의 basic vector \mathbf{v}_k 에 대해

$$T(\mathbf{v}_k) = \sum_{i=1}^{\ell} b_{ik} \mathbf{w}_i$$

이다. 따라서

$$\begin{aligned} (T \circ S)(\mathbf{u}_j) &= T\left(\sum_{k=1}^m a_{kj} \mathbf{v}_k\right) \\ &= \sum_{k=1}^m a_{kj} T(\mathbf{v}_k) \\ &= \sum_{k=1}^m a_{kj} \sum_{i=1}^{\ell} b_{ik} \mathbf{w}_i \\ &= \sum_{i=1}^{\ell} \left(\sum_{k=1}^m a_{kj} b_{ik} \right) \mathbf{w}_i \end{aligned}$$

이다. 이 표현을

$$(T \circ S)(\mathbf{u}_j) = \sum_{i=1}^{\ell} c_{ij} \mathbf{w}_i$$

와 비교하면

$$c_{ij} = \sum_{k=1}^m b_{ik}a_{kj} = (BA)_{ij}$$

가 되어

$$C = BA$$

를 얻는다. 이상을 요약하면 다음과 같은 정리를 얻을 수 있다.

Theorem 3.3 선형사상 S 와 T 의 특정 기저들에 대한 행렬표현이 A 와 B 인 경우, 동일한 기저들에 대해서 합성사상 $T \circ S$ 의 행렬표현은 BA 이다. 반대로 행렬 BA 에 대응되는 선형사상은 $T \circ S$ 이다.

3.8.2 Interpretable Linear Transformations

행렬 A 에 대해 $T(\mathbf{x}) = A\mathbf{x}$ 로 정의된 선형변환이 직관적으로 설명 가능한 경우를 살펴보자.

- **Scaling Matrix:** 행렬 $A = \alpha I = \begin{bmatrix} \alpha & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \alpha \end{bmatrix}$ 의 경우에는 $T(\mathbf{x}) = \alpha I \mathbf{x} = \alpha \mathbf{x}$ 가 되어 벡터 \mathbf{x} 의 각 좌표에 α 를 곱해주는 변환이다.

- **Rotation Matrix:** 2×2 행렬 $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ 의 경우에는 $T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$ 가 되어 벡터 \mathbf{x} 를 원점을 중심으로 반시계 방향으로 90° 회전시키는 변환이다. 이를 일반화하여 행렬

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

를 정의하자. 이 행렬에 대해 $R_\theta \mathbf{x}$ 는 \mathbf{x} 를 θ 만큼 회전이동시키는 변환이다.⁷ 정리 3.3을 이용하면 $R_\phi R_\theta$ 는 어떤 벡터를 θ 만큼 회전시키고 다시 ϕ 만큼 회전시키는 것이므로 결과적으로 $\theta + \phi$ 만큼 한번에 회전시키는 변환이고 이는 $R_{\theta+\phi}$ 로 표현된다. 따라서 행렬 간에

$$R_\phi R_\theta = R_{\theta+\phi}$$

이 성립함을 알 수 있다. 이 관계는 R_θ 의 정의와 삼각함수의 합 공식을 이용하여 쉽게 보일 수 있다. 따라서 $R_\theta R_{-\theta} = I$ 이므로 $R_\theta^{-1} = R_{-\theta}$ 이다. 또한 $R_{-\theta} = R_\theta^\top$ 임도 쉽게 알 수 있다. 이후에 살펴볼 R_θ 와 연관된 성질로 $R_\theta^{-1} = R_\theta^\top$ 이 있다. 어떤 행렬 A 가 $A^{-1} = A^\top$ 을 만족하면 orthogonal하다고 말한다.

A_v

- **Projection Matrix:** 행렬 $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ 의 경우에는 $T(\mathbf{x}) = P\mathbf{x} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$ 가 되어 벡터 \mathbf{x} 를 x 축에 정사영시키는 변환이다. 이를 일반화하여 원점을 지나며 x 축과 θ 의 각을 갖는 직선에 벡터를 정사영시키는

⁷ 이 행렬의 유도는 각자에게 맡긴다.

변환 T 의 행렬 P_θ 를 생각하자. $\mathbf{e}_1 = (1, 0)^\top$, $\mathbf{e}_2 = (0, 1)^\top$ 에 대해 $T(\mathbf{e}_1) = (\cos \theta \cos \theta, \sin \theta \cos \theta)^\top$ 이고 $T(\mathbf{e}_2) = (\cos \theta \sin \theta, \sin \theta \sin \theta)^\top$ 임을 유도할 수 있다. 따라서

$$\begin{aligned} T(\mathbf{x}) &= x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) \\ &= x_1 \begin{bmatrix} \cos \theta \cos \theta \\ \sin \theta \cos \theta \end{bmatrix} + x_2 \begin{bmatrix} \sin \theta \cos \theta \\ \sin \theta \sin \theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

이므로

$$P_\theta = \begin{bmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{bmatrix}$$

이다. 일반적으로 정사영한 벡터는 다시 정사영하더라도 변하지 않으므로 모든 정사영 행렬 P 는

$$P^2 = P$$

를 만족해야 한다. 이 성질을 위에서 계산한 행렬 P_θ 에서 확인할 수 있다. Fact 4.16에서 $n \times n$ 행렬 P 가 \mathbb{R}^n 공간의 정사영 행렬이기 위한 필요충분조건이 $P^2 = P$ 이면서 P 가 대칭임을 살펴본다. 마지막으로 언급할 사항은 P 가 정사영 행렬이면 $I - P$ 도 정사영 행렬이 된다는 것이다. P_θ 에 대해서 각자 확인해 보기 바란다.

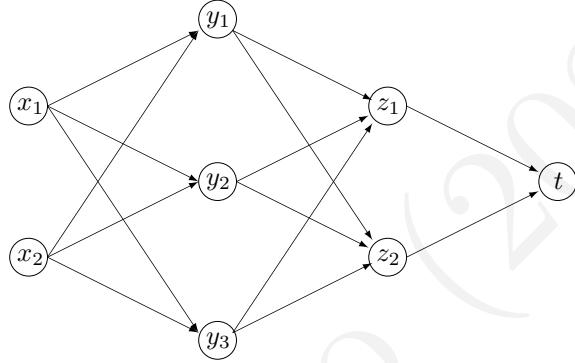
- **Reflection Matrix** 행렬 $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 의 경우에는 $T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$ 가 되어 벡터 \mathbf{x} 를 45° 기울기의 원점을 지나는 직선인 $y = x$ 에 대해 대칭이동시키는 변환이다.

이를 일반화하여 기울기 θ 이면서 원점을 지나는 직선에 대해 대칭이동시키는 변환 T 를 표현하는 행렬 H_θ 를 고려하자. 그러면 $T(\mathbf{x})$ 와 \mathbf{x} 의 중점인 벡터는 \mathbf{x} 를 대칭의 기준이되는 직선에 정사영시킨 벡터와 일치할 것이다. 즉 $\frac{1}{2}(\mathbf{x} + T(\mathbf{x})) = P_\theta \mathbf{x}$ 이다. 이를 정리하면 $T(\mathbf{x}) = 2P_\theta \mathbf{x} - \mathbf{x} = (2P_\theta - I)\mathbf{x}$ 를 만족한다. 따라서 $H_\theta = 2P_\theta - I$ 이다.

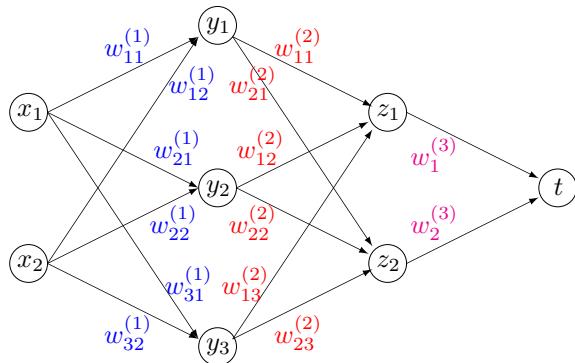
더 일반적으로 벡터공간 \mathbb{R}^n 의 어떤 subspace에 대한 대칭변환을 나타내는 reflection matrix를 생각해보자. 그 subspace상으로 정사영하는 변환을 나타내는 $n \times n$ projection matrix P 가 주어지면 (즉 대칭이고 $P^2 = P$ 를 만족하는) P 가 정사영하는 공간에 대한 대칭이동을 나타내는 행렬 H 는 앞에서 살펴본 2차원 공간의 예처럼 생각하여 $H = 2P - I$ 로 생각할 수 있다. 이 경우, $H^\top = H$ 이고 $H^2 = (2P - I)^2 = 4P^2 - 4P + I = I$ 이다. 반대로 $H^\top = H$ 이며 $H^2 = I$ 인 $n \times n$ 행렬 H 가 주어졌을 때 $P = \frac{1}{2}(H + I)$ 로 놓으면 $P^\top = P$ 이고 $P^2 = \frac{1}{4}(H^2 + 2H + I) = \frac{1}{4}(2H + 2I) = P$ 가 되어 P 가 정사영 행렬이라는 것을 알 수 있다. 이렇게 $H^\top = H$ 이면서 $H^2 = I$ 인 행렬들을 reflection matrix라고 한다. Projection matrix에서 언급된 것처럼 P 가 정사영 행렬일 때 $I - P$ 도 정사영 행렬이다. 따라서 정사영 행렬 P 에 대해서 $2P - I$ 뿐만 아니라 $2(I - P) - I = I - 2P$ 도 reflection matrix이다.

3.9 Application to Data Science: Neural Networks

기계학습의 인공신경망(neural network) 분야에서 선형/비선형 변환과 이 변환을 표현하는 행렬을 이용하여 많은 계산이 수행된다. 인공신경망에 의한 학습은 \mathbb{R}^d 의 부분집합인 입력 데이터에 반복적으로 선형변환과 비선형변환을 적용하여 출력된 값이 데이터의 관측값과 최대한 가까워지도록 선형변환을 수정하는 과정이다. Layer 구조를 갖는 신경망에서 한 layer의 선형변환은 network 구조로 시각화/모형화할 수 있고 비선형변환은 이 network의 node별로 적용된다. 다음과 같은 2차원 입력 데이터 $\mathbf{x} = (x_1, x_2)^\top \in \mathbb{R}^2$ 에 대한 인공신경망의 예를 보자.



이 인공신경망에서 계산의 흐름은 왼쪽에서 오른쪽 방향으로 두 개의 실수 (x_1, x_2) 를 입력으로 선형/비선형 변환을 거쳐 세 개의 실수 (y_1, y_2, y_3) 를 얻고 다시 이 세 실수를 변환하여 (z_1, z_2) , 마지막으로 한번 더 변환하여 최종 출력인 t 값이 계산된다. 그림에서 쉽게 알 수 있듯이 변수들이 이렇게 몇 개씩 구분되는 구조를 layer 구조라 하고 (x_1, x_2) 를 input layer, (y_1, y_2, y_3) 와 (z_1, z_2) 를 각각 hidden layer라 부르고 마지막에 node t 로 이루어진 output layer로 이루어진다. 이 layer 구조 덕분에 node가 매우 많은 인공신경망도 (특히 GPU를 활용하여) 효과적인 계산이 가능하다. 두 layer 간에 선형변환이 어떻게 이루어지는지 살펴보자. 먼저 (x_1, x_2) 에서 (y_1, y_2, y_3) 으로의 선형변환은 2차원 공간에서 3차원 공간으로의 선형변환이므로 Section 3.8.1의 설명에 따르면 3×2 행렬로 표현됨을 알 수 있다. 이 행렬을 $W^{(1)} = (w_{ij}^{(1)})$ 로 표현하자. 다음 layer 쌍인 (y_1, y_2, y_3) 에서 (z_1, z_2) 로의 선형변환은 2×3 행렬 $W^{(2)} = (w_{ij}^{(2)})$ 로 표현하고 마지막 layer 쌍인 (z_1, z_2) 에서 (t) 로의 선형변환은 1×2 행렬 $W^{(3)} = [w_1^{(3)} w_2^{(3)}]$ 로 표현하자. 예를 들면 $w_{ij}^{(1)}$ 는 x_j 가 y_i 에 미치는 영향의 선형적인 부분을 나타내므로 x_j 와 y_i 를 연결하는 화살표와 연관시켜야 한다. 이렇게 인공신경망의 화살표와 선형변환을 나타내는 행렬의 원소를 대응시키면 다음 그림과 같다. $w_{ij}^{(k)}$ 를 인공신경망의 가중치(weight) 또는 인공신경망을 활용한 학습(learning)의 매개변수(parameter)라고 부른다.



비선형적인 변환은 각 node에 가해진 선형적 영향을 모두 합친 값에 비선형 함수 $\sigma^{(k)}$ 를 취하여 계산된다. 각 노드의 비선형 변환을 나타내는 함수를 activation function이라고 부르는데 보통 동일한 layer에 속하는 node들은 동일한 activation function을 사용한다. 선형 변환이 이루어지고 비선형 변환이 이루어지기 직전의 값을 나타내는 변수명에 hat을 씌워서 $\hat{y}_i, \hat{z}_i, \hat{t}$ 로 표시하자.

이 예에 대해서 구체적으로 살펴보면 첫 번째와 두 번째 layer 간의 선형 변환은 파란색으로 표현된 parameter로 표현되는데

$$\hat{y}_1 = w_{11}^{(1)}x_1 + w_{12}^{(1)}x_2, \quad \hat{y}_2 = w_{21}^{(1)}x_1 + w_{22}^{(1)}x_2, \quad \hat{y}_3 = w_{31}^{(1)}x_1 + w_{32}^{(1)}x_2$$

이고 이 선형 변환의 결과에 activation function을 취하여 $y_1 = \sigma^{(1)}(\hat{y}_1), y_2 = \sigma^{(1)}(\hat{y}_2), y_3 = \sigma^{(1)}(\hat{y}_3)$ 로 표현할 수 있다. 벡터에 activation function을 취하는 경우, 벡터 함수의 미분과 적분처럼 벡터의 각 원소 별로 activation function을 취한 것으로 해석한다. 따라서 위의 계산을 행렬과 벡터로 표시하면

$$\hat{\mathbf{y}} = \mathbf{W}^{(1)}\mathbf{x} = \begin{bmatrix} w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \\ w_{31}^{(1)} & w_{32}^{(1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{y} = \sigma^{(1)}(\hat{\mathbf{y}})$$

로 나타낼 수 있다. 다음 layer 쪽에서도 유사하게

$\hat{z}_1 = w_{11}^{(2)}y_1 + w_{12}^{(2)}y_2 + w_{13}^{(2)}y_3, \quad \hat{z}_2 = w_{21}^{(2)}y_1 + w_{22}^{(2)}y_2 + w_{23}^{(2)}y_3$ 와 $z_1 = \sigma^{(2)}(\hat{z}_1), z_2 = \sigma^{(2)}(\hat{z}_2)$ 인데 이를 행렬과 벡터로 표현하면

$$\hat{\mathbf{z}} = \mathbf{W}^{(2)}\mathbf{y} = \begin{bmatrix} w_{11}^{(2)} & w_{12}^{(2)} & w_{13}^{(2)} \\ w_{21}^{(2)} & w_{22}^{(2)} & w_{23}^{(2)} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad \mathbf{z} = \sigma^{(2)}(\hat{\mathbf{z}})$$

이다. 마지막 layer 쪽에서는 $\hat{t} = w_1^{(3)}z_1 + w_2^{(3)}z_2$ 이고 $t = \sigma^{(3)}(\hat{t})$ 를 얻고

$$\hat{t} = \mathbf{W}^{(3)}\mathbf{z} = \begin{bmatrix} w_{11}^{(3)} & w_{12}^{(3)} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \quad t = \sigma^{(3)}(\hat{t})$$

로 행렬과 벡터를 사용하여 표현된다. 이 모든 식을 한 줄에 적으면

$$t = \sigma^{(3)}(\mathbf{W}^{(3)}\sigma^{(2)}(\mathbf{W}^{(2)}\sigma^{(1)}(\mathbf{W}^{(1)}\mathbf{x})))$$

이다.⁸ 인공신경망을 이용하여 학습한다는 것은 설정된 학습의 지표가 좋아지도록 $w_{ij}^{(k)}$ 를 수정하는 과정을 의미한다. 최근에 상용화된 AI서비스의 인공신경망은 parameter인 $w_{ij}^{(k)}$ 개수가 천억 개에서 조 단위인 경우까지 커지고 있다.

비선형 변환을 나타내는 activation function은 생물학의 S자 곡선과 유사한 sigmoid 함수 $(1 + e^{-x})^{-1}$ 가 초기에 사용되었는데 작은 크기의 인공신경망도 당시 계산환경에서는 학습이 쉽지 않았다. 인공신경망의 activation function으로 비선형 함수 중 계산이 가장 간단하다고 할 수 있는 한번만 꺠인 선형함수 $\max\{x, 0\}$ 를 사용한 것이 deep learning의 초반 폭발적인 발전이 가능하였던 요인 중 한 가지이다. 이 함수를 ReLU (Rectified Linear Unit)라고 부른다.

⁸이해가 쉽도록 bias라고 부르는 각 node에서 추가되는 parameter는 위의 표현에서 생략하였다. Bias는 node에 가해진 선형적 영향의 총합에 상수항처럼 더해지는 parameter로서 bias가 더해진 후에 activation function이 취해져서 최종적인 node의 값이 정해진다.

3.9.1 Flexibility of Neural Network Representations

직선 layer에 속한 node들의 출력값의 linear combination에 비선형함수인 activation function을 추가로 적용한 neural network은 얼마나 비선형적인 표현을 나타낼 수 있을까? Figure 3.1의 간단한 신경망으로 정의된 함수 $f(x, y; \Theta)$ 를 생각하자.

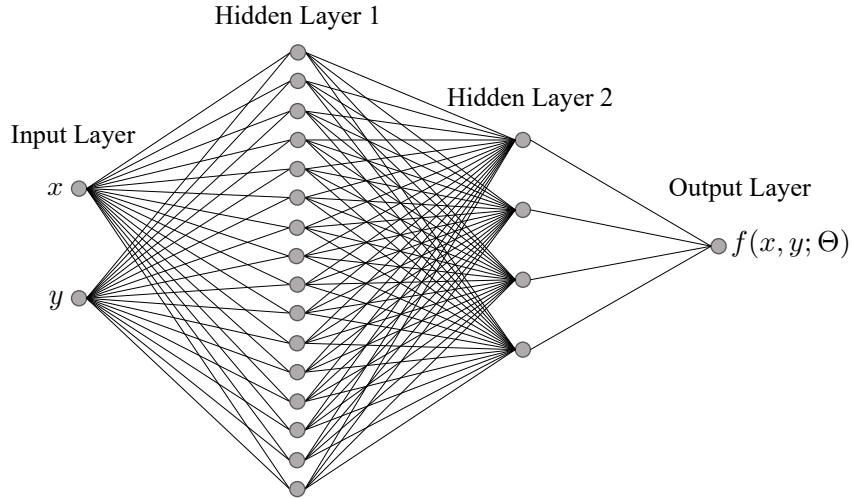


Figure 3.1: A simple neural network with 2 hidden layers and 121 learning parameters

두 변수 (x, y) 를 입력으로 16개의 node로 구성된 첫번째 hidden layer에 연결된 후에 4개의 노드로 이루어진 두번째 hidden layer를 거쳐서 output node값 $f(x, y; \Theta)$ 가 결정된다. 두개의 hidden layer에 속한 20개의 node들은 ReLU를 activation function으로 갖고 마지막 output node는 sigmoid function에 의해 activate 된다. 100개의 edge에 대응되는 100개의 weight와 input node를 제외한 21개의 node에서 더해지는 21개의 bias가 이 신경망의 parameter이며 $\Theta \in \mathbb{R}^{121}$ 로 표현하였다. 즉 Θ 가 바뀌면 동일한 (x, y) 에 대해서 $f(x, y; \Theta)$ 가 변하므로 함수 f 가 Θ 에 의해 결정되는 것이다. 4가지 종류의 실제 데이터 $\{(x_i, y_i, f_i)\}_{i=1}^n$ 에 대해서

$$\min_{\Theta \in \mathbb{R}^{121}} \sum_{i=1}^n \|f(x_i, y_i; \Theta) - f_i\|$$

를 풀어서 얻어진 4개의 $\Theta_1, \dots, \Theta_4$ 에 대해서 output node의 출력값인 $f(x, y; \Theta_j)$ 를 Figure 3.2에 도시하였다.

다시 한번 강조하지만 4개의 plot이 모두 Figure 3.1의 동일한 구조에서 4개의 다른 parameter로 정의된 함수값의 graph이다. 이 예제에서 단순한 ReLU의 nonlinearity로부터 아주 다양한 비선형 함수들을 얻을 수 있다는 것을 알 수 있다.

3.10 Application to Graph Theory

Section 2.7에서 유방향 또는 무방향 그래프(또는 네트워크)에 대해 정의된 두 가지 행렬–adjacency and incidence matrices–에 대해서 알아보았다. 이 절에서는 유방향 그래프의 incidence matrix에 의해 유도되는 부분공간에 대해서 살펴보자. 다음과 같은 간단한 그래프를 고려하자.

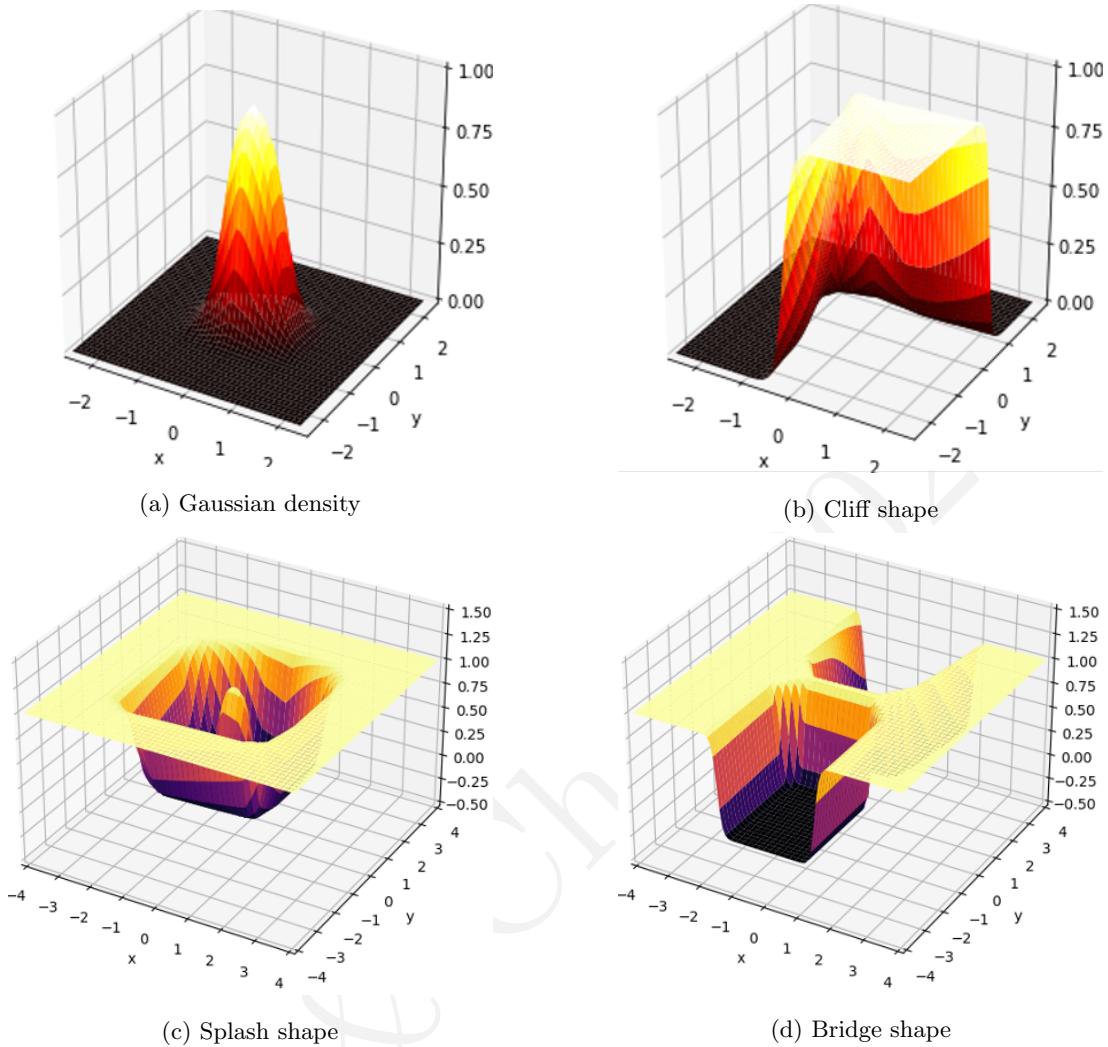
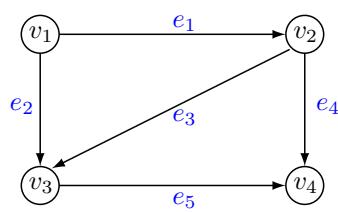


Figure 3.2: Various output representations by a simple neural network



	v_1	v_2	v_3	v_4	
e_1	-1	1			
e_2	-1		1		
e_3		-1	1		
e_4		-1		1	
e_5			-1	1	

이 그래프의 incidence matrix는 $A = \begin{matrix} & v_1 & v_2 & v_3 & v_4 \\ e_1 & -1 & 1 & & \\ e_2 & -1 & & 1 & \\ e_3 & & -1 & 1 & \\ e_4 & & -1 & & 1 \\ e_5 & & & -1 & 1 \end{matrix}$ 로 주어진다. $\mathbf{1} = (1, 1, 1, 1)^\top$ 이라 하고 편의 상 A 의 i 번째 행벡터를 \mathbf{r}_i 로 표시하자. 예제 그래프와 incidence matrix A 에 대해 다음과 같은 몇 가지 사실을 관찰할 수 있다.

1. $A\mathbf{1} = \mathbf{0}$ 즉 $\mathbf{1} \in \text{Null}(A)$ 이다. 이 성질은 위의 특정한 그래프뿐만 아니라 모든 유방향 그래프의 incidence matrix는 각 행에 한 개의 1과 한 개의 -1을 가지고 있기 때문에 이 등식이 성립한다.
2. $\mathbf{r}_1 - \mathbf{r}_2 + \mathbf{r}_3 = \mathbf{0}$ 의 관계가 있어서 행벡터 $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ 는 선형종속이고 $(1, -1, 1, 0, 0)^\top \in \text{Null}(A^\top)$ 이다. 그런데 그래프에서 arc의 방향을 무시하면 행벡터 $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ 에 대응되는 arc e_1, e_2, e_3 는 cycle 또는 loop를 형성한다. 이 성질은 모든 그래프에서 cycle이 존재하는 경우 그 cycle을 구성하는 arc들에 대해 성립한다.
3. 임의의 두 node가 arc에 의해 연결되는 그래프가 connected graph이고 그래프에 cycle이 없으면 acyclic graph라 부른다. 이 두 가지 성질을 모두 가지고 있는 acyclic connected graph를 tree라고 부른다. 이 tree를 구성하는 arc들의 행벡터들은 선형독립인 것을 보일 수 있다. 또한 n 개의 node에 대한 tree는 $n - 1$ 개의 arc로 구성된다는 것을 수학적 귀납법으로 보일 수 있다.
4. n 개의 node로 구성된 connected graph의 incidence matrix A 의 rank는 $n - 1$ 이 된다. $A\mathbf{1} = \mathbf{0}$ 이므로 $\text{rank } A \leq n - 1$ 이다. 최소한의 연결성만 같은 subgraph인 tree의 $n - 1$ 개의 arc에 대응하는 A 의 행벡터들은 선형독립이다. 따라서 $\text{rank } A \geq n - 1$ 을 얻는다.
5. $\text{rank } A = n - 1$ 이므로 $\dim \text{Null}(A) = 1$ 인데 $A\mathbf{1} = \mathbf{0}$ 이므로 $\text{Null}(A) = \text{span}\{\mathbf{1}\}$ 이다.

Kang & Cho (2024)

Chapter 4

Orthogonality and Projections

벡터공간 속의 두 벡터 간의 기하학적 관계를 어떻게 계량적으로 나타낼 수 있을까? 우리가 익숙한 이차원 평면 상에서 어떤 벡터의 좌측이라던지 영벡터를 기준으로 시계방향과 같은 기하학적 관념은 3차원 이상의 유클리드 공간이나 다향식을 모아 놓은 벡터공간에서 더이상 존재하지 않는다. 따라서 두 벡터를 비교할 때 두 벡터의 순서는 무의미하여 벡터의 기하학적 관계를 나타내는 양은 고려하는 두 벡터에 대해 대칭이어야 한다. 그리고 선형대수학 전반에 흐르는 선형성을 고려하면 기하학적 양은 비교하는 두 벡터 각각에 대해서 선형적인 것이 자연스럽다. 여기에 추가로 동일한 벡터에 대해서는 길이 성질을 만족하는 양의 제곱으로 해석할 수 있는데 이 양을 내적이라고 부른다. 어떤 내적에 몇 배를 해도 내적의 요건에는 문제가 없어서 어떤 벡터공간에 내적이 한개라도 존재하면 내적은 무한개가 존재하게 된다. 따라서 내적의 값이 0인 경우를 제외하고 다른 값들에 절대적인 의미를 부여하기 어렵다. 내적이 0이라는 것과 한 벡터의 내적이 길이의 제곱으로 해석되고 이 길이 하에서 내적이 0이라는 것과 두 벡터 간에 피타고拉斯 정리가 성립하는 것이 동치가 된다. 따라서 내적이 0인 경우에 두 벡터가 서로 수직하다고 기하학적으로 표현한다. 이차원 평면의 원점을 지나는 직선과 같이 벡터공간의 부분집합이 벡터공간의 요건을 만족하면 부분공간이라고 부른다. 그러면 수직 관계는 벡터와 부분공간, 두 부분공간 간의 관계로 확장가능하다. 어떤 벡터에서 부분공간에 수직한 벡터 방향으로 이동하면 부분공간 속의 한개의 벡터와 일치하는데 이를 정사영이라고 한다. 이때 부분공간이 선형독립인 벡터들, 또는 서로 수직인 벡터들을 기저로 생성될 때 내적을 이용하여 정사영을 표현할 수 있다. 특히 서로 수직인 벡터를 기저로 하는 경우 정사영의 계산이 유클리드 공간에서의 좌표 계산과 유사해지는데 여러 계산에서 유용하여 서로 수직인 벡터를 찾는 것이 중요하다. 다행히 Gram-Schmidt 방법을 따르면 서로 수직한 벡터를 차례로 찾을 수 있는데 이후에 많이 활용된다.

4.1 Inner Products

Definition 4.1 벡터공간 \mathbb{V} 의 임의의 두 벡터 \mathbf{v}_1 과 \mathbf{v}_2 간에 정의된 실수값을 갖는 함수

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle$$

가 다음과 같은 성질을 만족하면 내적(*inner product*)이라 부른다:

1. $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_2, \mathbf{v}_1 \rangle$;
2. $\langle c\mathbf{v}_1, \mathbf{v}_2 \rangle = c\langle \mathbf{v}_1, \mathbf{v}_2 \rangle$ for any real number c ;
3. $\langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_3 \rangle = \langle \mathbf{v}_1, \mathbf{v}_3 \rangle + \langle \mathbf{v}_2, \mathbf{v}_3 \rangle$ for any $\mathbf{v}_3 \in \mathbb{V}$;
4. $\langle \mathbf{v}_1, \mathbf{v}_1 \rangle > 0$ if and only if $\mathbf{v}_1 \neq \mathbf{0}$.

내적이 정의된 경우, 벡터 $\mathbf{v} \in \mathbb{V}$ 의 크기(*norm*)를 (실수의 절댓값 기호와 동일하게)

$$|\mathbf{v}| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

로 정의한다. 벡터 \mathbf{v} 의 크기가 1이면, 즉 $\langle \mathbf{v}, \mathbf{v} \rangle = 1$ 이면 \mathbf{v} 를 *unit vector*라고 부른다. 내적을 $\mathbf{v}_1 \bullet \mathbf{v}_2$ 로 표시하기도 한다.

정의 4.1의 2번으로부터 $\langle \mathbf{v}_1, \mathbf{0} \rangle = 0$ 임을 알 수 있고 1번과 3번을 고려하면 $\langle \mathbf{v}_1, \mathbf{v}_2 + \mathbf{v}_3 \rangle = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle + \langle \mathbf{v}_1, \mathbf{v}_3 \rangle$ 임도 쉽게 알 수 있다. 임의의 양의 상수 α 에 대해서 $\alpha \times \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$ 가 내적이 되므로 내적이 한 개 정의되면 무수히 많은 다른 내적이 존재한다. 또한, 임의의 두 내적 $\langle \cdot, \cdot \rangle_1$ 와 $\langle \cdot, \cdot \rangle_2$ 의 합인 $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle_1 + \langle \mathbf{v}_1, \mathbf{v}_2 \rangle_2$ 도 내적이다. Theorem 4.1에서는 가능한 내적을 어떻게 모두 파악할 수 있는지 살펴본다.

일반적으로 벡터공간 \mathbb{V} 상의 norm은 \mathbb{V} 에서 정의된 실수 값을 갖는 함수로서 이 함수를 f 라 하면

- $f(\mathbf{v}) \geq 0$ 이면서 $f(\mathbf{v}) = 0$ if and only if $\mathbf{v} \in \mathbb{V}$;
- 상수배: $f(c\mathbf{v}) = |c|f(\mathbf{v})$ for all $c \in \mathbb{R}$ and $\mathbf{v} \in \mathbb{V}$;
- 삼각부등식: $f(\mathbf{v} + \mathbf{w}) \leq f(\mathbf{v}) + f(\mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{V}$

를 만족하여야 한다. 내적의 정의와 Fact 4.2에 의해 내적에 의해 정의되는 norm이 진짜 norm이라는 것을 알 수 있다.

내적이 주어질 때 생각할 수 있는 첫 번째 시도로 고등학교 시절부터 자주 사용하던 코쉬-슈바르츠 부등식이 벡터공간에서도 성립한다는 결과를 생각해보자.

Fact 4.1 The Cauchy-Schwarz inequality holds between an inner product and the norm induced by the inner product:

$$|\langle \mathbf{v}_1, \mathbf{v}_2 \rangle| \leq |\mathbf{v}_1| \cdot |\mathbf{v}_2|.$$

Proof: For any real number $t \in \mathbb{R}$,

$$\begin{aligned}
 |t\mathbf{v}_1 + \mathbf{v}_2|^2 &= \langle t\mathbf{v}_1 + \mathbf{v}_2, t\mathbf{v}_1 + \mathbf{v}_2 \rangle \\
 &= \langle t\mathbf{v}_1 + \mathbf{v}_2, t\mathbf{v}_1 \rangle + \langle t\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 \rangle \\
 &= \langle t\mathbf{v}_1, t\mathbf{v}_1 \rangle + \langle \mathbf{v}_2, t\mathbf{v}_1 \rangle + \langle t\mathbf{v}_1, \mathbf{v}_2 \rangle + \langle \mathbf{v}_2, \mathbf{v}_2 \rangle \\
 &= t^2\langle \mathbf{v}_1, \mathbf{v}_1 \rangle + t\langle \mathbf{v}_2, \mathbf{v}_1 \rangle + t\langle \mathbf{v}_1, \mathbf{v}_2 \rangle + \langle \mathbf{v}_2, \mathbf{v}_2 \rangle \\
 &= t^2\langle \mathbf{v}_1, \mathbf{v}_1 \rangle + 2t\langle \mathbf{v}_1, \mathbf{v}_2 \rangle + \langle \mathbf{v}_2, \mathbf{v}_2 \rangle.
 \end{aligned}$$

Since $|t\mathbf{v}_1 + \mathbf{v}_2| \geq 0$ for any t , the quadratic equation on the right-hand side has no two different solutions, that is,

$$0 \geq \langle \mathbf{v}_1, \mathbf{v}_2 \rangle^2 - \langle \mathbf{v}_1, \mathbf{v}_1 \rangle \langle \mathbf{v}_2, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle^2 - |\mathbf{v}_1|^2 \cdot |\mathbf{v}_2|^2.$$

■

그리고 코쉬-슈바르츠 부등식을 얻으면 자연스럽게 다음과 같은 삼각부등식을 얻는다. 이 삼각부등식 덕분에 내적으로부터 정의되는 벡터의 크기가 자연스러워진다.

Fact 4.2 *The triangular inequality holds for the norm induced from the inner product $\langle \cdot, \cdot \rangle$:*

$$|\mathbf{v}_1 + \mathbf{v}_2| \leq |\mathbf{v}_1| + |\mathbf{v}_2|$$

and the positive homogeneity also holds:

$$|c\mathbf{v}| = c|\mathbf{v}| \text{ for any positive real number } c.$$

Proof:

$$\begin{aligned}
 |\mathbf{v}_1 + \mathbf{v}_2|^2 &= \langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2 \rangle \\
 &= \langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 \rangle + \langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 \rangle \\
 &= \langle \mathbf{v}_1, \mathbf{v}_1 \rangle + \langle \mathbf{v}_2, \mathbf{v}_1 \rangle + \langle \mathbf{v}_1, \mathbf{v}_2 \rangle + \langle \mathbf{v}_2, \mathbf{v}_2 \rangle \\
 &= \langle \mathbf{v}_1, \mathbf{v}_1 \rangle + 2\langle \mathbf{v}_1, \mathbf{v}_2 \rangle + \langle \mathbf{v}_2, \mathbf{v}_2 \rangle \\
 &\leq |\mathbf{v}_1|^2 + 2|\mathbf{v}_1| \cdot |\mathbf{v}_2| + |\mathbf{v}_2|^2 \\
 &= (|\mathbf{v}_1| + |\mathbf{v}_2|)^2.
 \end{aligned}$$

$$|c\mathbf{v}_1| = \sqrt{\langle c\mathbf{v}, c\mathbf{v} \rangle} = \sqrt{c^2 \langle \mathbf{v}, \mathbf{v} \rangle} = c\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

■

내적에 대한 첫번째 예로 다음과 같은 다항식을 모아 놓은 벡터공간을 생각해 보자.

Example 4.1 Let \mathbb{V} be a vector space of polynomials of degrees less than or equal to n . For two polynomials $f(t), g(t) \in \mathbb{V}$, we define an inner product

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$

For $f(t) = t$ and $g(t) = t^2$,

- $\int_{-1}^1 f(t)g(t)dt = \int_{-1}^1 g(t)f(t)dt$ is linear in f and g . $\int_{-1}^1 f(t)^2 dt = 0$ implies $f \equiv 0$ since f is a polynomial which is continuous. Hence, $\langle \cdot, \cdot \rangle$ is an inner product.
- $\langle f, g \rangle = \int_{-1}^1 t \cdot t^2 dt = \int_{-1}^1 t^3 dt = 0$.
- $|f| = \sqrt{\langle f, f \rangle} = \sqrt{\int_{-1}^1 t^2 dt} = \sqrt{\frac{1}{3}t^3 \Big|_{-1}^1} = \sqrt{\frac{2}{3}}$.
- $|g| = \sqrt{\langle g, g \rangle} = \sqrt{\int_{-1}^1 t^4 dt} = \sqrt{\frac{1}{5}t^5 \Big|_{-1}^1} = \sqrt{\frac{2}{5}}$.
- $|f - g| = \sqrt{|f|^2 - 2\langle f, g \rangle + |g|^2} = \sqrt{\frac{2}{3} - 2 \cdot 0 + \frac{2}{5}} = \sqrt{\frac{16}{15}}$.

■

벡터공간에 정의된 내적을 분석하자. 모든 벡터 쌍에 대해서 정의된 내적 값들은 너무 다양하여 일반적으로 분석하기 쉽지 않다. 대신에 선형독립인 벡터 쌍들의 내적 값들을 집합적으로 분석해보자.

Lemma 4.1 Assume that non-zero vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in a vector space \mathbb{V} are linearly independent. Define $a_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$ and a $k \times k$ matrix $A = (a_{ij})$. Then A is symmetric and invertible.

Proof: Symmetry comes from the definition of the inner product. To show the invertibility of A , it is enough to show that $A\mathbf{x} = \mathbf{0}$ has no non-trivial solution. Let $\mathbf{x} \in \mathbb{R}^k$ satisfy $A\mathbf{x} = \mathbf{0}$. The i -th equation of $A\mathbf{x} = \mathbf{0}$ is

$$a_{i1}x_1 + \cdots + a_{ik}x_k = x_1\langle \mathbf{v}_i, \mathbf{v}_1 \rangle + \cdots + x_k\langle \mathbf{v}_i, \mathbf{v}_k \rangle = 0.$$

Rearranging the equation using the linearity of inner product gives

$$\langle \mathbf{v}_i, x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k \rangle = 0 \text{ for } i = 1, \dots, k.$$

If we add the equations after multiplying x_i to i -th equation, the linearity again allows

$$0 = \sum_{i=1}^k x_i \langle \mathbf{v}_i, x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k \rangle = \langle x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k, x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k \rangle.$$

Then the definition of inner product imposes $x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$, which implies $\mathbf{x} = \mathbf{0}$ by the linear independence of \mathbf{v}_i 's. Hence the nullspace of A is $\{\mathbf{0}\}$ and the desired conclusion holds. ■

Characterization of Inner Products

선형독립인 벡터 쌍에 대한 분석한 Lemma 4.1을 바탕으로 기저를 구성하는 벡터 쌍에 대한 내적 값들로 내적을 기술하는 방법에 대해서 생각해 보자.

Let us fix a basis $\mathcal{B}_{\mathbb{V}} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ for \mathbb{V} . For any two vectors \mathbf{v} and \mathbf{w} in \mathbb{V} , there exist two unique vectors $\mathbf{x} = (x_1, \dots, x_n)^\top$ and $\mathbf{y} = (y_1, \dots, y_n)^\top$ in \mathbb{R}^n , respectively, such that

$$\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n \quad \text{and} \quad \mathbf{w} = y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n.$$

Then, the bilinearity of inner product implies

$$\langle \mathbf{v}, \mathbf{w} \rangle = \left\langle \sum_{i=1}^n x_i \mathbf{v}_i, \sum_{j=1}^n y_j \mathbf{v}_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle.$$

If we set $a_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$ and an $n \times n$ symmetric matrix $A = (a_{ij})$ as well, then

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j a_{ij} = \mathbf{x}^\top A \mathbf{y}.$$

In addition, note that $\mathbf{v} = \mathbf{0} \in \mathbb{V}$ if and only if $\mathbf{x} = \mathbf{0} \in \mathbb{R}^n$. Then, $\langle \mathbf{v}, \mathbf{v} \rangle = \mathbf{x}^\top A \mathbf{x} > 0$ if and only if $\mathbf{x} \neq \mathbf{0}$. Hence A can be characterized by

$$\text{an } n \times n \text{ real symmetric matrix such that } \mathbf{x}^\top A \mathbf{x} > 0 \text{ for any } \mathbf{x} \neq \mathbf{0}.$$

Therefore, a quadratic form taking positive values for any non-zero vectors characterizes an inner product. And the matrix appearing in the quadratic form also characterizes the inner product. This property is called a **positive definiteness** and formalized by Definition 4.2.¹

Definition 4.2 A square matrix A is **positive definite** if $\mathbf{x}^\top A \mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$.

Conversely, if we have a positive definite matrix A and define a bilinear function as

$$\left\langle \sum_{i=1}^n x_i \mathbf{v}_i, \sum_{j=1}^n y_j \mathbf{v}_j \right\rangle = \mathbf{x}^\top A \mathbf{y}, \tag{4.1}$$

it is an inner product by Lemma 4.2.

We can summarize the conclusion as the following theorem:

Theorem 4.1 An inner product in an n -dimensional vector space is characterized by an $n \times n$ symmetric positive definite matrix as in (4.1).

다음 lemma를 증명하면 Theorem 4.1의 증명이 완료된다.

Lemma 4.2 Let \mathbb{V} be a vector space and A be an $n \times n$ symmetric positive definite matrix. Fix an arbitrary basis of \mathbb{V} , say, $\mathcal{B}_{\mathbb{V}} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. For any two vectors \mathbf{v} and \mathbf{w} in \mathbb{V} , define $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{x}^\top A \mathbf{y}$ where $\mathbf{x} = (x_1, \dots, x_n)^\top$ and $\mathbf{y} = (y_1, \dots, y_n)^\top$ satisfy $\mathbf{v} = \sum_{i=1}^n x_i \mathbf{v}_i$ and $\mathbf{w} = \sum_{i=1}^n y_i \mathbf{v}_i$. Then, $\langle \mathbf{v}, \mathbf{w} \rangle$ is an inner product of \mathbb{V} .

¹이 개념에 대해서는 Chapter 7에서 자세하게 다룬다.

Proof: Let $\mathbf{v} = \sum_{i=1}^n x_i \mathbf{v}_i$, $\mathbf{w} = \sum_{i=1}^n y_i \mathbf{v}_i$, and $\mathbf{u} = \sum_{i=1}^n z_i \mathbf{v}_i$.

- $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{x}^\top A \mathbf{y} = (\mathbf{x}^\top A \mathbf{y})^\top = \mathbf{y}^\top A^\top \mathbf{x} = \mathbf{y}^\top A \mathbf{x} = \langle \mathbf{w}, \mathbf{v} \rangle$;
- $\langle c\mathbf{v}, \mathbf{w} \rangle = (c\mathbf{x})^\top A \mathbf{y} = c(\mathbf{x}^\top A \mathbf{y}) = c\langle \mathbf{w}, \mathbf{v} \rangle$;
- $\langle \mathbf{v} + \mathbf{u}, \mathbf{w} \rangle = (\mathbf{x} + \mathbf{z})^\top A \mathbf{y} = \mathbf{x}^\top A \mathbf{y} + \mathbf{z}^\top A \mathbf{y} = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
- $\langle \mathbf{v}, \mathbf{v} \rangle = \mathbf{x}^\top A \mathbf{x} > 0$ if $\mathbf{x} \neq \mathbf{0}$ and equivalently $\mathbf{v} \neq \mathbf{0}$.

■

특히 우리가 자주 사용하고 친숙한 유클리드 공간에서도 다양한 내적을 정의할 수 있고, 여러 내적 중에서 다음과 같이 정의되는 내적을 표준내적이라 부른다.

Definition 4.3 n 차원 유클리드 공간 \mathbb{R}^n 에서 두 벡터 $\mathbf{x} = (x_1, \dots, x_n)^\top$, $\mathbf{y} = (y_1, \dots, y_n)^\top$ 간의 표준내적 (standard inner product, dot product)을

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n = \mathbf{x}^\top \mathbf{y}$$

로 정의한다. 따라서 벡터의 크기 (norm)는 $|\mathbf{x}| = \sqrt{\mathbf{x}^\top \mathbf{x}} = \sqrt{x_1^2 + \cdots + x_n^2}$ 일 때 Euclidean norm이라고 부른다.

The standard inner product in \mathbb{R}^n corresponds to (4.1) with $A = I_n$. $\langle \mathbf{x}, \mathbf{y} \rangle = 2 \sum_{k=1}^n x_k y_k$, $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^n k x_k y_k$, and $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^n \frac{1}{k} x_k y_k$ are examples of inner products in \mathbb{R}^n .

Example 4.2 Let us see that the standard inner product is indeed an inner product in \mathbb{R}^n by showing that the dot product satisfies the properties of an inner product, following elementary arithmetic steps. Given $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ and $c \in \mathbb{R}$,

- $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y} = x_1 y_1 + \cdots + x_n y_n = y_1 x_1 + \cdots + y_n x_n = \mathbf{y}^\top \mathbf{x} = \langle \mathbf{y}, \mathbf{x} \rangle$;
- $\langle c\mathbf{x}, \mathbf{y} \rangle = (c\mathbf{x})^\top \mathbf{y} = (cx_1)y_1 + \cdots + (cx_n)y_n = c(x_1 y_1 + \cdots + x_n y_n) = c\langle \mathbf{x}, \mathbf{y} \rangle$;
- $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = (\mathbf{x} + \mathbf{y})^\top \mathbf{z} = (x_1 + y_1)z_1 + \cdots + (x_n + y_n)z_n = (x_1 z_1 + \cdots + x_n z_n) + (y_1 z_1 + \cdots + y_n z_n) = \mathbf{x}^\top \mathbf{z} + \mathbf{y}^\top \mathbf{z} = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$;
- $|\mathbf{x}|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^\top \mathbf{x} = x_1^2 + \cdots + x_n^2 \geq 0$. If $|\mathbf{x}| = 0$, all $x_i = 0$ and $\mathbf{x} = \mathbf{0}$. It is straightforward to show the converse.

■

\mathbb{R}^n 의 $\mathbf{0}$ 가 아닌 두 벡터 $\mathbf{x} = (x_1, \dots, x_n)^\top$, $\mathbf{y} = (y_1, \dots, y_n)^\top$ 와 원점을 연결하는 두 방향이 수직이라는 것을 어떻게 알 수 있을까? \mathbf{x} 와 \mathbf{y} 가 평행하지 않아서² 원점 $\mathbf{0}$ 과 \mathbf{x} , \mathbf{y} 를 지나는 평면이 존재한다고 가정하자.

²두 벡터가 평행하면 사이각은 0 또는 180도이다.

이 세 벡터를 포함하는 평면이라 함은 $\mathbf{x} - \mathbf{0}$ 과 $\mathbf{y} - \mathbf{0}$ 에 의해서 생성되는 벡터들의 집합 중 최소 차원인 2차원 부분공간인 $P = \{a\mathbf{x} + b\mathbf{y} : a, b \in \mathbb{R}\}$ 를 생각하는 것이다.³ 2차원 평면 P 상의 세 벡터 $\mathbf{0}$ 과 \mathbf{x} , \mathbf{y} 로 이루어진 삼각형의 세변의 유클리드 길이는 $|\mathbf{x}|$, $|\mathbf{y}|$, $|\mathbf{x} - \mathbf{y}|$ 이다. 2차원 기하의 결과를 이용하면 이 세변의 길이 간에 Pythagorean relation인 $|\mathbf{x}|^2 + |\mathbf{y}|^2 = |\mathbf{x} - \mathbf{y}|^2$ 성립한다면 원점과 \mathbf{x} , \mathbf{y} 를 연결하는 방향이 서로 수직이라는 것을 알 수 있다. 즉

$$\begin{aligned} 0 &= |\mathbf{x}|^2 + |\mathbf{y}|^2 - |\mathbf{x} - \mathbf{y}|^2 \\ &= x_1^2 + \cdots + x_n^2 + y_1^2 + \cdots + y_n^2 - (x_1 - y_1)^2 - \cdots - (x_n - y_n)^2 \\ &= 2(x_1y_1 + \cdots + x_ny_n) \\ &= 2\mathbf{x}^\top \mathbf{y} \end{aligned}$$

이므로 $\mathbf{x}^\top \mathbf{y} = 0$ 이 유클리드 거리에 대해서 두 벡터 \mathbf{x} 와 \mathbf{y} 가 수직일 필요충분 조건이다. 90도가 아닌 일반적인 사잇각은 어떻게 찾을 수 있을까? 다시 한번 2차원 기하의 결과를 이용하면 사잇각이 θ 일 때 세 변의 유클리드 거리간에는 $|\mathbf{x} - \mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2|\mathbf{x}||\mathbf{y}|\cos\theta$ 의 관계가 있다. 수직인 경우의 전개를 이용하면 $-2\mathbf{x}^\top \mathbf{y} = -2|\mathbf{x}||\mathbf{y}|\cos\theta$ 으로 \mathbf{x} 와 \mathbf{y} 의 사잇각은

$$\cos\theta = \frac{\mathbf{x}^\top \mathbf{y}}{|\mathbf{x}||\mathbf{y}|} \quad (4.2)$$

이다. Data science에서는 (4.2)의 cosine 값을 cosine similarity라고 부르며 두 벡터의 크기는 무시하고 방향의 유사성을 측정하는데 많이 활용된다.

4.2 Orthogonal Vectors and Subspaces

앞에서 두 벡터의 사잇각을 표준내적과 표준내적으로 유도된 유클리드 거리로 표현하고 수직인 조건이 표준내적이 0이라는 것을 유도했다. 본 절에서는 표준내적처럼 구체적으로 형태가 정의되지 않은 일반적인 내적에 대해서 벡터공간의 유일한 기하학적 개념인 수직(orthogonality)을 도입하고, 이로부터 상상할 수 없을 만큼 다양한 결과들을 도출한다.

³원점을 지나는 1차원 부분공간은 \mathbf{x} 와 \mathbf{y} 를 동시에 포함할 수 없다는 것을 보이면 P 가 2차원이라는 것은 당연하므로 최소 차원이라는 것을 보일 수 있다.

Definition 4.4 유한차원 벡터공간 \mathbb{V} 에서

1. 임의의 두 벡터 \mathbf{v}_1 과 \mathbf{v}_2 가

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$$

이면, 수직(orthogonal)하다고 말하고 $\mathbf{v}_1 \perp \mathbf{v}_2$ 로 표시한다.

2. 벡터들 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 에 대해 서로 다른 두 벡터들이 모두 수직이면, 즉 $\mathbf{v}_i \perp \mathbf{v}_j$ for $i \neq j$ 이면, 이 벡터들은 서로 수직 (mutually orthogonal)이라고 말한다.
3. unit vector들 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 가 mutually orthogonal하면 이 벡터들이 orthonormal하다고 말한다.
4. 어떤 basis의 basic vector들이 orthonormal하면 이 기저를 orthonormal basis라고 부른다.

일반적인 벡터공간에서 내적이 0이라는 수직의 정의를 우리가 이미 아는 사실과 어떻게 연관시킬 수 있을까? 직각삼각형의 Pythagorean relation이 성립하는 두 벡터는 수직이다라는 평면기하의 직관을 확장할 수 있다.

Fact 4.3 Show that two vectors \mathbf{v}_1 and \mathbf{v}_2 are orthogonal if and only if the Pythagorean relation holds:

$$|\mathbf{v}_1|^2 + |\mathbf{v}_2|^2 = |\mathbf{v}_1 + \mathbf{v}_2|^2.$$

Proof: Since $|\mathbf{v}_1 + \mathbf{v}_2|^2 = \langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_1 \rangle + 2\langle \mathbf{v}_1, \mathbf{v}_2 \rangle + \langle \mathbf{v}_2, \mathbf{v}_2 \rangle$ for any two vectors \mathbf{v}_1 and \mathbf{v}_2 , we have

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

which implies the conclusion. ■

그리고 서로 수직인 벡터들은 선형독립이라는 다음 결과도 자연스럽다.

Fact 4.4 If non-zero vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are mutually orthogonal, then they are linearly independent.

Proof: Consider $c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n = \mathbf{0}$ for some real c_i 's. Since \mathbf{v}_i 's are mutually orthogonal,

$$\begin{aligned} \langle \mathbf{v}_i, \mathbf{0} \rangle &= \langle \mathbf{v}_i, c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n \rangle \\ &= c_1\langle \mathbf{v}_i, \mathbf{v}_1 \rangle + \dots + c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle + \dots + c_n\langle \mathbf{v}_i, \mathbf{v}_n \rangle \\ &= c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle. \end{aligned}$$

Since $\langle \mathbf{v}_i, \mathbf{v}_i \rangle > 0$ and $\langle \mathbf{v}_i, \mathbf{0} \rangle = 0$, $c_i = 0$. ■

Fact 4.5 Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be an orthonormal basis for a vector space \mathbb{V} . For any vector $\mathbf{v} \in \mathbb{V}$, show that

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \dots + \langle \mathbf{v}, \mathbf{v}_n \rangle \mathbf{v}_n = \sum_{i=1}^n \langle \mathbf{v}, \mathbf{v}_i \rangle \mathbf{v}_i$$

and the representation is unique.

Proof: Since the basis span the space, we have $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \sum_{i=1}^n x_i\mathbf{v}_i$ for some real x_i 's. From the orthogonality and normality,

$$\langle \mathbf{v}, \mathbf{v}_j \rangle = \left\langle \sum_{i=1}^n x_i \mathbf{v}_i, \mathbf{v}_j \right\rangle = \sum_{i=1}^n x_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle = x_j \langle \mathbf{v}_j, \mathbf{v}_j \rangle = x_j, \quad \text{for each } j = 1, \dots, n$$

and we get the desired representation. ■

Fact 4.5으로부터 orthonormal한 기저가 존재하면 임의의 벡터가 basic vector와의 내적으로 완벽하게 표현됨을 알 수 있다. 즉 내적값 $\langle \mathbf{v}, \mathbf{v}_i \rangle$ 가 orthonormal 기저에 대한 좌표값의 역할을 하는 것이다.

Example 4.3 For the Euclidean vector space \mathbb{R}^n , define the standard basic vectors

$$\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^\top \in \mathbb{R}^n, \quad i = 1, \dots, n$$

whose i -th component is 1, and all others are 0. The standard basis in \mathbb{R}^n is $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$.

1. The standard basis in \mathbb{R}^n is orthonormal since $|\mathbf{e}_i|^2 = \mathbf{e}_i^\top \mathbf{e}_i = 1$, $\mathbf{e}_i^\top \mathbf{e}_j = 0$ for $i \neq j$;
2. For $\mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$, $x_i = \langle \mathbf{x}, \mathbf{e}_i \rangle = \mathbf{x}^\top \mathbf{e}_i$ and $\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n$, which confirms Fact 4.5 for the standard basis and the standard inner product in \mathbb{R}^n since $\mathbf{x}^\top \mathbf{e}_i = x_1 \times 0 + \dots + x_{i-1} \times 0 + x_i \times 1 + x_{i+1} \times 0 + \dots + x_n \times 0 = x_i$, $\mathbf{x} = (x_1, 0, \dots, 0)^\top + \dots + (0, \dots, 0, x_n)^\top = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n$. ■

수학에서 두 원소 간에 성립하는 성질이 두 집합의 모든 원소 쌍 간에도 만족하는 경우, 그 성질이 두 집합 간에 성립한다고 말하는 경우가 많다. 벡터 간의 수직도 동일한 방식으로 부분공간 간의 수직으로 확장할 수 있다.

Definition 4.5 벡터공간 \mathbb{V} 의 두 부분공간 \mathbb{U} 와 \mathbb{W} 에 대해서

$$\mathbf{u} \perp \mathbf{w} \quad \text{for all } \mathbf{u} \in \mathbb{U}, \mathbf{w} \in \mathbb{W}$$

이면 \mathbb{U} 와 \mathbb{W} 가 서로 수직 (orthogonal)이라고 하고 $\mathbb{U} \perp \mathbb{W}$ 로 표시한다.

Definition 4.6 벡터공간 \mathbb{V} 의 부분공간 \mathbb{W} 에 대한 orthogonal complement는

$$\mathbb{W}^\perp = \{\mathbf{v} \in \mathbb{V} : \mathbf{v} \perp \mathbf{w} \text{ for all } \mathbf{w} \in \mathbb{W}\}$$

로 정의한다.

내적이 선형성을 가지고 있어 \mathbb{W}^\perp 도 벡터공간이며 \mathbb{V} 의 부분공간이다.

Fact 4.6 \mathbb{W}^\perp is also a subspace of \mathbb{V} if \mathbb{W} is a subspace of a vector space \mathbb{V} .

Proof: $\mathbb{W}^\perp \subset \mathbb{V}$ 는 자명하고 $\mathbf{0} \in \mathbb{W}^\perp$ 이다. $\alpha \in \mathbb{R}$ 이고 $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{W}^\perp$ 라면 임의의 $\mathbf{w} \in \mathbb{W}$ 에 대해서 $\langle \alpha\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w} \rangle = \alpha\langle \mathbf{v}_1, \mathbf{w} \rangle + \langle \mathbf{v}_2, \mathbf{w} \rangle = 0$ 이어서 $\alpha\mathbf{v}_1 + \mathbf{v}_2 \in \mathbb{W}^\perp$ 이다. 따라서 \mathbb{W}^\perp 는 \mathbb{V} 의 subspace이다. ■

부분공간들이 서로 수직이라도 orthogonal complement 관계는 아닐 수도 있음에 유의하자. 예를 들면, 선형공간 \mathbb{R}^3 와 표준내적 하에서 부분공간인 $\mathbb{U} = \text{span}\{(1, 0, 0)\}$ 와 $\mathbb{W} = \text{span}\{(0, 1, 0)\}$ 는 서로 수직인 부분공간이지만 orthogonal complement $\mathbb{U}^\perp = \text{span}\{(0, 1, 0), (0, 0, 1)\}$ 는 \mathbb{W} 를 진부분집합으로 갖는다. 즉 $\mathbb{U} \perp \mathbb{W}$ 이지만 $\mathbb{U}^\perp \supsetneq \mathbb{W}$ 이다. Section 4.6에서 $\text{Col}(A^\top)^\perp = \text{Null}(A)$ 와 $\text{Null}(A)^\perp = \text{Col}(A^\top)$ 등과 같은 orthogonal complement의 구체적인 예를 알 수 있다.

수직한 두 벡터로 생성되는 두개의 일차원 부분공간의 예로부터 일반적으로 서로 수직한 두 부분공간은 원점을 제외하고는 겹치지 않는다는 사실을 알 수 있다.

Fact 4.7 Let \mathbb{W} be a subspace of a vector space \mathbb{V} , and \mathbb{W}^\perp be the orthogonal complement of \mathbb{W} in \mathbb{V} . Then, $\mathbb{W} \cap \mathbb{W}^\perp = \{\mathbf{0}\}$.

Proof: Let $\mathbf{w} \in \mathbb{W} \cap \mathbb{W}^\perp$. Since $\mathbf{w} \in \mathbb{W}$ and $\mathbf{w} \in \mathbb{W}^\perp$, we conclude $\langle \mathbf{w}, \mathbf{w} \rangle = 0$ by the definition of the orthogonal complement. The definition of the inner product implies that $\mathbf{w} = \mathbf{0}$. ■

이렇게 수직하여 겹치지 않는 두 부분공간의 원소들의 합으로 임의의 벡터를 표현하는 방법은 유일하다.

Theorem 4.2 Let \mathbb{W} be a subspace of a vector space \mathbb{V} , and \mathbb{W}^\perp be the orthogonal complement of \mathbb{W} in \mathbb{V} . If $\mathbf{v} \in \mathbb{V}$ has a decomposition of $\mathbf{v} = \mathbf{w} + \mathbf{z}$ where $\mathbf{w} \in \mathbb{W}$ and $\mathbf{z} \in \mathbb{W}^\perp$, then the decomposition is unique.

Proof: Assume that we have two decompositions of $\mathbf{v} = \mathbf{w}_1 + \mathbf{z}_1 = \mathbf{w}_2 + \mathbf{z}_2$ where $\mathbf{w}_i \in \mathbb{W}$ and $\mathbf{z}_i \in \mathbb{W}^\perp$. By re-arranging the two representations, we get $\mathbb{W} \ni \mathbf{w}_1 - \mathbf{w}_2 = \mathbf{z}_2 - \mathbf{z}_1 \in \mathbb{W}^\perp$, that is, $\mathbf{w}_1 - \mathbf{w}_2 = \mathbf{z}_2 - \mathbf{z}_1 \in \mathbb{W} \cap \mathbb{W}^\perp = \{\mathbf{0}\}$. Therefore, $\mathbf{w}_1 = \mathbf{w}_2$ and $\mathbf{z}_1 = \mathbf{z}_2$, which implies the uniqueness of the decomposition. ■

Vector space의 벡터들은 subspace의 벡터와 orthogonal complement의 벡터의 합으로 항상 표현 가능한 것을 Section 4.5에서 증명한다. 이 결과를 Theorem 4.2의 유일성과 합치면 vector space는 늘 subspace와 orthogonal complement의 벡터들의 합으로 유일하게 표현된다는 것을 알 수 있다.

이후에 자주 부딪치는 문제가 어떤 벡터 \mathbf{v} 가 $\mathbb{W} = \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ 에 수직한 지 여부이다. 이 경우 \mathbb{W} 의 모든 벡터 \mathbf{w} 에 대해서 $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ 를 확인하지 않고 \mathbf{w}_j 에 대해서만 $\langle \mathbf{v}, \mathbf{w}_j \rangle = 0$ 를 확인해도 된다.

Lemma 4.3 Let \mathbb{V} be a vector space and $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ be a set of some vectors in \mathbb{V} . Show that, for any $\mathbf{v} \in \mathbb{V}$,

$$\mathbf{v} \perp \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\} \quad \text{if and only if} \quad \mathbf{v} \perp \mathbf{w}_j \text{ for } j = 1, \dots, k.$$

Proof: “only if” part is clear since all $\mathbf{w}_j \in \mathbb{W} = \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$. For “if” direction, assume $\mathbf{w} \in \mathbb{W}$. Then \mathbf{w} should have a representation $\mathbf{w} = x_1\mathbf{w}_1 + \dots + x_k\mathbf{w}_k$ for some $x_1, \dots, x_k \in \mathbb{R}$. So

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, x_1\mathbf{w}_1 + \dots + x_k\mathbf{w}_k \rangle = x_1\langle \mathbf{v}, \mathbf{w}_1 \rangle + \dots + x_k\langle \mathbf{v}, \mathbf{w}_k \rangle = 0$$

and hence $\mathbf{v} \perp \mathbb{W}$. ■

위 결과에서 벡터 $\mathbf{w}_1, \dots, \mathbf{w}_k$ 에 아무런 조건도 없는 것에 유의하자.

4.3 Orthogonal Projection

4.3.1 Projection onto the Direction of a Vector

벡터공간 \mathbb{V} 의 고정된 벡터 $\mathbf{w} \neq \mathbf{0}$ 를 생각하자. 벡터 \mathbf{w} 를 포함하는 직선 상의 여러 점 중에서 다른 벡터 \mathbf{v} 로부터 거리가 제일 가까운 점을 찾아보자. 거리는 내적으로 정의된 벡터 크기인 norm을 사용하여 측정한다. 우선, \mathbf{w} 방향 직선 상의 점(벡터)들을 λ 라는 parameter를 사용하여

$$\{\lambda\mathbf{w} : \lambda \in \mathbb{R}\}$$

로 매개화하자. 따라서 벡터 \mathbf{v} 에서 벡터 $\lambda\mathbf{w}$ 까지의 거리는 $|\lambda\mathbf{w} - \mathbf{v}|$ 로 계산한다. 이를 내적을 이용하여 전개하면 다음과 같이 λ 에 대한 2차식으로 표현된다. 마지막 표현에서 세 개의 내적은 상수이므로

$$\begin{aligned} |\lambda\mathbf{w} - \mathbf{v}|^2 &= \langle \lambda\mathbf{w} - \mathbf{v}, \lambda\mathbf{w} - \mathbf{v} \rangle \\ &= \lambda^2 \langle \mathbf{w}, \mathbf{w} \rangle - 2\lambda \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \end{aligned}$$

가 최소가 되는 λ 값을 찾으면

$$\lambda^* = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle}$$

이다. 따라서 \mathbf{v} 와의 거리가 최소가 되는 \mathbf{w} 직선 상의 벡터는

$$\lambda^*\mathbf{w} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w} = \left\langle \mathbf{v}, \frac{1}{|\mathbf{w}|} \mathbf{w} \right\rangle \frac{1}{|\mathbf{w}|} \mathbf{w}$$

로 주어진다.⁴ 유클리드 공간에서의 직관은 최소거리를 나타내는 벡터 $\mathbf{v} - \lambda^*\mathbf{w}$ 와 \mathbf{w} 가 수직 (orthogonal) 이기를 기대한다. 실제로, 내적이 정의된 임의의 벡터공간에서

$$\langle \mathbf{v} - \lambda^*\mathbf{w}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle - \lambda^* \langle \mathbf{w}, \mathbf{w} \rangle = 0$$

이 됨을 알 수 있다. 그래서, 이 $\lambda^*\mathbf{w}$ 를 \mathbf{v} 의 \mathbf{w} 로의 정사영(orthogonal projection)이라고 부른다.⁵ 내적이 정의된 벡터공간에서는 정사영된 벡터는 거리가 최소가 되는 점과 일치함을 알 수 있다. 이 정사영을

$$\mathbf{p}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w}, \quad (4.3)$$

또는, \mathbf{w} 가 unit vector, 즉 $\langle \mathbf{w}, \mathbf{w} \rangle = 1$ 이라면

$$\mathbf{p}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} \rangle \mathbf{w} \quad (4.4)$$

⁴ 이 표현에서 $\frac{1}{|\mathbf{w}|} \mathbf{w}$ 는 길이가 1인 벡터임에 유의하자. 본래, 벡터 \mathbf{w} 는 방향을 규정하는 역할이지 크기는 중요하지 않았기 때문에 $\frac{1}{|\mathbf{w}|} \mathbf{w}$ 와 같이 길이를 normalize한 후에 사용해도 무방하다. 따라서 처음부터 $|\mathbf{w}| = \langle \mathbf{w}, \mathbf{w} \rangle = 1$ 인 벡터를 사용하면 표현이 간단해지고 해석이 편리해지는 장점이 있다.

⁵ 이 책에서는 orthogonal projection만을 다루므로 projection은 orthogonal이 생략된 것으로 이해하면 된다.

로 표현할 수 있는데, $\langle \mathbf{v}, \mathbf{w} \rangle$ 이 \mathbf{v} 에 대해 선형이므로 $\mathbf{p}(\cdot)$ 도 선형사상/변환임을 알 수 있다. 따라서 적절한 기저 하에서 정사영 변환은 행렬로 표현될 수 있다. 특히 정사영된 벡터를 같은 방향으로 다시 정사영하면

$$\mathbf{p}(\mathbf{p}(\mathbf{v})) = \frac{\langle \mathbf{p}(\mathbf{v}), \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w} = \frac{\langle \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w} = \mathbf{p}(\mathbf{v})$$

가 되어 처음 정사영된 벡터가 더 이상 변하지 않는 것을 알 수 있다.

한 벡터 방향으로의 정사영에 대한 결과를 부분공간에 대한 정사영으로 다음과 같이 확장할 수 있다.

Definition 4.7 An orthogonal projection is a linear transform which maps any vector to another vector in a subspace \mathbb{W} such that the direction connecting two vectors is orthogonal to the subspace \mathbb{W} . We denote the projection by $\mathbf{P}_{\mathbb{W}}$. A matrix is called an orthogonal projection matrix if the matrix represents a projection $\mathbf{P}_{\mathbb{W}}$ in the Euclidean vector space.

물론 수직하지 않은 방향으로 사영하는 경우도 있으나 대부분의 경우에는 정사영을 의미하므로 orthogonal을 생략하고 projection이라고만 기술해도 정사영을 의미하는 것으로 한다. 만약 부분공간에 수직이 아닌 방향으로 projection하는 경우에는 방향을 명시하여 혼란이 없도록 할 것이다.

Fact 4.8 $\mathbb{V} = \mathbb{R}^n$ 인 유클리드 공간에서 벡터 \mathbf{w} 방향으로의 정사영을 표시하는 행렬이

$$P = \frac{1}{\mathbf{w}^\top \mathbf{w}} \mathbf{w} \mathbf{w}^\top \quad (4.5)$$

으로 표현되고, 이 행렬이 대칭행렬이면서 $P^2 = P$ 를 만족하는 것을 보이시오.

Proof: 정사영할 벡터를 $\mathbf{v} \in \mathbb{R}^n$ 이라 하자. (4.3)식을 \mathbb{R}^n 에서 정리하면

$$\mathbf{p}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w} = \frac{1}{\mathbf{w}^\top \mathbf{w}} (\mathbf{v}^\top \mathbf{w}) \mathbf{w} = \frac{1}{\mathbf{w}^\top \mathbf{w}} \mathbf{w} (\mathbf{w}^\top \mathbf{v}) = \frac{1}{\mathbf{w}^\top \mathbf{w}} (\mathbf{w} \mathbf{w}^\top) \mathbf{v}$$

인데 마지막 항이 행렬 $\frac{1}{\mathbf{w}^\top \mathbf{w}} \mathbf{w} \mathbf{w}^\top$ 와 벡터 \mathbf{v} 의 곱이므로 정사영을 나타내는 행렬은 다음과 같은 rank-one 행렬이 된다:

$$P = \frac{1}{\mathbf{w}^\top \mathbf{w}} \mathbf{w} \mathbf{w}^\top.$$

P 가 대칭행렬인 것은 자명하고,

$$P^2 = \frac{1}{(\mathbf{w}^\top \mathbf{w})^2} (\mathbf{w} \mathbf{w}^\top) (\mathbf{w} \mathbf{w}^\top) = \frac{1}{(\mathbf{w}^\top \mathbf{w})^2} \mathbf{w} (\mathbf{w}^\top \mathbf{w}) \mathbf{w}^\top = \frac{\mathbf{w}^\top \mathbf{w}}{(\mathbf{w}^\top \mathbf{w})^2} \mathbf{w} \mathbf{w}^\top = P.$$

■

Example 4.4 [Householder matrix] Reflection matrix의 구체적인 예로 $\mathbf{v} \in \mathbb{R}^n$ 방향으로의 정사영을 나타내는 $P = \frac{1}{\mathbf{v}^\top \mathbf{v}} \mathbf{v} \mathbf{v}^\top$ 를 이용하여 구성된 reflection matrix

$$H = I - 2P = I - 2 \frac{1}{\mathbf{v}^\top \mathbf{v}} \mathbf{v} \mathbf{v}^\top$$

를 **Householder matrix**라고 부른다.

임의의 벡터 \mathbf{u} 에 대해 $\mathbf{v} = \mathbf{u} + |\mathbf{u}| \mathbf{e}_1$ 로 놓으면 $\mathbf{v}^\top \mathbf{u} = \mathbf{u}^\top \mathbf{u} + |\mathbf{u}| u_1$ 이므로

$$\mathbf{v}^\top \mathbf{v} = \mathbf{u}^\top \mathbf{u} + 2u_1|\mathbf{u}| + |\mathbf{u}|^2 = 2(\mathbf{u}^\top \mathbf{u} + |\mathbf{u}| u_1) = 2\mathbf{v}^\top \mathbf{u}$$

이다. 따라서 방향 \mathbf{v} 에 대해 구성된 Householder matrix는 벡터 \mathbf{u} 를

$$H\mathbf{u} = \left(I - 2 \frac{1}{\mathbf{v}^\top \mathbf{v}} \mathbf{v} \mathbf{v}^\top \right) \mathbf{u} = \mathbf{u} - 2 \frac{1}{\mathbf{v}^\top \mathbf{v}} \mathbf{v} \mathbf{v}^\top \mathbf{u} = \mathbf{u} - \mathbf{v} = -|\mathbf{u}| \mathbf{e}_1 = (-|\mathbf{u}|, 0, \dots, 0)^\top.$$

로 이동시킨다. 즉 \mathbf{u} 를 크기는 유지하면서 음수 방향의 x_1 축으로 옮기는 reflection transformation이다. 이러한 Householder transformation은 numerical linear algebra에서 많이 활용된다. ■

4.3.2 Projection onto a Subspace Spanned by Orthonormal Vectors

유한차원 벡터공간 \mathbb{V} 의 부분공간 \mathbb{W} 의 orthonormal basis $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ 를 가정하자. 즉 $\langle \mathbf{w}_j, \mathbf{w}_j \rangle = 1$ 이고 $i \neq j$ 인 경우에 $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = 0$ 이다. \mathbb{W} 상의 vector \mathbf{w} 를 basic vector들의 linear combination으로

$$\mathbf{w} = x_1 \mathbf{w}_1 + \cdots + x_k \mathbf{w}_k = \sum_{i=1}^k x_i \mathbf{w}_i \in \mathbb{W}$$

로 표현하자. 임의의 벡터 $\mathbf{v} \in \mathbb{V}$ 에 대해 $\mathbf{v} - \mathbf{w}$ 방향이 부분공간 \mathbb{W} 에 수직인 \mathbf{w} 를 찾아보자. \mathbf{w} 를 표현하는 basic vector들에 대한 좌표(가중치) x_i 를 어떻게 정할 수 있을까? 벡터공간에 주어진 inner product가 중요한 역할을 한다. Lemma 4.3에 의해 $\mathbf{v} - \mathbf{w}$ 방향이 부분공간 \mathbb{W} 에 수직한 것은 $\mathbf{v} - \mathbf{w}$ 가 모든 \mathbf{w}_j 들에 수직하다는 것과 동치이다. 따라서 $\langle \mathbf{w}_j, \mathbf{v} - \mathbf{w} \rangle = 0$ 으로부터

$$\left\langle \mathbf{w}_j, \mathbf{v} - \sum_{i=1}^k x_i \mathbf{w}_i \right\rangle = \langle \mathbf{w}_j, \mathbf{v} \rangle - \sum_{i=1}^k x_i \langle \mathbf{w}_j, \mathbf{w}_i \rangle = \langle \mathbf{w}_j, \mathbf{v} \rangle - x_j = 0$$

이어서 $x_j = \langle \mathbf{w}_j, \mathbf{v} \rangle$ 임을 알 수 있다. 즉

$$\mathbf{w}^* = \langle \mathbf{v}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \cdots + \langle \mathbf{v}, \mathbf{w}_k \rangle \mathbf{w}_k = \sum_{i=1}^k \langle \mathbf{v}, \mathbf{w}_i \rangle \mathbf{w}_i$$

로 놓으면 $\mathbf{v} - \mathbf{w}^*$ 가 부분공간 \mathbb{W} 의 모든 basic vector에 수직하고, 따라서 $\mathbf{v} - \mathbf{w}^*$ 가 \mathbb{W} 에 수직하다.

\mathbf{v} 로부터 \mathbf{w}^* 를 뺀다. 다음과 같이

$$\mathbf{v} = \mathbf{w}^* + (\mathbf{v} - \mathbf{w}^*)$$

로 분해되는데 $\mathbf{w}^* \in \mathbb{W}$ 이면서 $\mathbf{v} - \mathbf{w}^* \in \mathbb{W}^\perp$ 이므로 Theorem 4.2에 의해 이러한 분해가 유일하다. 즉 주어진 벡터 \mathbf{v} 로부터 수직 방향으로 만나는 \mathbb{W} 상의 벡터가 유일하게 결정되는데 이 벡터를 \mathbf{v} 의 부분공간 \mathbb{W} 상으로의 정사영(projection)이라고 부르고 $\mathbf{P}_{\mathbb{W}}(\mathbf{v})$ 로 표시한다. 따라서 다음과 같이 표시할 수 있다.

$$\mathbf{P}_{\mathbb{W}}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \cdots + \langle \mathbf{v}, \mathbf{w}_k \rangle \mathbf{w}_k = \sum_{i=1}^k \langle \mathbf{v}, \mathbf{w}_i \rangle \mathbf{w}_i \quad \text{where } \mathbb{W} = \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\}. \quad (4.6)$$

또한 정사영은 $\mathbf{P}_W(\mathbf{v}) \in W^\circ$ 이므로 $\mathbf{P}_W(\mathbf{P}_W(\mathbf{v})) = \mathbf{P}_W(\mathbf{v})$, 즉 $\mathbf{P}_W \circ \mathbf{P}_W = \mathbf{P}_W$ 이다. 정사영을 구성하는 $\langle \mathbf{v}, \mathbf{w}_i \rangle \mathbf{w}_i$ 는 \mathbf{v} 에 대해서 선형사상이므로 정사영 $\mathbf{P}_W(\mathbf{v})$ 도 \mathbf{v} 에 대해서 선형사상이다. 따라서 적절한 기저에 대해서 행렬로 표현될 수 있다. $k = 1$ 인 경우에는 (4.6)와 (4.4)가 일치한다.

Fact 4.9 유클리드 공간 $V = \mathbb{R}^n$ 에서 orthonormal한 벡터들 $\{\mathbf{w}_1, \dots, \mathbf{w}_k\} \subset \mathbb{R}^n$ 에 의해 span된 부분공간 W 상으로의 정사영을 \mathbf{P}_W 로 표시하자. \mathbf{w}_i 벡터를 열로하는 $n \times k$ 행렬을 Q 라 하자. Q 는 $Q^\top Q = I_k$ 를 만족한다. 이 경우, 선형사상인 정사영 \mathbf{P}_W 를 표준기저에 대해 표현하는 행렬 P 는

$$P = QQ^\top \quad (4.7)$$

임을 보이시오. 이 행렬은 대칭행렬이면서 $P^2 = P$ 를 만족한다.

Proof: 정사영할 벡터를 $\mathbf{v} \in \mathbb{R}^n$ 이라 하자. 표준내적 $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y}$ 를 사용하면서 행렬곱 형식으로 표현하면

$$\begin{aligned} \mathbf{P}_W(\mathbf{v}) &= \mathbf{w}_1 \langle \mathbf{w}_1, \mathbf{v} \rangle + \cdots + \mathbf{w}_k \langle \mathbf{w}_k, \mathbf{v} \rangle \\ &= \mathbf{w}_1 \mathbf{w}_1^\top \mathbf{v} + \cdots + \mathbf{w}_k \mathbf{w}_k^\top \mathbf{v} \\ &= (\mathbf{w}_1 \mathbf{w}_1^\top + \cdots + \mathbf{w}_k \mathbf{w}_k^\top) \mathbf{v} \\ &= QQ^\top \mathbf{v} \quad \text{by Lemma 3.5} \end{aligned}$$

를 얻는다. ■

4.3.3 Projection onto a Subspace Spanned by Independent Vectors

유한차원 벡터공간 V 의 부분공간 W 가 linearly independent인 벡터들 $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ 로 span된다라고 가정하자. 이 벡터들은 orthonormal하지 않을 수도 있다는 점에 유의하자. 임의의 벡터 $\mathbf{v} \in V$ 의 부분공간 W 상으로의 정사영은 $(\mathbf{v} - \mathbf{w}) \perp W$ 인 벡터 $\mathbf{w} \in W$ 이다. $W = \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ 이므로 Lemma 4.3에 의해

$$(\mathbf{v} - \mathbf{w}) \perp W \Leftrightarrow (\mathbf{v} - \mathbf{w}) \perp \mathbf{w}_i \quad \text{for all } i = 1, \dots, k$$

이고, 정사영 \mathbf{w} 는 적절한 $\mathbf{x} = (x_1, \dots, x_k)^\top \in \mathbb{R}^k$ 에 대해 $\mathbf{w} = x_1 \mathbf{w}_1 + \cdots + x_k \mathbf{w}_k$ 로 표현된다. 따라서

$$\begin{aligned} 0 &= \langle \mathbf{w}_i, \mathbf{w} - \mathbf{v} \rangle \\ &= \langle \mathbf{w}_i, x_1 \mathbf{w}_1 + \cdots + x_k \mathbf{w}_k - \mathbf{v} \rangle \\ &= \langle \mathbf{w}_i, x_1 \mathbf{w}_1 + \cdots + x_k \mathbf{w}_k \rangle - \langle \mathbf{w}_i, \mathbf{v} \rangle \\ &= x_1 \langle \mathbf{w}_i, \mathbf{w}_1 \rangle + \cdots + x_k \langle \mathbf{w}_i, \mathbf{w}_k \rangle - \langle \mathbf{w}_i, \mathbf{v} \rangle \end{aligned}$$

이므로 i, j 원소가 $b_{ij} = \langle \mathbf{w}_i, \mathbf{w}_j \rangle$ 인 $k \times k$ symmetric matrix $B = (b_{ij})$ 와 i 번째 원소가 $c_i = \langle \mathbf{w}_i, \mathbf{v} \rangle$ 인 벡터 $\mathbf{c} = (c_i)$ 에 대해서

$$B\mathbf{x} = \mathbf{c}$$

를 만족하는 \mathbf{x}^* 가 정사영을 결정한다. 특히 Lemma 4.1 때문에 B 는 가역행렬이므로

$$\mathbf{x}^* = B^{-1} \mathbf{c}$$

로 구할 수 있고, 정사영은

$$\mathbf{P}_{\mathbb{W}}(\mathbf{v}) = x_1^* \mathbf{w}_1 + \cdots + x_k^* \mathbf{w}_k \quad (4.8)$$

와 같이 간접적인 방식(implicit form)으로 표현할 수 있다.

만약 $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ 가 orthonormal하다면 linearly independent하므로 본 절의 방법으로 정사영을 구할 수 있다. 이 경우, $B = I_k$ 이고 $\mathbf{x}^* = \mathbf{c}$ 이므로 orthonormal한 경우에 유도한 정사영 식 (4.6)과 일치한다.

다음 연습문제에서는 유클리드 벡터공간 \mathbb{R}^n 에서 얻어지는 좀 더 구체적인 표현에 대해서 알아본다.

Fact 4.10 유클리드 공간 $\mathbb{V} = \mathbb{R}^n$ 에서 linearly independent한 벡터들 $\{\mathbf{w}_1, \dots, \mathbf{w}_k\} \subset \mathbb{R}^n$ 에 의해 span된 부분공간 \mathbb{W} 상으로의 정사영을 $\mathbf{P}_{\mathbb{W}}$ 로 표시하고 $\mathbf{P}_{\mathbb{W}}$ 를 표준기저에 대해서 표현하는 행렬을 P 라 하자. \mathbf{w}_i 벡터를 열로 하는 $n \times k$ 행렬을 A 라 하면

$$P = A(A^\top A)^{-1}A^\top \quad (4.9)$$

이다. 이 행렬은 대칭행렬이면서 $P^2 = P$ 를 만족한다. 또한, (4.5)와 (4.7)이 (4.9)의 특별한 경우이다.

Proof: $b_{ij} = \langle \mathbf{w}_i, \mathbf{w}_j \rangle = \mathbf{w}_i^\top \mathbf{w}_j$ 이므로 $B = (b_{ij}) = A^\top A$ 이고 B 는 가역행렬이다. 정사영하고자 하는 벡터 $\mathbf{v} \in \mathbb{R}^n$ 에 대해 $\mathbf{c} = A^\top \mathbf{v}$ 이므로 $\mathbf{x}^* = (A^\top A)^{-1}A^\top \mathbf{v}$ 가 되고 정사영은

$$\mathbf{P}_{\mathbb{W}}(\mathbf{v}) = A\mathbf{x}^* = A(A^\top A)^{-1}A^\top \mathbf{v}$$

로 표현되어 구하고자 하는 정사영 행렬은

$$P = A(A^\top A)^{-1}A^\top$$

로 표현된다. $A = \mathbf{e}_i$ 와 $B = I_k$ 로 놓으면 (4.5)와 (4.7)이 (4.9)의 특별한 경우라는 것을 알 수 있다. ■

Example 4.5 Let S be a 2-dimensional surface in \mathbb{R}^3 . Consider a point on S , $\mathbf{p} = (1, 2, -1)^\top \in S$. When two tangential vectors of S at \mathbf{p} are $(-1, 0, 1)^\top$ and $(1, 1, 0)^\top$, let us find the projection point of $(2, 3, 0)$ onto the tangential plane of S at \mathbf{p} . Assuming \mathbf{p} as the origin, the tangential subspace (plane) is spanned

by two tangential vectors. Set the 3×2 matrix A consisting of two tangential vectors: $A = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$

and the projected point shifted to the new origin is $\mathbf{b} = (2, 3, 0)^\top - (1, 2, -1)^\top = (1, 1, 1)^\top$. Hence, the projection matrix is

$$\begin{aligned} P &= A(A^\top A)^{-1}A^\top \\ &= \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \left(\begin{bmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\
&= \frac{1}{3} \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} \\
&= \frac{1}{3} \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}
\end{aligned}$$

and the projection point $P\mathbf{b} = (\frac{2}{3}, \frac{4}{3}, \frac{2}{3})^\top$. ■

4.4 How to Build an Orthonormal Basis: Gram-Schmidt Procedure

앞 절에서 orthonormal 벡터들이 span 하는 부분공간 상으로의 정사영을 계산하는 방법을 살펴보았다. 즉 orthonormal basis의 basic vector들 각각에 정사영한 벡터를 합치면 해당 기저로 생성되는 부분공간에 대한 정사영이 된다. 이 절에서는 이 정사영 계산법을 활용하여 orthonormal 벡터를 늘려가는 과정을 살펴본다.

$\mathcal{B}_k = \{\mathbf{w}_1, \dots, \mathbf{w}_k\} \subset \mathbb{W}$ 가 orthonormal하다고 가정하자. 만약 $\mathbb{W}_k = \text{span } \mathcal{B}_k$ 가 \mathbb{W} 전체가 아니라면, 즉 $\mathbb{W}_k \subsetneq \mathbb{W}$ 이라면 $\mathbf{w} \in \mathbb{W} \setminus \mathbb{W}_k$ 인 벡터가 적어도 한 개 이상 존재한다. 이 \mathbf{w} 의 \mathbb{W}_k 상의 정사영을 기저 \mathcal{B}_k 에 대해 표현하면 (4.6)에 의해

$$\mathbf{P}_{\mathbb{W}_k}(\mathbf{w}) = \sum_{i=1}^k \langle \mathbf{w}, \mathbf{w}_i \rangle \mathbf{w}_i \in \mathbb{W}_k$$

이고 $\mathbf{w} - \mathbf{P}_{\mathbb{W}_k}(\mathbf{w}) \in \mathbb{W}_k^\perp$ 이다. 그런데 $\mathbf{w} \notin \mathbb{W}_k$ 므로 $\mathbf{w} - \mathbf{P}_{\mathbb{W}_k}(\mathbf{w}) \neq \mathbf{0}$ 이다.

$$\mathbf{w}_{k+1} = \frac{1}{|\mathbf{w} - \mathbf{P}_{\mathbb{W}_k}(\mathbf{w})|} (\mathbf{w} - \mathbf{P}_{\mathbb{W}_k}(\mathbf{w}))$$

로 놓으면 $|\mathbf{w}_{k+1}| = 1$ 이다. 그리고 $\mathbf{w}_{k+1} \perp \mathcal{B}_k$ 이므로 $\mathcal{B}_{k+1} = \mathcal{B}_k \cup \{\mathbf{w}_{k+1}\}$ 은 orthonormal하게 된다. \mathcal{B}_{k+1} 이 span 하는 \mathbb{W}_{k+1} 은 \mathbb{W}_k 보다 dimension이 1이 큰 \mathbb{W} 의 부분공간이다. \mathbb{W} 가 유한차원 공간이므로 이 과정을 반복하면 \mathbb{W} 의 orthonormal한 기저를 찾을 수 있다. 이렇게 orthonormal한 벡터 집합을 늘려가는 과정을 **Gram-Schmidt procedure**라고 부른다. \mathbb{W} 에 영벡터가 아닌 벡터가 하나라도 있으면 이 벡터를 normalize하여 첫 번째 벡터 \mathbf{w}_1 으로 놓으면 Gram-Schmidt procedure를 시작할 수 있다. 따라서 다음과 같은 결론을 얻을 수 있다.

Fact 4.11 Let \mathbb{W} be a finite-dimensional subspace and \mathcal{B} be a set of orthonormal vectors in \mathbb{W} . Then, there exists an orthonormal basis of \mathbb{W} containing \mathcal{B} , and we can construct it explicitly.

4.4.1 Gram-Schmidt Procedure for Given Linearly Independent Vectors

k 개의 linearly independent한 벡터 $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ 가 주어졌다고 하자. 이 벡터들에 의해 span되는 공간을 $\mathbb{W} = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ 라 하면 이 공간의 차원은 $\dim \mathbb{W} = k$ 이다. 이전 절에서 유도한 Gram-Schmidt procedure와 동일한 다음과 같은 방법으로 특별한 성질(Fact 4.12)을 갖는 기저를 찾을 수 있다.

Set $\mathbf{w}_1 = \frac{1}{|\mathbf{a}_1|} \mathbf{a}_1$, $i = 1$;

1. A set of i orthonormal vectors, $\mathcal{B}_i = \{\mathbf{w}_1, \dots, \mathbf{w}_i\}$ with $i < k$,

$$\begin{aligned} 2. \text{ Compute } \mathbf{v}_{i+1} &= \mathbf{a}_{i+1} - \mathbf{P}_{\text{span } \mathcal{B}_i}(\mathbf{a}_{i+1}) \\ &= \mathbf{a}_{i+1} - (\langle \mathbf{a}_{i+1}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \dots + \langle \mathbf{a}_{i+1}, \mathbf{w}_i \rangle \mathbf{w}_i) \\ \mathbf{w}_{i+1} &= \frac{1}{|\mathbf{v}_{i+1}|} \mathbf{v}_{i+1} \end{aligned}$$

3. Update $\mathcal{B}_{i+1} = \mathcal{B}_i \cup \{\mathbf{w}_{i+1}\}$.

4. Set $i \leftarrow i + 1$. Repeat while $i < k$.

이 procedure에서 생성되는 \mathbf{w}_i 들은 다음과 같은 성질을 갖는다.

Fact 4.12 Let $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ be independent vectors. The above Gram-Schmidt procedure produces an orthonormal basis $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ of the subspace $\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ such that

$$\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_j\} = \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_j\} \quad \text{for } j = 1, \dots, k.$$

Furthermore,

$$\langle \mathbf{a}_j, \mathbf{w}_j \rangle \neq 0 \tag{4.10}$$

also holds for $j = 1, \dots, k$.

Proof: k 에 대해 수학적 귀납법을 적용하자. $k = 1$ 인 경우는 자명하다. $k \leq j$ 까지 성립한다고 가정하고 $k = j + 1$ 에 대해 증명하자. 그러면 수학적 귀납법 가정에 의해

$$\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_j\} = \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_j\} \tag{4.11}$$

이다.

- 2번 식에서 $|\mathbf{v}_{j+1}| = 0$ 이면 $\mathbf{v}_{j+1} = \mathbf{0}$ 이다. 그러면 $\mathbf{a}_{j+1} \in \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_j\}$ 인데 수학적 귀납법 가정(4.11)에 의해 $\mathbf{a}_{j+1} \in \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_j\}$ 이므로 \mathbf{a}_j 들의 독립성에 모순이어서 $|\mathbf{v}_{j+1}| \neq 0$ 이다.

- 2번 식에 $\mathbf{v}_{j+1} = |\mathbf{v}_{j+1}| \mathbf{w}_{j+1}$ 를 대입하고 다시 쓰면

$$\mathbf{a}_{j+1} = \langle \mathbf{a}_{j+1}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \cdots + \langle \mathbf{a}_{j+1}, \mathbf{w}_j \rangle \mathbf{w}_j + |\mathbf{v}_{j+1}| \mathbf{w}_{j+1}$$

이므로 $\mathbf{a}_{j+1} \in \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_{j+1}\}$ 이다.

- 2번 식에 $\mathbf{v}_{j+1} = |\mathbf{v}_{j+1}| \mathbf{w}_{j+1}$ 를 대입하고 $|\mathbf{v}_{j+1}|$ 로 양변을 나누면

$$\mathbf{w}_{j+1} = \frac{1}{|\mathbf{v}_{j+1}|} \mathbf{a}_{j+1} - \frac{1}{|\mathbf{v}_{j+1}|} (\langle \mathbf{a}_{j+1}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \cdots + \langle \mathbf{a}_{j+1}, \mathbf{w}_j \rangle \mathbf{w}_j)$$

이어서 \mathbf{w}_{j+1} 가 \mathbf{a}_{j+1} 와 $\mathbf{w}_1, \dots, \mathbf{w}_j$ 의 선형결합으로 표현된다. (4.11)에 의해서 $\mathbf{w}_{j+1} \in \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{j+1}\}$ 이다.

따라서 $\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{j+1}\} = \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_{j+1}\}$ 이다.

또한 2번 식의 양변에 \mathbf{w}_{j+1} 를 내적하면 $\mathbf{w}_{j+1} \perp \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_j\}$ 이므로 $\langle \mathbf{a}_{j+1}, \mathbf{w}_{j+1} \rangle = \langle \mathbf{v}_{j+1}, \mathbf{w}_{j+1} \rangle$ 이고 $\langle \mathbf{v}_{j+1}, \mathbf{w}_{j+1} \rangle = |\mathbf{v}_{j+1}|$ 이어서 $\langle \mathbf{a}_{j+1}, \mathbf{w}_{j+1} \rangle = |\mathbf{v}_{j+1}| > 0$ 이다. ■

4.4.2 Projection as a Distance-Minimization

\mathbb{V} 의 k 차원 부분공간으로 orthonormal한 기저 $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ 에 의해 span 되는 \mathbb{W} 를 생각하자. 이 기저를 포함하는 벡터공간 \mathbb{V} 의 기저를 (예를 들면, Fact 4.11의 Gram-Schmidt orthogonalization 등을 활용하여) 찾을 수 있다. 전체 공간의 orthonormal 기저가 $\{\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ 로 완성되었다고 하자. 그러면, \mathbb{W} 상으로 정사영하고자 하는 벡터를 \mathbf{u} 라 할 때, Fact 4.5에 의해 이 기저에 대한 벡터 \mathbf{u} 의 표현(좌표)은 다음과 같다:

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \cdots + \langle \mathbf{u}, \mathbf{w}_k \rangle \mathbf{w}_k + \langle \mathbf{u}, \mathbf{v}_{k+1} \rangle \mathbf{v}_{k+1} + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n.$$

부분공간 \mathbb{W} 의 임의의 벡터 \mathbf{w} 를 $\mathbf{w} = x_1 \mathbf{w}_1 + \cdots + x_k \mathbf{w}_k$ 로 표현하면,

$$\mathbf{u} - \mathbf{w} = (\langle \mathbf{u}, \mathbf{w}_1 \rangle - x_1) \mathbf{w}_1 + \cdots + (\langle \mathbf{u}, \mathbf{w}_k \rangle - x_k) \mathbf{w}_k + \langle \mathbf{u}, \mathbf{v}_{k+1} \rangle \mathbf{v}_{k+1} + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n$$

이므로

$$\begin{aligned} |\mathbf{u} - \mathbf{w}|^2 &= (\langle \mathbf{u}, \mathbf{w}_1 \rangle - x_1)^2 + \cdots + (\langle \mathbf{u}, \mathbf{w}_k \rangle - x_k)^2 + \langle \mathbf{u}, \mathbf{v}_{k+1} \rangle^2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle^2 \\ &\geq \langle \mathbf{u}, \mathbf{v}_{k+1} \rangle^2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle^2 \end{aligned}$$

이고 최솟값은 모든 i 에 대하여 $x_i = \langle \mathbf{u}, \mathbf{w}_i \rangle$ 일 때 얻어진다. 즉 $|\mathbf{u} - \mathbf{w}|$ 를 최소로 하는 부분공간 \mathbb{W} 의 벡터 \mathbf{w} 는

$$\mathbf{w} = \langle \mathbf{u}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \cdots + \langle \mathbf{u}, \mathbf{w}_k \rangle \mathbf{w}_k$$

로 주어지는데 (4.6)로부터 이 최소거리 벡터가 정사영 $\mathbf{P}_{\mathbb{W}}(\mathbf{u})$ 과 동일함을 알 수 있다.

4.5 Decomposition into Orthogonal Complements

모든 유한차원 벡터공간은 임의의 부분공간과 그 부분공간의 orthogonal complement의 합으로 표현된다. 따라서 임의의 벡터는 부분공간에 속하는 벡터와 orthogonal complement에 속하는 벡터의 합으로 표현된다. 그리고 이러한 분해는 늘 분해된 성분 벡터들이 유일하게 결정되어 더욱 유용하다.

Fact 4.13 Let \mathbb{V} be a finite-dimensional vector space and \mathbb{W} be a subspace of \mathbb{V} . For any $\mathbf{v} \in \mathbb{V}$, there exist unique $\mathbf{w} \in \mathbb{W}$ and $\mathbf{z} \in \mathbb{W}^\perp$ such that $\boxed{\mathbf{v} = \mathbf{w} + \mathbf{z}}$.

Proof: By Fact 4.11, we can obtain an orthonormal basis of \mathbb{W} . Then, through (4.6) we get the projection $\mathbf{w} := \mathbf{P}_{\mathbb{W}}(\mathbf{v})$ of \mathbf{v} onto \mathbb{W} , and $\mathbf{z} := \mathbf{v} - \mathbf{w} \in \mathbb{W}^\perp$. By Theorem 4.2, this representation is unique. ■

유한차원 벡터공간의 부분공간은 orthogonal complement를 두번 취하면 원래의 부분공간이 된다⁶는 다음의 결과가 유용한 경우가 종종 있다.

Fact 4.14 Let \mathbb{V} be a finite-dimensional vector space and \mathbb{W} be a subspace of \mathbb{V} . Then, $(\mathbb{W}^\perp)^\perp = \mathbb{W}$.

Proof: Let $\mathbf{w} \in \mathbb{W}$. Then, for any $\mathbf{z} \in \mathbb{W}^\perp$, $\langle \mathbf{w}, \mathbf{z} \rangle = 0$ by the definition of orthogonal complement. Therefore, $\mathbf{w} \in (\mathbb{W}^\perp)^\perp$, that is, $\mathbb{W} \subset (\mathbb{W}^\perp)^\perp$.

Conversely, assume $\mathbf{v} \in (\mathbb{W}^\perp)^\perp$. By Fact 4.13, $\mathbf{v} = \mathbf{w} + \mathbf{z}$ for some $\mathbf{w} \in \mathbb{W}$ and $\mathbf{z} \in \mathbb{W}^\perp$. Note that $\langle \mathbf{w}, \mathbf{z} \rangle = 0$. Since $\mathbf{v} \in (\mathbb{W}^\perp)^\perp$, we also have $\mathbf{v} \perp \mathbf{z}$. Then, $0 = \langle \mathbf{v}, \mathbf{z} \rangle = \langle \mathbf{w} + \mathbf{z}, \mathbf{z} \rangle = \langle \mathbf{w}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{z} \rangle = |\mathbf{z}|^2$. Hence, $\mathbf{z} = \mathbf{0}$ and $\mathbf{v} = \mathbf{w} \in \mathbb{W}$, which implies $(\mathbb{W}^\perp)^\perp \subset \mathbb{W}$. ■

4.6 Orthogonality in Euclidean Spaces \mathbb{R}^n

이 절에서는 모든 벡터들이 유한차원 유클리드 공간의 원소이고, 두 벡터 $\mathbf{x} = (x_1, \dots, x_n)^\top$ 과 $\mathbf{y} = (y_1, \dots, y_n)^\top$ 의 inner product는 유클리드 공간의 표준내적인

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n = \mathbf{x}^\top \mathbf{y}$$

의 경우에 대해서 생각한다.

먼저 행렬에 관한 4개의 기본공간 간의 수직관계에 대해 알아보자. 행공간과 영공간의 정의에서 다음 사실은 쉽게 유추할 수 있다.

Fact 4.15 Let A be an $m \times n$ matrix. The row space of A is orthogonal to the nullspace (in \mathbb{R}^n).

The column space of A is orthogonal to the left nullspace (in \mathbb{R}^m).

Proof: Let $\mathbf{y} \in \text{Null}(A^\top)$ and $\mathbf{b} \in \text{Col}(A)$. Then, there exists an $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{b} = A\mathbf{x}$. Since $\mathbf{y}^\top \mathbf{A} = \mathbf{0}$, $\mathbf{y}^\top \mathbf{b} = \mathbf{y}^\top (A\mathbf{x}) = (\mathbf{y}^\top \mathbf{A})\mathbf{x} = \mathbf{0}^\top \mathbf{x} = 0$. Therefore, $\mathbf{y} \perp \mathbf{b}$, that is, $\text{Null}(A^\top) \perp \text{Col}(A)$.

⁶무한차원 벡터공간에서는 성립하지 않을 수도 있는 성질이다.

Considering A^\top instead of A , we can see the orthogonality between the nullspace of A and the row space of A . ■

부분공간들이 서로 수직이라도 orthogonal complement 관계는 아닐 수도 있으므로 위의 Fact 4.15으로부터 $\text{Col}(A^\top) \perp \text{Null}(A)$ 는 알 수 있지만, $\text{Null}(A) = \text{Col}(A^\top)^\perp$ 또는 $\text{Null}(A)^\perp = \text{Col}(A^\top)$ 여부는 아직 알 수 없다.

4.6.1 Orthogonal Complements of Fundamental Subspaces

$m \times n$ 행렬 A 의 부분공간과 orthogonal complements인 부분공간 간의 관계에 대해서 살펴보자.

- $\boxed{\text{Null}(A) = \text{Col}(A^\top)^\perp}$

먼저 포함관계 $\text{Null}(A) \subset \text{Col}(A^\top)^\perp$ 를 살펴보자. $\mathbf{x} \in \text{Null}(A)$, 즉 $A\mathbf{x} = \mathbf{0}$ 라 하자. $A\mathbf{x}$ 의 i -번째 원소는 행렬 A 의 i -번째 행과 벡터 \mathbf{x} 간의 내적으로 이루어지므로 \mathbf{x} 는 A 의 모든 행벡터와 수직하고, 따라서 행공간의 모든 벡터들과 수직하다, 즉 $\text{Null}(A) \subset \text{Col}(A^\top)^\perp$ 이다.

반대로 $\mathbf{x} \in \text{Col}(A^\top)^\perp$ 를 가정하자. A 를 구성하는 모든 행들이 $\text{Col}(A^\top)$ 에 포함되므로 \mathbf{x} 는 A 의 모든 행들과 내적이 0이다, 즉 $A\mathbf{x} = \mathbf{0}$ 이어서 $\text{Col}(A^\top)^\perp \subset \text{Null}(A)$ 임을 알 수 있다.

- $\boxed{\text{Null}(A)^\perp = \text{Col}(A^\top)}$

Fact 4.14를 활용하면 $\text{Null}(A) = \text{Col}(A^\top)^\perp$ 와 $(\text{Col}(A^\top)^\perp)^\perp = \text{Col}(A^\top)$ 에서 $\text{Null}(A)^\perp = \text{Col}(A^\top)$ 를 얻는다.

이상의 결과에서 “The nullspace contains everything orthogonal to the row space” 이면서 동시에 ”The row space contains everything orthogonal to the nullspace” 임을 알 수 있다. A 와 A^\top 의 역할을 바꾸면 ”The left nullspace is the orthogonal complement of the column space”, 즉

Lemma 4.4 For any matrix A ,

$$\text{Null}(A^\top) = \text{Col}(A)^\perp \quad \text{and} \quad \text{Null}(A^\top)^\perp = \text{Col}(A).$$

을 얻는다. 이 결과는 선형대수학의 기본정리(Fundamental Theorem of Linear Algebra)의 일부이다.

Fact 4.9에서 직교기저로 생성되는 부분공간으로의 정사영하는 행렬을 살펴보았다. (4.7)의 정사영 행렬 P 는 대칭이고 $P^2 = P$ 를 만족하는데 Lemma 4.4에 의해 이 두 조건이 정사영 행렬을 정의함을 알 수 있다.

Fact 4.16 P is a matrix representing the orthogonal projection onto a subspace in the Euclidean vector space \mathbb{R}^n if and only if $P^2 = P$ and P are symmetric.

Solution: Let an operator project vectors orthogonally into a subspace \mathbb{W} . Let A be a matrix whose columns are the linearly independent vectors spanning \mathbb{W} . Then, by Fact 4.10, the matrix representation P of the projection onto $\text{Col}(A)$ is given as

$$P = A(A^\top A)^{-1}A^\top$$

which is symmetric and satisfies $P^2 = P$.

Conversely, assume that P is a symmetric matrix satisfying $P^2 = P$. Denote $\mathbb{W} = \text{Col}(P)$. Then, for an arbitrary vector \mathbf{v} ,

$$P^\top(\mathbf{v} - P\mathbf{v}) = P^\top(I - P)\mathbf{v} = P(I - P)\mathbf{v} = (P - P^2)\mathbf{v} = \mathbf{0}.$$

So, $(\mathbf{v} - P\mathbf{v}) \in \text{Null}(P^\top) = \text{Col}(P)^\perp = \mathbb{W}^\perp$ holds. That is, $(\mathbf{v} - P\mathbf{v}) \perp \mathbb{W}$ and $P\mathbf{v} \in \mathbb{W}$, which means that the operator corresponding to P is a projection onto the subspace $\mathbb{W} = \text{Col}(P)$. ■

4.7 Orthogonal Matrices

$m \times n$ 행렬 Q 의 열벡터들이 orthonormal하다고 가정하자. Orthonormal한 벡터들은 linearly independent 하므로 $m \geq n$ 이어야 한다. 그러면, $n \times n$ 행렬 $Q^\top Q$ 의 i, j 원소는 Q 의 열벡터 \mathbf{q}_i 와 \mathbf{q}_j 간의 내적이므로 대각 원소($i = j$)는 1이고 나머지($i \neq j$)는 모두 0이다. 즉

$$Q^\top Q = I_n$$

이어서 Q 의 left-inverse가 존재한다. 만약 $m = n$ 이라면 (Q 의 열벡터들이 \mathbb{R}^n 의 orthonormal basis를 이룬다면) $Q^{-1} = Q^\top$ 이 되고 이 경우에 Q 를 **orthogonal matrix**라고 부른다. 따라서 $QQ^{-1} = QQ^\top = I$ 므로 Q 의 열들이 orthonormal하면 Q 의 행들도 orthonormal함을 알 수 있다.

Example 4.6 Both the rotation matrix in \mathbb{R}^2 , $Q = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ and the permutation matrix in \mathbb{R}^3 ,

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \text{ are orthogonal matrices since}$$

$$Q^\top Q = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \cos^2 \theta + \sin^2 \theta & 0 \\ 0 & \sin^2 \theta + \cos^2 \theta \end{bmatrix} = I,$$

$$\text{and } P^\top P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I.$$

사실 (A.3)에서 임의의 $n \times n$ permutation matrix P 에 대해서 $PP^\top = I$ 성립함을 보인다. 따라서 모든 permutation matrix는 orthogonal하다.

4.7.1 QR Decomposition

선형독립인 열들로 이루어진 $m \times n$ 행렬 A 를 고려하자. 열들이 선형독립이므로 $m \geq n$ 이다. 행렬 A 를 $A = [\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_n]$ 로 나타내자. Section 4.4.1의 Gram-Schmidt procedure를 선형독립인 열 벡터들 $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ 에 적용하여 얻어지는 orthonormal basis를 $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$ 라고 하자. Fact 4.12에 의해 각 열 $\mathbf{a}_j, j = 1, \dots, n$ 은 $\mathbf{a}_j \in \text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_j\}$ 를 만족하므로

$$\mathbf{a}_j = \langle \mathbf{a}_j, \mathbf{q}_1 \rangle \mathbf{q}_1 + \dots + \langle \mathbf{a}_j, \mathbf{q}_j \rangle \mathbf{q}_j$$

로 표현되고, 특히 (4.10) 때문에 $\langle \mathbf{a}_j, \mathbf{q}_j \rangle \neq 0$ 이다. 따라서

$$A = [\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_n] = [\mathbf{q}_1 | \mathbf{q}_2 | \dots | \mathbf{q}_n] \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{q}_1 \rangle & \langle \mathbf{a}_2, \mathbf{q}_1 \rangle & \langle \mathbf{a}_3, \mathbf{q}_1 \rangle & \cdots & \langle \mathbf{a}_n, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{a}_2, \mathbf{q}_2 \rangle & \langle \mathbf{a}_3, \mathbf{q}_2 \rangle & \cdots & \langle \mathbf{a}_n, \mathbf{q}_2 \rangle \\ 0 & 0 & \langle \mathbf{a}_3, \mathbf{q}_3 \rangle & \cdots & \langle \mathbf{a}_n, \mathbf{q}_3 \rangle \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \langle \mathbf{a}_n, \mathbf{q}_n \rangle \end{bmatrix}$$

의 표현을 얻는데 $Q = [\mathbf{q}_1 | \mathbf{q}_2 | \dots | \mathbf{q}_n]$ 로 표시하고 Q 와 곱해지는 upper triangular 행렬을 R 로 표시하면, orthonormal한 열들로 이루어진 행렬 Q 와 upper triangular 행렬 R 에 대해 QR -decomposition이라 불리는

$$A = QR$$

의 표현을 얻는다. 여기서, R 은 영이 아닌 대각원소를 갖는 upper triangular matrix이므로 가역이다. 만약 $m = n$ 면 행렬 Q 가 orthogonal matrix가 된다.

4.7.2 Isometry induced by an Orthogonal Matrix

Orthogonal matrix Q 에 의한 linear transformation $T : \mathbf{x} \in \mathbb{R}^n \rightarrow T(\mathbf{x}) = Q\mathbf{x} \in \mathbb{R}^n$ 를 생각하자. $\langle \cdot, \cdot \rangle$ 이 \mathbb{R}^n 의 표준내적이라 하면

$$\langle T(\mathbf{x}), T(\mathbf{y}) \rangle = \langle Q\mathbf{x}, Q\mathbf{y} \rangle = (Q\mathbf{x})^\top Q\mathbf{y} = \mathbf{x}^\top Q^\top Q\mathbf{y} = \mathbf{x}^\top \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle$$

이 된다. 즉 임의의 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 에 대해서

$$\langle T(\mathbf{x}), T(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle \quad (4.12)$$

이 성립하여 orthogonal matrix를 통한 변환은 inner product를 보존한다. 특히 $\mathbf{x} = \mathbf{y}$ 인 경우를 생각하면

$$|T(\mathbf{x})|^2 = \langle T(\mathbf{x}), T(\mathbf{x}) \rangle = \langle \mathbf{x}, \mathbf{x} \rangle = |\mathbf{x}|^2$$

으로 표준내적에서 유도된 거리도 보존됨을 알 수 있다. 이렇게 거리를 보존하는 변환을 **isometry**라고 부른다. 반대로 어떤 linear transformation ℓ 이 $|\ell(\mathbf{x})| = |\mathbf{x}|$ 로 내적에서 유도된 거리를 보존하는 경우,

$$|\mathbf{x}|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + |\mathbf{y}|^2$$

$$\begin{aligned}
&= |\mathbf{x} + \mathbf{y}|^2 \\
&= |\ell(\mathbf{x} + \mathbf{y})|^2 \\
&= |\ell(\mathbf{x}) + \ell(\mathbf{y})|^2 \\
&= |\ell(\mathbf{x})|^2 + 2\langle \ell(\mathbf{x}), \ell(\mathbf{y}) \rangle + |\ell(\mathbf{y})|^2 \\
&= |\mathbf{x}|^2 + 2\langle \ell(\mathbf{x}), \ell(\mathbf{y}) \rangle + |\mathbf{y}|^2
\end{aligned}$$

이어서 $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \ell(\mathbf{x}), \ell(\mathbf{y}) \rangle$ 를 얻게되어 inner product가 보존된다. 즉 linear transformation의 경우, 내적이 보존되는 것과 내적에서 유도된 거리가 보존되는 것이 동치이다. 이런 기준으로 Section 3.8.2에서 살펴 본 linear transformation 중 isometry인 것들은 다음과 같다.

- Rotation matrix: 회전변환은 내적과 거리를 보존하므로 orthogonal할 것이라는 것을 예측할 수 있다. 실제로 \mathbb{R}^2 에서 rotation matrix $R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$ 에 대해 $R_\theta^\top R_\theta = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 이 되어 R_θ 가 orthogonal함을 알 수 있다.
- Reflection matrix: 대칭변환도 내적과 거리를 보존하므로 reflection matrix가 orthogonal하다는 것을 예측할 수 있다. 실제로 reflection matrix H 는 $H^\top = H$ 이고 $H^2 = I$ 이므로 $H^\top = H^{-1}$ 라는 것을 쉽게 알 수 있다.

4.8 Matrix Norms

행렬을 활용하다 보면 행렬의 크기(matrix norm)가 필요한 경우가 생긴다. 예를 들면 두 행렬이 비슷한지 궁금할 때 두 행렬의 차이를 행렬의 뺄셈(subtraction)으로 표현하고 이 차이를 나타내는 matrix norm이 작으면 원래 두 행렬이 비슷하다고 판단할 수 있다. 많은 종류의 matrix norm이 있는데 이 책에서는 Frobenius norm과 spectral norm 두 가지로 충분하다. Matrix norm을 설명하기 위해 필요한 추가적인 용어로 $n \times n$ 행렬 $A = (a_{ij})$ 의 trace는 대각 원소의 합, 즉

$$\text{trace } A = a_{11} + \cdots + a_{nn}$$

으로 정의한다.

- **Frobenius norm** : Euclidean vector norm을 이차원 행렬에 대해 단순하게 확장한 matrix norm으로

$$\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^d a_{ij}^2} = \sqrt{\text{trace}(AA^\top)} = \sqrt{\text{trace}(A^\top A)}$$

로 정의된다.⁷ 정의에서 쉽게 알 수 있듯이 $\|A\|_F = \|A^\top\|_F$ 이다. 직교하는 열들로 구성된 행렬 V (즉 $V^\top V = I$) 와 임의의 행렬 A 에 대해서

$$\|AV^\top\|_F^2 = \text{trace}(AV^\top(AV^\top)^\top) = \text{trace}(AV^\top VA^\top) = \text{trace}(AA^\top) = \|A\|_F^2$$

⁷ 임의의 행렬 A 와 B 에 대하여 AB 와 BA 가 정의되면 두 trace가 일치한다. 즉 $\text{trace}(AB) = \text{trace}(BA)$. A 가 $m \times n$ 행렬, B 가 $n \times m$ 행렬이라면 $\text{trace}(AB) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ji} = \sum_{j=1}^n \sum_{i=1}^m b_{ji} a_{ij} = \text{trace}(BA)$ 이 된다.

이다. 직교하는 열들로 구성된 행렬 U 에 대해서 $\|UA\|_F = \|A\|_F$ 임도 쉽게 알 수 있다.

- **Spectral norm :** 행렬이 선형 함수로서 단위벡터들의 크기를 얼마나 변화시키는지를 측정한 matrix norm으로

$$\max_{|\mathbf{x}| \leq 1} |A\mathbf{x}|$$

를 고려하자. 이 최대화 문제에서 $|A\mathbf{x}|$ 의 값을 최대로 하는 $|\mathbf{x}| \leq 1$ 인 \mathbf{x} 를 찾을 수 있는지는 자명하지 않다. 이러한 \mathbf{x} 의 존재 뿐만 아니라 실제로는 unit vector임을 Lemma B.3에서 증명한다. 따라서 spectral norm은

$$\|A\|_2 = \max_{|\mathbf{x}| \leq 1} |A\mathbf{x}| = \max_{|\mathbf{x}|=1} |A\mathbf{x}| \quad (4.13)$$

로 나타낼 수 있다. 회전변환이나 원점에 대한 절대칭을 나타내는 행렬의 경우에 spectral norm은 1이다. $\mathbf{0}$ 이 아닌 벡터 \mathbf{x} 에 대해 $|A(\frac{1}{|\mathbf{x}|}\mathbf{x})| \leq \|A\|_2$ 이므로 임의의 $\mathbf{x} \in \mathbb{R}^d$ 에 대해

$$|A\mathbf{x}| \leq \|A\|_2 |\mathbf{x}| \quad (4.14)$$

가 성립한다.

다음은 matrix norm에 관한 몇 가지 성질들이다.

Fact 4.17 For $n \times d$ matrix A and $d \times m$ matrix B ,

1. $\|AB\|_2 \leq \|A\|_2 \|B\|_2$;
2. $\|AB\|_F \leq \|A\|_F \|B\|_F$;
3. $\|A\|_2 \leq \|A\|_F$;
4. For orthogonal Q , $\|QA\|_2 = \|A\|_2$ and $\|QA\|_F = \|A\|_F$;
5. If $A = \mathbf{u}^\top$ or $A = \mathbf{u}$ for an n -vector \mathbf{u} , $\|A\|_2 = \|A\|_F = |\mathbf{u}|$.

Proof:

1. 임의의 \mathbf{x} 에 대하여 (4.14)에 의해 $|AB\mathbf{x}| \leq \|A\|_2 |B\mathbf{x}| \leq \|A\|_2 \|B\|_2 |\mathbf{x}|$

2. $a_{i\bullet} = (a_{i1}, \dots, a_{id})^\top$, $b_{\bullet j} = (b_{1j}, \dots, b_{dj})^\top$ 이라 하자. 그러면,

$$\|AB\|_F^2 = \sum_i \sum_j (a_{i\bullet}^\top b_{\bullet j})^2 \leq \sum_i \sum_j |a_{i\bullet}|^2 |b_{\bullet j}|^2 = \sum_i |a_{i\bullet}|^2 \sum_j |b_{\bullet j}|^2 = \|A\|_F^2 \|B\|_F^2.$$

3. For $|\mathbf{x}| \leq 1$, $|A\mathbf{x}|^2 = \sum_{i=1}^n (a_{i\bullet}^\top \mathbf{x})^2 \leq \sum_{i=1}^n |a_{i\bullet}|^2 |\mathbf{x}|^2 \leq \sum_{i=1}^n |a_{i\bullet}|^2 = \|A\|_F^2$.

4. $|QA\mathbf{x}| = \sqrt{\mathbf{x}^\top A^\top Q^\top QA\mathbf{x}} = \sqrt{\mathbf{x}^\top A^\top A\mathbf{x}} = |A\mathbf{x}|$, $\|QA\|_F^2 = \text{trace}(A^\top Q^\top QA) = \text{trace}(A^\top A) = \|A\|_F^2$.

5. $\|A\|_F = |\mathbf{u}|$ is clear. By Cauchy-Schwarz inequality, $|\mathbf{u}^\top \mathbf{x}| \leq |\mathbf{u}|$ for any n -vector \mathbf{x} with $|\mathbf{x}| = 1$. If we take $\mathbf{x} = \frac{1}{|\mathbf{u}|}\mathbf{u}$, then $A\mathbf{x} = \mathbf{u}^\top \mathbf{x} = \frac{1}{|\mathbf{u}|}\mathbf{u}^\top \mathbf{u} = |\mathbf{u}|$. Hence, $\|A\|_2 = |\mathbf{u}|$. For $A = \mathbf{u}$, the results are clear.

■

Fact 4.18 Every $m \times n$ rank-1 matrix can be represented as $\mathbf{u}\mathbf{v}^\top$ for some $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$. An upper bound on both the Frobenius and spectral norms of a rank-one matrix $\mathbf{u}\mathbf{v}^\top$ is $|\mathbf{u}||\mathbf{v}|$.

Proof: For any n -vector \mathbf{x} ,

$$\begin{aligned} |(\mathbf{u}\mathbf{v}^\top)\mathbf{x}| &= |\mathbf{u}(\mathbf{v}^\top \mathbf{x})| \\ &\leq \|\mathbf{u}\|\|\mathbf{v}^\top \mathbf{x}\| \quad (\text{by regarding } \mathbf{u} \text{ as an } m \times 1 \text{ matrix}) \\ &\leq \|\mathbf{u}\|\|\mathbf{v}^\top\|\|\mathbf{x}\| \quad (\text{by regarding } \mathbf{v}^\top \text{ as an } 1 \times n \text{ matrix}) \\ &= |\mathbf{u}||\mathbf{v}||\mathbf{x}| \quad (\text{by 5 of Fact 4.17}), \end{aligned}$$

regardless of the type of the norm $\|\cdot\|$.

■

4.9 Application to Data Science: Least Square and Projection

자연과학이나 공학에서 얻어진 데이터 중에서 설명변수 벡터 $\mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$ 과 종속변수 y 의 관계로 표현되는 경우가 많다. 이 관계가 적절한 함수들 $f_1, \dots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ 와 상수들 θ_j 에 대해서

$$y = \theta_1 f_1(\mathbf{x}) + \dots + \theta_k f_k(\mathbf{x})$$

로 나타나는 경우를 생각하자. 함수값 f_i 를 계산하는 과정에서는 오차가 없다고 가정하자. 따라서 관측된 하나의 데이터 $(\mathbf{x}, y) \in \mathbb{R}^n \times \mathbb{R}$ 에 대해 모형에 기술되는 값인 $(f_1(\mathbf{x}), \dots, f_k(\mathbf{x}))^\top \in \mathbb{R}^k$ 가 직접 관측된다고 가정하여도 특별한 문제가 없다. 즉 이렇게 함수가 적용된 값을 관측한다고 가정하면, 모형을 다음과 같은 선형모형으로 가정해도 충분하다: $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)^\top$ 에 대해

$$y = \theta_1 x_1 + \dots + \theta_n x_n = \boldsymbol{\theta}^\top \mathbf{x}.$$

그런데 현실에서 측정된 데이터 쌍을 $(\mathbf{a}_1, b_1), \dots, (\mathbf{a}_m, b_m)$ 라 하면 여러가지 이유로 측정 오차 ε_i 가 존재하며 이를 다음과 같은 가산적 형태(additive form)로 가정하자:

$$b_i = \boldsymbol{\theta}^\top \mathbf{a}_i + \varepsilon_i, \quad i = 1, \dots, m.$$

이 모형을 \mathbf{a}_i 가 i 번째 행인 $m \times n$ 데이터 행렬 A 와 벡터 $\mathbf{b} = (b_1, \dots, b_m)^\top$ 에 대해 오차를 포함하지 않은 모형을 표현하면 $A\boldsymbol{\theta} = \mathbf{b}$ 를 만족하는 $\boldsymbol{\theta}$ 를 구하는 선형연립방정식이지만 실제 관측에서 발생하는 오차 ε_i 때문에 데이터 쌍에 대해서 $b_i = \boldsymbol{\theta}^\top \mathbf{a}_i$ 가 되는 $\boldsymbol{\theta}$ 가 없을 수도 있다. 해가 되는 $\boldsymbol{\theta}$ 가 존재하지 않는다는 것은 $\mathbf{b} \notin \text{Col}(A)$ 인 경우이다.

차선책으로 θ 에 따라 달라지는 오차 $\varepsilon_i = \theta^\top \mathbf{a}_i - b_i$ 가 작게되는 θ 를 구하는 문제를 생각할 수 있다. 이 문제를

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^n}{\operatorname{argmin}} \sum_{i=1}^m \varepsilon_i^2 = \underset{\theta \in \mathbb{R}^n}{\operatorname{argmin}} \sum_{i=1}^m (\theta^\top \mathbf{a}_i - b_i)^2 = \underset{\theta \in \mathbb{R}^n}{\operatorname{argmin}} |A\theta - \mathbf{b}|^2 \quad (4.15)$$

로 모형화하자.⁸

4.9.1 Least Square as a Convex Quadratic Minimization

θ 에 관한 2차식이 최소가 되는 θ 를 정해보자. θ 에 대한 2차식을 전개하면

$$\begin{aligned} |A\theta - \mathbf{b}|^2 &= \langle A\theta - \mathbf{b}, A\theta - \mathbf{b} \rangle \\ &= \langle A\theta, A\theta \rangle - \langle A\theta, \mathbf{b} \rangle - \langle \mathbf{b}, A\theta \rangle + \langle \mathbf{b}, \mathbf{b} \rangle \\ &= (A\theta)^\top A\theta - 2\mathbf{b}^\top A\theta + |\mathbf{b}|^2 \\ &= \theta^\top A^\top A\theta - 2\mathbf{b}^\top A\theta + |\mathbf{b}|^2 \end{aligned}$$

이고, θ 에 대한 극점으로 gradient가 영벡터가 되는 점⁹을 Fact 4.19을 이용하여 계산하면

$$2A^\top A\theta - 2A^\top \mathbf{b} = \mathbf{0} \quad \text{또는} \quad A^\top A\theta = A^\top \mathbf{b}$$

이 되는데 이 식을 정규방정식(normal equation)이라고 부른다. 그런데 이런 유형의 문제들은 대부분 데이터의 개수 m 이 parameter의 개수 n 보다 훨씬 큰 경우 ($m \gg n$)가 많아 보통 A 의 rank가 n 이다. 만약 rank가 n 보다 작다면 데이터를 기술하는 모형에 collinearity가 존재하여 parameter 개수를 줄일 수 있으므로 rank를 n 으로 가정하여도 무방하다. A 의 rank가 n 이면 Fact 3.8에 의해 $n \times n$ 행렬 $A^\top A$ 의 rank도 n 이 되어 invertible하고 정규방정식은

$$\hat{\theta} = (A^\top A)^{-1} A^\top \mathbf{b}$$

를 해로 갖는다.

Fact 4.19 Let Q be an $n \times n$ matrix, \mathbf{b} an n -vector, and c a real number. A real-valued quadratic function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined as

$$f(\mathbf{x}) = \mathbf{x}^\top Q\mathbf{x} + \mathbf{b}^\top \mathbf{x} + c.$$

Then, the gradient of f is given as

$$\nabla f(\mathbf{x}) = Q\mathbf{x} + Q^\top \mathbf{x} + \mathbf{b}.$$

If the matrix Q is symmetric, then $\nabla f(\mathbf{x}) = 2Q\mathbf{x} + \mathbf{b}$.

Proof: Since $f(\mathbf{x})$ can be written as

$$f(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n q_{ij}x_i x_j + \sum_{i=1}^n b_i x_i + c = q_{kk}x_k^2 + \sum_{i \neq k} \sum_{j=1}^n q_{ij}x_i x_j + \sum_{j \neq k} q_{kj}x_k x_j + \sum_{i=1}^n b_i x_i + c,$$

⁸여기서 새롭게 등장하는 기호인 argmin 의 의미는 다음과 같다: 임의의 함수 f 에 대해서 $\operatorname{argmin}_{\mathbf{x} \in A} f(\mathbf{x})$ 는 A 상에서 함수 f 가 최소가 되는 A 의 원소를 의미한다. A 가 \mathbb{R}^n 과 같이 명확한 경우에는 생략하는 경우도 있다. 영어로는 **argument minimizing the function f** 이다. argmax 도 비슷하게 정의된다.

⁹오차의 제곱합인 $|A\theta - \mathbf{b}|^2$ 은 θ 에 대해 아래로 볼록(convex)인 함수이므로 gradient가 영벡터인 점에서 최솟값을 갖는다.

its partial derivative with respect to x_k is

$$\frac{\partial f}{\partial x_k}(x_1, \dots, x_n) = 2q_{kk}x_k + \sum_{i \neq k} q_{ik}x_i + \sum_{j \neq k} q_{kj}x_j + b_k = \sum_{i=1}^n q_{ik}x_i + \sum_{j=1}^n q_{kj}x_j + b_k = (Q^\top \mathbf{x})_k + (Q\mathbf{x})_k + b_k,$$

which shows the desired representation. For a symmetric Q , the conclusion is straightforward. ■

4.9.2 Equivalence of Least Square and Projection

최적화 문제 (4.15)의 해 $\hat{\theta}$ 로부터 $A\hat{\theta}$ 가 \mathbf{b} 와 거리가 제일 가깝기 때문에 $A\hat{\theta} = A(A^\top A)^{-1}A^\top \mathbf{b}$ 를 \mathbf{b} 에 대한 근사로 생각할 수 있다. 앞에서 살펴본 바와 같이 부분공간에 대한 정사영이 부분공간까지의 거리가 최소인 벡터와 동일한 것처럼 오차를 나타내는 벡터인 $\mathbf{b} - A\hat{\theta}$ 도 A 의 열공간 $\text{Col}(A)$ 에 수직한지 확인해보자. 그런데 Lemma 4.4에 의해 $\text{Col}(A)^\perp = \text{Null}(A^\top)$ 이므로 $\mathbf{b} - A\hat{\theta} \in \text{Null}(A^\top)$ 여부를 확인하면 된다. 곱해보면

$$A^\top(\mathbf{b} - A\hat{\theta}) = A^\top(I - A(A^\top A)^{-1}A^\top)\mathbf{b} = (A^\top - A^\top A(A^\top A)^{-1}A^\top)\mathbf{b} = \mathbf{0}$$

으로 오차벡터가 열공간에 수직함을 알 수 있다. 임의의 벡터와 열공간 $\text{Col}(A)$ 상의 가장 가까운 벡터는 동일한 벡터의 정사영과 일치하고, 가장 가까운 벡터를 대응시키는 변환을 표현하는 행렬은

$$A(A^\top A)^{-1}A^\top$$

로 정사영 행렬 (4.9)와 일치한다. 이런 의미에서 least square 문제는 projection과 동일하다.

Kang & Cho (2024)

Chapter 5

Singular Value Decomposition (SVD)

고차원 데이터들을 저차원 부분공간 상에 정사영하여 얻어지는 벡터를 원래 데이터의 저차원 근사로 생각할 수 있다. 즉 저차원 부분공간으로의 정사영이 데이터의 차원을 축소하는 한가지 방법이다. 데이터를 근사할 부분공간이 주어진 경우 이전 장에서 어떻게 정사영을 계산하는지 살펴보았다. 다음 단계로 어떤 부분공간이 데이터를 잘 근사하는 부분공간인지 생각해보자. 근사를 잘하는 기준이 달라지면 근사하는 부분공간도 변한다. 근사 정도에 대한 여러 기준 중에서 정사영과 원래 벡터 간의 잔차 벡터의 길이의 제곱을 기준으로 삼으면 피타고拉斯 정리를 활용할 수 있어서 여러 장점이 많다. 데이터를 나타내는 벡터를 행으로 쌓아 만든 행렬을 데이터 행렬이라고 부르면, 이렇게 잔차 길이의 제곱합을 최소화하는 부분공간을 찾는 것을 데이터 행렬의 singular value decomposition (SVD)이라고 부른다. 찾아진 부분공간을 기술하는 여러 기저 중에 SVD는 특별한 직교기저를 찾는다. 제약식이 단계별로 늘어나는 1차원 최적화 문제를 반복적으로 풀어서 구해지는 최적해가 부분공간의 직교기저를 이루는 벡터들이 된다. 정사영 잔차 길이의 제곱합을 최소로 하는 단위벡터가 첫번째 right singular vector이고 이때 잔차 길이의 제곱합이 첫번째 singular value의 제곱이 된다. 첫번째 right singular vector에 수직한 단위벡터로 제한하여 잔차 길이의 제곱합을 최소로 하는 1차원 최적화 문제의 최적해로부터 두번째 right singular vector와 singular value를 구할 수 있고 서로 수직한 left singular vector도 얻는다. 이를 반복하면 데이터 행렬에 대한 SVD의 모든 singular vector와 singular value를 찾을 수 있다. 이때 데이터 행렬은 rank가 최대가 아니어도 된다. SVD는 rank-1행렬의 합으로 표현하거나 singular value로 구성된 대각행렬과 singular vector로 열을 이루는 두 행렬의 곱으로 표현된다. 얻어진 SVD에서 singular value가 제일 큰 k 개의 rank-1행렬을 더하면 데이터 행렬의 rank- k 근사 행렬이 된다. 또한 데이터 행렬의 가역여부에 무관하게 근사적인 역행렬인 pseudoinverse를 SVD로부터 쉽게 얻는데 이는 선형 종속인 벡터들로 생성되는 부분공간 상으로의 정사영을 나타내거나 데이터 행렬의 rank가 최대가 아닐 때의 최소제곱법의 해를 제공한다. 특히 벡터의 수직 조건이 그 벡터들로 가중된 확률벡터들의 합이 확률적으로 상관이 없는 것임을 보이면 SVD는 principal components analysis (PCA)로 해석할 수 있다.

5.1 Variational Formulation for the Best-fit Subspaces

d 차원의 속성을 가진 n 개의 데이터가 각각 행을 이루는 $n \times d$ 행렬 A 를 가정하자. i 번째 행을 $\mathbf{a}_i \in \mathbb{R}^d$ 라 하면, 행렬 A 는 n 개의 원소를 가진 집합 $\{\mathbf{a}_i : i = 1, \dots, n\}$ 과 동일한 정보를 나타낸다. 데이터가 속하는 \mathbb{R}^d 공간에서 데이터 집합 $\{\mathbf{a}_i : i = 1, \dots, n\}$ 를 잘 표현하는 k 차원 부분공간 \mathbb{W} 를 찾는 문제를 고민해보자. 당연히 $k < d$ 로 가정한다. 잘 표현한다는 것을 여러 기준으로 생각해 볼 수 있는데, 정량적인 접근으로 데이터 벡터 \mathbf{a}_i 를 k 차원 부분공간으로 정사영할 때 벡터 크기의 제곱을 SVD는 사용한다. 즉

$$\mathbb{W}^* = \underset{\mathbb{W}: \dim(\mathbb{W}) \leq k}{\operatorname{argmin}} \sum_{i=1}^n |\mathbf{a}_i - \mathbf{P}_{\mathbb{W}}(\mathbf{a}_i)|^2 \quad (5.1)$$

을 통해 k 차원 부분공간들 중에서 “sum of squares of residuals” 기준으로 최적의 \mathbb{W}^* 를 정하는 것이다. 이 절에서는 벡터공간 \mathbb{R}^d 의 표준내적을 사용한다. 정사영을 공부할 때는 주어진 방향들로의 정사영을 어떻게 계산할지를 고민했지만 이 문제에서는 정사영할 방향을 어떻게 정할지 고민한다. 이 경우 자연스러운 질문이 몇 가지 있다.

1. 정사영 길이의 제곱합이 아닌 다른 기준은 어떨까? Machine learning 관점에서는 제곱합은 손실함수 $\sum_{i=1}^n \ell(\mathbf{a}_i, \mathbf{P}_{\mathbb{W}}(\mathbf{a}_i))$ 의 특별한 경우이다. 이후 전개과정에서 분명해지지만 제곱합의 경우 피타고拉斯 정리를 사용할 수 있어서 간결한 결과를 얻을 수 있다는 것이 장점이다. 즉

$$|\mathbf{a}_i|^2 = |\mathbf{P}_{\mathbb{W}}(\mathbf{a}_i)|^2 + |\mathbf{a}_i - \mathbf{P}_{\mathbb{W}}(\mathbf{a}_i)|^2$$

이고 $\sum_{i=1}^n |\mathbf{a}_i|^2$ 은 상수 이므로 (5.1)은

$$\mathbb{W}^* = \underset{\mathbb{W}: \dim(\mathbb{W}) \leq k}{\operatorname{argmax}} \sum_{i=1}^n |\mathbf{P}_{\mathbb{W}}(\mathbf{a}_i)|^2 \quad (5.2)$$

과 동치이고, 우리가 singular value decomposition (SVD)을 유도하는 과정에서는 (5.2)를 사용한다. 물론, 정사영 길이의 합을 포함해서 손실함수로 다른 선택도 가능하지만 간결한 결과를 얻을 수 있을지 고민해 보아야 한다.

2. k 차원 부분공간은 어떻게 기술할까? 임의의 linearly independent인 k 개의 벡터가 k 차원 부분공간을 기술할 수 있지만, 정사영 표현의 간결성을 고려하여 orthonormal한 k 개의 벡터를 사용하여 부분공간을 기술하는데 이 벡터들을 **right-singular vector**라고 부른다.

5.1.1 Best-fit 1-dimensional subspace

데이터 집합 $\{\mathbf{a}_i : 1 \leq i \leq n\}$ 을 제일 잘 표현하는 1차원 부분공간을 결정해 보자. 1차원 부분공간의 기저를 단위길이를 갖는 벡터로 제한하고, 이 기저에 의해 표현되는 부분(즉 정사영)의 크기는 내적으로 표현된다. 1차원 부분공간의 기저를 \mathbf{v} 라 하면, \mathbf{a}_i 의 정사영은 (4.4)에 의해

$$\mathbf{P}_{\operatorname{span}\{\mathbf{v}\}}(\mathbf{a}_i) = \langle \mathbf{a}_i, \mathbf{v} \rangle \mathbf{v}$$

이므로 데이터 벡터들이 기저 \mathbf{v} 에 의해 표현되는 부분(정사영)의 길이는 $\{|\langle \mathbf{a}_i, \mathbf{v} \rangle| : 1 \leq i \leq n\}$ 이 된다. 따라서 데이터 집합 전체의 표현력은

$$\sum_{i=1}^n |\mathbf{P}_{\text{span}\{\mathbf{v}\}}(\mathbf{a}_i)|^2 = \sum_{i=1}^n \langle \mathbf{a}_i, \mathbf{v} \rangle^2 = \sum_{i=1}^n (\mathbf{a}_i^\top \mathbf{v})^2$$

이고 이를 최대화 하고자 한다. 그런데 \mathbf{a}_i 가 행렬 A 의 i 번째 행벡터를 나타내므로 이 제곱합은 $|A\mathbf{v}|^2$ 과 같다. 따라서 $n \times d$ 인 행렬 A 의 행들을 제일 잘 표현하는 1차원 부분공간의 기저 \mathbf{v}_1 은

$$\mathbf{v}_1 = \underset{\mathbf{v} \in \mathbb{R}^d : |\mathbf{v}|=1}{\operatorname{argmax}} |A\mathbf{v}| = \underset{\mathbf{v} \in \mathbb{R}^d : |\mathbf{v}|=1}{\operatorname{argmax}} \sum_{i=1}^n \langle \mathbf{a}_i, \mathbf{v} \rangle^2 \quad (5.3)$$

로 표현된다. 이 최대화 문제의 해가 존재함을 Lemma B.3에서 증명하였다. 데이터행렬 A 가 단위행렬 I 인 경우를 생각하면 \mathbf{v}_1 이 임의의 단위벡터이면 되므로 실제 문제에서 \mathbf{v}_1 은 유일하지 않을 수도 있다는 것을 알 수 있다.

Vector Orthogonal to a Set of Vectors

2차원 이상의 최적 부분공간을 찾는 과정에서는 k 개의 orthonormal한 벡터들에 직교하는 $k+1$ 차원 부분공간의 벡터를 찾는 문제가 등장한다. 물론 $k < d$ 이다.

Lemma 5.1 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subset \mathbb{R}^d$ 가 orthonormal하고 \mathbb{W} 가 임의의 $k+1$ 차원 부분공간이면 적어도 한 개의 $\mathbf{0} \neq \mathbf{w} \in \mathbb{W}$ 에 대해서 $\langle \mathbf{v}_1, \mathbf{w} \rangle = \dots = \langle \mathbf{v}_k, \mathbf{w} \rangle = 0$ 이 된다. 즉 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 모두에 수직하지만 $\mathbf{0}$ 이 아닌 \mathbb{W} 의 벡터가 존재한다.

Proof: \mathbb{W} 의 기저를 $\{\mathbf{w}_1, \dots, \mathbf{w}_{k+1}\}$ 이라 하자. 그러면 $\mathbf{x} = (x_1, \dots, x_{k+1}) \in \mathbb{R}^{k+1}$ 에 대해 벡터 $\mathbf{w} = x_1\mathbf{w}_1 + \dots + x_{k+1}\mathbf{w}_{k+1} \in \mathbb{W}$ 를 대응시킬 수 있다. 이 대응관계를 고려하면 원래 찾고자 하는 수직인 벡터 \mathbf{w} 가 존재한다면

$$\langle \mathbf{v}_i, \mathbf{w} \rangle = x_1 \langle \mathbf{v}_i, \mathbf{w}_1 \rangle + \dots + x_{k+1} \langle \mathbf{v}_i, \mathbf{w}_{k+1} \rangle = 0, \quad i = 1, \dots, k$$

로 기술되는 \mathbf{x} 에 대한 연립방정식의 $\mathbf{0}$ 이 아닌 해에 대응되는 벡터이다. $k \times (k+1)$ 행렬

$$B = (\langle \mathbf{v}_i, \mathbf{w}_j \rangle)_{1 \leq i \leq k, 1 \leq j \leq k+1}$$

에 대해 $B\mathbf{x} = \mathbf{0}$ 를 만족하는 $\mathbf{0}$ 이 아닌 해를 찾는 문제이고, 이 문제는 미지수의 개수가 방정식의 개수보다 한 개 많으므로 Lemma 3.2에 의해 반드시 $\mathbf{0}$ 이 아닌 해를 갖는다. ■

5.1.2 Best-fit 2-dimensional subspace

데이터 집합 $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ 을 근사하는 2차원 부분공간은

$$(\mathbf{w}_1^*, \mathbf{w}_2^*) = \underset{\substack{\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^d : \\ |\mathbf{w}_1|=|\mathbf{w}_2|=1, \\ \langle \mathbf{w}_1, \mathbf{w}_2 \rangle=0}}{\operatorname{argmax}} \sum_{i=1}^n |\mathbf{P}_{\text{span}\{\mathbf{w}_1, \mathbf{w}_2\}}(\mathbf{a}_i)|^2 \quad (5.4)$$

을 풀어서 결정되는 2차원 부분공간 $\mathbb{W}^* = \text{span}\{\mathbf{w}_1^*, \mathbf{w}_2^*\}$ 이다. 하지만 이 문제를 그대로 풀면 \mathbf{w}_1^* 과 \mathbf{w}_2^* 간의 관계를 파악하는 것이 쉽지 않다. 그래서 2차원 문제를 그대로 풀지 말고 앞에서 찾은 1차원 부분공간을 생성하는 방향 \mathbf{v}_1 을 활용하여 \mathbf{w}_1 를 \mathbf{v}_1 로 대체하여 문제 (5.4)를 풀어보면 어떨까? 즉 다음과 같이 \mathbf{v}_1 을 구하고 이를 이용하여 \mathbf{v}_2 를 구하는 greedy한 방법을 고려해보자.

$$1. \mathbf{v}_1 = \underset{\mathbf{v} \in \mathbb{R}^d : |\mathbf{v}|=1}{\operatorname{argmax}} |A\mathbf{v}|$$

$$2. \mathbf{v}_2 = \underset{\substack{\mathbf{v} \in \mathbb{R}^d : |\mathbf{v}|=1, \\ \langle \mathbf{v}, \mathbf{v}_1 \rangle = 0}}{\operatorname{argmax}} |A\mathbf{v}|$$

2번 문제의 해가 존재한다는 사실을 Lemma B.4에서 확인할 수 있다. 어떤 알고리듬이 greedy하다는 것은 흔히 직관적인 접근을 반복적으로 실행하는 경우를 말한다. 위의 경우에는 이미 찾은 부분공간에 수직한 방향의 1차원 최대화 문제를 풀어서 원하는 벡터를 구한다는 뜻에서 부르는 말이다. 즉 1차원 근사를 풀고, 구해진 방향에 수직한 방향 중에서 데이터를 제일 많이 담아내는 두 번째 일차원 방향을 정하는 것이다. 이렇게 정해진 $\{\mathbf{v}_1, \mathbf{v}_2\}$ 에 의해 결정되는 2차원 부분공간이 \mathbb{W}^* 와 동일한 표현력을 갖는지를 확인하자. 우선, 정사영 표현 (4.6)로부터

$$\begin{aligned} \sum_{i=1}^n |\mathbf{P}_{\text{span}\{\mathbf{w}_1, \mathbf{w}_2\}}(\mathbf{a}_i)|^2 &= \sum_{i=1}^n |\langle \mathbf{a}_i, \mathbf{w}_1 \rangle \mathbf{w}_1 + \langle \mathbf{a}_i, \mathbf{w}_2 \rangle \mathbf{w}_2|^2 \\ &= \sum_{i=1}^n \langle \mathbf{a}_i, \mathbf{w}_1 \rangle^2 + \langle \mathbf{a}_i, \mathbf{w}_2 \rangle^2 \quad \text{since } \langle \mathbf{w}_1, \mathbf{w}_2 \rangle = 0 \\ &= |A\mathbf{w}_1|^2 + |A\mathbf{w}_2|^2 \end{aligned}$$

이므로 2차원 문제는

$$(\mathbf{w}_1^*, \mathbf{w}_2^*) = \underset{\substack{\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^d : \\ |\mathbf{w}_1|=|\mathbf{w}_2|=1, \\ \langle \mathbf{w}_1, \mathbf{w}_2 \rangle = 0}}{\operatorname{argmax}} |A\mathbf{w}_1|^2 + |A\mathbf{w}_2|^2$$

로 표현된다. 이 표현을 얻는 과정에서 \mathbf{w}_1 과 \mathbf{w}_2 의 orthogonality가 중요한 역할을 함에 유의하자. 이렇게 동시에 최적화한 결과로 얻어진 부분공간을 $\mathbb{W}^* = \text{span}\{\mathbf{w}_1^*, \mathbf{w}_2^*\}$ 라 하자.

이제, greedy방법의 첫 번째 1차원 문제를 풀어서 \mathbf{v}_1 을 구했다고 가정하자. 그러면 $k = 1$ 인 경우인 $\{\mathbf{v}_1\}$ 과 2차원 \mathbb{W}^* 에 대해 앞에서 증명한 Lemma 5.1을 적용하면 $\mathbf{0} \neq \mathbf{w} \in \mathbb{W}^*$ 가 존재하여 $\langle \mathbf{v}_1, \mathbf{w} \rangle = 0$ 이 되는데 이 벡터를 \mathbf{w}_2 로 놓자. 그리고 \mathbf{w}_2 와 수직인 단위벡터 \mathbf{w}_1 을 \mathbb{W}^* 에서 찾아서¹ $\{\mathbf{w}_1, \mathbf{w}_2\}$ 가 \mathbb{W}^* 의 기저가 된다. 그런데 \mathbf{w}_2 는 $\langle \mathbf{v}_1, \mathbf{w}_2 \rangle = 0$ 으로 greedy방법의 2번째 일차원 최적화문제의 가능해이므로 $|A\mathbf{w}_2| \leq |A\mathbf{v}_2|$ 이어야 한다. 그리고 \mathbf{w}_1 은 greedy방법 1번 최적화 문제의 가능해이므로 $|A\mathbf{w}_1| \leq |A\mathbf{v}_1|$ 이어서 greedy 방법으로 최적 2차원 부분공간을 찾을 수 있다.

5.1.3 Best-fit k -dimensional subspace

데이터 집합 $\{\mathbf{a}_i : 1 \leq i \leq n\}$ 을 제일 잘 표현하는 k 차원 부분공간은 최적화 문제

¹Gram-Schmidt orthogonalization을 활용하여

$$\begin{aligned}
 (\mathbf{w}_1^*, \dots, \mathbf{w}_k^*) &= \underset{\mathbf{w}_1, \dots, \mathbf{w}_k \in \mathbb{R}^d}{\operatorname{argmax}} \sum_{i=1}^n |\mathbf{P}_{\operatorname{span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\}}(\mathbf{a}_i)|^2 \\
 \text{s.t.} \quad |\mathbf{w}_i| &= 1 \quad \text{for } i = 1, \dots, k \\
 \langle \mathbf{w}_i, \mathbf{w}_j \rangle &= 0 \quad \text{for } i \neq j
 \end{aligned}$$

이 의해 $\mathbb{W}^* = \operatorname{span}\{\mathbf{w}_1^*, \dots, \mathbf{w}_k^*\}$ 로 결정할 수 있다. \mathbb{W}^* 만큼 $\{\mathbf{a}_i : 1 \leq i \leq n\}$ 를 잘 근사하는 부분공간을 결정하는 greedy방법을 생각하자.

Greedy procedure to find the best-fit k -dimensional subspace of $n \times d$ matrix A

1. Set the first vector (by breaking ties arbitrarily) as

$$\mathbf{v}_1 = \underset{\mathbf{v} \in \mathbb{R}^d: |\mathbf{v}|=1}{\operatorname{argmax}} |A\mathbf{v}|;$$

2. For $j = 2, \dots, k$: $\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}\}$ is already known and set

$$\mathbf{v}_j = \underset{\substack{\mathbf{v} \in \mathbb{R}^d: |\mathbf{v}|=1, \\ \langle \mathbf{v}, \mathbf{v}_1 \rangle = 0, \dots, \langle \mathbf{v}, \mathbf{v}_{j-1} \rangle = 0}}{\operatorname{argmax}} |A\mathbf{v}|. \quad (5.5)$$

(5.5)의 해가 존재하는 것을 Lemma B.4에서 확인할 수 있다. Greedy방법으로 얻어진 기저를 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 로 표시하자. 수학적 귀납법을 사용하기 위해서 $k - 1$ 차원 부분공간까지는 greedy방법이 최적 부분공간을 찾는다고 가정하자. 즉 A 에 대한 최적 $k - 1$ 차원 부분 공간의 기저 $\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}$ 을 greedy방법으로 찾았다고 하자. 그러면 Lemma 5.1에 의해 $\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}$ 각각에 수직인 \mathbb{W}^* 의 단위벡터 \mathbf{w}_k 가 존재하고 \mathbf{w}_k 를 포함하는 \mathbb{W}^* 의 기저 $\{\mathbf{w}_1, \dots, \mathbf{w}_{k-1}, \mathbf{w}_k\}$ 를 찾을 수 있다. 그러면, $\langle \mathbf{v}_1, \mathbf{w}_k \rangle = 0, \dots, \langle \mathbf{v}_{k-1}, \mathbf{w}_k \rangle = 0$ 이어서 \mathbf{w}_k 가 k 번째 최적화의 가능해 이므로 $|A\mathbf{w}_k| \leq |A\mathbf{v}_k|$ 이다. 그리고 수학적 귀납법 가정에 의해

$$\begin{aligned}
 |A\mathbf{w}_1|^2 + \cdots + |A\mathbf{w}_{k-1}|^2 &= \sum_{i=1}^n |\mathbf{P}_{\operatorname{span}\{\mathbf{w}_1, \dots, \mathbf{w}_{k-1}\}}(\mathbf{a}_i)|^2 \\
 &\leq \sum_{i=1}^n |\mathbf{P}_{\operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}}(\mathbf{a}_i)|^2 = |A\mathbf{v}_1|^2 + \cdots + |A\mathbf{v}_{k-1}|^2
 \end{aligned}$$

이어서

$$\begin{aligned}
 \sum_{i=1}^n |\mathbf{P}_{\operatorname{span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\}}(\mathbf{a}_i)|^2 &= |A\mathbf{w}_1|^2 + \cdots + |A\mathbf{w}_k|^2 \\
 &\leq |A\mathbf{v}_1|^2 + \cdots + |A\mathbf{v}_k|^2
 \end{aligned}$$

을 얻게된다. 그런데 $\mathbb{W}^* = \operatorname{span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ 이므로

$$\sum_{i=1}^n |\mathbf{P}_{\mathbb{W}^*}(\mathbf{a}_i)|^2 \leq |A\mathbf{v}_1|^2 + \cdots + |A\mathbf{v}_k|^2$$

이 되어 greedy 방법으로 최적 k 차원 부분공간을 찾을 수 있다.

위에 기술한 i -번째 일차원 최적화 문제에서 얻은 단위 벡터 \mathbf{v}_i 를 i -번째 **right singular vector**라고 부르고

$$\sigma_i = |A\mathbf{v}_i|$$

를 i -번째 **singular value**라고 부른다. Greedy 방법의 특성으로부터 모든 i 에 대해서 $\sigma_i \geq \sigma_{i+1}$ 임을 알 수 있고 σ_{r+1} 이 0이되면 singular vector를 찾는 과정이 끝나고

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1}$$

를 얻는다. $i = 1, \dots, r$ 에 대하여 단위벡터 \mathbf{u}_i 를

$$\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$$

로 정의하면

$$A\mathbf{v}_i = \sigma_i \mathbf{u}_i$$

이고 \mathbf{u}_i 를 **left singular vector**라고 부른다. 이상의 결과를 정리하여 greedy 방법이 최적 부분공간을 찾는다는 다음 정리를 얻을 수 있다.

Theorem 5.1 $n \times d$ 행렬 A 에 대해서 right singular vector가 $\mathbf{v}_1, \dots, \mathbf{v}_r$ 이라 하자. 그러면, $1 \leq k \leq r$ 인 $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 는 A 의 행들을 (오차의 제곱합 기준으로) 가장 잘 근사하는 k 차원 부분공간이다.

만약 $r < d$ 에 대해서 $\sigma_r > \sigma_{r+1} = 0$ 이라면 $|A\mathbf{v}_{r+1}| = 0$ 이고 모든 \mathbf{a}_i 들이 $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ 에 의해 형성되는 부분공간에 포함된다. 이 경우, 모든 i 에 대해서 $\mathbf{a}_i = \langle \mathbf{a}_i, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle \mathbf{a}_i, \mathbf{v}_r \rangle \mathbf{v}_r$ 으로 $|\mathbf{a}_i|^2 = |\langle \mathbf{a}_i, \mathbf{v}_1 \rangle|^2 + \cdots + |\langle \mathbf{a}_i, \mathbf{v}_r \rangle|^2$ 되어

$$\|A\|_F^2 = \sum_{i=1}^n |\mathbf{a}_i|^2 = |A\mathbf{v}_1|^2 + \cdots + |A\mathbf{v}_r|^2 = \sigma_1^2 + \cdots + \sigma_r^2 \quad (5.6)$$

이 됨을 알 수 있다.

$\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ 을 포함하는 부분공간의 최소 차원 r 보다 작은 k 에 대해 k 차원 부분공간으로 데이터를 설명하고자 하는 경우를 생각하자. 이 경우에는 $|A\mathbf{v}_{k+1}| \neq 0$ 이지만 $|A\mathbf{v}_{k+1}|$ 값이 충분히 작아지면 데이터들이 k 개의 대표적 데이터타입, 즉 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 의 조합으로 상당 부분 설명된다고 가정하고 여러 분석을 수행한다.

5.2 Orthogonality of Left Singular Vectors

$A\mathbf{v}_i$ 들의 직교성은 \mathbf{u}_i 들의 직교성과 동치이다. \mathbf{u}_i 들이 orthogonal하지 않다고 가정하자. \mathbf{u}_i 를 다른 \mathbf{u}_j 와 직교하지 않는 가장 작은 index의 벡터라고 하자. 즉 $i < j$ 이고 $\langle \mathbf{u}_i, \mathbf{u}_j \rangle \neq 0$ 이다. 특히 \mathbf{u}_j 의 부호를 조정하여 $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta > 0$ 이라고 하자. 충분히 작은 양수 $\epsilon > 0$ 에 대해서 단위벡터

$$\mathbf{v}'_i = \frac{\mathbf{v}_i + \epsilon \mathbf{v}_j}{|\mathbf{v}_i + \epsilon \mathbf{v}_j|} = \frac{1}{\sqrt{1 + \epsilon^2}} (\mathbf{v}_i + \epsilon \mathbf{v}_j)$$

를 정의한다. 그러면

$$A\mathbf{v}'_i = \frac{1}{\sqrt{1 + \epsilon^2}} (\sigma_i \mathbf{u}_i + \epsilon \sigma_j \mathbf{u}_j)$$

이 고

$$\begin{aligned}
 |A\mathbf{v}'_i|^2 &= \frac{1}{1+\epsilon^2}(\sigma_i^2 + 2\epsilon\delta\sigma_i\sigma_j + \epsilon^2\sigma_j^2) \\
 &> (1-\epsilon^2)(\sigma_i^2 + 2\epsilon\delta\sigma_i\sigma_j + \epsilon^2\sigma_j^2) \\
 &> (1-\epsilon^2)(\sigma_i^2 + 2\epsilon\delta\sigma_i\sigma_j) \\
 &= \sigma_i^2 + \epsilon(-\epsilon\sigma_i^2 + 2(1-\epsilon^2)\delta\sigma_i\sigma_j) \\
 &> \sigma_i^2
 \end{aligned}$$

이다. 마지막 부등식은 충분히 작은 ϵ 에 대해서는 $2(1-\epsilon^2)\delta\sigma_i\sigma_j > \epsilon\sigma_i^2$ 이라는 사실을 활용하였다. 그런데 \mathbf{v}_i 와 \mathbf{v}_j 는 $\ell < i$ 인 모든 \mathbf{v}_ℓ 과 직교하므로 \mathbf{v}'_i 는 greedy 방법의 i 번째 최적화 문제의 가능해이고 최적 목적함수 값인 $\sigma_i = |A\mathbf{v}_i|$ 보다 큰 목적함수 값을 가지므로 모순이다. 따라서 \mathbf{u}_i 들, 즉 $A\mathbf{v}_i$ 들은 orthogonal하다.

Lemma 5.2 *The left singular vectors \mathbf{u}_i 's of a matrix A defined as, for each right-singular vector \mathbf{v}_i ,*

$$\mathbf{u}_i = \frac{1}{|A\mathbf{v}_i|} A\mathbf{v}_i$$

are orthogonal to each other.

5.3 Representation of SVD in Various Forms

$\sigma_1, \mathbf{v}_1, \mathbf{u}_1$ 가 $A\mathbf{v}_1 = \sigma_1\mathbf{u}_1$ 를 만족하는 singular triplet이라고 하자. 이 정보 관점에서 A 를 어떻게 표현할 수 있을까? 간단하면서 singular triplet으로만 표현된 행렬 A_1 이 singular value relation $A_1\mathbf{v}_1 = \sigma_1\mathbf{u}_1$ 을 만족 할 수 있을까? 행렬이 간단하다는 것을 rank가 적다는 것으로 판단하는 것도 한가지 방법이다. 그러면 rank 가 1인 A_1 이 존재할까? 의외로 간단하게

$$A_1 = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top$$

로 적으면 $A_1\mathbf{v}_1 = \sigma_1\mathbf{u}_1$ 을 만족한다. 물론 A_1 은 rank-one이다. 행렬 A 의 다른 singular triplet을 $\sigma_2, \mathbf{v}_2, \mathbf{u}_2$ 라 하자. \mathbf{v}_1 과 \mathbf{v}_2 는 orthonormal하고 \mathbf{u}_1 과 \mathbf{u}_2 도 orthonormal하다. 두 번째 triplet에 대해서도 rank-one 행렬 $\sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top$ 은 singular value relation을 만족한다. 두 rank-one 행렬을 더하여

$$A_2 = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top$$

로 놓자. 그러면 $A_2\mathbf{v}_1 = \sigma_1\mathbf{u}_1$ 이고 $A_2\mathbf{v}_2 = \sigma_2\mathbf{u}_2$ 가 됨을 orthonormality로부터 쉽게 알 수 있다. 그리고 A_2 는 rank가 2임을 Fact 5.4에서 알 수 있다. 이를 얻어진 rank A 개 만큼의 singular triplet에 대해 확장하여

$$A = \sum_{i=1}^{\text{rank } A} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$$

가 될 것으로 추측할 수 있다. 이 표현을 포함한 여러 성질을 살펴보자.

$n \times d$ 행렬 A 에 대해 $\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = 0$ 이라 하자.

• $A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$

$\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ 를 포함하는 \mathbb{R}^d 의 기저 $\{\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_d\}$ 를 생각하자. 기저이므로 기저벡터들은 모두 orthonormal하다. \mathbb{R}^d 의 임의의 벡터를 $\mathbf{x} = x_1 \mathbf{v}_1 + \dots + x_d \mathbf{v}_d$ 로 표현하자. 그러면, $\sigma_{r+1} = 0$ 이므로 $j = r+1, \dots, d$ 에 대해 $A\mathbf{v}_j = \mathbf{0}$ 이고, 모든 i 에 대해 $x_i = \mathbf{v}_i^\top \mathbf{x}$ 이므로

$$A\mathbf{x} = x_1 A\mathbf{v}_1 + \dots + x_r A\mathbf{v}_r = x_1 \sigma_1 \mathbf{u}_1 + \dots + x_r \sigma_r \mathbf{u}_r = \left(\sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \right) \mathbf{x}$$

가 \mathbb{R}^d 의 모든 벡터 \mathbf{x} 에 대해 성립하고 따라서 $A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$ 로 표현된다.

• $A = UDV^\top$ or $AV = UD$

여기서, $U^\top U = I_r = V^\top V$, D 는 양의 대각행렬이다.

$$\begin{bmatrix} A \\ n \times d \end{bmatrix} = \begin{bmatrix} U \\ n \times r \end{bmatrix} \begin{bmatrix} D \\ r \times r \end{bmatrix} \begin{bmatrix} V^\top \\ r \times d \end{bmatrix}.$$

Lemma 3.5을 행렬 UD 와 V 에 대해 적용해 보자. D 가 $\sigma_1, \dots, \sigma_r$ 를 대각원소로 갖는 대각행렬이므로 UD 는 각 열이 $\sigma_i \mathbf{u}_i$ 인 행렬이므로 (3.5)에 의해 $UDV^\top = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$ 이 된다. 즉 $A = UDV^\top$ 이다.

Singular Value Decomposition (SVD) Any $n \times d$ matrix A (with $r = \text{rank } A$) has a representation of

$$A = UDV^\top = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \quad (5.7)$$

where $n \times r$ matrix U and $d \times r$ matrix V are orthogonal matrices, and $D = \text{diag}(\sigma_1, \dots, \sigma_r)$ is a diagonal matrix with diagonal elements $\sigma_1 \geq \dots \geq \sigma_r > 0$. \mathbf{u}_i and \mathbf{v}_i are i -th column vectors of U and V , respectively. Note that $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ and $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ are orthonormal sets of vectors, respectively.

- $A^\top \mathbf{u}_i = \sigma_i \mathbf{v}_i$: 행렬 A 와 A^\top 는 동일한 singular value를 갖고, A 의 right(left) singular vector는 A^\top 의 left(right) singular vector가 된다. 이유는 $A^\top = \sum_{j=1}^r \sigma_j \mathbf{v}_j \mathbf{u}_j^\top$ 이므로 $A^\top \mathbf{u}_i = \sum_{j=1}^r \sigma_j \mathbf{v}_j \mathbf{u}_j^\top \mathbf{u}_i = \sigma_i \mathbf{v}_i$ 이기 때문이다.

- $AA^\top \mathbf{u}_i = \sigma_i^2 \mathbf{u}_i$, $A^\top A \mathbf{v}_i = \sigma_i^2 \mathbf{v}_i$: $(AA^\top)\mathbf{u}_i = A(A^\top \mathbf{u}_i) = A(\sigma_i \mathbf{v}_i) = \sigma_i A\mathbf{v} = \sigma_i^2 \mathbf{u}_i$ 이므로 $(\sigma_i^2, \mathbf{u}_i)$ 는 행렬 AA^\top 의 eigenpair이고, 마찬가지로 $(\sigma_i^2, \mathbf{v}_i)$ 는 행렬 $A^\top A$ 의 eigenpair이다.

- $A^\top = \sum_{i=1}^r \sigma_i \mathbf{v}_i \mathbf{u}_i^\top$
- $A^\top = VDU^\top$ 단, $U^\top U = I_r$, $V^\top V = I_r$, $D : r \times r$ 양의 원소를 갖는 대각행렬
- $A^\top \mathbf{u}_i = \sigma_i \mathbf{v}_i$
- $AA^\top \mathbf{u}_i = \sigma_i^2 \mathbf{u}_i$, $A^\top A \mathbf{v}_i = \sigma_i^2 \mathbf{v}_i$

특히 첫 번째 singular vector를 구하는 문제를 잘 살펴보면 다음 결과를 알 수 있다.

Fact 5.1 $\|A\|_2 = \sigma_1$

Proof: 첫 번째 singular vector를 구하는 최대화 문제 (5.3)과 spectral norm을 측정하는 최대화 문제 (4.13)가 동일하므로 $\sigma_1 = |A\mathbf{v}_1| = \|A\|_2$ 를 얻는다. ■

5.4 Properties of $\sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top$

$\alpha_1 \geq \dots \geq \alpha_k > 0$ 인 실수들과 orthonormal한 $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \subset \mathbb{R}^n$ 과 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subset \mathbb{R}^d$ 에 대해 rank-one 행렬 $\alpha_i \mathbf{u}_i \mathbf{v}_i^\top$ 들의 합으로 정의된 행렬

$$\sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top \quad (5.8)$$

을 생각하자. 이 행렬은 SVD와 관계없이 만들어졌을 수도 있다. 이 행렬에 대해서 다음 결과들을 얻는다.

Fact 5.2 $(\alpha_i, \mathbf{v}_i, \mathbf{u}_i)$ is i -th singular triplet of $\sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top$.

Proof: $A = \sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top$ 로 표시하자. 그리고 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 를 orthogonal하게 확장하여 \mathbb{R}^d 의 orthonormal basis인 $\{\mathbf{v}_1, \dots, \mathbf{v}_d\}$ 를 만들자. $\mathbf{v} \in \mathbb{R}^d$ 이고 $|\mathbf{v}| = 1$ 인 벡터를 $\mathbf{v} = x_1 \mathbf{v}_1 + \dots + x_d \mathbf{v}_d$ 로 표시하면

$$A\mathbf{v} = \left(\sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top \right) \mathbf{v} = \sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top \mathbf{v} = \sum_{i=1}^k \alpha_i x_i \mathbf{u}_i, \quad |A\mathbf{v}|^2 = \sum_{i=1}^k \alpha_i^2 x_i^2$$

이므로 $|\mathbf{v}| = 1$ 을 만족하기 위해서 추가되는 $x_1^2 + \dots + x_d^2 = 1$ 제약 하에서 $|A\mathbf{v}|$ 을 최대로 하는 x_i 값은 $x_1 = 1, x_2 = \dots = x_d = 0$, 즉 $\mathbf{v} = \mathbf{v}_1$ 이 된다. 따라서 A 의 first right singular vector는 \mathbf{v}_1 이 된다. 그리고 $A\mathbf{v}_1 = \alpha_1 \mathbf{u}_1$ 이고 $\alpha_1 = |A\mathbf{v}_1|$ 도 성립하여 α_1 이 first singular value, \mathbf{u}_1 이 first left singular vector가 된다. 두 번째 singular vector에 대해서는 $\langle \mathbf{v}, \mathbf{v}_1 \rangle = 0$ 을 추가하면 $\mathbf{v} = x_2 \mathbf{v}_2 + \dots + x_d \mathbf{v}_d$ 인 \mathbf{v} 만을 고려하므로 $\alpha_2, \mathbf{v}_2, \mathbf{u}_2$ 가 second singular value와 벡터들이 된다. 그 다음 순서의 singular value와 벡터들도 마찬가지로 얻을 수 있다. ■

Fact 5.3 $\left\| \sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top \right\|_2 = \alpha_1$, $\left\| \sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top \right\|_F = \sqrt{\alpha_1^2 + \dots + \alpha_k^2}$.

Proof: Fact 5.1과 Fact 5.2에 의해 성립한다. ■

Fact 5.4 $\text{rank} \left(\sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top \right) = k$.

Proof: $A = \sum_{i=1}^k \alpha_i \mathbf{u}_i \mathbf{v}_i^\top$ 의 null space를 알아보자. A 는 $n \times d$ 행렬이므로 $\mathbf{v} \in \mathbb{R}^d$ 인 벡터가 $A\mathbf{v} = \mathbf{0}$ 인 경우를 찾으면 된다. 적절한 방법으로 orthonormal한 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 를 확장하여 $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_d\}$ 가 \mathbb{R}^d 의 기저라 하자. 임의의 $\mathbf{v} \in \mathbb{R}^d$ 는 $(x_1, \dots, x_d)^\top \in \mathbb{R}^d$ 에 대해서 $\mathbf{v} = \sum_{i=1}^d x_i \mathbf{v}_i$ 로 표현되어

$$A\mathbf{v} = \sum_{i=1}^d x_i A\mathbf{v}_i = \sum_{i=1}^d x_i \left(\sum_{j=1}^k \alpha_j \mathbf{u}_j \mathbf{v}_j^\top \right) \mathbf{v}_i = \sum_{j=1}^k \sum_{i=1}^d x_i \alpha_j \mathbf{u}_j \mathbf{v}_j^\top \mathbf{v}_i = \sum_{i=1}^k x_i \alpha_i \mathbf{u}_i$$

이므로 $A\mathbf{v} = \mathbf{0}$ 이기 위한 필요충분조건은 $x_1 = \dots = x_k = 0$ 이다. 즉 $\text{Null}(A) = \text{span}\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_d\}$ 이어서 $\dim \text{Null}(A) = d - k$ 이다. 그러면 fundamental theorem of linear algebra (3.4)에 의해

$$\text{rank } A = d - \dim \text{Null}(A) = k$$

이다. ■

Fact 5.5 For any matrix A , $\|A\|_2 \leq \|A\|_F$. If $\text{rank}(A) = r$, $\|A\|_F \leq \sqrt{r} \|A\|_2$. For rank-one matrices, two norms coincide.

Proof: $\|A\|_2 \leq \|A\|_F$ 은 이미 Fact 4.17에서 증명되었으나 SVD를 이용한 간단한 증명을 추가로 생각해보자. $n \times d$ 행렬 A 의 SVD가 $A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$ 로 표현된다고 하자. Fact 5.3에 의해

$$\|A\|_2 = \sigma_1 \leq \sqrt{\sigma_1^2 + \dots + \sigma_r^2} = \|A\|_F$$

이다. 또한

$$\|A\|_F = \sqrt{\sigma_1^2 + \dots + \sigma_r^2} \leq \sqrt{r\sigma_1^2} = \sqrt{r} \sigma_1 = \sqrt{r} \|A\|_2$$

이 된다. ■

Fact 5.6 Let A be an $n \times n$ invertible matrix. Assume that an SVD of A is given as

$$U\Sigma V^\top = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^\top.$$

Then all $\sigma_i > 0$ and

$$A^{-1} = V\Sigma^{-1}U^\top = \sum_{j=1}^n \sigma_j^{-1} \mathbf{v}_j \mathbf{u}_j^\top. \quad (5.9)$$

Proof: Invertibility of A implies $\text{rank } A = n$. So Fact 5.4 implies all $\sigma_i > 0$. Then, Σ^{-1} and $\sum_{j=1}^n \sigma_j^{-1} \mathbf{v}_j \mathbf{u}_j^\top$ are well-defined. Furthermore, U and V are orthogonal matrices. So

$$A^{-1} = (U\Sigma V^\top)^{-1} = (V^\top)^{-1} \Sigma^{-1} U^{-1} = V\Sigma^{-1} U^\top = \sum_{j=1}^n \sigma_j^{-1} \mathbf{v}_j \mathbf{u}_j^\top.$$

■

Section 5.8에서 정사각이 아니거나 invertible하지 않은 일반적인 행렬에 대해서 역행렬과 유사한 pseudoinverse를 도입하는데 (5.9)와 매우 비슷하게 표현된다.

Example 5.1 Let A be an $n \times n$ invertible matrix. If an SVD of A is given as $\sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$, where $\sigma_1 \geq \dots \geq \sigma_n > 0$. Let us find $\|A^{-1}\|_2 \|A\|_2$. Facts 5.3 and 5.6 imply $\|A\|_2 = \sigma_1$ and $\|A^{-1}\|_2 = \sigma_n^{-1}$. Therefore, $\|A^{-1}\|_2 \|A\|_2 = \sigma_1 \sigma_n^{-1}$. ■

$\|A^{-1}\|_2 \|A\|_2$ is called the *condition number* of a matrix A and denoted by $\kappa(A)$, which is important in numerical linear algebra. If you are interested in this concept, we recommend that you consult any numerical analysis book.

5.5 Spectral Decomposition of Symmetric Matrices

Rank가 r 인 $n \times n$ 대칭행렬 A 의 singular value가 $\sigma_1 \geq \dots \geq \sigma_r > 0 = \sigma_{r+1}$ 이라 하자. 대칭행렬이므로 right-/left-singular vector들이 모두 \mathbb{R}^n -벡터이다. 처음 $(j-1)$ 개의 singular triplet이 알려졌다고 가정하고 이를 $(\sigma_1, \mathbf{v}_1, \mathbf{u}_1), \dots, (\sigma_{j-1}, \mathbf{v}_{j-1}, \mathbf{u}_{j-1})$ 으로 표시하자. 특히 이미 얻은 singular vector들이 $\mathbf{u}_i = \mathbf{v}_i$ 또는 $-\mathbf{v}_i$ 라고 가정하자. Greedy SVD procedure (5.5)의 j 번째 최적화 문제를 풀어서 얻어진 A 의 singular triplet이 $(\sigma_j, \mathbf{v}, \mathbf{u})$ 라고 하자. \mathbf{v} 는 (5.5)에 feasible하므로 제약식인 $\mathbf{v} \perp \{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}\}$ 를 만족하고 Lemma 5.2에 의해 $\mathbf{u} \perp \{\mathbf{u}_1, \dots, \mathbf{u}_{j-1}\}$ 이다. 특히 $\mathbf{u}_i = \pm \mathbf{v}_i$ 를 가정하였으므로

$$(\mathbf{v} + \mathbf{u}) \perp \{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}\} \quad (5.10)$$

이다. $A = A^\top$ 이므로 $A\mathbf{v} = \sigma_j \mathbf{u}$ 이자 동시에 $A\mathbf{u} = \sigma_j \mathbf{v}$ 이다. 다음 두 가지 경우가 가능하다:

- $\mathbf{v} + \mathbf{u} \neq \mathbf{0}$: $\mathbf{v}_j = \frac{1}{|\mathbf{v} + \mathbf{u}|}(\mathbf{v} + \mathbf{u})$ 으로 놓으면 $A\mathbf{v}_j = \sigma_j \mathbf{v}_j$ 이다. \mathbf{v}_j 는 $|A\mathbf{v}_j| = \sigma_j$ 이면서 (5.5)에 feasible하므로 (5.5)의 optimal solution이다. 따라서 \mathbf{v}_j 는 j -번째 right-singular vector이고 left-singular vector는 $\mathbf{u}_j = \mathbf{v}_j$ 이다;
- $\mathbf{v} + \mathbf{u} = \mathbf{0}$: $\mathbf{v}_j = \mathbf{v}$ 로 놓으면 \mathbf{v}_j 는 j -번째 right-singular vector이다. $\mathbf{v} = -\mathbf{u}$ 므로 $A\mathbf{v} = \sigma_j \mathbf{u} = -\sigma_j \mathbf{v}$ 이다. 그러면 left-singular vector는 $\mathbf{u}_j = -\mathbf{v}_j$ 이다.

따라서 어느 경우에나 $\mathbf{u}_j = \pm \mathbf{v}_j$ 인 j -번째 singular triplet $(\sigma_j, \mathbf{v}_j, \mathbf{u}_j)$ 를 얻는다. 이 변형된 procedure를 r 번 수행하면 모든 right-singular vector와 left-singular vector가 parallel한 r 개의 singular triplet을 얻는다. 그리고 $\mathbf{v}_i = \mathbf{u}_i$ 이면 $\lambda_i = \sigma_i$ 로 놓고 $\mathbf{v}_i = -\mathbf{u}_i$ 인 경우에는 $\lambda_i = -\sigma_i$ 로 놓으면 $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ 가 모든 i 에 대해 성립된다. 정사각행렬에 대해서 left singular vector와 right singular vector가 일치하는 특별한 경우에 singular value와 singular vector에 다음과 같이 특별한 이름을 붙인다.²

Definition 5.1 For a square matrix A , a scalar λ and a vector \mathbf{v} are **eigenvalue** and **eigenvector** of A , respectively, if the following vector equation

$$A\mathbf{v} = \lambda \mathbf{v} \quad (5.11)$$

holds. The pair (λ, \mathbf{v}) is called an **eigenpair**.

²Chapter 9에서 eigenvalue와 eigenvector에 관해 자세하게 공부하고 그 이전에는 (5.11)로 정의된 eigenvalue와 eigenvector의 관계만 필요하다.

위에서 유도한 바와 같이 left singular vector와 right singular vector가 일치하는 특별한 형태의 SVD를 통해 대칭행렬에 대한 eigenpair $(\lambda_1, \mathbf{v}_1), \dots, (\lambda_r, \mathbf{v}_r)$ 을 얻고, 이로부터 rank-one matrix의 합 형태 표현인

$$A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top = \sum_{i=1}^r \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$$

을 얻는다. r 개의 eigenvector를 Gram-Schmidt procedure를 통해 확장하여 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 을 얻고 $\lambda_{r+1} = \dots = \lambda_n = 0$ 으로 놓으면 n 개의 eigenpair $(\lambda_1, \mathbf{v}_1), \dots, (\lambda_n, \mathbf{v}_n)$ 를 얻게 되고 eigenvalue들이 singular value 이거나 부호만 바꾸면 되므로 모두 실수이다. 이렇게 얻어진 행렬의 분해를 **real spectral decomposition**이라고 부른다.

Theorem 5.2 (Real Spectral Decomposition) *Let A be a real symmetric matrix. Then, A can be represented as*

$$A = V \Lambda V^\top = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top, \quad (5.12)$$

where V is an orthogonal matrix with orthonormal columns $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, $|\mathbf{v}_i| = 1$ and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. λ_i 's and \mathbf{v}_i 's are eigenvalues and eigenvectors of A , respectively.

이 정리에 대한 SVD를 이용하지 않는 직접적인 증명이 Section 9.5에서 주어지는 데 수학적으로 매우 흥미로운 증명이므로 꼭 읽어보기를 권한다.

Linear transform 관점에서 유용한 대칭행렬의 표현을 얻어보자. Theorem 5.2에서 동일한 λ_i 에 대응되는 orthonormal한 eigenvector가 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 의 k 개일 때 $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 상으로의 정사영을 $\mathbf{P}_{A, \lambda_i}$ 로 나타내자. 즉

$$\mathbf{P}_{A, \lambda_i} = \sum_{j=1}^k \mathbf{v}_j \mathbf{v}_j^\top = \mathbf{P}_{\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}} \quad (5.13)$$

이다.³ 그러면 $\mathbf{P}_{A, \lambda_i}$ 는 projection이므로

- $\mathbf{P}_{A, \lambda_i}^2 = \mathbf{P}_{A, \lambda_i}$;
- $\mathbf{P}_{A, \lambda_i}^\top = \mathbf{P}_{A, \lambda_i}$;
- $i \neq j$ 일 때 $\mathbf{P}_{A, \lambda_i} \mathbf{P}_{A, \lambda_j} = \mathbf{0}$

가 성립한다. 따라서 서로 다른 r 개의 eigenvalue가 존재할 때 real spectral decomposition (5.12)는 다음과 같이 정사영의 가중 합인

$$A = \sum_{i=1}^r \lambda_i \mathbf{P}_{A, \lambda_i} \quad (5.14)$$

³ 이 표현에서는 linear transform과 이를 표현하는 행렬을 같은 것으로 취급하고 있음에 유의하자.

로도 표현된다. 이러한 분해의 간단한 예를 생각해보자.

$$\begin{aligned}
 A &= \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \\
 &= 2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
 &= 2 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^\top + 2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}^\top + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}^\top + 3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}^\top \\
 &= 2 \left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^\top + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}^\top \right) + 3 \left(\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}^\top + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}^\top \right) \\
 &= 2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
 &= 2 \mathbf{P}_{A,2} + 3 \mathbf{P}_{A,3}
 \end{aligned}$$

인데 3번째 등식에서 orthonormal 벡터를 다르게 놓더라도

$$\begin{aligned}
 A &= 2 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{bmatrix}^\top + 2 \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{bmatrix}^\top + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}^\top + 3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}^\top \\
 &= 2 \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
 &= 2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
 \end{aligned}$$

$$= 2 \mathbf{P}_{A,2} + 3 \mathbf{P}_{A,3}$$

로 정사영 표현은 동일하다.

반대 방향으로 대칭행렬의 eigenvalue들로부터 SVD를 얻는 방법을 살펴보자. $n \times n$ symmetric matrix A 의 모든 eigenvector들 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 이 orthonormal한 eigenpair $(\lambda_i, \mathbf{v}_i)$, $i = 1, \dots, n$ 을 알려졌다고 하자. 이 eigenpair로부터 singular value decomposition을 얻을 수 있다. Eigenvector들 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 이 orthonormal하므로 $A = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$ 으로 표현된다. 그리고 $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_r| > 0 = \lambda_{r+1} = \dots = \lambda_n$ 로 n 개의 eigenvalue 중 r 개만 0이 아니라고 하자. 그러면, $\sigma_i = |\lambda_i|$, $\mathbf{u}_i = \text{sign}(\lambda_i) \mathbf{v}_i$, $i = 1, \dots, r$ 로 놓으면 $A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$ 로 표현되어 A 의 singular value decomposition을 얻게 된다. 다음은 대칭행렬에 관한 matrix norm의 성질들이다.

Fact 5.7 If a symmetric matrix A has eigenvalues $\lambda_1, \dots, \lambda_n$, then

$$\|A\|_2 = \max_{1 \leq i \leq n} |\lambda_i| = \max_{\|\mathbf{x}\|=1} |\mathbf{x}^\top A \mathbf{x}| \quad \text{and} \quad \|A\|_F = \sqrt{\lambda_1^2 + \dots + \lambda_n^2}.$$

Proof: $A = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$ 로 표시하자. 그리고 편의상 $|\lambda_1| = \max_{1 \leq i \leq n} |\lambda_i|$ 이라고 가정하자. 크기가 1인 \mathbf{x} 는 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 이 orthonormal하므로 $\sum_{i=1}^n x_i^2 = 1$ 인 x_i 가 존재하여 $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{v}_i$ 이다. 그러면, $A\mathbf{x} = \sum_{i=1}^n x_i \lambda_i \mathbf{v}_i$ 이고 $|A\mathbf{x}|^2 = \sum_{i=1}^n \lambda_i^2 x_i^2 \leq \sum_{i=1}^n \lambda_1^2 x_i^2 \leq \lambda_1^2$ 이다. 따라서 $\|A\|_2 \leq |\lambda_1|$ 이 성립한다. 또한, $|\mathbf{x}| = 1$ 이라면 $|\mathbf{x}^\top A \mathbf{x}| \leq \|\mathbf{x}\| |A\mathbf{x}| = |A\mathbf{x}|$ 이므로 $\max_{\|\mathbf{x}\|=1} |\mathbf{x}^\top A \mathbf{x}| \leq \|A\|_2$ 이다. 그런데 $\max_{\|\mathbf{x}\|=1} |\mathbf{x}^\top A \mathbf{x}| \geq |\mathbf{v}_1^\top A \mathbf{v}_1| = |\lambda_1|$ 이므로 첫 번째 등식관계가 성립한다. $\|A\|_F = \|V \Lambda V^\top\|_F = \|\Lambda\|_F = \sqrt{\lambda_1^2 + \dots + \lambda_n^2}$. ■

5.6 Relationship between Singular Values and Eigenvalues

실수 성분으로 이루어진 asymmetric matrix를 고려하자. 정사각행렬이 아니거나 또는 정사각행렬이더라도 대칭 행렬이 아니면 eigenvalue가 복소수이거나 독립인 eigenvector가 충분하게 존재하지 않을 수 있다. 이에 반해 SVD는 모든 $n \times d$ 행렬에 대해 rank 만큼의 singular triplet이 존재한다. 또한 같은 행렬에 대해 singular value는 늘 양수이지만 대칭행렬을 포함한 모든 행렬은 eigenvalue가 음수일 수도 있다. 따라서 두 가지 분해가 동일하지는 않다. 하지만 Section 5.5에서 보였듯이 SVD를 찾는 greedy procedure (5.5)를 약간 변경하여 대칭행렬에 대한 특별한 형태의 SVD를 얻으면 대칭행렬의 eigenpair를 singular triplet에 대응시킬 수 있다. 실제 응용에서 제일 흔한 경우로 정사각행렬이 아니거나 대칭행렬이 아닌 A 에 대해서 $A^\top A$ 또는 AA^\top 의 eigenpair와 A 의 singular triplet 간의 대응관계를 활용할 수 있다.

Singular triplets of A and eigenpairs of $A^\top A$

A 를 rank가 r 인 $n \times d$ 행렬이라 하자. A 의 singular triplet $(\sigma, \mathbf{v}, \mathbf{u})$ 에 대해서 $A\mathbf{v} = \sigma\mathbf{u}$ 와 $A^\top \mathbf{u} = \sigma\mathbf{v}$ 가 성립하여 이미 살펴본 바와 같이 (σ^2, \mathbf{v}) 와 (σ^2, \mathbf{u}) 는 각각 $A^\top A$ 와 AA^\top 의 eigenpair가 된다.

반대로 $A^\top A$ 의 eigenpair (λ, \mathbf{v}) , $\lambda \neq 0$, $|\mathbf{v}| = 1$ 을 생각해 보자. $A^\top A\mathbf{v} = \lambda\mathbf{v}$ 이므로 양변의 좌측에 \mathbf{v}^\top 을 곱하면 $|\mathbf{v}^\top A\mathbf{v}|^2 = \lambda|\mathbf{v}|^2 = \lambda$ 가 되어 $\lambda \neq 0$ 이라면 $\lambda > 0$ 임을 알 수 있다. $\mathbf{u} = \frac{1}{\sqrt{\lambda}} A\mathbf{v}$ 로 놓으면 $|\mathbf{u}| = 1$ 인 것을 직전의 계산으로부터 알 수 있다. 그리고 $A^\top \mathbf{u} = \sqrt{\lambda}\mathbf{v}$ 가 성립한다. Spectral decomposition 정리 Theorem

5.2로부터 대칭행렬인 $A^\top A$ 의 orthonormal한 eigenvector들을 얻을 수 있다. 따라서 $A^\top A$ 의 모든 eigenpair들이 orthonormal하다고 가정하여도 된다. $(\bar{\lambda}, \bar{\mathbf{v}})$ 가 $\bar{\mathbf{v}}^\top \mathbf{v} = 0$ 인 다른 eigenpair라 하자. 이 eigenpair에 대해 동일하게 $\bar{\mathbf{u}} = \frac{1}{\sqrt{\bar{\lambda}}} A\bar{\mathbf{v}}$ 로 정의하면

$$\bar{\mathbf{u}}^\top \mathbf{u} = \frac{1}{\sqrt{\bar{\lambda}\lambda}} \bar{\mathbf{v}}^\top A^\top A \mathbf{v} = \frac{\sqrt{\bar{\lambda}}}{\sqrt{\lambda}} \bar{\mathbf{v}}^\top \mathbf{v} = 0$$

이 되어 \mathbf{u} 도 orthonormal한 벡터들이다. Rank가 r 이므로 r 개의 양수 λ_i 와 orthonormal한 \mathbf{v}_i 와 \mathbf{u}_i 로 구성된 triplet $(\sqrt{\lambda_i}, \mathbf{v}_i, \mathbf{u}_i)$ 를 얻을 수 있다. 이 triplet으로부터 유도되는 rank-one행렬의 합 $\sum_{i=1}^r \sqrt{\lambda_i} \mathbf{u}_i \mathbf{v}_i^\top$ 와 행렬 A 는 임의의 벡터를 행렬의 우측에 곱할 때 동일한 벡터값을 갖으므로

$$A = \sum_{i=1}^r \sqrt{\lambda_i} \mathbf{u}_i \mathbf{v}_i^\top$$

가 되고 r 개의 $(\sqrt{\lambda_i}, \mathbf{v}_i, \mathbf{u}_i)$ 가 A 의 singular triplet이 된다. 많은 상용 소프트웨어들은 singular value를 AA^\top 이나 $A^\top A$ 의 eigenvalue계산을 통해서 구한다. 하지만, singular vector와 eigenvector가 일의적으로 결정되지 않으므로 구해진 분해도 일의적으로 결정되지는 않는다.

Lemma 5.3 Assume that A is a real matrix of arbitrary size. Then the square of any singular value and right singular vector of A is an eigenpair of $A^\top A$, and the left singular vector is an eigenvector of AA^\top . Conversely, if $A^\top A$ admits eigenvalues and orthonormal eigenvectors, then the square roots of eigenvalues are singular values of A , and the eigenvectors form right singular vectors.

5.6.1 Symmetrization

$m \times n$ 행렬 A 의 symmetrization $s(A)$ 를 $(m+n) \times (m+n)$ 행렬

$$s(A) = \begin{bmatrix} \mathbf{0} & A \\ A^\top & \mathbf{0} \end{bmatrix} \quad (5.15)$$

로 정의하자. 이름이 의미 하듯이 $s(A)$ 는 대칭 행렬이다. 또한 $s(A+B) = s(A) + s(B)$ 로 symmetrization이 행렬의 덧셈을 보존하는 것을 쉽게 알 수 있다. Symmetrization을 통해서 크기가 커진 행렬에 벡터를 맞추기 위해 \mathbb{R}^m -vector \mathbf{u} 와 \mathbb{R}^n -vector \mathbf{v} 에 대해서 $\mathbf{w} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$ 와 $\hat{\mathbf{w}} = \begin{bmatrix} -\mathbf{u} \\ \mathbf{v} \end{bmatrix}$ 로 놓자. 그러면

- $\sigma > 0$, $|\mathbf{u}| = 1$, $|\mathbf{v}| = 1$ 에 대해서 $(\sigma, \mathbf{v}, \mathbf{u})$ 가 A 의 singular triplet일 때

$$s(A)\mathbf{w} = \begin{bmatrix} A\mathbf{v} \\ A^\top \mathbf{u} \end{bmatrix} = \begin{bmatrix} \sigma\mathbf{u} \\ \sigma\mathbf{v} \end{bmatrix} = \sigma\mathbf{w}$$

이므로 σ 가 symmetrization $s(A)$ 의 eigenvalue가 되고 \mathbf{w} 가 대응되는 eigenvector이다.

- 반대로 $\lambda \neq 0$ 일 때 (λ, \mathbf{w}) 가 $s(A)$ 의 eigenpair라 하자. 그러면

$$s(A)\mathbf{w} = \begin{bmatrix} A\mathbf{v} \\ A^\top \mathbf{u} \end{bmatrix} = \lambda \mathbf{w} = \begin{bmatrix} \lambda \mathbf{u} \\ \lambda \mathbf{v} \end{bmatrix}$$

이어서 $A\mathbf{v} = \lambda \mathbf{u}$ 와 $A^\top \mathbf{u} = \lambda \mathbf{v}$ 가 동시에 성립한다. 이 때 eigenvector \mathbf{w} 를 구성하는 두 벡터는 영벡터가 아니다. 만약 \mathbf{v} 나 \mathbf{u} 가 영벡터이면 나머지 한 벡터도 영벡터이어야 하므로 $\mathbf{w} = \mathbf{0}$ 이 되어 모순이다. 따라서 eigenvalue λ^2 에 대해서 \mathbf{v} 가 $A^\top A$ 의 eigenvector이므로 Lemma 5.3에 의해 $|\lambda|$ 이 A 의 singular value가 된다.

- 또한 (λ, \mathbf{w}) 가 $s(A)$ 의 eigenpair인 경우 $A\mathbf{v} = \lambda \mathbf{u}$ 이고 $A^\top \mathbf{u} = \lambda \mathbf{v}$ 이므로

$$s(A)\hat{\mathbf{w}} = \begin{bmatrix} A\mathbf{v} \\ A^\top(-\mathbf{u}) \end{bmatrix} = \begin{bmatrix} (-\lambda)(-\mathbf{u}) \\ (-\lambda)\mathbf{v} \end{bmatrix} = (-\lambda)\hat{\mathbf{w}}$$

이어서 $(-\lambda, \hat{\mathbf{w}})$ 도 $s(A)$ 의 eigenpair가 된다.

이상을 정리하면 다음과 같은 결과를 얻는다.

Lemma 5.4 Any singular value of a matrix A is an eigenvalue of its symmetrization $s(A)$. The symmetrization $s(A)$ has its eigenvalues in both signs; that is, if λ is an eigenvalue of $s(A)$, then both $\pm|\lambda|$ are eigenvalues of $s(A)$. Hence, a one-to-one correspondence exists between the singular values of A and the positive eigenvalues of $s(A)$.

Lemma 5.4는 대칭행렬의 eigenvalue에 대해 얻어진 결과를 임의의 행렬의 singular value에 대한 결과로 바꿀 때 유용하다.

5.7 Low rank approximation

행렬 A 의 singular value decomposition

$$A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top, \quad \sigma_1 \geq \dots \geq \sigma_r > 0 = \sigma_{r+1}$$

라 하자. 합으로 표현된 행렬의 표현에서 제일 큰 k 개의 singular value에 대응하는 left/right singular vector 들로 구성된 rank-one 행렬 만을 더하여

$$A_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$$

로 표시하자. 그러면, $A - A_k = \sum_{i=k+1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$ 이다. 이 경우 Fact 5.4에 의해 A 의 rank는 r , A_k 의 rank는 k 이고 $A - A_k$ 의 rank는 $r - k$ 이다.

Low Rank Approximation: Spectral Norm

Fact 5.8 rank k 이하인 임의의 행렬 B 에 대해서 $\|A - A_k\|_2 \leq \|A - B\|_2$ 이다

Proof:

- $A - A_k = \sum_{i=k+1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$, $\sigma_{k+1} \geq \dots \geq \sigma_r > 0$ 이므로 Fact 5.3에 의해

$$\|A - A_k\|_2 = \sigma_{k+1}$$

$k < d - 1$ 일 때 성립한다.

- B 의 rank가 k 이하이므로 B 의 null space의 rank는 $d - k$ 이상이다. $\{\mathbf{v}_1, \dots, \mathbf{v}_{k+1}\}$ 를 A 의 제일 큰 $k + 1$ 개의 singular value의 짝인 right singular vector라 하자. 그러면 $(d - k) + (k + 1) > d$ 이므로 $\text{Null}(B) \cap \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k+1}\}$ 의 차원이 1 이상이고, 따라서 0이 아닌 단위벡터 \mathbf{z} 가 교집합에 포함된다. 그러면 $B\mathbf{z} = \mathbf{0}$ 이고, $\mathbf{z} = \sum_{j=1}^{k+1} \langle \mathbf{z}, \mathbf{v}_j \rangle \mathbf{v}_j$ 으로

$$A\mathbf{z} = \left(\sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \right) \left(\sum_{j=1}^{k+1} \langle \mathbf{z}, \mathbf{v}_j \rangle \mathbf{v}_j \right) = \sum_{i=1}^{k+1} \sigma_i \langle \mathbf{z}, \mathbf{v}_i \rangle \mathbf{u}_i$$

이다. 따라서

$$\begin{aligned} \|A - B\|_2 &\geq |(A - B)\mathbf{z}| && (\text{spectral norm 정의에 의해서}) \\ &= |A\mathbf{z}| && (\text{by } B\mathbf{z} = \mathbf{0}) \\ &= \left(\sum_{i=1}^{k+1} \sigma_i^2 \langle \mathbf{z}, \mathbf{v}_i \rangle^2 \right)^{1/2} && \left(\text{by } A\mathbf{z} = \sum_{i=1}^{k+1} \sigma_i \langle \mathbf{z}, \mathbf{v}_i \rangle \mathbf{u}_i \right) \\ &\geq \sigma_{k+1} \left(\sum_{i=1}^{k+1} \langle \mathbf{z}, \mathbf{v}_i \rangle^2 \right)^{1/2} && (\text{by } \sigma_1 \geq \dots \geq \sigma_k \geq \sigma_{k+1}) \\ &= \sigma_{k+1} && \left(\text{by } \sum_{i=1}^{k+1} \langle \mathbf{z}, \mathbf{v}_i \rangle^2 = |\mathbf{z}|^2 = 1 \right) \\ &= \|A - A_k\|_2. \end{aligned}$$

즉 rank k 이하인 임의의 행렬 B 에 대해서 $\|A - A_k\|_2 \leq \|A - B\|_2$ 가 성립한다. ■

Low Rank Approximation: Frobenius Norm

Fact 5.9 rank k 이하인 임의의 행렬 B 에 대해서 $\|A - A_k\|_F \leq \|A - B\|_F$ 이다

Proof: B 가 rank k 이하인 $n \times d$ 행렬 중 $\|A - B\|_F$ 를 최소로 한다고 가정하고 행렬 B 의 행들을 $\mathbf{b}_1, \dots, \mathbf{b}_n$ 라 하자. 행렬 A 의 i 번째 행 \mathbf{a}_i 의 부분공간 $\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ 로의 정사영이 행렬 B 의 i 번째 행 \mathbf{b}_i 와 같지 않다고 하자. 즉 $\mathbf{P}_{\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}}(\mathbf{a}_i) \neq \mathbf{b}_i$ 이라 하자. 그러면, 정사영의 유일성과 최소 거리 성질에 의해

$|\mathbf{a}_i - \mathbf{b}_i| > |\mathbf{a}_i - \mathbf{P}_{\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}}(\mathbf{a}_i)|$ 이고, $\|A - B\|_F^2 = \sum_{i=1}^n |\mathbf{a}_i - \mathbf{b}_i|^2$ 이므로 행렬 B 의 i 번째 행을 \mathbf{b}_i 에서 $\mathbf{P}_{\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}}(\mathbf{a}_i)$ 로 바꾼 행렬을 B' 이라 하면 $\|A - B\|_F > \|A - B'\|_F$ 이다. 그런데 $\mathbf{P}_{\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}}(\mathbf{a}_i)$ 는 B 의 행들의 조합으로 표시되므로 B' 의 rank는 B 의 rank보다 크지 않다. 따라서 Frobenius norm이 최소라는 가정에 모순이므로 $\mathbf{b}_i = \mathbf{P}_{\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}}(\mathbf{a}_i)$ 이다. 즉

$$\min_{B: \text{rank}(B) \leq k} \|A - B\|_F^2 = \min_{B: \text{rank}(B) \leq k} \sum_{i=1}^n |\mathbf{a}_i - \mathbf{P}_{\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}}(\mathbf{a}_i)|^2$$

이 성립한다. 그런데 rank가 k 이하인 행렬 B 에 대한 $\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ 를 찾는 것은 차원이 k 이하인 부분 공간을 찾는 것과 동일하여

$$\min_{B: \text{rank}(B) \leq k} \sum_{i=1}^n |\mathbf{a}_i - \mathbf{P}_{\text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}}(\mathbf{a}_i)|^2 = \min_{\mathbb{B}: \dim(\mathbb{B}) \leq k} \sum_{i=1}^n |\mathbf{a}_i - \mathbf{P}_{\mathbb{B}}(\mathbf{a}_i)|^2$$

이 되며 부분공간은 기저로 규정되므로

$$\min_{\mathbb{B}: \dim(\mathbb{B}) \leq k} \sum_{i=1}^n |\mathbf{a}_i - \mathbf{P}_{\mathbb{B}}(\mathbf{a}_i)|^2 = \min_{\substack{\mathbf{v}_1, \dots, \mathbf{v}_k: \\ \text{orthonormal}}} \sum_{i=1}^n |\mathbf{a}_i - \mathbf{P}_{\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}}(\mathbf{a}_i)|^2$$

가 성립함을 알 수 있다. 이 등식을 모두 연결하면

$$\begin{aligned} \min_{B: \text{rank}(B) \leq k} \|A - B\|_F^2 &= \min_{\substack{\mathbf{v}_1, \dots, \mathbf{v}_k: \\ \text{orthonormal}}} \sum_{i=1}^n |\mathbf{a}_i - \mathbf{P}_{\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}}(\mathbf{a}_i)|^2 \\ &= \min_{\substack{\mathbf{v}_1, \dots, \mathbf{v}_k: \\ \text{orthonormal}}} \sum_{i=1}^n |\mathbf{a}_i|^2 - |\mathbf{P}_{\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}}(\mathbf{a}_i)|^2 \\ &= \sum_{i=1}^n |\mathbf{a}_i|^2 - \max_{\substack{\mathbf{v}_1, \dots, \mathbf{v}_k: \\ \text{orthonormal}}} \sum_{i=1}^n |\mathbf{P}_{\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}}(\mathbf{a}_i)|^2 \\ &= \|A\|_F^2 - \max_{\substack{\mathbf{v}_1, \dots, \mathbf{v}_k: \\ \text{orthonormal}}} |A\mathbf{v}_1|^2 + \dots + |A\mathbf{v}_k|^2 \\ &= (\sigma_1^2 + \dots + \sigma_r^2) - (\sigma_1^2 + \dots + \sigma_k^2) \quad \text{by (5.6)} \\ &= \sigma_{k+1}^2 + \dots + \sigma_r^2 \\ &= \|A - A_k\|_F^2 \quad \text{by (5.6)} \end{aligned}$$

이 된다. 따라서 rank가 k 이하인 임의의 행렬 B 에 대하여 $\|A - A_k\|_F \leq \|A - B\|_F$ 이다. ■

5.8 Pseudoinverse

Fact 5.6에서 가역인 정방행렬의 역행렬을 (5.9)와 같이 SVD 형식으로 표현하였다. 본 절에서는 정사각행렬이 아니거나 invertible하지 않은 행렬도 포함한 일반적인 행렬에 대해서 역행렬과 비슷한 역할을 하는 pseudoinverse (또는, Moore-Penrose generalized inverse)라고 불리는 행렬을 (5.9)와 유사하게 정의한다.

Definition 5.2 (Pseudoinverse) Let A be an $n \times d$ matrix of rank r and $A = U\Sigma V^\top$ be the compact^a singular value decomposition of A . Then, the $d \times n$ matrix

$$A^+ = V\Sigma^{-1}U^\top \quad (5.16)$$

is the **pseudoinverse** of A . Note that $\Sigma^{-1} = \text{diag}\left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}\right)$ where $\sigma_1, \dots, \sigma_r$ are the non-zero singular values of A .

^a $n \times d$ 행렬 A 의 singular value decomposition $A = U\Sigma V^\top$ 은 unique하지 않다. $\text{rank}(A) = r$ 일 때, 각각 r 개의 left/right singular vector를 열로 갖는 U 와 V , 그리고 r 개의 singular value를 대각원소로 갖는 Σ 로 이루어진 경우에 특별히 *compact* singular value decomposition이라 부른다.

(5.16)과 (5.9)의 유사성을 확인해보기 바란다. U 와 V 행렬은 orthonormal한 열로 구성되어 있지만 정방행렬이 아닐 수도 있어서 우리가 이용할 수 있는 사실은 $U^\top U = I_r = V^\top V$ 뿐이다. 이로부터

$$AA^+ = U\Sigma V^\top V\Sigma^{-1}U^\top = UU^\top, \quad A^+A = V\Sigma^{-1}U^\top U\Sigma V^\top = VV^\top$$

이므로 각각 대칭행렬임을 알 수 있다. 그리고

$$AA^+A = UU^\top U\Sigma V^\top = U\Sigma V^\top = A, \quad A^+AA^+ = VV^\top V\Sigma^{-1}U^\top = V\Sigma^{-1}U^\top = A^+$$

도 쉽게 알 수 있다. 이 성질들을 모아서 Penrose identities라고 부른다.

Definition 5.3 (Penrose identities) An $n \times d$ matrix A and a $d \times n$ matrix B satisfy the **Penrose identities** if A and B satisfy the following identities:

- (a) $(AB)^\top = AB$ and $(BA)^\top = BA$;
- (b) $ABA = A$;
- (c) $BAB = B$.

임의의 행렬 A 에 대해 Penrose identities를 만족하는 행렬 B 는 유일함을 보일 수 있고, 이로부터 pseudoinverse의 정의와 동치라는 것을 보일 수 있다.

Fact 5.10 The matrix satisfying the Penrose identities is unique.

Proof: 행렬 A 에 대해서 B 와 C 가 Penrose identities를 만족한다고 가정하자. 세 가지 성질을 적절하게 사용하여 다음과 같이 첫 두 줄을 통해 BAB 를 BAC 로 바꾸고 비슷하게 그 다음 두 줄에서 BAC 를 CAC 로 바꾼다:

$$\begin{aligned} B &= BAB = B(AB)^\top = BB^\top A^\top = BB^\top(ACA)^\top = BB^\top A^\top(AC)^\top \\ &= B(AB)^\top AC = BABAC = (BAB)AC = BAC \end{aligned}$$

$$\begin{aligned}
 &= (BA)^\top C = A^\top B^\top C = (ACA)^\top B^\top C = (CA)^\top A^\top B^\top C \\
 &= (CA)^\top (BA)^\top C = CABAC = C(ABA)C = CAC \\
 &= C.
 \end{aligned}$$

즉 (a), (b), (c)를 모두 만족하는 행렬은 유일하다는 것을 알 수 있다. ■

Fact 5.11 A $d \times n$ matrix is a pseudoinverse of A if and only if the matrix satisfies Penrose identities.

Proof: “only if” 방향은 이미 보였다. 그리고 “only if” 방향에 의해 A^+ 가 Penrose identities를 만족하는데, Fact 5.10에 의해 Penrose identities를 만족하는 행렬이 유일하므로 Penrose identities를 만족하는 행렬은 pseudoinverse이다. ■

그리고 어떻게든 구해진 pseudoinverse를 마음 편하게 사용하기 위해서는 pseudoinverse의 유일성이 필요하다.

Theorem 5.3 Pseudoinverse is unique.

Proof: Fact 5.10와 5.11에 의해 증명된다. ■

Fact 5.12 If the rank of an $n \times d$ matrix A is d , then $A^+ = (A^\top A)^{-1} A^\top$.

Proof: $(A^\top A)^{-1} A^\top$ 이 Penrose identities를 만족하므로 Fact 5.11에 의해 pseudoinverse이다. ■

Pseudoinverse와 관련되어 여러 군데에서 유용한

$$A^\top (AA^+ - I) = V\Sigma U^\top (UU^\top - I) = \mathbf{0}$$

를 이용하면 임의의 벡터 \mathbf{x} 와 \mathbf{b} 에 대해

$$\begin{aligned}
 |A\mathbf{x} - \mathbf{b}|^2 &= |A\mathbf{x} - AA^+\mathbf{b} + AA^+\mathbf{b} - \mathbf{b}|^2 \\
 &= |A(\mathbf{x} - A^+\mathbf{b}) + (AA^+\mathbf{b} - \mathbf{b})|^2 \\
 &= |A(\mathbf{x} - A^+\mathbf{b})|^2 + 2(\mathbf{x} - A^+\mathbf{b})^\top A^\top (AA^+ - I)\mathbf{b} + |AA^+\mathbf{b} - \mathbf{b}|^2 \\
 &= |A(\mathbf{x} - A^+\mathbf{b})|^2 + 0 + |AA^+\mathbf{b} - \mathbf{b}|^2 \\
 &\geq |AA^+\mathbf{b} - \mathbf{b}|^2
 \end{aligned}$$

이다. 요약하면 임의의 행렬 A 와 임의의 벡터 \mathbf{b} 에 대해서

$$|A\mathbf{x} - \mathbf{b}| \geq |AA^+\mathbf{b} - \mathbf{b}| \quad \text{for all } \mathbf{x} \in \mathbb{R}^n \quad (5.17)$$

이 성립한다. $A^- = (A^+)^\top$ 라고 놓으면 비슷하게 $A^\top A^- = (A^+ A)^\top = VV^\top$ 으로

$$A(A^\top A^- - I) = U\Sigma V^\top (VV^\top - I) = \mathbf{0}$$

이다. 적절한 크기의 \mathbf{y} 와 \mathbf{c} 에 대해서

$$\begin{aligned} |A^\top \mathbf{y} - \mathbf{c}|^2 &= |A^\top(\mathbf{y} - A^- \mathbf{c}) + (A^\top A^- \mathbf{c} - \mathbf{c})|^2 \\ &= |A^\top(\mathbf{y} - A^- \mathbf{c})|^2 + 2(\mathbf{y} - A^- \mathbf{c})^\top A(A^\top A^- - I)\mathbf{c} + |A^\top A^- \mathbf{c} - \mathbf{c}|^2 \\ &= |A^\top(\mathbf{y} - A^- \mathbf{c})|^2 + 0 + |A^\top A^- \mathbf{c} - \mathbf{c}|^2 \\ &\geq |A^\top A^- \mathbf{c} - \mathbf{c}|^2 \end{aligned}$$

이므로 임의의 행렬 A 와 임의의 벡터 \mathbf{c} 에 대해서

$$|\mathbf{y}^\top A - \mathbf{c}^\top| \geq |\mathbf{c}^\top A^+ A - \mathbf{c}^\top| \quad \text{for all } \mathbf{y} \in \mathbb{R}^m \quad (5.18)$$

이 성립한다. 이 결과를 활용하면 다음과 같이 의미심장한 pseudoinverse의 characterization을 얻는다.

Theorem 5.4 *Pseudoinverse of A is a matrix X that minimizes $\|AX - I_n\|_F$, that is,*

$$A^+ = \underset{X : d \times n \text{ matrix}}{\operatorname{argmin}} \|AX - I_n\|_F = \underset{Y : d \times n \text{ matrix}}{\operatorname{argmin}} \|YA - I_d\|_F \quad (5.19)$$

Proof: j -번째 표준 기저벡터를 \mathbf{e}_j 로 표시하자. 식 (5.17)의 \mathbf{x} 대신에 임의의 행렬 X 의 j -번째 열 $X\mathbf{e}_j$ 를 대입하고 \mathbf{b} 를 \mathbf{e}_j 로 놓으면

$$|AX\mathbf{e}_j - \mathbf{e}_j| \geq |AA^+\mathbf{e}_j - \mathbf{e}_j|$$

이므로

$$\|AX - I_n\|_F^2 = \sum_{j=1}^n |AX\mathbf{e}_j - \mathbf{e}_j|^2 \geq \sum_{j=1}^n |AA^+\mathbf{e}_j - \mathbf{e}_j|^2 = \|AA^+ - I_n\|_F^2$$

이 성립한다. 비슷하게 (5.18)으로부터

$$|\mathbf{e}_i^\top YA - \mathbf{e}_i^\top| \geq |\mathbf{e}_i^\top A^+ A - \mathbf{e}_i^\top|$$

이므로

$$\|YA - I_d\|_F^2 = \sum_{i=1}^d |\mathbf{e}_i^\top YA - \mathbf{e}_i^\top|^2 \geq \sum_{i=1}^d |\mathbf{e}_i^\top A^+ A - \mathbf{e}_i^\top|^2 = \|A^+ A - I_d\|_F^2.$$

■

5.8.1 Generalized Projection and Least Squares

일반화된 projection 또는 least square 문제를 생각해 보자. 일반화된 porjection이라 함은 정사영되는 부분 공간을 span하는 벡터들이 linearly dependent 할 수도 있다는 의미이다. 행렬 $A = [\mathbf{a}_1 | \cdots | \mathbf{a}_n]$ 의 행보다 열의 개수가 적고 full rank가 아닐 수도 있는 경우이다. 이 경우, A 의 열공간 $\text{Col}(A)$ 에 \mathbf{b} 를 정사영하거나 $A\mathbf{x} = \mathbf{b}$ 를 근사적으로 풀기 위해서는 $|A\mathbf{x} - \mathbf{b}|$ 를 최소로 만드는 \mathbf{x} 를 구해야 한다. 위에서 유도한 부등식

(5.17)에 의해 $\mathbf{x} = A^+ \mathbf{b}$ 가 $|A\mathbf{x} - \mathbf{b}|$ 를 최소로 만드는 벡터이고 Pythagorean relation에 의해 $\mathbf{b} - AA^+ \mathbf{b}$ 와 $AA^+ \mathbf{b}$ 가 수직임을 알 수 있다. 특히 A 와 A^+ 가 Penrose identity를 만족하므로

$$(\mathbf{b} - AA^+ \mathbf{b})^\top AA^+ \mathbf{b} = \mathbf{b}^\top (I - AA^+) AA^+ \mathbf{b} = \mathbf{b}^\top (AA^+ - AA^+ AA^+) \mathbf{b} = \mathbf{b}^\top (AA^+ - AA^+) \mathbf{b} = \mathbf{0}$$

으로 $\mathbf{b} - AA^+ \mathbf{b}$ 와 $AA^+ \mathbf{b}$ 가 수직함을 직접 보일 수도 있다. 따라서 $A^+ \mathbf{b}$ 가 least square 문제의 해가 되는 동시에 $\text{Col}(A) = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ 상으로의 \mathbf{b} 의 정사영은

$$\mathbf{P}_{\text{Col}(A)}(\mathbf{b}) = AA^+ \mathbf{b} \quad (5.20)$$

이 되고 정사영을 나타내는 행렬은 AA^+ 이다.

행렬 A 의 rank가 d 라고 하면 Fact 5.12에 의해 least square 해가 $(A^\top A)^{-1} A^\top \mathbf{b}$ 가 되고 projection은 $A(A^\top A)^{-1} A^\top \mathbf{b}$ 로 이전에 얻은 결과와 일치한다.

본 장의록에서 $n \times d$ 행렬 $A = [\mathbf{a}_1 | \dots | \mathbf{a}_n]$ 에 대해서 얻은 정사영에 관한 결과를 정리하면 다음과 같다.

- 행렬 A 에 대해 특별한 조건이 없는 경우

$$\mathbf{P}_{\text{Col}(A)}(\mathbf{x}) = AA^+ \mathbf{x} \quad \text{and} \quad A^+ ;$$

- 행렬 A 의 열 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 이 선형독립이면 (즉 $\text{rank } A = d$ 이면)

$$\mathbf{P}_{\text{Col}(A)}(\mathbf{x}) = A(A^\top A)^{-1} A^\top \mathbf{x} \quad \text{and} \quad A^+ = (A^\top A)^{-1} A^\top ;$$

- 행렬 A 의 열 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 이 orthonormal하면

$$\mathbf{P}_{\text{Col}(A)}(\mathbf{x}) = AA^\top \mathbf{x} \quad \text{and} \quad A^+ = A^\top ;$$

- 행렬 A 의 열이 \mathbf{v} 한 개이면

$$\mathbf{P}_{\text{Col}(A)}(\mathbf{x}) = \frac{1}{\mathbf{v}^\top \mathbf{v}} \mathbf{v} \mathbf{v}^\top \mathbf{x} \quad \text{and} \quad A^+ = \frac{1}{\mathbf{v}^\top \mathbf{v}} \mathbf{v}^\top .$$

5.9 How to obtain SVD

행렬이 (5.8) 형태나 또는 유사하게 주어진 경우에는 Fact 5.2를 활용하여 singular value와 vector를 얻을 수 있는지 확인한다. 구체적인 숫자로 행렬이 주어졌던 문자로 주어졌던 마찬가지이다. 예를 들면, 행렬이 rank-one인 경우에는 바로 Fact 5.2를 적용할 수 있다. 주어진 $n \times d$ 행렬 A 에 대해서 이러한 hint를 발견하지 못한 경우에는, AA^\top 과 $A^\top A$ 중에서 작은 크기의 행렬에 대해 symmetric eigenvalue problem을 푼다. 이렇게 얻어진 eigenvalue와 eigenvector는 right- 또는 left-singular value와 vector가 된다. 예를 들어 $n > d$ 이면 $A^\top A$ 가 둘 중에서 작은 대칭행렬이고 eigenvalue λ 와 eigenvector \mathbf{v} 가 얻었다면, $\sqrt{\lambda}$ 가 singular value이고 \mathbf{v} 가 right-singular vector이다. 그리고 $A\mathbf{v}$ 를 normalize하면 left-singular vector를 얻는다.

Centering Data

우리가 주어진 데이터를 가장 잘 표현하는 k -차원 공간을 구할 때, k -차원 부분공간을 찾는지, 아니면 k -차원 affine space를 찾고 있는지 결정해야 한다.⁴ 후자의 경우에 SVD를 계산하기 전에 데이터 점들의 평균을 각 데이터 점들에서 빼서 데이터의 centroid가 원점이 되도록 한 후에 SVD계산을 수행하여야 한다.

Fact 5.13 *The k -dimensional affine space that minimizes the sum of squared perpendicular distances to the data points must pass through the centroid of the points.*

Proof: Affine space를 $\mathbb{S} = \{\mathbf{v}_0 + \sum_{j=1}^k c_j \mathbf{v}_j : c_1, \dots, c_k \in \mathbb{R}\}$ 로 표현하자. 여기서, $\mathbf{v}_1, \dots, \mathbf{v}_k$ 는 orthonormal 벡터들이다. 데이터는 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 이고 \mathbf{v}_0 는 0과 제일 가까운 \mathbb{S} 의 점으로 \mathbf{v}_j 와 수직하다. \mathbb{S} 에서 \mathbf{a}_i 와 가장 가까운 점, 즉 \mathbf{a}_i 를 \mathbb{S} 에 정사영(projection)한 점을 $\mathbf{v}_0 + \sum_{j=1}^k c_j^* \mathbf{v}_j$ 라 하면, 두 점의 차를 나타내는 벡터는 \mathbb{S} 에 수직하다. 즉 모든 ℓ 에 대해 $\langle \mathbf{a}_i - \mathbf{v}_0 - \sum_{j=1}^k c_j^* \mathbf{v}_j, \mathbf{v}_\ell \rangle = 0$ 이다. 이를 전개하면 $c_\ell^* = \langle \mathbf{a}_i - \mathbf{v}_0, \mathbf{v}_\ell \rangle = \langle \mathbf{a}_i, \mathbf{v}_\ell \rangle$ 임을 알 수 있다. 데이터 점들의 \mathbb{S} 에 대한 수직거리의 제곱합은

$$\begin{aligned}
 \sum_{i=1}^n \text{dist}(\mathbf{a}_i, \mathbb{S})^2 &= \sum_{i=1}^n \left| \mathbf{a}_i - \mathbf{v}_0 - \sum_{j=1}^k c_j^* \mathbf{v}_j \right|^2 \\
 &= \sum_{i=1}^n \left\{ |\mathbf{a}_i - \mathbf{v}_0|^2 - 2 \left\langle \mathbf{a}_i - \mathbf{v}_0, \sum_{j=1}^k c_j^* \mathbf{v}_j \right\rangle + \left| \sum_{j=1}^k c_j^* \mathbf{v}_j \right|^2 \right\} \\
 &= \sum_{i=1}^n \left\{ |\mathbf{a}_i - \mathbf{v}_0|^2 - 2 \sum_{j=1}^k c_j^* \langle \mathbf{a}_i - \mathbf{v}_0, \mathbf{v}_j \rangle + \sum_{j=1}^k (c_j^*)^2 \right\} \\
 &= \sum_{i=1}^n \left\{ |\mathbf{a}_i - \mathbf{v}_0|^2 - \sum_{j=1}^k (c_j^*)^2 \right\} \\
 &= \sum_{i=1}^n \left\{ |\mathbf{a}_i|^2 - 2 \langle \mathbf{v}_0, \mathbf{a}_i \rangle + |\mathbf{v}_0|^2 - \sum_{j=1}^k \langle \mathbf{a}_i, \mathbf{v}_j \rangle^2 \right\} \\
 &= n|\mathbf{v}_0|^2 - 2 \sum_{i=1}^n \langle \mathbf{v}_0, \mathbf{a}_i \rangle + \sum_{i=1}^n \left(|\mathbf{a}_i|^2 - \sum_{j=1}^k \langle \mathbf{a}_i, \mathbf{v}_j \rangle^2 \right) \\
 &= n \left| \mathbf{v}_0 - \frac{1}{n} \sum_{i=1}^n \mathbf{a}_i \right|^2 - \frac{1}{n} \sum_{i=1}^n \left| \mathbf{a}_i \right|^2 + \sum_{i=1}^n \left(|\mathbf{a}_i|^2 - \sum_{j=1}^k \langle \mathbf{a}_i, \mathbf{v}_j \rangle^2 \right) \\
 &\geq -\frac{1}{n} \left| \sum_{i=1}^n \mathbf{a}_i \right|^2 + \sum_{i=1}^n \left(|\mathbf{a}_i|^2 - \sum_{j=1}^k \langle \mathbf{a}_i, \mathbf{v}_j \rangle^2 \right)
 \end{aligned}$$

으로 $\mathbf{v}_0 = \frac{1}{n} \sum_{i=1}^n \mathbf{a}_i$ 일 때, 즉 affine space \mathbb{S} 가 데이터 점들의 centroid를 지나갈 때 수직거리의 제곱합이 최소가 된다. ■

⁴Affine space는 linear space를 원점에서 평행이동시켜서 얻어진 집합이다. 보통 $\{\mathbf{v}_0 + x_1 \mathbf{v}_1 + \dots + x_k \mathbf{v}_k : (x_1, \dots, x_k) \in \mathbb{R}^k\}$ 로 표현된다. Affine space에서는 임의의 linear combination은 보전되지 않지만 affine combination은 보전된다. 즉 \mathbf{v}_1 과 \mathbf{v}_2 의 affine space에 속할 때 임의의 실수 λ 에 대해서 $\lambda \mathbf{v}_1 + (1 - \lambda) \mathbf{v}_2$ 가 다시 affine space에 포함된다.

5.10 Application to Statistics: PCA from SVD

Principal Components Analysis (PCA)

확률벡터⁵ \mathbf{X} 의 주성분(principal component)이라 함은 \mathbf{X} 의 변동성을 주로 설명할 수 있는 확률벡터 성분들의 조합, 즉 X_j 들의 선형조합을 말한다. 선형조합의 성분별 비중을 나타내는 벡터를 $\mathbf{v} = (v_1, \dots, v_d)^\top \in \mathbb{R}^d$ 라 하면 $\mathbf{v}^\top \mathbf{X} = \sum_{j=1}^d v_j X_j$ 는 확률변수가 된다. 그리고 확률변수의 변동성을 분산으로 생각하고자 한다. 그러면 첫 번째 주성분은 $|\mathbf{v}| = 1$ 인 조합 중, $\mathbf{v}^\top \mathbf{X}$ 의 분산을 최대로 하는 조합을 찾는 문제가 된다.

우선 \mathbf{X} 의 평균을 빼주는 보정을 통해 평균이 $\mathbf{0}$ 이라고 가정하자.⁶ 그러면 $\mathbf{v}^\top \mathbf{X}$ 는 평균이 0 인 확률변수가 되고, 분산은

$$\text{Var}(\mathbf{v}^\top \mathbf{X}) = \mathbb{E}[(\mathbf{v}^\top \mathbf{X})^2] = \mathbb{E}[\mathbf{v}^\top \mathbf{X}(\mathbf{v}^\top \mathbf{X})^\top] = \mathbb{E}[\mathbf{v}^\top \mathbf{X} \mathbf{X}^\top \mathbf{v}] = \mathbf{v}^\top \mathbb{E}[\mathbf{X} \mathbf{X}^\top] \mathbf{v} = \mathbf{v}^\top \Sigma \mathbf{v}$$

이 된다.

- 첫 번째 주성분은

$$\underset{|\mathbf{v}|=1}{\operatorname{argmax}} \text{Var}(\mathbf{v}^\top \mathbf{X}) = \underset{|\mathbf{v}|=1}{\operatorname{argmax}} \mathbf{v}^\top \Sigma \mathbf{v}$$

으로 실제 분석 시에는 Σ 가 주어지지 않고, \mathbf{X} 의 iid 관측으로 생각할 수 있는 $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ 을 얻게된다. 따라서 표본 공분산 $\hat{\Sigma}$ 를 추정하여 사용하게 된다. SVD 표현과 일관성을 유지하기 위해, i 번째 관측 \mathbf{X}_i 의 관측된 벡터 \mathbf{x}_i 를 i 번째 행으로 구성한 $n \times d$ 데이터 행렬 A 를 고려하자. 그러면,

$$\hat{\Sigma} = \frac{1}{n-1} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top \triangleq \frac{1}{n-1} A^\top A$$

이 된다. 여기서 ☆등식이 성립됨을 각자 확인해 보자. 그리고 $\mathbf{v}^\top A^\top A \mathbf{v} = |A\mathbf{v}|^2$ 임을 고려하면

$$\underset{|\mathbf{v}|=1}{\operatorname{argmax}} \mathbf{v}^\top \hat{\Sigma} \mathbf{v} = \underset{|\mathbf{v}|=1}{\operatorname{argmax}} |A\mathbf{v}|$$

가 되어 첫 번째 주성분 벡터는 A 행렬의 첫 번째 right singular vector \mathbf{v}_1 과 일치함을 알 수 있다. 그리고 첫 번째 주성분에 의해 표현(설명)되는 변동량(분산)은 $\frac{\sigma_1^2}{n-1}$ 이다.

- 두 번째 주성분 \mathbf{v}_* 는 첫 번째 주성분 확률변수 $\mathbf{v}_1^\top \mathbf{X}$ 와 uncorrelated인⁷ 성분 중 분산 $\text{Var}(\mathbf{v}_*^\top \mathbf{X})$ 이 최대가 되도록 정한다. 즉 통계학 관점에서는

$$\mathbf{v}_* = \underset{|\mathbf{v}|=1, \text{Cov}(\mathbf{v}^\top \mathbf{X}, \mathbf{v}_1^\top \mathbf{X})=0}{\operatorname{argmax}} \text{Var}(\mathbf{v}^\top \mathbf{X})$$

를 푸는 것이다. 이 문제의 제약식에서, $\text{Cov}(\mathbf{v}^\top \mathbf{X}, \mathbf{v}_1^\top \mathbf{X}) = \mathbf{v}^\top \Sigma \mathbf{v}_1 = 0$ 이다. 데이터 분석에서는

$$\mathbf{v}^\top \hat{\Sigma} \mathbf{v}_1 = 0 \Leftrightarrow \mathbf{v}^\top A^\top A \mathbf{v}_1 = \langle A\mathbf{v}, A\mathbf{v}_1 \rangle = 0$$

⁵ Appendix D에 확률벡터에 대한 간략한 설명이 있다.

⁶ 실제 주성분 분석을 수행할 때도 첫 번째 과정으로 표본평균을 $\mathbf{0}$ 으로 조정하는 작업을 수행한다.

⁷ $\text{Cov}(\mathbf{v}_*^\top \mathbf{X}, \mathbf{v}_1^\top \mathbf{X}) = 0$

이라는 조건으로 바뀐다. A 의 SVD를

$$A = \sum_{j=1}^d \sigma_j \mathbf{u}_j \mathbf{v}_j^\top$$

로 표현하자. 단, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0 = \sigma_{r+1} = \dots = \sigma_d$ 이라 하자. 임의의 $\mathbf{v} \in \mathbb{R}^d$ 는 계수벡터 $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)^\top$ 에 대해 $\mathbf{v} = \sum_{j=1}^d \alpha_j \mathbf{v}_j$ 로 표현할 수 있다. 그런데 $A\mathbf{v}_j = \sigma_j \mathbf{u}_j$ 이고 \mathbf{u}_j 들은 서로 수직하므로 $\langle A\mathbf{v}, A\mathbf{v}_1 \rangle = 0$ 이라면 $\langle \sum_{j=1}^d \alpha_j \sigma_j \mathbf{u}_j, \sigma_1 \mathbf{u}_1 \rangle = \alpha_1 \sigma_1^2 = 0$ 이어서 $\alpha_1 = 0$ 임을 알게 된다. 즉 두 확률벡터 $\mathbf{v}_1^\top \mathbf{X}$ 와 $\mathbf{v}^\top \mathbf{X}$ 가 uncorrelated일 필요충분조건이 $\langle \mathbf{v}, \mathbf{v}_1 \rangle = 0$, 즉 \mathbf{v} 과 \mathbf{v}_1 의 수직관계임을 알 수 있다. 즉

$$\mathbf{v}^\top \hat{\Sigma} \mathbf{v}_1 = 0 \Leftrightarrow \langle \mathbf{v}, \mathbf{v}_1 \rangle = 0.$$

따라서 두 번째 주성분을 구하는 문제는

$$\mathbf{v}_* = \underset{|\mathbf{v}|=1, \langle \mathbf{v}, \mathbf{v}_1 \rangle = 0}{\operatorname{argmax}} |A\mathbf{v}|$$

가 되어, 두 번째 주성분은 A 의 두 번째 right singular vector \mathbf{v}_2 와 일치하게 되고 두 번째 주성분 벡터에 의해 설명되는 분산은 $\frac{\sigma_2^2}{n-1}$ 이 된다.

- 이러한 과정은 singular value가 양수인 $\operatorname{rank} A = r$ 번째까지 반복하여 최대 r 개의 주성분을 찾을 수 있다. $r+1$ 번째 주성분도 찾을 수 있으나 설명되는 변동량은 $\sigma_{r+1}^2 = 0$ 으로 무의미한 성분이된다.
- $k \leq r$ 인 k 개의 주성분에 의해 설명되는 총분산은 $\frac{1}{n-1}(\sigma_1^2 + \dots + \sigma_k^2)$ 이다. 또한 SVD의 best-fit 성질에 의해 k 개의 주성분으로 설명할 수 있는 최대 분산량임을 알 수 있다. 흔히, $0 < \alpha < 1$ 에 대해서

$$\frac{\sigma_1^2 + \dots + \sigma_k^2}{\sigma_1^2 + \dots + \sigma_k^2 + \sigma_{k+1}^2 + \dots + \sigma_r^2} \geq \alpha$$

인 k 중 최솟값을 택하고 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 를 가장 효과적으로 전체 변동성의 $100\alpha\%$ 를 설명하는 주성분들이라고 말한다.

이상의 결과를 통계학자들은 다음과 같은 정리로 요약한다.

Theorem 5.5 Let $\operatorname{Cov}(\mathbf{X}, \mathbf{X}) = \boldsymbol{\Sigma}$ have eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_r$ with corresponding eigenvalues $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_r^2 > 0$. Then:

- (i) The j -th PC(principal component) is $\mathbf{v}_j^\top \mathbf{X} = v_{j,1} X_1 + \dots + v_{j,d} X_d$ for $j = 1, \dots, r$.
- (ii) The variance of j -th PC is $\operatorname{Var}(\mathbf{v}_j^\top \mathbf{X}) = \mathbf{v}_j^\top \boldsymbol{\Sigma} \mathbf{v}_j = \sigma_j^2$.
- (iii) The covariance between two PCs is uncorrelated, i.e. $\operatorname{Cov}(\mathbf{v}_j^\top \mathbf{X}, \mathbf{v}_k^\top \mathbf{X}) = \mathbf{v}_j^\top \boldsymbol{\Sigma} \mathbf{v}_k = 0$ for $j \neq k$.

$\boldsymbol{\Sigma} = A^\top A$ 이고 A 의 SVD가 $A = \sum_{j=1}^r \sigma_j \mathbf{u}_j \mathbf{v}_j^\top$ 인 경우를 생각하며 정리를 읽으면 별도의 증명이 필요없다.

Kang & Cho (2024)

Chapter 6

SVD in Real Life

6.1 Compression of Images via SVD

SVD의 여러 성질을 활용한 좋은 예가 image compression이다. 다음 Figure 6.1과 같은 컬러 이미지는 보통 각 픽셀의 색상을 (r, g, b) 세개의 숫자로 표현한다. 따라서 가로 세로 각각 n 개와 d 개의 픽셀로 구성된 컬러 이미지 파일은 $n \times d$ 행렬 세개로 표현할 수 있다. 대부분의 plotting 함수들은 이 세 행렬을 입력으로 컬러 이미지를 표시한다. 색을 나타내는 행렬의 각 행은 이미지를 구성하는 매우 폭이 좁은 가로 줄을 나타낸다. 실제 이미지에서 이러한 색의 연속을 더하는 것을 색칠하는 방식으로 이해하기는 어렵지만 3원색의 강도를 나타내는 수로 표현된 경우에는 실수 벡터로 생각하여 벡터간의 연산을 수행할 수 있다.



Figure 6.1: Mona Lisa in 687×1024 pixels.

각 색깔에 대한 행렬을 $A^{(1)}, A^{(2)}, A^{(3)}$ 로 나타내고 각 행렬의 rank가 r_1, r_2, r_3 라 가정하자. 편의상 각 행렬

들의 열별 합이 0이라 가정하자. Singular value decomposition은 양수 $\sigma_i^{(j)}$ 와 벡터 $\mathbf{u}_i^{(j)}, \mathbf{v}_i^{(j)}$ 에 대해서

$$A^{(j)} = \sum_{i=1}^{r_j} \sigma_i^{(j)} \mathbf{u}_i^{(j)} \mathbf{v}_i^{(j) \top}$$

와 같이 rank-one matrix의 합으로 A 를 표현해 준다. 그리고 $k \leq \min\{r_1, r_2, r_3\}$ 일 때 $A^{(j)}$ 의 **rank** k approximation

$$A^{(j)}_k = \sum_{i=1}^k \sigma_i^{(j)} \mathbf{u}_i^{(j)} \mathbf{v}_i^{(j) \top}$$

로 행렬 $A^{(j)}$ 를 근사한다. 다음 그림 6.2은 687×1024 픽셀로 표현된 Mona Lisa 그림을 $k = 3, 8, 18, 23, 34$ 에 대해 $(A^{(1)}_k, A^{(2)}_k, A^{(3)}_k)$ 로 근사된 image를 보여준다.

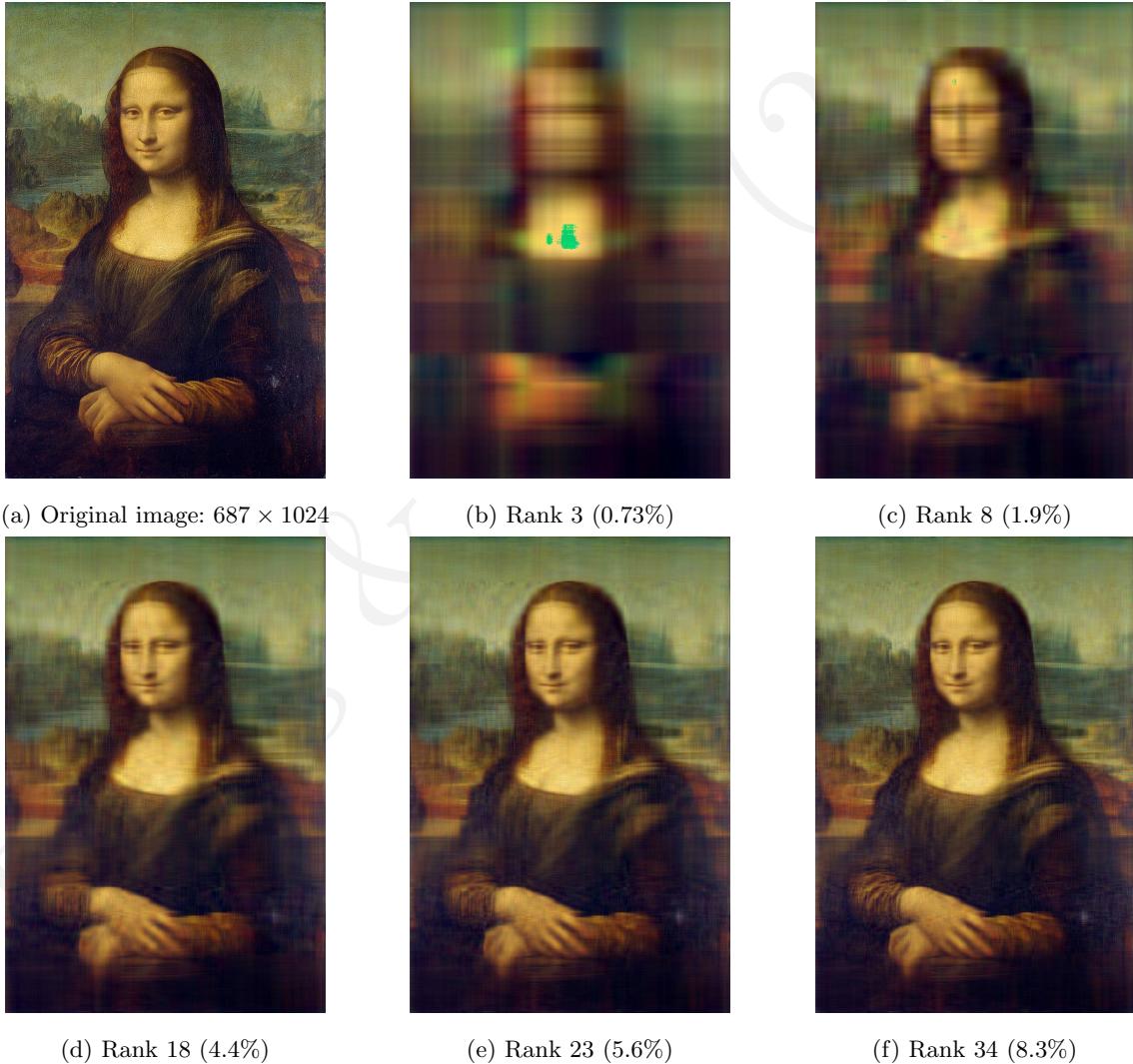


Figure 6.2: Low-rank approximations of the Mona Lisa (The percentage is the portion of data in use.)

Figure 6.2f는 원래 데이터의 8.3%만을 사용함에도 거의 원래 이미지처럼 느껴지는 것을 볼 수 있다.

마지막으로 언급할 내용은 앞에서 언급한 3원색 행렬들의 열별 합이 0이라는 가정이 맞지 않는 경우, 각 열에서 열별 평균을 빼서 SVD를 수행하고 low rank approximation에 열별 평균을 더하면 된다. 즉 열들의 평균을 나타내는 벡터

$$\boldsymbol{\mu}_j = \frac{1}{n} A^{(j)\top} \mathbf{1}_n$$

를 계산하여 열 별로 빼준 행렬

$$A^{(j)} - \mathbf{1}_n \boldsymbol{\mu}_j^\top$$

에 SVD를 적용한다. 그리고 얻어진 rank- k approximation

$$A^{(j)} - \mathbf{1}_n \boldsymbol{\mu}_j^\top \approx A^{(j)}_k = \sum_{i=1}^k \sigma_i^{(j)} \mathbf{u}_i^{(j)} \mathbf{v}_i^{(j)\top}$$

으로부터 원래 행렬의 근사

$$A^{(j)} \approx \mathbf{1}_n \boldsymbol{\mu}_j^\top + A^{(j)}_k = \mathbf{1}_n \boldsymbol{\mu}_j^\top + \sum_{i=1}^k \sigma_i^{(j)} \mathbf{u}_i^{(j)} \mathbf{v}_i^{(j)\top}$$

를 얻는다. 평균벡터를 보정한 후에 SVD를 적용하는 이유는 Section 5.9를 참조하기 바란다.

6.2 Left-singular Vectors as the Coordinates of Embedding Vectors in the Latent Space

d 차원 벡터공간의 데이터 $\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \subset \mathbb{R}^d$ 를 근사하는 k 차원 부분공간을 찾는 문제는 부분공간의 기저를 구성하는 orthonormal한 k 개의 d -차원 벡터를 찾는 문제이다. 찾고자 하는 부분공간의 basic vector가 열이 되도록 구성한 $d \times k$ unknown matrix V 중에 $\sum_{i=1}^n |V^\top \mathbf{a}_i|^2$ 을 최대로 만드는 V 를 찾는 문제로 표현할 수도 있다:

$$V^* = \underset{\substack{V: d \times k \\ V^\top V = I_k}}{\operatorname{argmax}} \sum_{i=1}^n |V^\top \mathbf{a}_i|^2. \quad (6.1)$$

$V^\top V = I_k$ 라는 조건을 이용하면

$$\begin{aligned} |\mathbf{a}_i - VV^\top \mathbf{a}_i|^2 &= (\mathbf{a}_i - VV^\top \mathbf{a}_i)^\top (\mathbf{a}_i - VV^\top \mathbf{a}_i) \\ &= (\mathbf{a}_i^\top - \mathbf{a}_i^\top VV^\top)(\mathbf{a}_i - VV^\top \mathbf{a}_i) \\ &= \mathbf{a}_i^\top \mathbf{a}_i - 2\mathbf{a}_i^\top VV^\top \mathbf{a}_i + \mathbf{a}_i^\top VV^\top VV^\top \mathbf{a}_i \\ &= |\mathbf{a}_i|^2 - \mathbf{a}_i^\top VV^\top \mathbf{a}_i \\ &= |\mathbf{a}_i|^2 - |V^\top \mathbf{a}_i|^2 \end{aligned}$$

가 된다. 그런데 데이터가 주어진 경우 $\sum_{i=1}^n |\mathbf{a}_i|^2$ 이 V 에 무관한 상수이므로 $\sum_{i=1}^n |V^\top \mathbf{a}_i|^2$ 을 최대화하는 V 를 찾는 것은 $\sum_{i=1}^n |\mathbf{a}_i - VV^\top \mathbf{a}_i|^2$ 를 최소화하는 V 를 찾는 것과 동치이다. 따라서 k 차원 부분공간을 찾는 원래 문제 (6.1)은

$$V^* = \underset{\substack{V: d \times k \\ V^\top V = I_k}}{\operatorname{argmax}} \sum_{i=1}^n |\mathbf{a}_i - VV^\top \mathbf{a}_i|^2 \quad (6.2)$$

와 동치가 되는데 두 번째 formulation (6.2)은 d 차원 데이터 벡터 \mathbf{a}_i 들을 $k \times d$ 행렬 V^\top 에 곱하여 k 차원으로 축소한 뒤에 $d \times k$ 행렬 V 를 곱하여 다시 d 차원으로 복구한 벡터들과 원래 벡터와의 차이를 최소로 하는 reconstruction 문제를 푸는 것으로 생각할 수 있으며¹ 이렇게 고차원 데이터가 V^\top 에 의해 변환/근사된 벡터들이 담겨있는 k 차원 부분공간을 **latent space**라고 부른다. 물론 orthonormal한 벡터들로 span되는 부분공간으로의 정사영을 나타내는 행렬 표현 (4.7)를 생각하면 (6.2)의 VV^\top 항이 자연스러우나 정사영 관점만으로는 행렬 V 를 통한 reconstruction을 설명하기가 쉽지 않다.

데이터 \mathbf{a}_i 를 행으로 쌓은 $n \times d$ 데이터 행렬을 A 로 표시하고 미지의 $d \times k$ 행렬 V 의 열들을 \mathbf{v}_j 로 표시하면

$$\sum_{i=1}^n |V^\top \mathbf{a}_i|^2 = \|AV\|_F^2 = \sum_{j=1}^k |A\mathbf{v}_j|^2$$

이므로 (6.1)의 해 벡터인 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 에 의해 span되는 부분공간과 A 행렬에서 singular value가 제일 큰 k 개의 singular vector에 의해 span되는 부분공간은 squared residual distance 측면에서 동일한 dimension reduction 효과가 있다. 하지만 (6.1)의 해 벡터들은 singular vector들이 갖는 순서 $|A\mathbf{v}_i| \geq |A\mathbf{v}_{i+1}|$ 가 지켜지지 않을 수도 있다. 반면에 A 의 singular value가 제일 큰 k 개의 singular vector들은 (6.1)의 해 벡터가 되고,

¹기계학습의 autoencoder와 유사하게 V^\top 이 encoder, V 가 decoder의 역할을 하는 것으로 생각할 수 있다.

따라서 singular vector들이 (6.2)와 같이 데이터를 효과적으로 reconstruct하는 부분공간을 기술하는 것도 알 수 있다.

데이터 \mathbf{a}_i 가 latent space로 변환/근사된 벡터를 \mathbf{b}_i 라 하자. 흔히 \mathbf{b}_i 를 \mathbf{a}_i 의 **embedding vector**라고 부른다. Latent space로의 변환이 V^\top 에 의해 이루어지므로 $\mathbf{b}_i = V^\top \mathbf{a}_i$ 이다. 행렬 A 의 j -번째 left-singular vector \mathbf{u}_j 의 i -번째 원소는 $(\mathbf{u}_j)_i = (A\mathbf{v}_j)_i = \mathbf{a}_i^\top \mathbf{v}_j = (V^\top \mathbf{a}_i)_j = (\mathbf{b}_i)_j$ 이어서 embedding vector \mathbf{b}_i 의 j -번째 원소이다. 즉 n -차원인 left-singular vector \mathbf{u}_j 는 latent space 속으로 embed된 n 개 데이터들을 나타내는 vector들의 j -번째 좌표를 나타내는 것으로 생각할 수 있다. 즉 $U = [\mathbf{u}_1 | \mathbf{u}_2 | \dots | \mathbf{u}_k]$ 의 i -번째 행이 \mathbf{b}_i 이다. 근사 공간의 차원 k 가 잘 선택되었으면 embedding vector \mathbf{b}_i 를 잘 분석하면 데이터 \mathbf{a}_i 에 대해 깊이 이해할 수 있다.

6.3 Approximation of MNIST Images via SVD

MNIST 데이터(Modified National Institute of Standards and Technology database)는 손으로 쓴 숫자들로 이루어진 데이터베이스로서 다양한 이미지 처리 시스템과 기계 학습 알고리듬의 학습 및 테스트에 널리 사용된다. MNIST 데이터는 0에서 9까지의 값을 나타내는 60,000개의 트레이닝 이미지와 10,000개의 테스트 이미지로 구성되었다. Figure 6.3에서 160개의 예시 이미지를 볼 수 있다.



Figure 6.3: MNIST Data Set

MNIST의 각 이미지는 28×28 pixel값으로 표현된다. 3을 나타내는 어느 MNIST 이미지²의 plot과 대응되는 28×28 행렬이 Figure 6.4에 예시되어 있다.



(a) Image of 3

(b) 28×28 matrix representing the left image

Figure 6.4: An Example of MNIST Image and its Matrix Representation

이러한 이미지를 수리적으로 분석하는 경우, 사람이 익숙한 6.4a 보다는 6.4b의 행렬을 $28 \times 28 = 784$ 차원 벡터로 reshape하여 여러 분석 도구를 적용한다. 계산은 784차원 벡터를 이용하지만 결과를 사람에게 보여줄 때는 오른쪽의 행렬보다는 왼쪽의 이미지가 훨씬 인간적이다. 따라서 이후에 벡터를 나타낼 때 6.4a와 같이 이미지를 통해 표현할 것이다.

앞에서 언급한 바와 같이 MNIST는 70,000개의 784차원 벡터로 구성된 데이터이므로 각 이미지를 행으로

²MNIST 데이터베이스에서 8번째 이미지이다.

쌓아서 $70,000 \times 784$ 행렬 A 를 구성하여 SVD를 수행하였다. $n = 70,000, d = 784$ 로 놓자. 모든 이미지(A 의 모든 행)의 평균이 Figure 6.5에 도시되어 있다. 참고로 python numpy linear algebra 패키지의 rank계산 함수인 rank_matrix로 계산하면 $\text{rank } A = 713$ 을 얻는다.

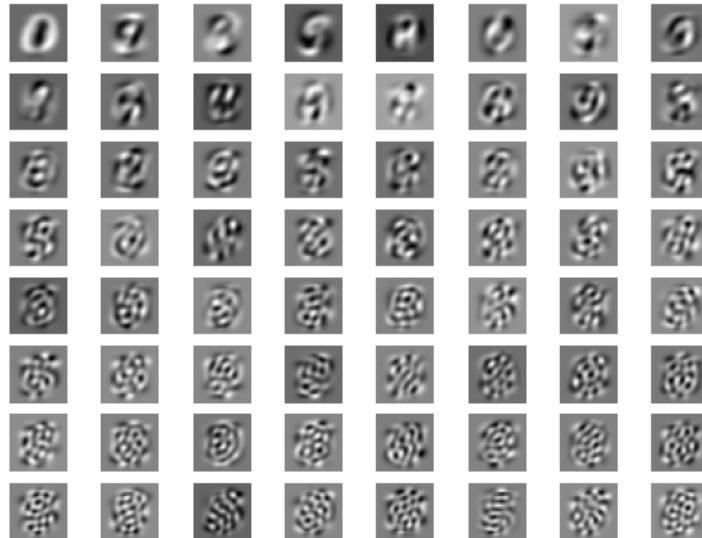
$$\frac{1}{n} A^\top \mathbf{1}_n = \begin{matrix} \text{Figure 6.5: The Mean of MNIST} \\ \text{A grayscale image showing a blurry, circular pattern centered in a dark square, representing the mean of the MNIST dataset.} \end{matrix}$$

Figure 6.5: The Mean of MNIST

MNIST 데이터 행렬 A 에서 이 평균을 뺀 행렬

$$A - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^\top A$$

에 SVD를 적용하자. SVD결과에서 제일 큰 singular value 64개에 대응되는 784차원 right singular vector를 도시하면 Figure 6.6과 같다.

Figure 6.6: 64 Leading Right Singular Vectors of A

첫번째 벡터는 중심이 검게 채워진 0의 이미지이다. 이 첫번째를 제외한 나머지 벡터들은 데이터의 여러 특징을 한번에 요약하는 right singular vector의 특징 상 직관적으로 이해하기 어려운 값을 갖고 따라서 해당 이미지도 시각적으로 파악하기 어렵다. 하지만 전반적으로 점점 면적이 작아지는 흑백 영역들의 개수가 많아지는 것을 살필 수 있다.

○ right singular vector들을 기저벡터로 MNIST의 각 이미지(A 의 각 행)를 표현하는 left singular vector 중 첫번째와 두번째로 큰 singular value에 대응되는 $\mathbf{u}_1 \in \mathbb{R}^n$ 과 $\mathbf{u}_2 \in \mathbb{R}^n$ 을 n 개의 2차원 좌표값 $(u_{i1}, u_{i2}), i = 1, \dots, n$ 으로 도시하면 Figure 6.7와 같다. 이 그림에서 시각적으로 이미지가 유사한 7과 9, 6과 8, 3과 5를 나타내는 이미지들이 많이 겹치는 것을 파악할 수 있다. 또한 0을 나타내는 이미지는 첫번째 right singular vector로부터 짐작할 수 있듯이 주로 x 축 근방에서 제일 큰 x 값을 갖는 영역인 $(0.01, 0)$ 근방에 많이 위치한다.

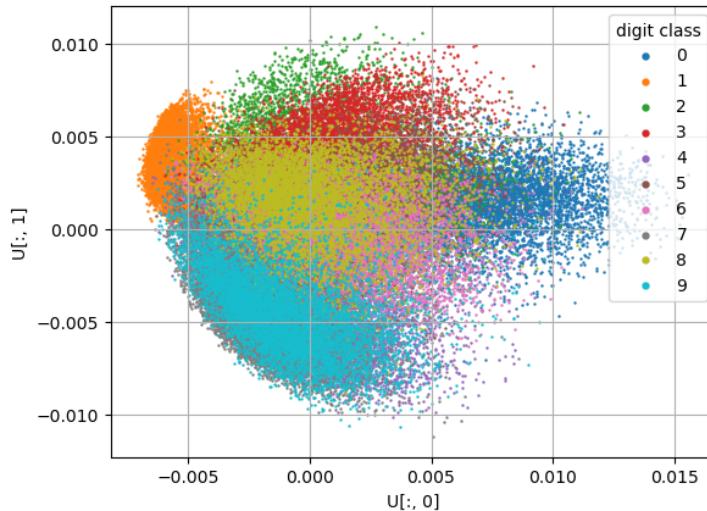


Figure 6.7: Two Leading Left Singular Vectors of A

다른 숫자를 나타내는 이미지들도 추가적인 분석이 가능한지 각자 살펴보면서 SVD의 의미를 더 깊게 이해하기 바란다. 다음 실험으로 random하게 추출된 8개 이미지를 singular value가 제일 큰 2개, 22개, 42개, 62개, 82개의 right singular vector로 low rank approximation을 한 결과를 Figure 6.8에 도시하였다.

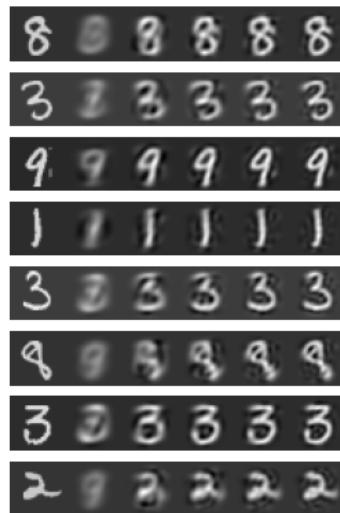


Figure 6.8: More Singular Vectors, More Accurate Approximation

첫번째 열은 이미지 원본이고 두번째 열이 원쪽의 원본을 2개의 leading right singular vector들로 근사한 것

이고 나머지 열들도 20개씩 추가된 right singular vector들로 근사한 이미지이다. 총 700여개의 right singular vector 중 42개만으로도 상당히 근사를 잘 하는 것을 볼 수 있다.

6.3.1 Geometry of MNIST Images in the Latent Space of Variational Auto-Encoder

MNIST 이미지들의 2차원 근사표현에 해당하는 Figure 6.7에는 숫자들의 영역에 겹치는 부분이 아직도 많은 것을 볼 수 있다. 이러한 제약은 SVD가 선형적인 분석이라는 점에서 극복이 어렵다. 그럼에도 10개의 숫자를 표현하는데 784차원은 너무 많다고 믿으며 데이터를 비선형적으로 변환하여 저차원에서 분류를 시도하는 기계학습 기법 중 한가지가 variational auto-encoder (VAE)이다. VAE에서는 Section 3.9.1에서 본 것처럼 신경망이 비선형 함수들을 다양하게 표현한다는 사실을 바탕으로 784개의 node로 구성된 input layer로부터 작은 숫자인 d 개의 output node를 갖는 output layer로 출력하는 encoding network와 반대로 d 개의 input node로부터 784개의 output node로 이루어진 decoding network를 통해 원래 MNIST 이미지를 근사적으로 복구한다. 이 과정에서 784차원 이미지가 d 차원으로 요약되어도 decoding과정에서 원래 이미지가 거의 복구되며 이 d 차원 공간을 latent space라고 부른다. MNIST 이미지들에 대해서 극단적으로 $d = 2$ 로 놓아서 2 차원 latent space로 압축했을 때 이미지들의 latent space상에서의 위치를 표현하면 아래 Figure 6.9와 같다.

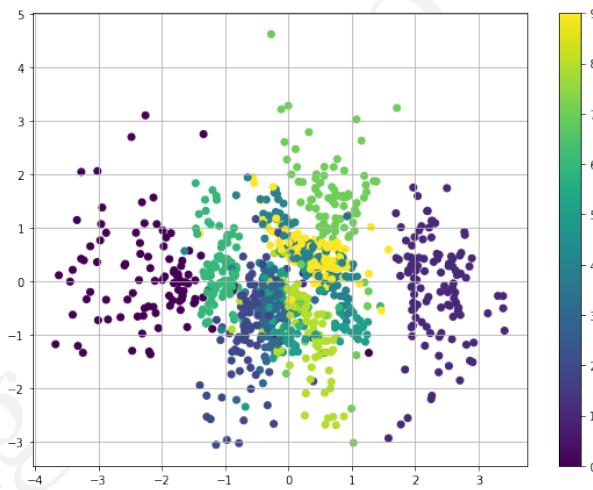


Figure 6.9: Latent space of VAE

6.4 Word Embeddings via SVD

기계학습에서 오랫동안 연구해 온 분야 중 하나가 자연어처리 (Natural Language Processing, NLP)이다. 우리가 여러 편집기를 사용할 때 일상적으로 도움을 받는 철자법 확인부터 최근 급격하게 발전한 언어 번역, 문장의 의미 분석, 질문에 답하기 등을 컴퓨터로 처리할 때 단어를 컴퓨터에 인식시키는 word embedding이 필요하다. 분석 이전에 벡터를 대응시킬 때는 단어 간의 관계에 대한 지식이 없으므로 최대한 모든 단어를 동등하게 n 차원 공간의 벡터에 대응시키는 것이 합리적이다. 예로서 one-hot vector로 대응시키는 경우 차원 인 n 이 백만 이상이 되는 어려움이 있다. 따라서 dimension reduction을 생각하는데 적절하게 구성된 행렬에 대해서 SVD를 적용하는 방법이 많이 개발되었다.

흔히 word-document matrix 또는 word-word co-occurrence matrix라고 불리는 각 행이 각각의 단어에 대응하는 행렬을 구성한다. 그리고 각 열은 document나 단어를 나타내며 원소들은 그 행에 대응하는 단어의 출현 빈도와 관련된 양으로 정의한다. 이 행렬을 A 라 표시하자. 데이터 행렬 A 의 SVD를

$$A = \sum_{\ell=1}^n \sigma_\ell \mathbf{u}_\ell \mathbf{v}_\ell^\top$$

로 표시하자. 대개는 A 의 열의 개수가 행의 개수보다 크므로 rank-one 행렬 n 개의 합으로 표시하였다. 그리고 $\sum_{i=1}^k \sigma_i^2 / \sum_{i=1}^n \sigma_i^2$ 이 충분히 크게 되는 rank k 근사 행렬인 $\sum_{\ell=1}^k \sigma_\ell \mathbf{u}_\ell \mathbf{v}_\ell^\top$ 를 선택한다. k 개 left singular vector로 $n \times k$ 행렬 $U = [\mathbf{u}_1 | \dots | \mathbf{u}_k]$ 를 구성하면 이 행렬의 i 번째 행벡터 $(u_{i1}, \dots, u_{ik})^\top \in \mathbb{R}^k$ 이 i 번째 단어의 embedding이 된다. 보통 $k \ll n$ 이다.

이해를 돋기위해 다음과 같이 3개의 문서가 7개의 단어로 이루어진 간단한 예를 살펴보자.

Doc1: I enjoy coding.

Doc2: I like NLP.

Doc3: I like deep learning.

Word-Document Matrix

n 개의 word로 이루어진 d 개의 document에 대해서 i 번째 단어가 j 번째 document에 p 번 사용되었으면 $x_{ij} = p/d$ 로 놓고 사용되지 않았으면 0으로 놓은 $n \times d$ 행렬 $X = (x_{ij})$ 를 word-document matrix라고 부른다.

	Doc1	Doc2	Doc3
I	1	1	1
enjoy	1	0	0
like	0	1	1
coding	1	0	0
NLP	0	1	0
deep	0	0	1
learning	0	0	1

위의 예에 대한 word-document matrix는 이다.

Word-Word Co-Occurrence Matrix

또 다른 예로 word-word co-occurrence matrix도 있다. 이 행렬의 정의에는 window size라고 부르는 문장에서 좌우 몇 단어 이내에 위치할 때 인접한 것으로 판단할지 기준이 필요하다. 즉 window size가 w 일 때, 문장들에서 i 번째 단어의 위치와 w 단어 이내에 단어 j 가 위치할 때마다 x_{ij} 를 1씩 증가시켜서 x_{ij} 가 단어 i 와 단어 j 가 문장들 속에서 w 단어 이내에 위치한 총 횟수일 때 $X = (x_{ij})$ 로 정의한다. 이 word-word co-occurrence matrix가 symmetric하다는 것은 정의의 대칭성 때문에 쉽게 알 수 있다. 위의 예에 대한 window size $w = 2$ 인 경우

	I	enjoy	like	coding	NLP	deep	learning	
I	0	1	2	1	1	1	0	
enjoy	1	0	0	1	0	0	0	
like	2	0	0	0	1	1	1	이
coding	1	1	0	0	0	0	0	
NLP	1	0	1	0	0	0	0	
deep	1	0	1	0	0	0	1	
learning	0	0	1	0	0	1	0	

다. 이 행렬의 SVD에서 두 개의 left-singular vector로 각 단어의 좌표를 도시하면 다음과 같다.

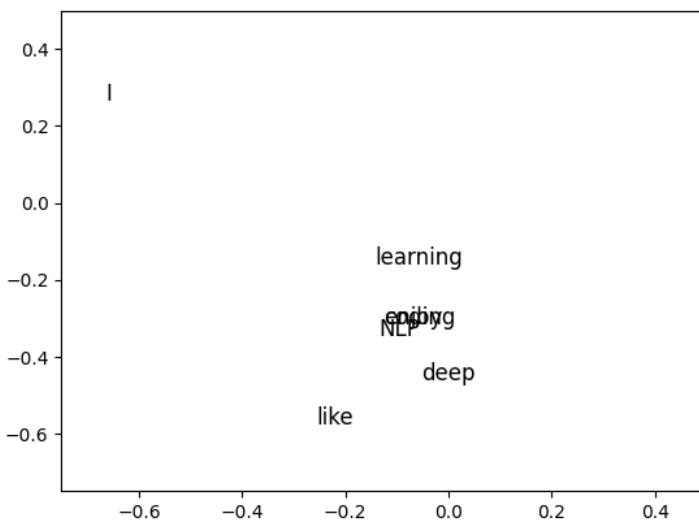


Figure 6.10: Word embedding via left-singular vectors

6.5 Decomposition of Financial Time-Series via SVD

2010년부터 2017년까지의 국채 수익률 곡선(yield curve)를 주 1회 도시해보면 Figure 6.11과 같다. 곡선들은 매주 조금씩 변화하며 만기가 다른 국채들의 시장 가격으로부터 수익률 곡선이 결정된다. 관측된 수익률 곡선들을 분해 또는 잘 설명할 수 있는 곡선 모양들을 SVD를 통해 찾아보자.

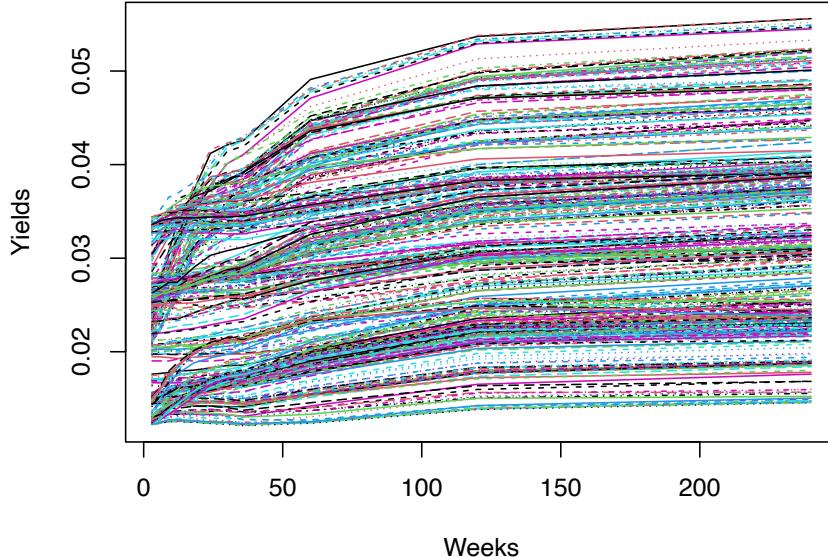


Figure 6.11: 2010년부터 2017년까지의 주별 yield curve들

관측된 수익률 곡선을 모은 행렬을 A 라 하자. A 의 열은 채권의 만기(3개월, 6개월, 9개월, 1년, 1년 6개월, 2년, 2년 6개월, 3년, 5년, 10년, 20년)에서의 수익률을 의미하고 행은 특정 주에서 관측된 수익률 곡선을 나타낸다. 즉 A 의 각 행은 해당 주에 관찰된 수익률 곡선을 나타내는 벡터로 각 열에 대응되는 만기에서의 수익률을 의미한다.

데이터 행렬 A 의 열별 평균벡터 $\mu = \frac{1}{n}A^\top \mathbf{1}_n \in \mathbb{R}^d$ 을 모든 행에서 빼서 열별 평균이 0으로 조정된 행렬

$$\bar{A} = A - \mathbf{1}_n \mu^\top$$

를 분석하자. \bar{A} 의 SVD를 $\bar{A} = UDV^\top$ 라하면 \bar{A} 의 i -번째 행은 $\sum_{k=1}^r u_{ik}\sigma_k \mathbf{v}_k$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$ 이고 $|\mathbf{v}_k| = 1$, $k = 1, \dots, r$ 이다. 2010년부터 2017년까지의 국채 수익률 행렬 A 의 경우, 첫 번째 right-singular vector에 의해 전체 변동의 99.42%가 설명되고 처음 세 개의 right-singular vector들에 의해 총 변동의 99.99%가 설명된다.

이 세 개의 right-singular vector를 도시하면 Figure 6.12a와 같다. 처음 3개의 σ_k 값이 유독 크므로 \bar{A} 의 i -번째 행을

$$\sum_{k=1}^r u_{ik}\sigma_k \mathbf{v}_k \approx u_{i1}\sigma_1 \mathbf{v}_1 + u_{i2}\sigma_2 \mathbf{v}_2 + u_{i3}\sigma_3 \mathbf{v}_3$$

로 근사하더라도 상당히 유사한 곡선이 된다. 첫 번째 right-singular vector는 Figure 6.13에서 빨간 곡선이고 전체적인 yield curve의 값을 결정하므로 level factor라 부른다. 두 번째 right-singular vector인 녹색 곡선은

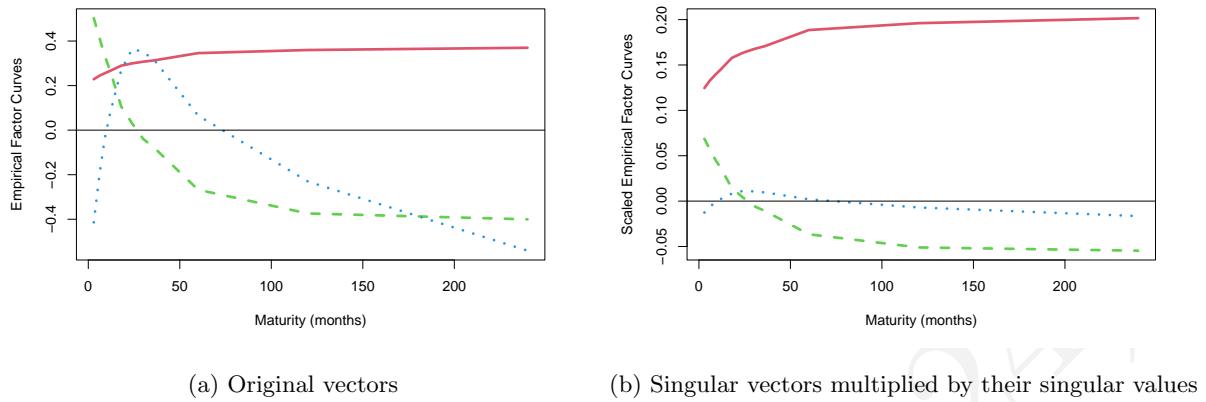


Figure 6.12: Right-singular vectors of leading three singular values

초반의 기울기에 영향을 끼치므로 slope factor로 해석되고, 세 번째 벡터인 파란 곡선은 만기 5년 부근의 곡률에 관련되어 curvature factor라고 불린다. 이러한 경험적 관찰에서 해당 분야 연구자들이 다음과 같은 세 가지 함수 형태를 제안하여 현실 데이터를 이러한 세 함수의 조합으로 표현하기 위한 관련 모수의 추정을 연구하였다. 세 함수형태는

$$1, \quad \frac{1 - e^{-\lambda\tau}}{\lambda\tau}, \quad \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}$$

로서 Figure 6.13에 도시되어 있다.

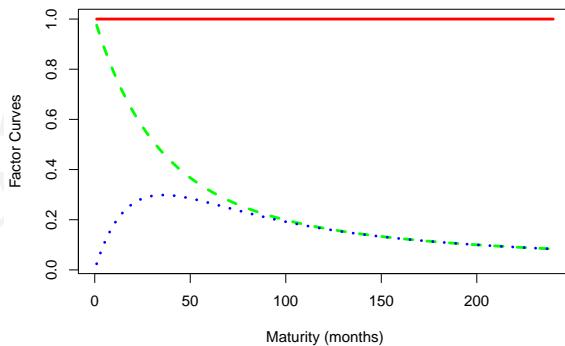


Figure 6.13: Three factor curves of DNS for $\lambda = 0.05$

시점 t 에서 만기 τ 인 국채의 수익률을 $y_t(\tau)$ 로 표시하자. 재무분석에서 L_t, S_t, C_t 와 λ 를 잘 정하여 함수

$$Y_t(\tau) = L_t \times 1 + S_t \times \frac{1 - e^{-\lambda\tau}}{\lambda\tau} + C_t \times \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$$

가 $y_t(\tau)$ 를 잘 근사하도록 정하는 것을 (Dynamic) Nelson-Siegel방법이라 부른다. 각 시점별로 λ 를 각각 정하는 경우 Nelson-Siegel 방법이라 부르고, 시점에 무관하게 최적의 λ 를 정하고 각 시점 t 별로 L_t, S_t, C_t

를 정하는 것을 Dynamic Nelson-Siegel (DNS) 방법이라고 부른다. 만약 analytic curve들을 대응되는 empirical curve라 할 수 있는 right-singular vector $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 로 대치하면 analytic curve들의 계수 벡터인 $\{(L_t, S_t, C_t) : t = 1, \dots, n\}$ 에 $\{(u_{t1}, u_{t2}, u_{t3}) : t = 1, \dots, n\}$ 이 대응된다. 즉 $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ 는 수익률 곡선들이 right-singular vector \mathbf{v}_i 들의 어떤 조합으로 표현되는지 알려주는 \mathbf{u}_i 들 중에서 제일 큰 세 singular value에 대응하며 dynamic Nelson-Siegel 방법의 계수들에 해당한다. Figure 6.14에서 $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ 를 시계열로 도시하였다. 수익률의 전반적인 level을 나타내는 첫번째 factor의 계수인 \mathbf{u}_1 에 해당하는 red curve를 보면 관측기간 전반에 걸쳐 수익률이 감소함을 알 수 있다. 이 세 factor의 흐름을 살펴보면 시간에 따라 수익률이 어떻게 변해 갔는지를 짐작할 수 있다.

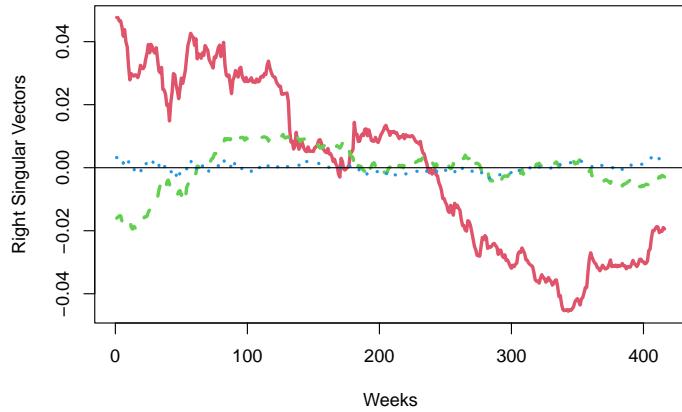


Figure 6.14: $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ plot

수익률의 큰 흐름을 나타내는 $\sigma_1 \mathbf{v}_1$ 의 계수인 u_{t1} 이 음수여도 실제 수익률은 SVD분석 전에 뺀 평균 μ 를 다시 더하면 양수가 될 수 있다. 즉

$$\mu + u_{t1} \sigma_1 \mathbf{v}_1 + u_{t2} \sigma_2 \mathbf{v}_2 + u_{t3} \sigma_3 \mathbf{v}_3$$

은 보통 양수이다.

Chapter 7

Positive Definite Matrices

우리가 인지하지 못하더라도 Positive definite 행렬은 널리 사용되고 있다. 공분산 행렬이라고 불리는 positive definite 행렬은 다양한 확률적 관측들의 상관 정도를 행렬로 표시해준다. 많은 공학과 데이터 사이언스의 결과들이 parameter의 최적화가 필요한데 이때 Hessian 행렬이 positive definite인 것이 최적화 알고리듬의 수행에 중요한 역할을 한다. Positive definite 행렬은 고차원 타원을 통해 볼록성과 연관이 깊고 데이터 분석에서 다양한 아이디어를 제공한다. 정사각행렬이 positive definite하다는 것은 임의의 벡터에 대한 이차 형식이 항상 양수가 되는 경우이다. 그리고 사칙연산과 거듭제곱이 가능하다는 면에서 정사각행렬과 실수가 비슷한데 positive definite 행렬은 양의 실수와 비슷한 면이 많다.

7.1 Positive (Semi-)Definite Matrices

Positive (semi-)definiteness는 다변량분포의 공분산 행렬이나 미분가능한 convex 함수의 이계 미분 행렬이 가지는 성질로 여러 분야에서 자주 활용된다. 또한 벡터 공간의 임의의 내적도 positive definite 행렬과 1-1로 대응된다(Theorem 4.1). Definition 4.2과 유사하게 positive semi-definite matrix도 다변수 이차함수의 계수 행렬일 때 원점에서 최솟값을 갖도록 다음과 같이 정의한다.

Definition 7.1 A square matrix A is positive semi-definite if $\mathbf{x}^\top A \mathbf{x} \geq 0$ for all \mathbf{x} .

다음은 positive semi-definiteness를 규정하는 동치 조건이다.

Fact 7.1 For a symmetric matrix A , the followings are equivalent:

1. For all \mathbf{x} , $\mathbf{x}^\top A \mathbf{x} \geq 0$;
2. All eigenvalues are nonnegative;
3. $A = B^\top B$ for some matrix B .

Proof: Real spectral theorem에 의해 대칭행렬인 A 는 실수로 이루어진 eigenvalue λ_i 와 벡터 \mathbf{v}_i 와 행렬 $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, $V = [\mathbf{v}_1 | \mathbf{v}_2 | \cdots | \mathbf{v}_n]$ 에 대해서

$$A = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top = V \Lambda V^\top$$

로 표현된다. 여기서 \mathbf{v}_i 는 orthonormal하고 V 는 orthogonal하다.

(1) \Rightarrow (2) 모든 j 에 대해서 $\mathbf{v}_j^\top A \mathbf{v}_j = \lambda_j \geq 0$ 이다

(2) \Rightarrow (3) Λ 의 대각원소가 0보다 크거나 같은 실수로 이루어졌으므로 $\Lambda^{1/2} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$ 이 잘 정의된다. $A = V \Lambda V^\top = V \Lambda^{1/2} \Lambda^{1/2} V^\top = B^\top B$.

(3) \Rightarrow (1) $\mathbf{x}^\top A \mathbf{x} = \mathbf{x}^\top B^\top B \mathbf{x} = |B\mathbf{x}|^2 \geq 0$. ■

Positive semi-definite matrix와 유사한 positive definite matrix에 대한 동치 결과는 다음과 같다.

Fact 7.2 For a $d \times d$ symmetric matrix A , the followings are equivalent:

1. For all $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x}^\top A \mathbf{x} > 0$;
2. All eigenvalues are positive;
3. $A = B^\top B$ for some invertible matrix B .

Proof: Fact 7.1의 증명에서처럼 Real spectral theorem에 의해 $A = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top = V \Lambda V^\top$ 이다.

(1) \Rightarrow (2) 모든 j 에 대해서 $\mathbf{v}_j^\top A \mathbf{v}_j = \lambda_j > 0$ 이다

(2) \Rightarrow (3) Λ 의 대각원소가 0보다 큰 실수로 이루어졌으므로 $\Lambda^{1/2} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$ 이 잘 정의된다. $A = V \Lambda V^\top = V \Lambda^{1/2} \Lambda^{1/2} V^\top = B^\top B$ 에서 B 는 orthogonal행렬 V 의 행을 양수로 scaling한 행렬이므로 가역이다.

(3) \Rightarrow (1) $\mathbf{x}^\top A \mathbf{x} = \mathbf{x}^\top B^\top B \mathbf{x} = |B\mathbf{x}|^2 > 0$ if $\mathbf{x} \neq \mathbf{0}$ since B is invertible. ■

Fact 7.3 If A is symmetric and positive definite, then A is invertible, and A^{-1} is also positive definite.

Proof: Fact 7.2에 의해 A 는 가역행렬 B 에 대해 $B^\top B$ 로 표현된다. 가역행렬의 곱도 가역이므로 A 는 가역이다. 그리고 $A^{-1} = B^{-1}(B^{-1})^\top$ 이므로 A^{-1} 도 가역이다. ■

Example 7.1 For a positive (semi-)definite matrix $A = (a_{ij})$,

1. all diagonal entries are positive, i.e., $a_{ii} > (\geq) 0$ for all i since $\mathbf{x} = \mathbf{e}_i$ leads to $\mathbf{x}^\top A \mathbf{x} = a_{ii}$ which has to be greater than (or equal to) 0.
2. leading $k \times k$ block $(a_{ij})_{\substack{1 \leq i \leq k \\ 1 \leq j \leq k}}$ of A called principal minor is also positive (semi-)definite for all k . We can see it by letting $\mathbf{x} = (x_1, \dots, x_k, 0, \dots, 0)^\top$.

$A \nmid$ asymmetric square matrix라고 하자. 임의의 벡터 \mathbf{x} 에 대해

$$\mathbf{x}^\top A \mathbf{x} = \frac{\mathbf{x}^\top A \mathbf{x} + \mathbf{x}^\top A^\top \mathbf{x}}{2} = \mathbf{x}^\top \left(\frac{1}{2}(A + A^\top) \right) \mathbf{x}$$

이므로 A 의 positive (semi-)definiteness는 symmetric matrix인 $\frac{1}{2}(A + A^\top)$ 의 positive (semi-)definiteness와 동일하다. 따라서 symmetric positive definite matrix를 살펴보는 것이 일반적인 행렬에 대한 조건을 포함하는 것으로 생각할 수 있다. 그래서 positive definiteness의 정의에 행렬의 대칭성을 포함하여도 대부분의 경우에 문제가 되지 않는다.

Convexity of Quadratic Forms

함수 $f : \mathbb{R}^m \rightarrow \mathbb{R}$ 가 convex라고 하자. 그리고 임의의 선형사상 $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 을 고려하자. 그러면

$$(f \circ \ell)(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) = f(\alpha \ell(\mathbf{x}) + (1 - \alpha) \ell(\mathbf{y})) \leq \alpha(f \circ \ell)(\mathbf{x}) + (1 - \alpha)(f \circ \ell)(\mathbf{y})$$

이므로 $f \circ \ell$ 은 convex이다. 특히 어떤 $m \times n$ 행렬 A 에 대하여 $\ell(\mathbf{x}) = A\mathbf{x}$ 를 적용하면 $f(A\mathbf{x})$ 는 \mathbf{x} 에 대하여 convex이다.

Theorem 7.1 If A is a positive semi-definite matrix, then $\mathbf{x}^\top A \mathbf{x}$ is a convex function in \mathbf{x} .

Proof: Without losing generality, we may assume that A is symmetric. Then, by item 3 of Fact 7.1, we get

$$A = B^\top B$$

for some $m \times n$ matrix. Consider a function $f(\mathbf{y}) = \mathbf{y}^\top \mathbf{y} = \sum_{i=1}^m y_i^2$. We know that each y_i^2 is convex, and the sum of convex functions is also convex. Hence, f is convex. Then, $f(B\mathbf{x}) = \mathbf{x}^\top B^\top B \mathbf{x} = \mathbf{x}^\top A \mathbf{x}$ is convex following the above discussion. ■

7.2 The Square Root of Positive Semi-definite Matrices

임의의 양의 실수 x 는 적당한 양의 실수 y 에 대해서 $x = y^2$ 이 되는데 이러한 y 는 유일하다. Positive definite 행렬도 동일한 성질을 만족한다. 즉 임의의 positive definite matrix A 는 적당한 positive definite matrix B 에 대해서 $A = B^2$ 을 만족한다. 이 조건은 Fact 7.2의 full rank matrix B 에 대한 $A = B^\top B$ 와는 B 가 대칭이 아닌 한 다른 성질임에 유의하자. A 가 symmetric positive (semi-)definite matrix라고 가정하자. 그러면, real spectral decomposition에 의해 적절한 orthogonal matrix V 와 non-negative diagonal matrix Λ 에 대해

$$A = V^\top \Lambda V$$

로 표현된다. $\Lambda^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_i})$ 로 표시하자. 그러면, $\Lambda^{\frac{1}{2}}$ 과 Λ 는 동일한 positive definiteness를 갖고 $\Lambda = \Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}$ 이다. 그런데 $VV^\top = I$ 이므로 $B = V^\top \Lambda^{\frac{1}{2}}V$ 로 놓으면

$$A = V^\top \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} V = V^\top \Lambda^{\frac{1}{2}} V V^\top \Lambda^{\frac{1}{2}} V = BB = B^2$$

이면서 A 와 B 는 동일한 positive definiteness를 보인다.

Fact 7.4 For a positive (semi-)definite matrix A , there exists a unique positive (semi-)definite matrix B such that $A = B^2 = B^\top B$. We denote this B as $A^{\frac{1}{2}}$.

Proof: Uniqueness만 보이면 된다. Real spectral decomposition의 projection 형태인 (5.14)을 활용하자. 먼저 A 에 (5.14)을 적용하여 $A = \sum_{i=1}^r \lambda_i \mathbf{P}_{A,\lambda_i}$ 를 얻었다고 하자. 그리고 symmetric한 B 에 대해서 $B = \sum_{j=1}^s \mu_j \mathbf{P}_{B,\mu_j}$ 를 얻었다고 하자. 정사영이므로 $\mathbf{P}_{B,\mu_j}^2 = \mathbf{P}_{B,\mu_j}$ 이고, $\mathbf{P}_{B,\mu_j}^\top = \mathbf{P}_{B,\mu_j}$ 이며 $j \neq k$ 인 경우 $\mathbf{P}_{B,\mu_j} \mathbf{P}_{B,\mu_k} = \mathbf{0}$ 이다. 따라서 $B^2 = (\sum_{j=1}^s \mu_j \mathbf{P}_{B,\mu_j})^2 = \sum_{j=1}^s \mu_j^2 \mathbf{P}_{B,\mu_j}$ 이 된다. $A = B^2$ 이려면 eigenvalue의 개수에 대해서 $r = s$ 이어야 하고 적당한 i 에 대해서 $\mu_j = \sqrt{\lambda_i}$ 이고 $\mathbf{P}_{B,\mu_j} = \mathbf{P}_{A,\lambda_i}$ 이다. 따라서

$$B = \sum_{i=1}^r \sqrt{\lambda_i} \mathbf{P}_{A,\lambda_i}$$

이어야 하므로 B 가 unique하다. ■

7.3 Cholesky Factorization of Positive Definite Matrices

Positive definite matrix A 를 고려하자. Fact 7.2에 의해 적절한 가역행렬 B 가 존재하여 $A = B^\top B$ 로 표현된다. B 에 대해 QR 분해를 적용하면 적절한 orthogonal matrix Q 와 양의 대각원소를 갖는 upper triangular matrix R 에 대해 $B = QR$ 로 표현된다. $Q^\top Q = I$ 를 이용하면

$$A = B^\top B = (QR)^\top QR = R^\top Q^\top QR = R^\top R$$

을 얻는다. 이러한 표현을 Cholesky decomposition이라고 부른다.

Fact 7.5 (Cholesky decomposition) For a positive definite matrix A , there exists a unique upper triangular matrix R with positive diagonal entries such that $A = R^\top R$.

Proof: Uniqueness만 보이면 된다. 양의 대각원소를 갖는 상삼각행렬 $R = (r_{ij})$ 과 $S = (s_{ij})$ 가 $A = R^\top R = S^\top S$ 라고 가정하자. 이로부터

$$(S^{-1})^\top R^\top = SR^{-1}$$

를 얻는다. 상삼각행렬의 역행렬은 상삼각행렬이고 상(하)삼각행렬 끼리의 곱은 상(하)삼각행렬이므로 이식의 좌변은 하삼각행렬이고 우변은 상삼각행렬이다. 따라서 SR^{-1} 은 대각행렬이 된다. 특히 $(S^{-1})^\top R^\top = \text{diag}(r_{ii}/s_{ii})$ 이고 $SR^{-1} = \text{diag}(s_{ii}/r_{ii})$ 이므로 모든 i 에 대해

$$\frac{r_{ii}}{s_{ii}} = \frac{s_{ii}}{r_{ii}}$$

인데 R 과 S 의 대각원소가 양수이므로 $s_{ii} = r_{ii}$ 를 얻고 이로부터 $SR^{-1} = I$, 즉 $S = R$ 을 얻는다. ■

An Algorithm for Cholesky Decomposition

$R = (r_{ij})$ 일 때, $A = R^\top R$ 가 되는 R 을 찾는 방법을 생각해 보자. 먼저 $R^\top R$ 은 대칭행렬이고 첫 번째 행은 $r_{11}r_{1j}$ 으로 구성되어 있다. $0 < a_{11} = r_{11}^2$ 에서 $r_{11} = \sqrt{a_{11}}$ 이다. 그러면, $a_{1j} = \sqrt{a_{11}} r_{1j}$ 로부터 행의 나머지 원소 r_{1j} 를 결정할 수 있다. 그리고 이미 r_{12} 를 알고 있으므로 $0 < a_{22} = r_{12}^2 + r_{22}^2$ 에서 r_{22} 를 결정할 수 있다. 두 번째 행의 나머지 원소들은 $a_{2j} = r_{12}r_{1j} + r_{22}r_{2j}$ 에서 r_{2j} 를 제외한 나머지 원소들은 모두 알고 있으므로 r_{2j} 를 정할 수 있다. 이러한 방법으로 나머지 행들의 원소도 모두 결정할 수 있고, 이렇게 얻어진 행렬은 Cholesky 분해의 유일성에 의해 우리가 원하는 행렬임을 알 수 있다.

7.4 Variational Characterization of Symmetric Eigenvalues

$n \times n$ symmetric matrix A 는 n 개의 실수 eigenvalue를 갖는다. 이 n 개의 eigenvalue를 크기 순으로 다음과 같아

$$\lambda_{\max}(A) = \lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A) = \lambda_{\min}(A)$$

로 표시하자. 대칭행렬의 eigenvalue는 *Rayleigh quotient*로 불리는

$$\frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}}, \quad \mathbf{x} \neq \mathbf{0}$$

의 최대/최솟값과 연관지어 생각할 수 있다. 그런데 임의의 $\mathbf{x} \neq \mathbf{0}$ 에 대해서 $\mathbf{y} = \frac{1}{\|\mathbf{x}\|} \mathbf{x}$ 로 놓으면

$$\frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = \mathbf{y}^\top A \mathbf{y}$$

이므로 Rayleigh quotient에 관한 서술은 단위 벡터에 대한 이차 형식으로 바꾸어 서술해도 된다는 점에 유의하자.

모든 eigenvector들이 아닌 일부 벡터에 의해 span되는 부분공간 상의 Rayleigh quotient값은 부분공간을 span하는데 사용한 eigenvector의 짹인 eigenvalue들 중에서 최대/최소를 갖는다는 다음 결과는 간단하지만 이후에 자주 활용된다.

Lemma 7.1 Consider an $n \times n$ symmetric matrix A and its real spectral decomposition $A = \sum_{i=1}^n \lambda_i(A) \mathbf{v}_i \mathbf{v}_i^\top$. Let $1 \leq p \leq q \leq n$. Then, for any non-zero vector $\mathbf{x} \in \text{span}\{\mathbf{v}_p, \mathbf{v}_{p+1}, \dots, \mathbf{v}_q\}$,

$$\lambda_q(A) \leq \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \lambda_p(A).$$

Proof: $\mathbf{x} = \sum_{i=p}^q x_i \mathbf{v}_i$ 이라 하자. 그러면

$$\mathbf{x}^\top A \mathbf{x} = \left(\sum_{i=p}^q x_i \mathbf{v}_i^\top \right) \left(\sum_{j=1}^n \lambda_j(A) \mathbf{v}_j \mathbf{v}_j^\top \right) \left(\sum_{k=p}^q x_k \mathbf{v}_k \right) = \sum_{i=p}^q \lambda_i(A) x_i^2$$

인데, $\lambda_p(A) \geq \lambda_i(A) \geq \lambda_q(A)$ 이고 $\mathbf{x}^\top \mathbf{x} = \sum_{i=p}^q x_i^2$ 이므로

$$\lambda_q(A) \mathbf{x}^\top \mathbf{x} \leq \sum_{i=p}^q \lambda_i(A) x_i^2 \leq \lambda_p(A) \mathbf{x}^\top \mathbf{x}$$

이어서 원하는 결과를 얻는다. ■

o) Lemma의 특별한 경우로 다음 결과를 얻는다.

Theorem 7.2 (Rayleigh quotients) *For any $n \times n$ symmetric matrix A ,*

$$\lambda_{\min}(A) \leq \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \lambda_{\max}(A) \quad \text{for all } \mathbf{x} \neq \mathbf{0}.$$

Moreover,

$$\lambda_{\max}(A) = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \quad \text{and} \quad \lambda_{\min}(A) = \min_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$$

and the maximum and minimum are attained for $\mathbf{x} = \mathbf{v}_1$ and for $\mathbf{x} = \mathbf{v}_n$, respectively, where \mathbf{v}_1 (resp. \mathbf{v}_n) is the unit-norm eigenvector of A associated with its largest (resp. smallest) eigenvalue of A .

Proof: Real spectral decomposition에 의해 대칭행렬 A 는 적절한 orthonormal vector들 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 에 대해서

$$A = \sum_{i=1}^n \lambda_i(A) \mathbf{v}_i \mathbf{v}_i^\top$$

으로 표현 가능하다. Lemma 7.1에서 $p = 1, q = n$ 인 경우에

$$\lambda_1(A) \mathbf{x}^\top \mathbf{x} \geq \mathbf{x}^\top A \mathbf{x} \geq \lambda_n(A) \mathbf{x}^\top \mathbf{x}$$

를 얻게 되어 $\lambda_1(A)$ 와 $\lambda_n(A)$ 가 Rayleigh quotients의 최댓값과 최솟값이라 것을 알 수 있다. 특히 $\mathbf{v}_1^\top A \mathbf{v}_1 = \lambda_1(A)$ 와 $\mathbf{v}_n^\top A \mathbf{v}_n = \lambda_n(A)$ 를 쉽게 확인할 수 있다. ■

Lemma 7.1와 Theorem 7.2를 더 일반화하여 k 차원 부분공간에서의 Rayleigh quotient의 최대/최솟값에 대한 다음과 같은 결과들을 얻을 수 있다.

Lemma 7.2 (Poincare inequality) *Consider an $n \times n$ symmetric matrix A and a k -dimensional subspace \mathbb{W} of \mathbb{R}^n . Then,*

$$\min_{\substack{\mathbf{x} \in \mathbb{W} \\ \mathbf{x} \neq \mathbf{0}}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \lambda_k(A), \quad \max_{\substack{\mathbf{y} \in \mathbb{W} \\ \mathbf{y} \neq \mathbf{0}}} \frac{\mathbf{y}^\top A \mathbf{y}}{\mathbf{y}^\top \mathbf{y}} \geq \lambda_{n-k+1}(A).$$

Proof: A 의 real spectral decomposition o) $A = \sum_{i=1}^n \lambda_i(A) \mathbf{v}_i \mathbf{v}_i^\top$ 로 주어졌다고 하자. $\mathbb{W}' = \text{span}(\{\mathbf{v}_k, \dots, \mathbf{v}_n\})$ 으로 정의하면 $\dim(\mathbb{W}') = n - k + 1$ 이므로 $\mathbb{W} \cap \mathbb{W}' = \{\mathbf{0}\}$ 이라면 $\dim \text{span}(\mathbb{W} \cup \mathbb{W}') = n + 1$ 이 되어 모순이다. 따라서 $\mathbf{x} \in \mathbb{W} \cap \mathbb{W}'$ 인 $\mathbf{x} \neq \mathbf{0}$ 가 존재한다. 그러면, $\mathbf{x} \in \mathbb{W}'$ 이므로 Lemma 7.1에 의해 $\frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \lambda_k(A)$ 인데 $\mathbf{x} \in \mathbb{W}$ 이므로 첫 번째 부등식을 얻는다. 다른 부등식은 $\mathbb{W}' = \text{span}(\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n-k+1}\})$ 로 놓고 동일한 방식으로 전개하면 얻을 수 있다. ■

이상의 결과로부터 대칭 행렬의 모든 eigenvalue가 적절한 제약 하에서의 Rayleigh quotient의 minimax 와 maxmin으로 표현된다는 매우 유용한 결과를 얻는다.

Theorem 7.3 (Minimax Principle) Consider an $n \times n$ symmetric matrix A . Then,

$$\lambda_k(A) = \max_{\dim \mathbb{W}=k} \min_{\substack{\mathbf{x} \in \mathbb{W} \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \quad (7.1)$$

$$= \min_{\dim \mathbb{W}=n-k+1} \max_{\substack{\mathbf{x} \in \mathbb{W} \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \quad (7.2)$$

for $1 \leq k \leq n$.

Proof: A 의 real spectral decomposition이 $A = \sum_{i=1}^n \lambda_i(A) \mathbf{v}_i \mathbf{v}_i^\top$ 로 주어졌다고 하자. Lemma 7.2의 첫 번째 부등식에 의해 $\dim \mathbb{W} = k$ 인 모든 부분공간 \mathbb{W} 에 대해서

$$\min_{\substack{\mathbf{x} \in \mathbb{W} \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \lambda_k(A)$$

이다. 따라서 Rayleigh quotient의 최솟값이 $\lambda_k(A)$ 인 k 차원 부분공간을 찾으면 (7.1)이 증명된다. 그런데 $\mathbb{W}^* = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 인 경우에는 $\mathbf{0} \neq \mathbf{x} \in \mathbb{W}^*$ 인 벡터들에 대해서 Lemma 7.1에 의해 $\frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \geq \lambda_k(A)$ 인데 $\mathbf{v}_k^\top A \mathbf{v}_k = \lambda_k(A)$ 이므로 $\min_{\substack{\mathbf{x} \in \mathbb{W}^* \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = \lambda_k(A)$ 가 된다. 즉 k 차원 부분공간 \mathbb{W}^* 상에서 Rayleigh quotient의 최솟값이 $\lambda_k(A)$ 이므로

$$\lambda_k(A) = \max_{\dim \mathbb{W}=k} \min_{\substack{\mathbf{x} \in \mathbb{W} \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$$

가 성립한다. Lemma 7.2의 나머지 부등식을 이용하면 두 번째 등식을 증명할 수 있다. ■

Theorem 7.3에 의해

$$\lambda_{\max}(A) = \lambda_1(A) = \max_{|\mathbf{x}|=1} \mathbf{x}^\top A \mathbf{x}, \quad \lambda_{\min}(A) = \lambda_n(A) = \min_{|\mathbf{x}|=1} \mathbf{x}^\top A \mathbf{x} \quad (7.3)$$

임을 알 수 있고, A 대신에 $A^\top A$ 를 대입하면

$$\lambda_1(A^\top A) = \max_{|\mathbf{x}|=1} |A\mathbf{x}|^2, \quad \lambda_n(A^\top A) = \min_{|\mathbf{x}|=1} |A\mathbf{x}|^2$$

또는

$$\sqrt{\lambda_1(A^\top A)} = \max_{|\mathbf{x}|=1} |A\mathbf{x}| = \|A\|_2, \quad \sqrt{\lambda_n(A^\top A)} = \min_{|\mathbf{x}|=1} |A\mathbf{x}|$$

가 된다.

7.4.1 Eigenvalues and Singular Values of Matrix Sums

Minimax principle의 중요한 응용으로 Weyl의 부등식으로 불리는 행렬 합의 eigenvalue에 대한 다음 결과가 있다.

Theorem 7.4 (Weyl's inequality of eigenvalues) Let A and B be $n \times n$ symmetric matrices with eigenvalues $\lambda_i(A)$ and $\lambda_i(B)$, respectively. Then

$$\lambda_{k+\ell+1}(A+B) \leq \lambda_{k+1}(A) + \lambda_{\ell+1}(B)$$

for $k, \ell = 0, 1, 2, \dots$

Proof: Let us bound $\lambda_{k+\ell+1}(A+B)$ in terms of the eigenvalues of A and B . By the minimax principle (Theorem 7.3), we have

$$\lambda_{k+\ell+1}(A+B) = \min_{\dim \mathbb{W}=n-k-\ell} \max_{\substack{\mathbf{x} \in \mathbb{W} \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top (A+B)\mathbf{x}}{\mathbf{x}^\top \mathbf{x}}.$$

Again by the minimax principle, we can find subspaces \mathbb{W}_A and \mathbb{W}_B of \mathbb{R}^n of dimensions $n-k$ and $n-\ell$, respectively, such that

$$\lambda_{k+1}(A) = \max_{\substack{\mathbf{x} \in \mathbb{W}_A \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top A\mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \quad \text{and} \quad \lambda_{\ell+1}(B) = \max_{\substack{\mathbf{x} \in \mathbb{W}_B \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top B\mathbf{x}}{\mathbf{x}^\top \mathbf{x}}.$$

If we let $\mathbb{W}_1 = \mathbb{W}_A \cap \mathbb{W}_B$ be their intersection, then it is clear that

$$\max_{\substack{\mathbf{x} \in \mathbb{W}_1 \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top A\mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \lambda_{k+1}(A) \quad \text{and} \quad \max_{\substack{\mathbf{x} \in \mathbb{W}_1 \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^\top B\mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \lambda_{\ell+1}(B).$$

This intersection \mathbb{W}_1 has dimension at least $n-k-\ell$, and let \mathbb{W}_2 be any $(n-k-\ell)$ -dimensional subspace of \mathbb{W}_1 . Then we have

$$\begin{aligned} \lambda_{k+\ell+1}(A+B) &\leq \max_{\mathbf{x} \in \mathbb{W}_2, |\mathbf{x}|=1} \mathbf{x}^\top (A+B)\mathbf{x} \\ &\leq \max_{\mathbf{x} \in \mathbb{W}_1, |\mathbf{x}|=1} \mathbf{x}^\top (A+B)\mathbf{x} \quad (\text{since } \mathbb{W}_2 \subset \mathbb{W}_1) \\ &\leq \max_{\mathbf{x} \in \mathbb{W}_1, |\mathbf{x}|=1} \mathbf{x}^\top A\mathbf{x} + \max_{\mathbf{x} \in \mathbb{W}_1, |\mathbf{x}|=1} \mathbf{x}^\top B\mathbf{x} \\ &\leq \lambda_{k+1}(A) + \lambda_{\ell+1}(B). \end{aligned}$$

■

대칭 행렬의 eigenvalue에 관한 결과는 symmetrization (5.15)에 의해 임의의 행렬의 singular value에 대한 결과로 바꿀 수 있다. 따라서 Theorem 7.4도 임의의 행렬의 singular value에 대한 결과로 기술할 수 있다.

Theorem 7.5 (Weyl's inequality of singular Values) Let A and B be matrices of same size with singular values $\sigma_i(A)$ and $\sigma_i(B)$, respectively. Then

$$\sigma_{k+\ell+1}(A+B) \leq \sigma_{k+1}(A) + \sigma_{\ell+1}(B)$$

for $k, \ell = 0, 1, 2, \dots$

Proof: Lemma 5.4에 의해 $\sigma_i(A) = \lambda_i(s(A))$, $\sigma_i(B) = \lambda_i(s(B))$, $\sigma_i(A + B) = \lambda_i(s(A + B))$ 이다. Symmetrization은 symmetric하므로 Theorem 7.4에 의해서

$$\sigma_{k+\ell+1}(A+B) = \lambda_{k+\ell+1}(s(A+B)) = \lambda_{k+\ell+1}(s(A)+s(B)) \leq \lambda_{k+1}(s(A))+\lambda_{\ell+1}(s(B)) = \sigma_{k+1}(A)+\sigma_{\ell+1}(B)$$

를 얻는다. ■

이 결과는 행렬 A 와 B 의 $k+1$ 번째 singular value와 $\ell+1$ 번째 singular value가 매우 작거나 0이라면 두 행렬의 rank를 k 와 ℓ 로 생각할 수 있는데, 이 경우 두 행렬의 합인 $A+B$ 의 rank가 $k+\ell$ 이하라는 것은 쉽게 알 수 있고 따라서 $k+\ell+1$ 번째 singular value도 매우 작거나 0이어야 한다는 사실과 일치한다.

Theorem 7.5에서 $\ell=1$ 인 경우에 $\sigma_{k+1}(A+B) \leq \sigma_{k+1}(A) + \sigma_1(B)$ 을 얻는다. 행렬 A 와 B 를 적절하게 바꾸면 다음과 같은 결과를 얻는데 이 결과는 여러 가지로 해석된다.

Fact 7.6 (Bound of additive perturbation) Let A and B be $m \times n$ matrices. Then, for $1 \leq k \leq \min\{m, n\}$,

$$|\sigma_k(A) - \sigma_k(B)| \leq \sigma_1(A - B) = \|A - B\|_2.$$

Proof: Theorem 7.5에서 $\ell=1$ 인 경우, A 와 B 를 적절하게 놓으면

$$\sigma_k(A) \leq \sigma_k(B) + \sigma_1(A - B), \quad \sigma_k(B) \leq \sigma_k(A) + \sigma_1(B - A)$$

를 얻는다. 그리고 $\sigma_1(\cdot) = \|\cdot\|_2$ 이므로 원하는 결과를 얻는다. ■

두 대칭행렬을 합한 행렬의 eigenvalue의 상/하한에 관한 다음 부등식도 유용하다.

Fact 7.7 Let A and B be $n \times n$ symmetric matrices. Then, for $1 \leq k \leq n$,

$$\lambda_k(A) + \lambda_{\min}(B) \leq \lambda_k(A + B) \leq \lambda_k(A) + \lambda_{\max}(B).$$

Proof: (7.3)에 의해 모든 $\mathbf{x} \neq \mathbf{0}$ 에 대해 $\lambda_n(B) \leq \frac{\mathbf{x}^\top B \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \lambda_1(B)$ 가 성립한다. 따라서 모든 $\mathbf{x} \neq \mathbf{0}$ 에 대해

$$\frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} + \lambda_n(B) \leq \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} + \frac{\mathbf{x}^\top B \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = \frac{\mathbf{x}^\top (A + B) \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} + \lambda_1(B)$$

가 성립하고, 이로부터

$$\max_{\dim \mathbb{W}=k} \min_{\substack{\mathbf{x} \in \mathbb{W} \\ \mathbf{x} \neq \mathbf{0}}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} + \lambda_n(B) \leq \max_{\dim \mathbb{W}=k} \min_{\substack{\mathbf{x} \in \mathbb{W} \\ \mathbf{x} \neq \mathbf{0}}} \frac{\mathbf{x}^\top (A + B) \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq \max_{\dim \mathbb{W}=k} \min_{\substack{\mathbf{x} \in \mathbb{W} \\ \mathbf{x} \neq \mathbf{0}}} \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} + \lambda_1(B)$$

를 얻는다. Theorem 7.3에 의해 원하는 부등식이 성립한다. ■

물론 Theorem 7.4를 이용해서 위 결과의 두 번째 부등식을 증명할 수도 있지만 첫 번째 부등식의 증명이 필요하여 두 번째 부등식까지 한꺼번에 증명하였다. 이 결과에서 행렬 B 가 positive (semi-)definite인 경우를 생각하면, 대칭행렬에 positive (semi-)definite 행렬을 더하는 경우 eigenvalue는 증가한다는 것을 알 수 있다.

Fact 7.8 Let A and B be $n \times n$ symmetric matrices. Let $1 \leq k \leq n$. If B is positive semi-definite, then

$$\lambda_k(A) \leq \lambda_k(A + B).$$

If B is positive definite, then

$$\lambda_k(A) < \lambda_k(A + B).$$

If we add a symmetric positive semi-definite matrix of rank-1 to a symmetric matrix, we can further obtain an upper bound on the eigenvalues as follows,

$$\lambda_k(A) \leq \lambda_k(A + \mathbf{q}\mathbf{q}^\top) \leq \lambda_k(A) + |\mathbf{q}|^2$$

for all $1 \leq k \leq n$.

Proof: Since B is positive semi-definite, $\lambda_{\min}(B) \geq 0$. Hence, Fact 7.7 implies the first result. If B is positive definite, $\lambda_{\min}(B) > 0$ implies the second result. A symmetric positive semi-definite matrix of rank-1 is always in a form of $\mathbf{q}\mathbf{q}^\top$ for a non-zero vector \mathbf{q} . It is easy to see that $\lambda_{\min}(\mathbf{q}\mathbf{q}^\top) = 0$ and $\lambda_{\max}(\mathbf{q}\mathbf{q}^\top) = |\mathbf{q}|^2$. Fact 7.7 then implies the third result. ■

대칭 행렬에 rank-one positive semi-definite 행렬을 더할 때 eigenvalue의 변화를 좀 더 체계적으로 분석하면 다음의 interlacing 결과를 얻는다.

Theorem 7.6 (Eigenvalue Interlacing) Let A be an $n \times n$ symmetric matrix and B an $n \times n$ symmetric positive semi-definite matrix of rank-one. Then,

$$\lambda_{k+1}(A + B) \leq \lambda_k(A) \leq \lambda_k(A + B), \quad \text{for all } k = 1, \dots, n - 1$$

and

$$\lambda_{k+1}(A) \leq \lambda_k(A - B) \leq \lambda_k(A), \quad \text{for all } k = 1, \dots, n - 1.$$

Proof: Real spectral decomposition에 의해 대칭행렬 $A+B$ 와 A 는 적절한 orthonormal vector들 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 과 $\mathbf{u}_1, \dots, \mathbf{u}_n$ 에 대해서

$$A + B = \sum_{i=1}^n \lambda_i(A + B) \mathbf{v}_i \mathbf{v}_i^\top \quad A = \sum_{i=1}^n \lambda_i(A) \mathbf{u}_i \mathbf{u}_i^\top$$

으로 표현 가능하다. 다음과 같이 \mathbb{R}^n 의 부분공간 3개를 정의하자: $\mathbb{U} = \text{span}\{\mathbf{u}_k, \dots, \mathbf{u}_n\}$, $\mathbb{V} = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k+1}\}$, 그리고 $\mathbb{W} = \text{null } B$. $\dim \mathbb{U} + \dim \mathbb{V} = n + 2$ 이므로 $\dim(\mathbb{U} \cap \mathbb{V}) \geq 2$ 이다. 그리고 $\dim \mathbb{W} = n - \text{rank}(B) = n - 1$ 이므로 $\dim(\mathbb{U} \cap \mathbb{V} \cap \mathbb{W}) \geq 1$ 이어야 한다. 즉 어떤 $\bar{\mathbf{x}}$ 가 $\mathbf{0} \neq \bar{\mathbf{x}} \in \mathbb{U} \cap \mathbb{V} \cap \mathbb{W}$ 를 만족하므로 $|\bar{\mathbf{x}}| = 1$ 로 가정하자. 그러면 $\bar{\mathbf{x}}$ 는 $B\bar{\mathbf{x}} = \mathbf{0}$ 이면서 $\bar{\mathbf{x}} \in \mathbb{V}$ 이고 $\bar{\mathbf{x}} \in \mathbb{U}$ 이다. 그러면

$$\begin{aligned} \lambda_{k+1}(A + B) &\leq \bar{\mathbf{x}}^\top (A + B) \bar{\mathbf{x}} && \text{by Lemma 7.1 and } \bar{\mathbf{x}} \in \mathbb{V} \\ &= \bar{\mathbf{x}}^\top A \bar{\mathbf{x}} && \text{by } B\bar{\mathbf{x}} = \mathbf{0} \end{aligned}$$

$$\leq \lambda_k(A) \quad \text{by Lemma 7.1 and } \bar{\mathbf{x}} \in \mathbb{U}$$

를 얻는다. 이 결과를 Fact 7.8과 결합하면 첫 줄의 두 부등식을 얻는다. 두 번째 줄은 첫째 줄의 A 대신에 $A - B$ 를 대입하면 얻어진다. ■

7.5 Ellipsoidal Geometry of Positive Definite Matrices

$n \times n$ symmetric positive definite matrix A 로 정의할 수 있는 기하학적 대상으로 ellipsoid가 있다. 다음과 같이 이차 부등식으로 정의된 집합

$$\mathcal{E} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^\top A^{-1} \mathbf{x} \leq 1\}$$

을 생각하자. 중심인 \mathbf{a} 인 ellipsoid $\{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} - \mathbf{a})^\top A^{-1}(\mathbf{x} - \mathbf{a}) \leq 1\}$ 는 \mathcal{E} 를 평행이동시키면 얻어지므로 본 절에서는 \mathcal{E} 에 대해서만 살펴본다. Fact 7.3에 의해 A^{-1} 도 positive definite이고, real spectral decomposition에 의해 적절한 orthonormal vector들 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 에 대해서

$$A^{-1} = \sum_{i=1}^n \lambda_i(A)^{-1} \mathbf{v}_i \mathbf{v}_i^\top$$

로 표현된다. 임의의 $\mathbf{x} \in \mathbb{R}^n$ 은 $y_i = \langle \mathbf{x}, \mathbf{v}_i \rangle$ 에 대해서 $\mathbf{x} = \sum_{i=1}^n y_i \mathbf{v}_i$ 로 표시되며

$$\mathbf{x}^\top A^{-1} \mathbf{x} = \sum_{i=1}^n \frac{y_i^2}{\lambda_i(A)}$$

이다. 따라서 orthonormal basis인 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 을 기준으로 표현하면

$$\mathcal{E} = \left\{ \mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^n \frac{y_i^2}{\lambda_i(A)} \leq 1, \mathbf{x} = \sum_{i=1}^n y_i \mathbf{v}_i \right\} \quad (7.4)$$

를 얻게된다. 이 표현을 ellipsoid의 각 축이 표준 기저 방향과 일치하는 우리가 익숙한 ellipsoid의 표현으로 바꾸어 보자. $\mathcal{E}' = \{\mathbf{y} \in \mathbb{R}^n : \sum_{i=1}^n y_i^2 / \lambda_i(A) \leq 1\}$ 에 \mathbf{e}_i 를 \mathbf{v}_i 로 대응시키는 선형 변환을 적용하면 \mathcal{E} 가 얻어진다. 이 ellipsoid의 i -번째로 긴 축은 길이가 $2\sqrt{\lambda_i(A)}$ 이고 방향은 \mathbf{v}_i 이다. 두 기저들 모두 직교좌표계이므로 $\text{vol}(\mathcal{E}) = \text{vol}(\mathcal{E}')$ 임을 직관적으로 알 수 있다. n 차원 공간의 단위구를 $B_n = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^\top \mathbf{x} \leq 1\}$ 로 표시하자. \mathbf{e}_i 를 $\sqrt{\lambda_i(A)}\mathbf{e}_i$ 에 대응시키는 선형변환에 의해 단위구 B_n 이 ellipsoid \mathcal{E}' 으로 변환된다. 서로 직교하는 방향으로 길이 1인 \mathbf{e}_i 가 길이 $\sqrt{\lambda_i(A)}$ 인 $\sqrt{\lambda_i(A)}\mathbf{e}_i$ 에 대응되므로 \mathcal{E}' 의 volume은 단위구 대비 $\prod_{i=1}^n \sqrt{\lambda_i(A)}$ 배만큼 변화한다. 따라서

$$\text{vol}(\mathcal{E}) = \text{vol}(\mathcal{E}') = \sqrt{\prod_{i=1}^n \lambda_i(A)} \text{vol}(B_n) \quad (7.5)$$

를 얻는다. 이 표현에서 eigenvalue들의 곱을 determinant로 표현한 결과를 Fact 9.9에서 얻는다.

연관된 개념으로 Mahalanobis distance를 생각해 보자. Positive definite matrix A 에 대해서 $\mathbf{x}^\top A \mathbf{y}$ 나 $\mathbf{x}^\top A^{-1} \mathbf{y}$ 은 Lemma 4.2에 의해 내적이 되고 $f(\mathbf{x}) = \sqrt{\mathbf{x}^\top A^{-1} \mathbf{x}}$ 는 norm이다. 따라서 $f(\mathbf{x} - \mathbf{y})$ 는 \mathbf{x} 와 \mathbf{y} 간의 거리로 생각할 수 있다. 즉

$$d(\mathbf{x}, \mathbf{y}) = f(\mathbf{x} - \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y})^\top A^{-1}(\mathbf{x} - \mathbf{y})}$$

로 놓으면 거리가 되는데 이 거리를 positive definite matrix A 에 대한 **Mahalanobis distance**라고 부르며 data science에서 벡터들 간의 covariance를 고려한 거리로 이해할 수 있다. 따라서 ellipsoid는 Mahalanobis distance가 동일한 벡터들의 집합으로 이해할 수 있다. Multivariate Gaussian density의 exponent도 Mahalanobis distance로 표현된다.

7.6 Application to Data Science: Kernel Trick in Machine Learning

본 절은 논문 [3]에 기반하여 기술되었다. 전통적인 기계학습 분야에서 수학적으로 잘 분석된 문제 중 하나가 분류 문제(classification)이다. 유클리드 공간의 데이터를 두 개의 그룹으로 선형 함수¹를 기준으로 분류하는 문제를 생각하자. 데이터는

$$\{(\mathbf{z}_i, \ell_i) \in \mathbb{R}^N \times \{+1, -1\} : i = 1, \dots, n\}$$

가 주어졌다고 하자. 각 그룹의 index를 모아서 $\mathcal{I}_+ = \{i : \ell_i = +1\}$ 와 $\mathcal{I}_- \{i : \ell_i = -1\}$ 로 나타내고 개수를 $n_+ = |\mathcal{I}_+|$, $n_- = |\mathcal{I}_-|$ 로 표시하자. 각 그룹의 중심점은

$$\boldsymbol{\mu}_+ = \frac{1}{n_+} \sum_{i \in \mathcal{I}_+} \mathbf{z}_i, \quad \boldsymbol{\mu}_- = \frac{1}{n_-} \sum_{i \in \mathcal{I}_-} \mathbf{z}_i$$

로 놓자.

새로운 관측 \mathbf{z} 에 대한 분류 ℓ 은 두 그룹 중심의 평균 $\boldsymbol{\mu} = \frac{1}{2}(\boldsymbol{\mu}_+ + \boldsymbol{\mu}_-)$ 과 두 그룹 중심의 차이 방향 $\boldsymbol{\mu}_+ - \boldsymbol{\mu}_-$ 을 기준으로 내적이 0보다 큰지 작은지로 결정한다. 즉

$$\ell = \text{sign}\langle \mathbf{z} - \boldsymbol{\mu}, \boldsymbol{\mu}_+ - \boldsymbol{\mu}_- \rangle$$

로 정하는 것이 선형 함수를 이용하는 선형 분류의 전형적인 방법이다. 판단의 기준이 되는 함수를 자세히 살펴보면 $b = \frac{1}{2}(\langle \boldsymbol{\mu}_+, \boldsymbol{\mu}_+ \rangle - \langle \boldsymbol{\mu}_-, \boldsymbol{\mu}_- \rangle) = \frac{1}{2}(|\boldsymbol{\mu}_+|^2 - |\boldsymbol{\mu}_-|^2)$ 인 상수일 때

$$\begin{aligned} \langle \mathbf{z} - \boldsymbol{\mu}, \boldsymbol{\mu}_+ - \boldsymbol{\mu}_- \rangle &= \langle \mathbf{z}, \boldsymbol{\mu}_+ - \boldsymbol{\mu}_- \rangle - \frac{1}{2} \langle \boldsymbol{\mu}_+ + \boldsymbol{\mu}_-, \boldsymbol{\mu}_+ - \boldsymbol{\mu}_- \rangle \\ &= \langle \mathbf{z}, \boldsymbol{\mu}_+ \rangle - \langle \mathbf{z}, \boldsymbol{\mu}_- \rangle - \frac{1}{2} (\langle \boldsymbol{\mu}_+, \boldsymbol{\mu}_+ \rangle - \langle \boldsymbol{\mu}_-, \boldsymbol{\mu}_- \rangle) \\ &= \frac{1}{n_+} \sum_{i \in \mathcal{I}_+} \langle \mathbf{z}, \mathbf{z}_i \rangle - \frac{1}{n_-} \sum_{i \in \mathcal{I}_-} \langle \mathbf{z}, \mathbf{z}_i \rangle - b \end{aligned}$$

와 같이 내적으로 표현된다.

Kernel trick은 데이터가 이렇게 선형적 기준으로 깔끔하게 분류가 되지 않는 데이터를 다루기 위해서 고안되었다. 데이터 $\{(\mathbf{x}_i, \ell_i) \in \mathbb{R}^d \times \{+1, -1\} : i = 1, \dots, n\}$ 의 $\mathbf{x}_i \in \mathbb{R}^d$ 벡터들은 선형 함수로 분류가 잘 안되지만, **feature space**라고 불리는 고차원 공간으로 비선형 변환 $\psi : \mathbb{R}^d \rightarrow \mathbb{R}^N$ 에 의해 embedding된 $\psi(\mathbf{x}_i)$ 벡터들은 선형 함수에 의해 분류가 잘 되는 경우를 고려해 보자. 보통 d 보다 N 이 더 크며 심지어 N

¹더 정확하게는 선형 함수에 상수항을 더한 affine function을 의미한다.

이 무한대인 경우도 있다. 이렇게 비선형 변환 ψ 에 의해 변환된 $\mathbf{z}_i = \psi(\mathbf{x}_i)$ 에 대해 feature space에서 위의 linear classification recipe를 따르면

$$\ell = \text{sign} \left(\frac{1}{n_+} \sum_{i \in \mathcal{I}_+} \langle \psi(\mathbf{x}), \psi(\mathbf{x}_i) \rangle - \frac{1}{n_-} \sum_{i \in \mathcal{I}_-} \langle \psi(\mathbf{x}), \psi(\mathbf{x}_i) \rangle - b \right)$$

이고 b 도 $\langle \psi(\mathbf{x}_i), \psi(\mathbf{x}_j) \rangle$ 들로 표현된다. 이렇게 비선형 변환을 통해 고차원 공간에 embedding된 벡터들의 내적

$$K(\mathbf{x}, \mathbf{y}) = \langle \psi(\mathbf{x}), \psi(\mathbf{y}) \rangle \quad (7.6)$$

을 kernel이라 부르고 $K(\cdot, \cdot)$ 로 나타낸다. 따라서

$$\ell = \text{sign} \left(\frac{1}{n_+} \sum_{i \in \mathcal{I}_+} K(\mathbf{x}, \mathbf{x}_i) - \frac{1}{n_-} \sum_{i \in \mathcal{I}_-} K(\mathbf{x}, \mathbf{x}_i) - b \right)$$

로 간결하게 표현된다. 다음 두 가지 예에서 볼 수 있듯이 $\psi(\mathbf{x})$ 와 그 내적의 계산량에 비해 $K(\cdot, \cdot)$ 의 계산이 간단한 경우에 kernel trick이 유용하다. 이 계산량의 차이는 kernel $K(\cdot, \cdot)$ 가 간단한 함수가 되도록 비선형 변환을 정하기 때문이다.

(7.6)로 표현되는 kernel은 내적의 성질 때문에 $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{y}, \mathbf{x})$ 로 대칭이고 $K(\mathbf{x}_i, \mathbf{x}_j)$ 를 원소로 구성된 행렬은 Theorem 4.1에 의해 positive definite이다. 그래서 이러한 kernel을 **symmetric positive definite kernel**이라고 부른다.

Kernel을 분석하기 위해 몇 가지 표현을 도입한다. 유한 차원 유클리드 공간의 표준 내적과 유사하게 무한 차원 공간에서 $(a_k) = (a_1, a_2, \dots) \in \mathbb{R}^\infty$ 와 $(b_k) = (b_1, b_2, \dots) \in \mathbb{R}^\infty$ 의 내적을

$$\langle (a_k), (b_k) \rangle = \sum_{k=1}^{\infty} a_k b_k$$

로 정한다. 그리고 기호 ;는 좌우에 있는 두 대상을 세로로 적층하는 것으로 약속한다. 즉 두 벡터 \mathbf{u} 와 \mathbf{v} 에 대해서 $(\mathbf{u}; \mathbf{v}) = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$ 이다. 또한 자연수를 성분으로 갖는 벡터 $\mathbf{j} = (j_1, j_2, \dots, j_k)^\top \in \{1, \dots, n\}^k$ 가 필요하다.

Polynomial Kernel

두 벡터 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 에 대해

$$\begin{aligned} (\mathbf{x}^\top \mathbf{y})^k &= \left(\sum_{j=1}^n x_j y_j \right) \cdots \left(\sum_{j=1}^n x_j y_j \right) \\ &= \sum_{\mathbf{j} \in \{1, \dots, n\}^k} \prod_{i=1}^k x_{j_i} y_{j_i} \\ &= \sum_{\mathbf{j} \in \{1, \dots, n\}^k} \prod_{i=1}^k x_{j_i} \prod_{i=1}^k y_{j_i} \end{aligned}$$

인데 $\{1, \dots, n\}^k$ 의 모든 벡터를 $\mathbf{j}^{(1)}, \dots, \mathbf{j}^{(n^k)}$ 로 나열하면 $\{\mathbf{j}^{(1)}, \dots, \mathbf{j}^{(n^k)}\} = \{1, \dots, n\}^k$ 이다. 이 때 $(\mathbf{j}^{(p)})_i = j^{(p)}_i \in \{1, \dots, n\}$ 이다. 따라서

$$(\mathbf{x}^\top \mathbf{y})^k = \sum_{p=1}^{n^k} \prod_{i=1}^k x_{(\mathbf{j}^{(p)})_i} \prod_{i=1}^k y_{(\mathbf{j}^{(p)})_i} = \sum_{p=1}^{n^k} \prod_{i=1}^k x_{j^{(p)}_i} \prod_{i=1}^k y_{j^{(p)}_i}$$

마지막 식의 n^k 개의 합을 두 고차원 벡터의 표준 내적으로 해석하기 위해, 비선형 변환은 n 차원 공간의 벡터를 n^k 차원으로 embedding해야 하므로 $\psi_k : \mathbb{R}^n \rightarrow \mathbb{R}^{n^k}$,

$$\psi_k(\mathbf{x}) = \left(\prod_{i=1}^k x_{j^{(1)}_i}, \prod_{i=1}^k x_{j^{(2)}_i}, \dots, \prod_{i=1}^k x_{j^{(n^k)}_i} \right)^\top \in \mathbb{R}^{n^k} \quad (7.7)$$

로 정의하면

$$(\mathbf{x}^\top \mathbf{y})^k = \langle \psi_k(\mathbf{x}), \psi_k(\mathbf{y}) \rangle$$

이 성립한다. $\{1, \dots, n\}^k$ 를 $\mathbf{j}^{(p)}$ 로 나열하는 순서가 다양하므로 ψ_k 의 표현은 유일하지 않다. 정리하면 n 차원 공간의 k 차 다행식 kernel인

$$K(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^\top \mathbf{y})^k$$

는 n^k 차 공간 벡터의 내적

$$\langle \psi_k(\mathbf{x}), \psi_k(\mathbf{y}) \rangle$$

로 해석할 수 있다.

$\mathbf{x}' = (1; \mathbf{x})$ $\mathbf{y}' = (1; \mathbf{y})$ 라 하면

$$(1 + \mathbf{x}'^\top \mathbf{y}')^k = (\mathbf{x}'^\top \mathbf{y}')^k = \langle \psi_{k+1}(\mathbf{x}'), \psi_{k+1}(\mathbf{y}') \rangle$$

이어서 내적에 양의 상수를 더하여 거듭 제곱한 경우에도 동일한 형태의 embedding으로 표현된다. 이 때 \mathbf{x}' 는 \mathbb{R}^{n+1} 공간 속의 저 차원 manifold인 첫 번째 좌표가 1인 평면 상에 존재한다.

Gaussian Kernel

Kernel trick 중 가장 많이 활용되는 Gaussian kernel의 표준적인 형태는 1차원에서 $K(x, y) = e^{-\frac{1}{2}(x-y)^2}$ 이다. 이 함수를 전개하면

$$\begin{aligned} e^{-\frac{1}{2}(x-y)^2} &= e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}y^2} e^{xy} \\ &= e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}y^2} \sum_{k=0}^{\infty} \frac{1}{k!} (xy)^k \\ &= \sum_{k=0}^{\infty} e^{-\frac{1}{2}x^2} \frac{x^k}{\sqrt{k!}} \times e^{-\frac{1}{2}y^2} \frac{y^k}{\sqrt{k!}} \end{aligned}$$

인데 마지막 표현을 두 수열의 내적으로 해석하기 위해 $\psi_{G1} : \mathbb{R} \rightarrow \mathbb{R}^\infty$ 를

$$\psi_{G1}(x) = e^{-\frac{1}{2}x^2} (1; x; \frac{x^2}{\sqrt{2!}}; \dots; \frac{x^k}{\sqrt{k!}}; \dots) \in \mathbb{R}^\infty$$

로 정의한다. 그러면

$$K(x, y) = \langle \psi_{G1}(x), \psi_{G1}(y) \rangle$$

가 된다.

두 벡터 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 에 대해 $K(\mathbf{x}, \mathbf{y}) = e^{-\frac{1}{2}|\mathbf{x}-\mathbf{y}|^2}$ 은 n 차원 Gaussian kernel로 확장해 보자. 1차원에서 전개한 것과 유사하게 전개한다. 이때 다항식 kernel (7.7)을 이용하면

$$\begin{aligned} e^{-\frac{1}{2}|\mathbf{x}-\mathbf{y}|^2} &= e^{-\frac{1}{2}|\mathbf{x}|^2} e^{-\frac{1}{2}|\mathbf{y}|^2} e^{\mathbf{x}^\top \mathbf{y}} \\ &= e^{-\frac{1}{2}|\mathbf{x}|^2} e^{-\frac{1}{2}|\mathbf{y}|^2} \sum_{k=0} \frac{1}{k!} (\mathbf{x}^\top \mathbf{y})^k \\ &= e^{-\frac{1}{2}|\mathbf{x}|^2} e^{-\frac{1}{2}|\mathbf{y}|^2} \sum_{k=0} \frac{1}{k!} \langle \psi_k(\mathbf{x}), \psi_k(\mathbf{y}) \rangle \\ &= \sum_{k=0} \left\langle e^{-\frac{1}{2}|\mathbf{x}|^2} \frac{1}{\sqrt{k!}} \psi_k(\mathbf{x}), e^{-\frac{1}{2}|\mathbf{y}|^2} \frac{1}{\sqrt{k!}} \psi_k(\mathbf{y}) \right\rangle \end{aligned}$$

이다. 마지막 합을 무한 차원 내적으로 표현하기 위해 $\psi_{Gk} : \mathbb{R}^n \rightarrow \mathbb{R}^\infty$ 를

$$\psi_{Gk}(\mathbf{x}) = e^{-\frac{1}{2}|\mathbf{x}|^2} \left(1; \psi_1(\mathbf{x}); \frac{1}{\sqrt{2!}} \psi_2(\mathbf{x}); \frac{1}{\sqrt{3!}} \psi_3(\mathbf{x}); \dots \right) \in \mathbb{R}^\infty$$

로 정의하면²

$$K(\mathbf{x}, \mathbf{y}) = \langle \psi_{Gk}(\mathbf{x}), \psi_{Gk}(\mathbf{y}) \rangle$$

으로 해석할 수 있다.

Kernel trick은 이론적으로 잘 개발되어 다양하게 활용되고 있다. 자세한 내용은 관련 문헌을 참고하기 바란다.

² ψ_{Gk} 의 앞 쪽 일부만 보면 $\left(1; \psi_1(\mathbf{x}); \frac{1}{\sqrt{2!}} \psi_2(\mathbf{x}); \frac{1}{\sqrt{3!}} \psi_3(\mathbf{x}) \right)$ 은 $\mathbb{R}^{1+n^1+n^2+n^3}$ 으로 차원이 매우 빨리 증가하고 있다.

Kang & Cho (2024)

Chapter 8

Determinants

n 차원 유클리드 공간에 n 개의 벡터가 있을 때, 이 벡터들을 원점에 인접한 모서리로 하는 n 차원 평행($2n$)면체의 부피를 생각해보자. 우리는 예전에 1차원 길이를 쟈 수 있을 때, 2차원 평행사변형의 면적을 한 변의 길이와 이 변에 수직한 방향으로 측정한 인접 변의 길이를 곱하여 면적으로 생각하였다. $n - 1$ 차원 평행($2(n - 1)$)면체의 부피를 어떻게 계산하는지 이미 약속되었다고 하자. 유사하게 n 차원 평행($2n$)면체의 한 모서리를 제외한 $n - 1$ 개 모서리로 만들어지는 $n - 1$ 차원 평행($2(n - 1)$)면체의 부피에 제외된 모서리의 인접한 $n - 1$ 개 모서리에 수직한 방향 길이를 곱하여 계산하는 것으로 약속하자. 그러면 모든 차원의 유클리드 공간에서 부피를 쟈 수 있다. 만약 한 모서리의 길이가 상수배되면 부피도 동일하게 상수배된다. 두 n 차원 평행($2n$)면체가 $n - 1$ 개 모서리를 공유하고 나머지 한 모서리만 다를 때, 두 평행($2n$)면체 부피의 합은 서로 다른 모서리를 벡터로 더하여 다른 $n - 1$ 개 모서리와 함께 만든 평행($2n$)면체의 부피와 같다. 따라서 n 차원 정($2n$)면체의 부피는 각 모서리를 나타내는 벡터들에 대해 선형이다. 그리고 n 차원 평행($2n$)면체의 두 모서리가 서로 바뀌어도 부피는 동일하다. 마지막으로 n 차원 정($2n$)면체 $\{\mathbf{x} \in \mathbb{R}^n : 0 \leq x_i \leq 1 \text{ for all } i = 1, \dots, n\}$ 의 부피는 1을 n 번 곱하여 1이 된다. 이 성질들은 n 차원 정($2n$)면체 부피가 가진 자연스러운 성질인데, 다른 성질은 모두 유지하면서 두 모서리를 서로 바꾸면 부호가 바뀌는 실수값을 갖는 함수는 유일하다라는 것을 곧 증명하고 determinant라 부른다. n 개의 \mathbb{R}^n -벡터를 행으로 쌓아서 만든 행렬을 A 라 하면 $\det(A)$ 가 determinant로서 실수값을 갖는다. n 차원 평행($2n$)면체의 n 개의 모서리 벡터를 행으로 쌓은 행렬이 A 이면 $|\det(A)|$ 가 부피를 나타낸다. \det 는 여러 유용한 성질을 가지고 있다.

8.1 Definition and Properties

Definition 8.1 $n \times n$ 정사각행렬에 대해 정의된 함수 \det 가 다음과 같은 성질을 만족하면 행렬식 (Determinant)라고 부른다:

1. 단위 행렬 I 에 대해 $\det(I) = 1$;
2. 행렬의 임의의 두 행을 교환하면 \det 값의 부호가 바뀐다. 즉 행렬 A 의 어느 두 행을 바꾼 행렬을 \bar{A} 이라 하면 $\det(A) = -\det(\bar{A})$ 이다;
3. 행렬의 각각의 행에 대해 선형이다. 즉 행렬 A 의 어떤 행이 $\mathbf{b} + \alpha\mathbf{c}$ 일 때, A 와 나머지 행들은 동일하고 해당 행만 \mathbf{b} 인 행렬과 \mathbf{c} 인 행렬을 각각 B 와 C 라 하면 $\det(A) = \det(B) + \alpha \det(C)$ 이다^a.

^a2번과 행렬식이 행렬의 첫 번째 행에 대해 선형이라는 가정을 조합하면 3번을 얻을 수 있다. 따라서 최소로 정의하기를 원하는 경우, 3번을 “행렬의 첫 번째 행에 대해 선형이다”라고 바꾸어도 무방하다.

정의 8.1의 세 가지 간단한 성질에 의해 다음과 같은 행렬식의 여러 성질을 유도할 수 있다.

Fact 8.1 *If two rows of A are equal, then $\det(A) = 0$.*

Proof: 동일한 A 의 두 행을 교환해도 행렬이 변하지 않으므로 A 는 A 의 두 행을 교환한 행렬로 생각할 수 있다. 따라서 $\det(A) = -\det(A)$ 가 되어 $\det(A) = 0$ 이다. ■

Fact 8.2 *Adding a multiple of one row to another row leaves the same determinant.*

Proof: 행렬 B 의 i 번째 행을 \mathbf{b} 라 하고 j 번째 행을 \mathbf{c} 라 하자. $i \neq j$ 라 하자. B 와 동일한데 i 번째 행만 \mathbf{c} 로 바꾼 행렬을 C 라 하고 B 와 동일한데 i 번째 행만 $\mathbf{b} + \alpha\mathbf{c}$ 로 바꾼 행렬을 A 라 하자. 행렬 C 의 i 행과 j 행이 동일하므로 Fact 8.1에 의해 $\det(C) = 0$ 이고 정의 8.1의 3번에 의해 $\det(A) = \det(B) + \alpha \det(C) = \det(B)$ 가 된다. ■

Fact 8.3 *If A has a row of zeros, then $\det(A) = 0$.*

Proof: 정의 8.1의 3번에서 $\alpha = -1$, $\mathbf{b} = \mathbf{c}$ 로 놓으면 $B = C$ 이므로 한 행이 영벡터인 행렬의 행렬식은 0이 됨을 알 수 있다. ■

Fact 8.4 *If A is triangular, then $\det(A)$ is the product of the diagonal entries.*

Proof: A 가 상삼각행렬이라 가정하자. 만약 $a_{nn} = 0$ 이라면 A 의 n 번째 행이 영벡터이므로 $\det(A) = 0$ 이 되어 명제가 참이다. $a_{nn} \neq 0$ 이라 하자. Fact 8.2에 의해 마지막 행의 상수배를 다른 행에 더하는 것은 행렬식의 값을 변화시키지 않는다. 그런데 상삼각행렬의 마지막 행은 a_{nn} 만 0이 아니므로 a_{nn} 을 제외한 A 의 마지막 열의 다른 원소들을 0으로 바꾸어도 행렬식이 변하지 않는다. 이 행렬을 A_{n-1} 이라 하자. 같은 과정을 $a_{(n-1)(n-1)}$ 에 대해 반복하면 마지막 두 행과 두 열에서 대각선 상의 두 원소만 0이 아니면서 원래 A 의 행렬식의 값은 동일한 행렬을 얻게된다. 이 과정을 반복하면 A 의 대각원소만으로 이루어진 대각행렬을

D 로 표시하면 $\det(A) = \det(D)$ 가 된다. 그리고 정의 8.1의 3번과 1번에 의해 $\det(D)$ 는 대각원소들의 곱과 같아지므로 명제가 증명된다. ■

Fact 8.4의 증명에서 알 수 있듯이 대각행렬의 대각원소 중에 0인 원소가 있으면 그 대각행렬의 행렬식은 0이된다. 일반적으로 비가역 정사각행렬의 행렬식은 0이 된다.

Fact 8.5 A is singular if and only if $\det(A) = 0$.

Proof: A 의 LU-decomposition을 생각하자. 즉 적당한 permutaiton 행렬 P 에 대해서 row-echelon form을 얻는 행연산을 나타내는 하삼각행렬 \tilde{L} 와 상삼각행렬 U 가 존재하여 $\tilde{L}PA = U$ 로 표현된다. 여기서, 하삼각행렬 \tilde{L} 는 어떤 행의 상수배를 다른 행에 더하는 행연산들을 나타내므로 Fact 8.2에 의해 $\det(\tilde{L}PA) = \det(PA)$ 이다. 또한 P 는 A 의 행들을 교환하는 연산들을 나타내는 permutation matrix이므로 Definition 8.1의 2번에 의해 $\det(PA) = \pm \det A$ 이다. 따라서 $\det(\tilde{L}PA) = \pm \det A = \det U$ 이므로 $\det A = \pm \det U$ 가 A 의 singularity와 관계없이 항상 성립한다. 이미 앞 장에서 알아본 바와 같이 정사각행렬 A 가 invertible이기 위한 필요충분조건은 row echelon form U 의 대각원소들이 모두 0인 아닌 pivot 원소들인 것이다. 그런데 이 조건은 Fact 8.4에 의해 $\det(U) \neq 0$ 과 동치이다. 따라서 명제가 성립한다. ■

아직 Definition 8.1를 만족하는 함수가 실제로 존재하는지, 존재한다면 유일한지 등을 생각해보자 않았다. Definition 8.1의 2, 3번을 만족하는 임의의 함수를 f 라 하자. 지금까지 증명된 determinant의 성질에 관한 Fact들 중 Fact 8.1, 8.2, 8.3은 Definition 8.1의 1번을 사용하지 않았으므로 방금 가정한 함수 f 에 대해서도 적용가능하다. A' 을 행렬 A 에서 첫 행을 제외한 두 번째 행부터 n 번째 행으로 이루어진 $(n-1) \times n$ 행렬로 놓자. 그리고 첫 행에서 j 번째 원소만 남기고 나머지 원소는 0으로 놓은 $n \times n$ 행렬 $A_j = \begin{bmatrix} a_{1j} \mathbf{e}_j^\top \\ \vdots \\ A' \end{bmatrix}$ 를 구성하자. A 의 첫 번째 행은 $a_{11} \mathbf{e}_1^\top + a_{12} \mathbf{e}_2^\top + \cdots + a_{1n} \mathbf{e}_n^\top$ 이므로 Definition 8.1의 세 번째 성질에 의해 $f(A) = f(A_1) + \cdots + f(A_n)$ 이다. 두 번째 행에 대해서도 동일한 방법으로 전개하자. A'' 을 행렬 A 에서 첫 두 행을 제외한 세 번째 행부터 n 번째 행으로 이루어진 $(n-2) \times n$ 행렬이라면 $A_{jk} = \begin{bmatrix} a_{1j} \mathbf{e}_j^\top \\ a_{2k} \mathbf{e}_k^\top \\ \vdots \\ A'' \end{bmatrix}$ 에 대해서 $f(A_j) = f(A_{j1}) + \cdots + f(A_{jn})$ 로 전개된다. 이때 $f(A_{jj}) = a_{1j} a_{2j} f\left(\begin{bmatrix} \mathbf{e}_j^\top \\ \mathbf{e}_j^\top \\ \vdots \\ A'' \end{bmatrix}\right)$ 의 첫 두 행이 동일하므로 $f(A_{jj}) = 0$ 이 된다. 이 과정을 모든 행에 대해 n 번 반복하여 k 번째 행에서 j_k 번째 원소를 선택하면 $A_{j_1 j_2 \dots j_n}$ 은 k 번째 행이 $a_{kj_k} \mathbf{e}_{j_k}^\top$ 인 행렬이고 위에서 살펴본 것처럼 어느 두 행 k, ℓ 에 대해서 $j_k = j_\ell$ 이면 $f(A_{j_1 j_2 \dots j_n}) = 0$ 이 된다. 이렇게 자명하게 0이 되는 조합을 제거하면 우리가 고려할 $A_{j_1 j_2 \dots j_n}$ 은 모두 다른 j_k 로 이루어진 경우 뿐이다. 이러한 경우는 j_1, j_2, \dots, j_n 은 $1, 2, \dots, n$ 의 순서를 바꾸어 얻어진다. 즉 고등학교에서 배웠던 순열(permutation)이 된다. 보통은 순열을 σ 로 표현하며 $\sigma(i)$ 가 i 번째에 대응되는 자연수를 의미한다. 따라서 $\sigma(k) = j_k$ 인 순열을 생각하면 $A_{j_1 j_2 \dots j_n} = A_{\sigma(1)\sigma(2)\dots\sigma(n)}$ 이다. 순열 기호를 이용하여 f 를 표현하면

$$f(A) = \sum_{\sigma: \text{permutation of } \{1, \dots, n\}} f(A_{\sigma(1)\sigma(2)\dots\sigma(n)})$$

이다. 순열 σ 에 대해 permutation matrix를 $P_\sigma = \begin{bmatrix} \mathbf{e}_{\sigma(1)}^\top \\ \vdots \\ \mathbf{e}_{\sigma(n)}^\top \end{bmatrix}$ 로 정의하자. 본 강의록에서 증명하지 않지만 현대대수학에서 증명되는 내용으로 P_σ 의 두 행을 교환하여 단위행렬 I 로 바꾸는 방법은 여러가지이고 교환의 횟수도 일의적이지 않지만 교환의 횟수는 항상 짝수이거나 홀수라는 사실이 있다. 그래서 permutation을 even permutation과 odd permutation으로 구분하고 $\text{sign}(\sigma) = 1$ 또는 -1 의 값을 갖는 sign 함수를 순열에 대해서 정의할 수 있다. 이 결과를 이용하면

$$\begin{aligned} f(A_{\sigma(1)\sigma(2)\dots\sigma(n)}) &= f\left(\begin{bmatrix} a_{1\sigma(1)}\mathbf{e}_{\sigma(1)}^\top \\ a_{2\sigma(2)}\mathbf{e}_{\sigma(2)}^\top \\ \vdots \\ a_{n\sigma(n)}\mathbf{e}_{\sigma(n)}^\top \end{bmatrix}\right) = a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}f\left(\begin{bmatrix} \mathbf{e}_{\sigma(1)}^\top \\ \mathbf{e}_{\sigma(2)}^\top \\ \vdots \\ \mathbf{e}_{\sigma(n)}^\top \end{bmatrix}\right) \\ &= a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}f(P_\sigma) = \text{sign}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}f(I) \end{aligned}$$

이므로 $f(A)$ 의 전개 꼴 표현을

$$f(A) = \sum_{\sigma: \text{permutation of } \{1, \dots, n\}} \text{sign}(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}f(I) \quad (8.1)$$

와 같이 나타낼 수 있다. 다시 한번 강조하면 (8.1)은 Definition 8.1의 2번과 3번 만을 사용하여 전개된 것으로 \det 함수도 만족하는 식이다. 반대로 (8.1)로 정의된 함수 f 는 Definition 8.1의 2번과 3번을 만족함을 확인할 수 있다.

Theorem 8.1 *The function satisfying Definition 8.1 is unique.*

Proof: (8.1)로 정의된 함수 f 는 Definition 8.1의 2번과 3번을 만족하므로 추가적으로 $f(I) = 1$ 로 정의해주면 Definition 8.1의 세 가지 규칙을 모두 만족하는 함수이다. 따라서 Definition 8.1를 모두 만족하는 함수가 적어도 한 개는 존재하고 이 함수를 \det 로 표시하자. g 가 Definition 8.1의 2번과 3번을 만족하는 함수라 가정하자. 그리고 임의의 $n \times n$ 행렬 A 에 대해 다음과 같이

$$h(A) = g(A) - \det(A)g(I)$$

h 를 정의하자. 그러면 $h(I) = g(I) - \det(I)g(I) = 0$ 이 된다. 그런데 함수 g 와 \det 가 모두 Definition 8.1의 2번과 3번을 만족하므로 h 도 2번과 3번을 만족하고, 따라서 (8.1)의 꼴로

$$h(A) = \sum_{\sigma: \text{permutation of } \{1, \dots, n\}} \text{sign}(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}h(I)$$

와 같이 전개된다. 이 전개 표현에서 $h(I) = 0$ 이므로 모든 행렬에 대해서 $h(A) = 0$ 이 됨을 알 수 있다. 즉 $g(A) = \det(A)g(I)$ 의 관계가 모든 행렬에 A 에 대해서 성립한다. $g(I) = 1$ 이면 $g \equiv \det$ 이다. 따라서 Definition 8.1을 만족하는 함수는 유일하게 존재한다. ■

o] Theorem은 determinant의 존재와 유일성 외에도 증명과정에서 (8.1)의 전개 꼴인

$$\det A = \sum_{\sigma: \text{permutation of } \{1, \dots, n\}} \text{sign}(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} \quad (8.2)$$

를 제공하는데 이 (8.2) 식은 이후에 이론적인 전개나 작은 크기의 행렬의 determinant를 계산할 때 유용하다. Theorem 8.1의 중요한 응용으로 determinant가 행렬곱과 잘 어울리는 것을 알 수 있다.

Fact 8.6 For two square matrix A and B of same sizes, $\det(AB) = \det(A)\det(B)$.

Proof: 주어진 행렬 A 와 B 를 고정하자. B 가 singular하면 증명할 등식의 양변이 모두 0이되어 등식이 성립 한다. 따라서 B 가 non-singular한 경우만 살펴보면 되므로 $\det B \neq 0$ 이라 가정하자. 임의의 $n \times n$ 정사각행렬 C 에 대해 실수값을 대응시키는 함수

$$\rho(C) = \frac{\det(CB)}{\det(B)}$$

를 고려하자.

1. $C = I$ 이면 $\rho(I) = \frac{\det(IB)}{\det(B)} = 1$;
2. CB 의 각 행은 C 의 행과 행렬 B 의 곱으로 결정되므로 C 의 i 행과 j 행을 교환하면 CB 의 해당 행들이 교환된다. 따라서 $\det(CB)$ 의 부호가 바뀌는데 $\det(B)$ 의 부호는 그대로이므로 $\rho(C)$ 의 부호는 바뀐다;
3. C 의 어떤 행이 $\mathbf{b}^\top + \alpha\mathbf{c}^\top$ 이면 CB 의 해당 행은 $\mathbf{b}^\top B + \alpha\mathbf{c}^\top B$ 이 되고 분모는 그대로이므로 ρ 는 각 행들에 대해 선형이 됨을 알 수 있다.

따라서 새로운 함수 ρ 는 정의 8.1를 만족하므로 Theorem 8.1에 의해 행렬식임을 알 수 있다. 즉 모든 행렬에 대해 $\rho(C) = \det(C)$ 이다. 따라서 $C = A$ 인 경우

$$\rho(A) = \frac{\det(AB)}{\det(B)} = \det(A)$$

이므로 곱의 행렬식은 각각의 행렬식의 곱과 같음을 알 수 있다. ■

만약 A 가 가역행렬이면 Fact 8.5에 의해 $\det(A) \neq 0$ 이고 $AA^{-1} = I$ 에 대해 Fact 8.6를 적용하면

$$\det(A^{-1}) = \det(A)^{-1} = \frac{1}{\det(A)}$$

임을 알 수 있다.

Fact 8.7 For a square matrix A , $\det(A^\top) = \det(A)$.

Proof: Consider the decomposition $PA = LDU$ where D is a diagonal matrix, and U is an upper triangular matrix whose diagonal entries are 1's. As before, L is a lower triangular matrix whose diagonals are also 1's. Then $\det(P)\det(A) = \det(PA) = \det(LDU) = \det(L)\det(D)\det(U) = \det(D)$ since $\det(L) = \det(U) = 1$ by Fact 8.4. Noticing $D^\top = D$, $\det(A^\top)\det(P^\top) = \det(A^\top P^\top) = \det(U^\top D^\top L^\top) = \det(D)$. Since the permutation matrix P is obtained by exchanging the rows of I , $\det(P) = \pm 1$ by the second property of Definition 8.1. From $PP^\top = I$, $\det(P)\det(P^\top) = 1$ and $\det(P) = \det(P^\top)$. The transposition doesn't change the determinant. ■

8.2 Formulas for the Determinant

행렬의 determinant를 표현하는 (8.2) 식을 작은 크기의 행렬에 대해 구체적으로 계산해보자. 먼저 2×2 행렬

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 를 살펴보자. $(a, b) = (a, 0) + (0, b)$ 와 $(c, d) = (c, 0) + (0, d)$ 를 생각하여

$$\begin{aligned}\det(A) &= \det \begin{bmatrix} a & 0 \\ c & d \end{bmatrix} + \det \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} \\ &= \det \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} + \det \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} + \det \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix} \\ &= 0 + \det \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} - \det \begin{bmatrix} c & 0 \\ 0 & b \end{bmatrix} + 0 \\ &= ad - bc\end{aligned}$$

로 계산할 수 있다. 같은 방식으로 3×3 행렬

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

의 행렬식도 계산해 보자. 다음과 같이 전개할 수 있는데, 빨강, 파랑, 보라, 흥색 부분의 의미나 관계를 생각해 보기 바란다.

$$\begin{aligned}\det(A) &= \det \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ &= \det \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ &\quad + \det \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ &\quad + \det \begin{bmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & 0 & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ &= \det \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & 0 & 0 \end{bmatrix} + \det \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{bmatrix} + \det \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{bmatrix}\end{aligned}$$

$$\begin{aligned}
& + \det \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{bmatrix} + \det \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & a_{32} & 0 \end{bmatrix} + \det \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \\
& + \det \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{bmatrix} + \det \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{bmatrix} + \det \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \\
& + \det \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ a_{31} & 0 & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \\
& + \det \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & a_{22} & 0 \\ 0 & a_{32} & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \\
& + \det \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \\
& + \det \begin{bmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ a_{31} & 0 & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \\
& + \det \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ 0 & a_{32} & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \\
& + \det \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \\
& = a_{11}a_{22}a_{33} \times \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + a_{11}a_{23}a_{32} \times \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} + a_{12}a_{21}a_{33} \times \det \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
& \quad + a_{12}a_{23}a_{31} \times \det \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} + a_{13}a_{21}a_{32} \times \det \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + a_{13}a_{22}a_{31} \times \det \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\
& = \sum_{\sigma: \text{permutation of } \{1,2,3\}} a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)} \det(P_\sigma) \\
& = a_{11}a_{22}a_{33} \times 1 + a_{11}a_{23}a_{32} \times (-1) + a_{12}a_{21}a_{33} \times (-1) \\
& \quad + a_{12}a_{23}a_{31} \times (-1)^2 + a_{13}a_{21}a_{32} \times (-1)^2 + a_{13}a_{22}a_{31} \times (-1) \\
& = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})
\end{aligned}$$

$$\begin{aligned}
&= a_{11} \times \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \times \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \times \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\
&= \sum_{j=1}^n (-1)^{1+j} a_{1j} \det(A_{1j}).
\end{aligned}$$

여기서 A_{1j} 는 원래 행렬 A 에서 1행과 j 열을 제거한 $(n-1) \times (n-1)$ 부분행렬(submatrix)을 의미하며 마지막 등식은 3×3 행렬의 경우를 $n \times n$ 크기로 증명없이 확장한 것이다. 이 부분은 받아들이되 첫째 행에 대한 전개나 임의의 행에 대한 전개나 같은 값을 준다는 다음 사실을 증명하자.

$$\sum_{j=1}^n (-1)^{1+j} a_{1j} \det(A_{1j}) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}).$$

일반적으로 A_{ij} 는 원래 행렬 A 에서 i 행과 j 열을 제거한 $(n-1) \times (n-1)$ 부분행렬(submatrix)을 의미한다. 먼저 행렬 A 의 1행과 i 행을 맞바꾼 행렬을 \bar{A} 로 표시하자. 그러면 $\det(A) = -\det(\bar{A})$ 이다. 그리고 \bar{A}_{1j} 는 \bar{A} 의 1행과 j 열을 제거한 $(n-1) \times (n-1)$ 행렬인데 행렬 A 기준으로는 i 번째 행을 제외한 $(n-1)$ 개 행을 $2, \dots, (i-1), 1, (i+1), \dots, n$ 의 순서로 행으로 나열하고 j 번째 열을 제거한 행렬이다. A 의 1번 행이 있는 \bar{A}_{1j} 의 $(i-1)$ 번째 행을 $(i-2)$ 번의 행교환을 통해 첫번째 행까지 이동시킨 행렬은 행렬 A 의 행번호 기준으로 $1, 2, \dots, (i-1), (i+1), \dots, n$ 번째 행을 나열하고 j 번째 열을 제거한 A_{ij} 와 동일하다. 행교환 때마다 부호가 바뀌므로 $\det(\bar{A}_{1j}) = (-1)^{i-2} \det(A_{ij})$ 이다. 이를 대입하면

$$\det(\bar{A}) = \sum_{j=1}^n (-1)^{1+j} a_{ij} \det(\bar{A}_{1j}) = \sum_{j=1}^n (-1)^{1+j} a_{ij} (-1)^{i-2} \det(A_{ij}) = \sum_{j=1}^n (-1)^{i+j-1} a_{ij} \det(A_{ij})$$

이어서 원하던 결과를 얻는다. 그리고 $C_{ij} = (-1)^{i+j} \det(A_{ij})$ 를 cofactor라고 부르며 이를 활용하면 임의의 i 에 대해서 다음과 같은 cofactor expansion

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}) = \sum_{j=1}^n a_{ij} C_{ij}.$$

를 얻는다.

Fact 8.8 Assume that A is an $n \times n$ square matrix A . For any i ,

$$\det(A) = \sum_{j=1}^n a_{ij} C_{ij} = \sum_{j=1}^n a_{ji} C_{ji}.$$

Proof: $\det(A) = \sum_{j=1}^n a_{ij} C_{ij}$ 은 이미 보였다. 이를 A^\top 에 적용하면 $(-1)^{i+j} \det((A_{ji})^\top) = (-1)^{i+j} \det(A_{ji}) = C_{ji}$ 이므로

$$\det(A^\top) = \sum_{j=1}^n (A^\top)_{ij} (-1)^{i+j} \det((A_{ji})^\top) = \sum_{j=1}^n a_{ji} C_{ji}.$$



즉 cofactor expansion은 임의의 행 뿐만 아니라 임의의 열에 대해서도 적용할 수 있다.

Fact 8.9 $\sum_{j=1}^n a_{kj} C_{ij} = 0$ for $i \neq k$.

Proof: A 의 i 번째 행을 k 번째 행으로 바꾼 행렬을 B 라 하자. 즉 B 의 i 번째 행과 k 번째 행은 동일하여 $\det(B) = 0$ 이다. A 와 B 의 원소들은 i 번째 행을 제외하면 모두 동일하여 두 행렬의 i 번째 행에 대한 cofactor는 일치하고 B 의 cofactor도 C_{ij} 이다. 따라서 행렬 B 의 i 번째 행에 대한 cofactor expansion은

$$0 = \det(B) = \sum_{j=1}^n b_{ij} C_{ij} = \sum_{j=1}^n a_{kj} C_{ij}$$

이므로 증명된다. ■

Example 8.1 Let us compute the determinant of (the second difference matrix)

$$A = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}.$$

For an appropriate elementary matrix L , $A = LD = L \begin{bmatrix} 2 & -1 & & & \\ & \frac{3}{2} & -1 & & \\ & & \ddots & \ddots & \\ & & & \frac{n}{n-1} & -1 \\ & & & & \frac{n+1}{n} \end{bmatrix}$, and therefore, $\det A = 2 \times \frac{3}{2} \times \cdots \times \frac{n}{n-1} \times \frac{n+1}{n} = n + 1$. ■

8.2.1 Determinant of Block Matrices

먼저 가장 간단한 square block matrix로서 A_{11} 과 A_{22} 가 각각 (다른 크기일 수도 있는) 정사각행렬일 때

$$A = \begin{bmatrix} A_{11} & \mathbf{0} \\ \mathbf{0} & A_{22} \end{bmatrix}$$

의 determinant를 계산해보자. 이미 알고있는 사실로부터 힌트를 얻기 위해 block의 크기가 1×1 으로 생각할 수도 있는 diagonal matrix $\begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}$ 를 생각하면 determinant는 $a_{11}a_{22} = \det([a_{11}])\det([a_{22}])$ 이다. 이로부터

$$\det(A) = \det(A_{11})\det(A_{22})$$

로 추측하는 것이 자연스럽다. A_{11} 과 A_{22} 의 크기가 각각 n_1 과 n_2 라 하자. 그러면, A 는 $(n_1 + n_2) \times (n_1 + n_2)$ 행렬이다.

$\{1, \dots, n_1 + n_2\}$ 의 permutation 을 σ 로 표시하고 그중에서 $\{1, \dots, n_1\}$ 만 permute하고 $\{n_1 + 1, \dots, n_1 + n_2\}$ 상에서는 identity인 permutation을 σ_1 으로 나타내자. 반대로 $\{n_1 + 1, \dots, n_1 + n_2\}$ 만 permute하고 $\{1, \dots, n_1\}$ 상에서는 identity인 permutation을 σ_2 로 나타내자. σ 에 대응되는 permutation matrix가 P_σ 일 때 $\text{sign}(\sigma) = \det P_\sigma$ 이다. 몇 가지 관찰할 수 있는 점은 $\sigma_1 \circ \sigma_2$ 는 $\{1, \dots, n_1 + n_2\}$ 상의 permutation인데 앞쪽 n_1 개와 뒤쪽 n_2 개 내에서 각각 permutation이 이루어지므로 $\sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1$ 이고, 역으로 이러한 permutation은 반드시 $\sigma_1 \circ \sigma_2$ 꼴로 표현된다. 여기서 행렬 A 와 관련된 중요한 관찰을 해보자. $\sigma_1 \circ \sigma_2$ 로 표현되지 않는 permutation σ 는 반드시 $i \leq n_1$ 이면서 $\sigma(i) > n_1$ 이거나 $i > n_1$ 이면서 $\sigma(i) \leq n_1$ 인 i 를 갖고 있다. 이러한 $i, \sigma(i)$ 는 A_{11} 과 A_{22} 의 영역을 벗어난 위치로서 $a_{i\sigma(i)} = 0$ 이다. 따라서 block matrix의 determinant를 계산하는 합에서 0이 아닐수도 있어 반드시 합해야하는 permutation은 $\sigma_1 \circ \sigma_2$ 형태 만으로 제한해도 된다. 아래 표현을 간단히 나타내기 위해 $\mathcal{S} = \text{set of permutation of } \{1, \dots, n_1 + n_2\}$, $\mathcal{S}_1 = \text{set of permutation of } \{1, \dots, n_1\}$, $\mathcal{S}_2 = \text{set of permutation of } \{n_1 + 1, \dots, n_1 + n_2\}$ 이라 하자.

$$\begin{aligned}\det(A) &= \sum_{\sigma \in \mathcal{S}} a_{1\sigma(1)} \cdots a_{(n_1+n_2)\sigma(n_1+n_2)} \text{sign}(\sigma) \\ &= \sum_{\sigma_1 \in \mathcal{S}_1, \sigma_2 \in \mathcal{S}_2} a_{1\sigma_1 \circ \sigma_2(1)} \cdots a_{(n_1+n_2)\sigma_1 \circ \sigma_2(n_1+n_2)} \text{sign}(\sigma_1 \circ \sigma_2) \\ &= \sum_{\sigma_1 \in \mathcal{S}_1} \sum_{\sigma_2 \in \mathcal{S}_2} a_{1\sigma_1(1)} \cdots a_{n_1\sigma_1(n_1)} a_{(n_1+1)\sigma_2(n_1+1)} \cdots a_{(n_1+n_2)\sigma_2(n_1+n_2)} \text{sign}(\sigma_1) \text{sign}(\sigma_2) \\ &= \sum_{\sigma_1 \in \mathcal{S}_1} a_{1\sigma_1(1)} \cdots a_{n_1\sigma_1(n_1)} \text{sign}(\sigma_1) \sum_{\sigma_2 \in \mathcal{S}_2} a_{(n_1+1)\sigma_2(n_1+1)} \cdots a_{(n_1+n_2)\sigma_2(n_1+n_2)} \text{sign}(\sigma_2) \\ &= \det(A_{11}) \det(A_{22})\end{aligned}$$

임을 알 수 있다.

다음 단계로 다음과 같은 block lower triangular matrix를 고려하자:

$$A = \begin{bmatrix} A_{11} & \mathbf{0} \\ A_{21} & A_{22} \end{bmatrix}.$$

일반 행렬의 경우 determinant가 대각원소의 곱으로 표현되므로 block matrix의 경우도

$$\det(A) = \det(A_{11}) \det(A_{22})$$

도 추측할 수 있다. 먼저 $\det(A_{11}) = 0$ 이면 $\text{rank}(A_{11}) < n_1$ 이므로 $\text{rank}(A) < n_1 + n_2$ 가 되어 $\det(A) = 0$ 이 되어 추측한 등식이 성립한다. A_{11} 이 invertible한 경우를 가정하자. Gaussian elimination을 수행하는 것을 상상하면

$$\begin{bmatrix} I_{11} & \mathbf{0} \\ -A_{21}A_{11}^{-1} & I_{22} \end{bmatrix} \begin{bmatrix} A_{11} & \mathbf{0} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & \mathbf{0} \\ \mathbf{0} & A_{22} \end{bmatrix}$$

를 얻는다. 우변의 행렬의 determinant는 $\det(A_{11}) \det(A_{22})$ 임을 이미 알았다. 좌변의 첫 번째 행렬은 lower triangular block matrix이긴 하지만 diagonal block들이 각각 identity matrix들이므로 보통의 lower triangular matrix이고 determinant가 대각원소들을 모두 곱한 1이다. 이 관찰을 합하면 우리가 원하던 결과를 얻는다.

마지막으로 가장 일반적인 block matrix를 고려하자. 먼저 A_{11} 이 invertible한 경우를 고려한다. A_{21} 을 소거하기 위한 Gaussian elimination을 수행하는 과정을 표현하면 다음과 같다.

$$\begin{bmatrix} A_{11}^{-1} & \mathbf{0} \\ -A_{21}A_{11}^{-1} & I_{22} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_{11} & A_{11}^{-1}A_{12} \\ \mathbf{0} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}$$

따라서 양변의 determinant를 비교하면

$$\det(A_{11}^{-1})\det(A) = \det(A_{22} - A_{21}A_{11}^{-1}A_{12})$$

를 얻고 역행렬의 determinant가 역수관계 임을 이용하면

$$\det(A) = \det(A_{11})\det(A_{22} - A_{21}A_{11}^{-1}A_{12}) \quad (8.3)$$

를 얻는데 $A_{22} - A_{21}A_{11}^{-1}A_{12}$ 는 Schur complement라는 것에 유의하자. 2×2 행렬의 행렬식 $a_{11}a_{22} - a_{21}a_{12} = a_{11}(a_{22} - a_{21}a_{11}^{-1}a_{12})$ 와 비교해보기 바란다.

8.2.2 Matrix Determinant Lemma

직전에 공부한 block matrix의 determinant를 응용하여 얻을 수 있는 결과로 matrix determinant lemma가 있다. Section 8.4에서 역행렬을 이미 아는 기준 행렬에 낮은 rank를 갖는 두 행렬의 곱을 더하여 변경하는 경우에 역행렬이 어떻게 변경되는지 설명하는 Sherman-Morrison 또는 Woodbury formula를 유도한다. 이 때 지금 공부할 matrix determinant lemma가 중요한 역할을 한다.

다음 등식의 행렬들을 어떻게 생각해 낼 수 있는지 직관적으로 설명할 수 없지만 다음과 같은 세 block matrix의 곱을 계산해보자.

$$\begin{bmatrix} I_n & \mathbf{0} \\ V^\top & I_k \end{bmatrix} \begin{bmatrix} I_n + UV^\top & U \\ \mathbf{0} & I_k \end{bmatrix} \begin{bmatrix} I_n & \mathbf{0} \\ -V^\top & I_k \end{bmatrix}.$$

여기서, U 와 V 는 $n \times k$ matrix들이다. 세 행렬 중 앞쪽 두 행렬의 곱이 $\begin{bmatrix} I_n + UV^\top & U \\ V^\top + V^\top UV^\top & I_k + V^\top U \end{bmatrix}$ 이므로 세 행렬의 곱은

$$\begin{bmatrix} I_n + UV^\top & U \\ V^\top + V^\top UV^\top & I_k + V^\top U \end{bmatrix} \begin{bmatrix} I_n & \mathbf{0} \\ -V^\top & I_k \end{bmatrix} = \begin{bmatrix} I_n & U \\ \mathbf{0} & I_k + V^\top U \end{bmatrix}$$

로 정리된다. 따라서 처음 세 행렬의 곱의 determinant와 마지막 행렬의 determinant가 일치해야 하므로

$$\det(I_n + UV^\top) = \det(I_k + V^\top U) \quad (8.4)$$

임을 알 수 있다. 이 결과를 이용하면 다음과 같은 일반적인 결과를 얻는다.

Theorem 8.2 (Matrix Determinant Lemma) *Let A be an $n \times n$ invertible matrix, and U and V be $n \times k$ matrices. Then,*

$$\det(A + UV^\top) = \det(A)\det(I_k + V^\top A^{-1}U). \quad (8.5)$$

Proof: $A + UV^\top = A(I_n + A^{-1}UV^\top)$ 이므로 식 (8.4)의 U 대신에 $A^{-1}U$ 를 넣으면 다음과 같이 전개할 수 있다.

$$\begin{aligned}\det(A + UV^\top) &= \det(A(I_n + A^{-1}UV^\top)) \\ &= \det(A) \det(I_n + A^{-1}UV^\top) \\ &= \det(A) \det(I_k + V^\top A^{-1}U) \quad \text{by plugging } A^{-1}U \text{ into } U \text{ in (8.4).}\end{aligned}$$

■

이 결과는 determinant 값을 실제로 계산하는 경우에도 활용되지만 $A + UV^\top$ 의 inverse가 행렬 $I_k + V^\top A^{-1}U$ 가 invertible한 경우에만 존재한다는 사실로도 자주 활용된다. $k = 1$ 인 경우에는 U, V 가 \mathbb{R}^n -vector \mathbf{u}, \mathbf{v} 가 되고, matrix determinant formula가

$$\det(A + \mathbf{u}\mathbf{v}^\top) = \det(A) \det(1 + \mathbf{v}^\top A^{-1}\mathbf{u}) \quad (8.6)$$

로 표현된다.

8.3 Applications of Determinant

8.3.1 Volume of a Parallelograms in \mathbb{R}^n

\mathbb{R}^n 공간에 있는 n 개의 벡터 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 을 고려하자. 원점과 이 벡터들로 구성된 고차원 평행사변형(hyper-parallelogram)의 부피 V 를 계산해보자. 만약 벡터들이 linearly independent하지 않다면 n 차원 기준으로는 volume을 갖지 못한다. 따라서 n 개의 벡터가 linearly independent하다고 가정하자.

- $\mathbf{a}_1, \dots, \mathbf{a}_n$ 이 orthogonal하다고 가정하자. 이 벡터들을 열로 갖는 행렬을 $A = [\mathbf{a}_1 | \dots | \mathbf{a}_n]$ 로 표시하자. 벡터들이 orthogonal하므로 계산하고자 하는 volume V 는 각 벡터들의 크기의 곱이다. 즉 $V = \prod_{i=1}^n |\mathbf{a}_i|$ 이다. 그런데

$$A^\top A = \begin{bmatrix} |\mathbf{a}_1|^2 & 0 & 0 & \cdots & 0 \\ 0 & |\mathbf{a}_2|^2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & |\mathbf{a}_{n-1}|^2 & 0 \\ 0 & 0 & \cdots & 0 & |\mathbf{a}_n|^2 \end{bmatrix}$$

이므로 Fact 8.7에 의해

$$\det(A)^2 = \det(A^\top A) = \prod_{i=1}^n |\mathbf{a}_i|^2 = V^2$$

가 되어 $V = |\det(A)|$ 임을 알 수 있다.

- $\mathbf{a}_1, \dots, \mathbf{a}_n$ 의 linear independence만 가정하자. 앞에서처럼 $A = [\mathbf{a}_1 | \dots | \mathbf{a}_n]$ 로 표시하자. 이미 배운 QR-decomposition을 A 에 적용하여 $A = QR$, Q 는 orthogonal 행렬, R 은 상삼각행렬이라 하자. $(i-1)$ 차원 부분공간 $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{i-1}\}$ 속에서 $\{\mathbf{a}_1, \dots, \mathbf{a}_{i-1}\}$ 로 구성된 $(i-1)$ 차원 parallelogram

의 부피가 계산되었을 때, $\{\mathbf{a}_1, \dots, \mathbf{a}_i\}$ 로 구성된 i 차원 parallelogram의 부피에 대한 \mathbf{a}_i 의 기여는 $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{i-1}\}$ 에 수직한 방향으로의 성분이다. 상삼각행렬 R 의 i 번째 열은 orthonormal한 Q 의 열로 구성된 기저의 basic vector $\{\mathbf{q}_1, \dots, \mathbf{q}_i\}$ 에 대한 좌표이므로 \mathbf{a}_i 가 $(i-1)$ 차원 부피에 추가적으로 기여한 높이는 \mathbf{q}_i 에 대한 좌표인 $R_{ii} = \langle \mathbf{a}_i, \mathbf{q}_i \rangle$ 의 절댓값이다. 따라서 Fact 8.4에 의해

$$V = \left| \prod_{i=1}^n R_{ii} \right| = |\det(R)|$$

임을 알 수 있다. 그런데 $A^\top A = R^\top Q^\top QR = R^\top R$ 이어서 $\det(A)^2 = \det(R)^2$ 이고 $V = |\det(A)|$ 가 된다.

8.3.2 Computation of A^{-1}

2×2 행렬 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 의 행렬식에서 관찰할 수 있는 또 다른 결과는

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix}$$

이다. 여기서 C_{ij} 는 cofactor이다. 양변의 행렬 왼쪽에 A 를 곱하여 다시 써보면

$$I = \frac{1}{\det(A)} A \begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix}$$

를 얻는다. Cofactor matrix를 $C = (C_{ij})$ 로 놓자. 그러면, 2×2 행렬의 경우에

$$A^{-1} = \frac{1}{\det(A)} C^\top$$

임을 알 수 있다. 이를 일반화하기 위해 앞에서 관찰한 임의의 i 번째 행에 대한 cofactor expansion

$$\det(A) = \sum_{j=1}^n a_{ij} C_{ij} = [a_{i1}, \dots, a_{in}] [C_{i1}, \dots, C_{in}]^\top$$

와 Fact 8.9를 이용하여

$$\begin{aligned} AC^\top &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix} = \begin{bmatrix} \det(A) & 0 & \dots & 0 \\ 0 & \det(A) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \det(A) \end{bmatrix} \\ &= \det(A)I \end{aligned}$$

을 얻을 수 있다. 따라서 $n \times n$ 행렬 A 에 대해서 $\det(A) \neq 0$ 이면

$$A^{-1} = \frac{1}{\det(A)} C^\top \quad (8.7)$$

임을 알 수 있다.

8.3.3 Cramer's Rule: Solution of $Ax = b$

A 가 가역행렬일 때 역행렬 공식 (8.7)를 바탕으로 다음과 같은 상상을 할 수 있다. A 의 j 번째 열을 \mathbf{a}_j 라 하자.

$$\begin{aligned}\mathbf{x} &= A^{-1}\mathbf{b} \\ &= \frac{1}{\det(A)} C^\top \mathbf{b} \\ &= \frac{1}{\det(A)} \left(\sum_{i=1}^n C_{ij} b_i \right)\end{aligned}$$

인데 행렬 A 의 j 열을 \mathbf{b} 로 바꾸어 determinant를 계산하면 $\sum_{i=1}^n C_{ij} b_i$ 이 된다. 따라서

$$x_j = \frac{1}{\det(A)} \det [\mathbf{a}_1 | \dots | \mathbf{a}_{j-1} | \mathbf{b} | \mathbf{a}_{j+1} | \dots | \mathbf{a}_n]$$

로 표현되는데 이를 Cramer's rule이라 부른다. Cramer's rule은 매우 간결한 표현으로 수학적 사고에는 매우 유용하나, 실제 선형연립방정식의 해를 구하는데 사용하기에는 계산이 매우 비효율적이다.

8.4 Sherman-Morrison and Woodbury Formulas

A 가 $n \times n$ invertible matrix일 때 $n \times k$ 행렬 U 와 V 에 대해 $A + UV^\top$ 의 inverse를 생각해보자. 중요한 점은 A 의 역행렬을 이미 알고 있을 때 이를 활용하여 새로운 행렬의 역행렬을 표현하는 것이다. Matrix determinant lemma (8.5)에 의해 $A + UV^\top$ 의 inverse는 $I_k + V^\top A^{-1} U$ 가 invertible한 경우에만 존재한다.

먼저 $I_n + UV^\top$ 의 inverse를 찾아보자. (쉽게 생각할 수 있는 아이디어는 아니지만) inverse 형태로 적당한 $k \times k$ matrix B 에 대해 $I_n + UBV^\top$ 를 시도해보자. 그러면,

$$(I_n + UV^\top)(I_n + UBV^\top) = I_n + UV^\top + UBV^\top + UV^\top UBV^\top = I_n + U(I_k + B + V^\top UB)V^\top$$

이므로 $I_k + B + V^\top UB = \mathbf{0} \nmid I_n + UV^\top$ 의 inverse를 갖기 위한 충분조건 임을 알 수 있다. 즉

$$-I_k = B + V^\top UB = (I_k + V^\top U)B$$

로부터

$$B = -(I_k + V^\top U)^{-1}$$

를 얻는다. 그런데 matrix determinant lemma에 의해 $I_k + V^\top U$ 가 invertible한 것이 $I_n + UV^\top$ 가 invertible하기 위한 필요충분조건이므로

$$(I_n + UV^\top)^{-1} = I_n - U(I_k + V^\top U)^{-1}V^\top \quad (8.8)$$

를 얻는다. 이제 $A + UV^\top$ 의 inverse를 구해보자.

$$\begin{aligned}(A + UV^\top)^{-1} &= (A(I_n + A^{-1}UV^\top))^{-1} \\ &= (I_n + A^{-1}UV^\top)^{-1}A^{-1} \\ &= (I_n - A^{-1}U(I_k + V^\top A^{-1}U)^{-1}V^\top)A^{-1}\end{aligned}$$

$$= A^{-1} - A^{-1}U(I_k + V^\top A^{-1}U)^{-1}V^\top A^{-1}$$

이므로 다음과 같은 결과를 얻을 수 있다. 이 결과를 A^{-1} 를 아는 경우에 $n \times n$ 행렬의 역행렬 계산이 $k \times k$ 행렬의 역행렬 계산으로 줄어든다는 측면으로 해석할 수도 있다.

Theorem 8.3 (Woodbury Formula) *Let A be an $n \times n$ invertible matrix, and U and V be $n \times k$ matrices. Then, $I_k + V^\top A^{-1}U$ is invertible if and only if $A + UV^\top$ is invertible. In such cases,*

$$(A + UV^\top)^{-1} = A^{-1} - A^{-1}U(I_k + V^\top A^{-1}U)^{-1}V^\top A^{-1}. \quad (8.9)$$

Proof: 두 조건이 필요충분조건임은 matrix determinant lemma (8.5)에 의해 자명하다. 그리고 inverse 표현은 위에서 얻었다. ■

이 정리는 $I_k + V^\top A^{-1}U$ 의 invertibility가 $A + UV^\top$ 가 역행렬을 가질 충분조건 중의 하나라는 것을 의미한다.

응용에서는 $k = 1$ 인 경우가 매우 중요하다. 이 경우에는 U 와 V 가 \mathbb{R}^n -vector인 \mathbf{u} 와 \mathbf{v} 가 되고 $A + \mathbf{u}\mathbf{v}^\top$ 가 invertible하기 위한 필요충분조건이 $1 + \mathbf{v}^\top A^{-1}\mathbf{u} \neq 0$ 이 된다.

Corollary 8.1 (Sherman-Morrison Formula) *Let A be an $n \times n$ invertible matrix, and \mathbf{u} and \mathbf{v} be \mathbb{R}^n -vectors. Then, $1 + \mathbf{v}^\top A^{-1}\mathbf{u} \neq 0$ if and only if $A + \mathbf{u}\mathbf{v}^\top$ is invertible. In such cases,*

$$(A + \mathbf{u}\mathbf{v}^\top)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{u}\mathbf{v}^\top A^{-1}}{1 + \mathbf{v}^\top A^{-1}\mathbf{u}}. \quad (8.10)$$

8.5 Application to Optimization: Rank-one update of Hessians

대부분의 학습 과정은 학습 모형의 parameter에 대한 최적화 과정으로 이해할 수 있다. 학습의 수준을 나타내는 함수를 최소화 또는 최대화하는 parameter를 찾아가는 과정이다. 이 parameter를 $\mathbf{w} \in \mathbb{R}^n$ 으로 나타내자. Section 3.9에서 neural network의 parameter $w_{ij}^{(k)}$ 가 좋은 예이다. 그리고 학습의 수준을 나타내는 함수를 $f(\mathbf{w})$ 로 표시하고 loss function이라고 부르는데 함수 값이 작을수록 좋아서 $\mathbf{w}^* = \operatorname{argmin}_{\mathbf{w}} f(\mathbf{w})$ 의 해를 구하는 과정이 학습이다. 학습을 포함하여 어떤 함수 f 를 추가적인 제약없이 최소화하는 일반적인 최적화 문제에 대해 셀 수 없을 정도로 많은 알고리듬이 제안되었다. 본 절에서는 그 중의 한 가지인 **symmetric rank-one update** 방법을 소개한다. 대부분의 비선형 최적화 알고리듬은 gradient가 영벡터가 되는 $\nabla f(\mathbf{w}) = \mathbf{0}$ 를 만족하는 \mathbf{w}^* 로 수렴하는 것을 목표로 한다. f 가 실수 값을 갖는 충분히 미분 가능한 다변수 함수일 때 다음과 같은 일차 미분인 gradient vector ∇f 와 이차 미분인 Hessian matrix $\nabla^2 f$ 을 생각하자:

$$\nabla f = \begin{pmatrix} \frac{\partial}{\partial w_1} f \\ \vdots \\ \frac{\partial}{\partial w_n} f \end{pmatrix}, \quad H = \nabla^2 f = \begin{bmatrix} \frac{\partial}{\partial w_1} \frac{\partial}{\partial w_1} f & \cdots & \frac{\partial}{\partial w_1} \frac{\partial}{\partial w_n} f \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial w_n} \frac{\partial}{\partial w_1} f & \cdots & \frac{\partial}{\partial w_n} \frac{\partial}{\partial w_n} f \end{bmatrix}.$$

다변수 함수 f 의 \mathbf{w}_k 주변에서의 2차 테일러 근사는 다음과 같이 주어진다:

$$f(\mathbf{w}) \approx f(\mathbf{w}_k) + \nabla f(\mathbf{w}_k)^\top (\mathbf{w} - \mathbf{w}_k) + \frac{1}{2} (\mathbf{w} - \mathbf{w}_k)^\top H(\mathbf{w}_k) (\mathbf{w} - \mathbf{w}_k).$$

대부분의 함수 f 에 대해서 $\frac{\partial}{\partial w_i} \frac{\partial}{\partial w_j} f(\mathbf{w}) = \frac{\partial}{\partial w_j} \frac{\partial}{\partial w_i} f(\mathbf{w})$ 이므로 $H(\mathbf{w})$ 는 symmetric 이고 Fact 4.19에 의해 f 의 근사에 대한 gradient는

$$\nabla f(\mathbf{w}) \approx \nabla f(\mathbf{w}_k) + H(\mathbf{w}_k)(\mathbf{w} - \mathbf{w}_k)$$

이라고 생각해도 무방하다. 따라서 $\nabla f(\mathbf{w}) = \mathbf{0}$ 인 \mathbf{w} 를 gradient의 근사식에서 찾기 위해서 secant condition

$$\nabla f(\mathbf{w}_k) + H(\mathbf{w}_k)(\mathbf{w} - \mathbf{w}_k) = \mathbf{0}$$

를 풀면 $\nabla f(\mathbf{w}) = \mathbf{0}$ 인 \mathbf{w} 의 근사값이 $\mathbf{w} = \mathbf{w}_k - H(\mathbf{w}_k)^{-1} \nabla f(\mathbf{w}_k)$ 로 주어진다. 근사식에서 계산된 벡터이므로 대부분 $\nabla f(\mathbf{w}) \neq \mathbf{0}$ 이다. 따라서 근사식에서 계산된 벡터를 다음 단계의 벡터 \mathbf{w}_{k+1} 로 정하여 다음과 같은 iteration을 통해 원하는 벡터 \mathbf{w}^* 로 수렴하기를 기대한다:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - H(\mathbf{w}_k)^{-1} \nabla f(\mathbf{w}_k). \quad (8.11)$$

이러한 방법을 Newton-Raphson 방법이라고 부른다. 적절한 조건에서 \mathbf{w}_k 가 \mathbf{w}^* 로 수렴하는 것을 보일 수 있고 모든 계산을 정확히 할 수 있는 경우 수렴 속도도 빠르다. 그런데 parameter가 많아지면 $n \times n$ 행렬인 Hessian matrix의 역행렬을 매번 구하는 것은 너무 많은 계산을 필요로 한다. Gradient 계산은 할 수 있지만 2차 미분인 Hessian $H(\mathbf{w}_k)$ 를 얻을 수 없다는 가정 하에서 Newton-Rapson 방법과 유사한 iterative한 방법을 찾아보자.

(8.11)에서 이차 미분에 해당하는 행렬 $H(\mathbf{w}_k)$ 을 비슷한 값을 갖는 대칭행렬 B_k 로 대체하고

$$\mathbf{w}_{k+1} = \mathbf{w}_k - B_k^{-1} \nabla f(\mathbf{w}_k) \quad (8.12)$$

에 의해 다음 parameter \mathbf{w}_{k+1} 을 정한 상황에서 2차 미분 정보없이 새로운 B_{k+1} 이 \mathbf{w}_{k+1} 에서 2차 미분 $H(\mathbf{w}_{k+1})$ 과 비슷해지기를 바라며 gradient ∇f 의 \mathbf{w}_{k+1} 에서의 1차 근사식인

$$\nabla f(\mathbf{w}_{k+1}) - \nabla f(\mathbf{w}_k) = B_{k+1}(\mathbf{w}_{k+1} - \mathbf{w}_k) \quad (8.13)$$

를 만족하는 B_{k+1} 을 찾는다. (8.13)은 n 개의 방정식으로 대칭인 $n \times n$ 행렬을 찾는 문제이므로 당연히 여러 개의 행렬 해를 갖는데 이렇게 찾아진 B_{k+1} 을 바탕으로 구성된 여러 최적화 알고리듬을 통칭하여 quasi-Newton methods라고 부른다. 이 중 한가지가 B_k 에 rank가 1인 행렬을 더하는 symmetric rank-one 방법이다. $\Delta \mathbf{w}_k = \mathbf{w}_{k+1} - \mathbf{w}_k$, $\mathbf{d}_k = \nabla f(\mathbf{w}_{k+1}) - \nabla f(\mathbf{w}_k)$ 의 기호로 만족해야 할 quasi-Newton update식 (8.13)을 표현하면 $\mathbf{d}_k = B_{k+1} \Delta \mathbf{w}_k$ 인데 symmetric rank-one 방법은

$$B_{k+1} = B_k + \frac{1}{(\mathbf{d}_k - B_k \Delta \mathbf{w}_k)^\top \Delta \mathbf{w}_k} (\mathbf{d}_k - B_k \Delta \mathbf{w}_k)(\mathbf{d}_k - B_k \Delta \mathbf{w}_k)^\top \quad (8.14)$$

로 다음 대칭행렬 B_{k+1} 을 결정한다. B_{k+1} 은 (8.13)을 만족한다는 것을 확인해보기 바란다. 매 iteration에서 parameter를 수정하기 위해 (8.12)을 풀어야 하므로 B_{k+1}^{-1} 를 알아야 한다. Sherman-Morrison formula (8.10)를 (8.14)에 적용하면

$$B_{k+1}^{-1} = B_k^{-1} - \frac{1}{(\Delta \mathbf{w}_k - B_k^{-1} \mathbf{d}_k)^\top \mathbf{d}_k} (\Delta \mathbf{w}_k - B_k^{-1} \mathbf{d}_k)(\Delta \mathbf{w}_k - B_k^{-1} \mathbf{d}_k)^\top \quad (8.15)$$

를 얻는다. B_k^{-1} 를 알고 있다면 B_{k+1}^{-1} 로 update하는 계산은 단순하게 rank-one 행렬을 더하는 작업이어서 iteration에 필요한 계산이 많지 않게 된다. 실제 알고리듬은 B_k^{-1} 만을 유지하여 (8.14)를 사용하지 않지만 update식을 해석하는 측면에서는 (8.15)보다는 (8.14)이 편리하고 이 두 update를 연결하는 과정에 Sherman-Morrison formula가 요긴하게 쓰인다.

Kang & Cho (2024)

Chapter 9

Eigenvalues and Eigenvectors

정의 5.1에서 이미 eigenvalue와 eigenvector를 소개하였지만 자세한 공부는 지금껏 미루었다. 이 책에서 지금까지 eigenvalue와 eigenvector에 대해서 기본적인 정의 외에는 다른 결과를 이용한 적이 없었음을 확인해보기 바란다.

어떤 변환이 간단하면서 설명이 가능한 선형변환이 있는가? 3.8.2절에서 살펴본 바와 같이 확대/축소를 하는 scaling 변환이 가장 간단하면서도 변환의 역할이 명확하였다. 그러면 일반적인 행렬 A 가 나타내는 선형변환을 scaling변환으로 설명할 수 있을까? 당연히 공간의 모든 벡터에 대해서는 scaling변환으로 해석하는 것이 불가능할 것이다. 그럼에도 일반적인 변환도 일차원 부분공간 상으로 제한하면 scaling변환으로 해석할 수 있을까라는 희망을 가져보자. 어떤 scaling 상수 λ 에 대해서 적절한 벡터 \mathbf{v} 가 생성하는 일차원 부분공간 상에서 $A = \lambda I$ 가 성립하면 된다. 즉 모든 실수 x 에 대해서 부분공간 상의 벡터 $x\mathbf{v}$ 들이

$$xA\mathbf{v} = A(x\mathbf{v}) = (\lambda I)(x\mathbf{v}) = x(\lambda I\mathbf{v}) = x\lambda\mathbf{v}$$

또는 $A\mathbf{v} = \lambda\mathbf{v}$ 를 만족하는 λ 와 \mathbf{v} 를 찾으면 된다. 이 λ 를 eigenvalue, \mathbf{v} 를 eigenvector, 한꺼번에 (λ, \mathbf{v}) 를 eigenpair라고 부른다. 하나의 eigenvalue에 한 개 이상의 선형독립인 eigenvector가 위의 식을 만족할 수도 있다. 이 경우 eigenvector들에 의해 생성되는 부분공간 상에서 $A = \lambda I$ 와 같이 scaling 변환으로 설명할 수 있다. 이 모든 설명은 복소수 eigenvalue에 대해서도 성립한다.

모르는 변수인 λ 와 \mathbf{v} 가 우변에 곱해져 있어서 이 식을 방정식으로 풀기는 어렵다. 하지만 $A\mathbf{v} = \lambda\mathbf{v}$ 를 $(A - \lambda I)\mathbf{v} = \mathbf{0}$ 로 다시 쓰면 \mathbf{v} 는 모르더라도 $A - \lambda I$ 가 가역이 아니라는 것은 알 수 있고 따라서 $\det(A - \lambda I) = 0$ 이 된다는 것을 알 수 있다. 즉 $\det(A - \lambda I) = 0$ 의 해를 찾으면 eigenvalue이고 이 n 차 방정식을 characteristic equation이라고 부르는데 실제 eigenvalue를 구하는 데 이용하기보다는 이론적인 연구에 활용이 많이 된다. 이렇게 eigenvalue λ 를 찾았으면 짹이 되는 eigenvector는 $\text{Null}(A - \lambda I)$ 의 기저를 찾으면 된다.

만약 $n \times n$ 행렬의 선형독립인 eigenvector를 n 개 찾으면 이 eigenvector를 기저로 행렬을 표현하면 대각행렬이 되는데 이를 diagonalization이라고 부른다. n 개 보다 적은 eigenvalue만 존재하는 경우 적절한 기저에 대해서 대각행렬은 아니지만 대각행렬과 유사한 Jordan form으로 표현되는데 11장에서 자세하게 설명한다.

9.1 Examples of Eigen Analysis

2×2 행렬 예를 통해 eigenpair의 다양한 경우를 살펴볼 수 있다.

1. Unstructured real matrix with two real eigenvalues : $A = \begin{bmatrix} 3 & 2 \\ -2 & -2 \end{bmatrix}$

$\det(A - \lambda I) = (3 - \lambda)(-2 - \lambda) + 4 = \lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1)$ 이므로 $\lambda = 2$ 와 -1 이다.

$A - \lambda I = \begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix}$ 와 $\begin{bmatrix} 4 & 2 \\ -2 & -1 \end{bmatrix}$ 이므로 eigenvector는 $(2, -1)^T$ 와 $(1, -2)^T$ 를 얻게 된다.

2. Real diagonal matrix : $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$

$\det(A - \lambda I) = (a - \lambda)(b - \lambda)$ 이므로 $\lambda = a$ 와 b 이다.

- $a \neq b$: $A - \lambda I = \begin{bmatrix} 0 & 0 \\ 0 & b - a \end{bmatrix}$ 와 $\begin{bmatrix} a - b & 0 \\ 0 & 0 \end{bmatrix}$ 이므로 eigenvector는 $(1, 0)^T$ 와 $(0, 1)^T$ 를 얻게 된다.

- $a = b$: $A - \lambda I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ 이므로 임의의 linearly independent한 두 벡터가 eigenvector가 된다.

3. Real upper triangular matrix : $A = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix}, c \neq 0$

$\det(A - \lambda I) = (a - \lambda)(b - \lambda)$ 이므로 $\lambda = a$ 와 b 이다.

- $a \neq b$: $A - \lambda I = \begin{bmatrix} 0 & c \\ 0 & b - a \end{bmatrix}$ 와 $\begin{bmatrix} a - b & c \\ 0 & 0 \end{bmatrix}$ 이므로 eigenvector는 $(1, 0)^T$ 와 $(1, \frac{b-a}{c})^T$ 를 얻게 된다.

- $a = b$: $A - \lambda I = \begin{bmatrix} 0 & c \\ 0 & 0 \end{bmatrix}$ 인데 이 행렬의 eigenvector는 $(1, 0)^T$ 의 상수배인 모든 벡터들이다. 즉 이 행렬은 중근인 eigenvalue를 갖지만 eigenvector는 한 개만 갖는다. 실수 eigenvalue를 갖지만 중근의 개수(algebraic multiplicity) 만큼의 eigenvector를 갖지 못하는 실수로 이루어진 비대칭 행렬의 예이다.

4. Real matrix with two complex eigenvalues : $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

$\det(A - \lambda I) = \lambda^2 + 1$ 이므로 $\lambda = i$ 와 $-i$ 이다.

$A - \lambda I = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix}$ 와 $\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}$ 이므로 eigenvector는 $(-i, 1)^T$ 와 $(i, 1)^T$ 를 얻게 된다.

5. I 나 0 이 아닌 non-trivial projection matrix P satisfying $P^2 = P$: $P\mathbf{v} = \lambda\mathbf{v}$ 인 (λ, \mathbf{v}) 를 생각하자.

이 식의 양변에 P 를 곱하면 $P^2\mathbf{v} = \lambda P\mathbf{v}$ 인데 좌변에는 $P^2 = P$ 를 적용하고 우변에는 $P\mathbf{v} = \lambda\mathbf{v}$ 를 적용하면 $P\mathbf{v} = \lambda^2\mathbf{v}$ 를 얻는다. 따라서 $(\lambda^2 - \lambda)\mathbf{v} = \mathbf{0}$ 를 얻게되고 $\lambda = 0$ 또는 1 이 된다. Projection

행렬은 $0 < \text{rank}(P) < n$ 이므로 $P\mathbf{v} = \mathbf{0}$ 인 null space의 벡터가 eigenvalue 0에 대한 eigenvector가 된다. 또한, P 가 정사영하는 부분공간의 벡터에 대해서는 $P\mathbf{v} = \mathbf{v}$ 가 되므로 이 벡터가 eigenvalue 1에 대한 eigenvector가 된다.

이상의 예를 통해 eigenpair는 거의 모든 조합이 가능하다는 것을 알 수 있다.

9.2 Some Properties of Eigenpairs

먼저 서로 다른 eigenvalue들과 짹인 eigenvector들의 성질에 대해서 살펴본다.

Fact 9.1 If eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ correspond to different eigenvalues $\lambda_1, \dots, \lambda_k$, then the eigenvectors are linearly independent.

Proof: $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 가 linearly dependent라고 가정하자. 그러면, 최소 한 개 이상 0이 아닌 scalar들 $c_1, \dots, c_k \in \mathbb{R}$ 이 존재하여 $c_1\mathbf{v}_1 + \dots + c_k\mathbf{v}_k = \mathbf{0}$ 이 된다. 필요하면 index를 바꾸어 $c_1 \neq 0$ 이라 가정하자. 이 등식의 양변에 행렬 A 를 곱하면 $A(c_1\mathbf{v}_1 + \dots + c_k\mathbf{v}_k) = c_1\lambda_1\mathbf{v}_1 + \dots + c_k\lambda_k\mathbf{v}_k = \mathbf{0}$ 를 얻는다. 이렇게 얻어진 등식에서 원래 등식의 양변에 λ_k 을 곱해서 빼면 $(c_1\lambda_1\mathbf{v}_1 + \dots + c_{k-1}\lambda_{k-1}\mathbf{v}_{k-1} + c_k\lambda_k\mathbf{v}_k) - (c_1\lambda_k\mathbf{v}_1 + \dots + c_{k-1}\lambda_k\mathbf{v}_{k-1} + c_k\lambda_k\mathbf{v}_k) = c_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + \dots + c_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{v}_{k-1} = \mathbf{0}$ 이다. 이때 $c_1(\lambda_1 - \lambda_k) \neq 0$ 이므로 $\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}$ 이 linearly dependent하게 되어 vector 개수가 한 개 적어진 것을 제외하면 처음과 상황이 동일해진다. 따라서 이 과정을 $(k-2)$ 번 더 반복하면 $c'_1 = c_1(\lambda_1 - \lambda_k) \cdots (\lambda_1 - \lambda_3)(\lambda_1 - \lambda_2)$ 에 대해서 $c'_1\mathbf{v}_1 = \mathbf{0}$ 을 얻는다. 그러면, $c'_1 = 0$, 즉 $c_1 = 0$ 을 얻게되어 모순이 된다. 따라서 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 는 linearly independent하다. ■

행렬의 determinant와 trace, eigenvalue간에는 다음과 같은 관계들이 있다.

Fact 9.2 Let $A = (a_{ij})$ be an $n \times n$ matrix and $\lambda_1, \dots, \lambda_n$ be its eigenvalues. Then

- $\det(A) = \lambda_1\lambda_2 \cdots \lambda_n = \prod_{i=1}^n \lambda_i$;
- $\text{trace}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \lambda_1 + \lambda_2 + \cdots + \lambda_n = \sum_{i=1}^n \lambda_i$.

Proof: (8.2)에 의해 $\det(A - \lambda I)$ 는 λ 에 관한 n 차 다항식이다. 특히 n 차 항의 계수는 $(-1)^n$ 이다. 이 n 차 다항식의 근이 eigenvalue이므로

$$\det(A - \lambda I) = (-1)^n(\lambda - \lambda_1) \cdots (\lambda - \lambda_n) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda) \quad (9.1)$$

이다.

- (9.1)에서 $\lambda = 0$ 으로 놓으면 $\det(A) = \lambda_1 \cdots \lambda_n$ 을 얻는다.
- (9.1)에서 λ^{n-1} 의 계수는 $(-1)^{n-1}(\lambda_1 + \cdots + \lambda_n)$ 이다. 또한, $A - \lambda I$ 에서 λ 를 포함하는 원소는 대각선 상의 $a_{ii} - \lambda$ 들이다. 첫 번째 행에 대한 cofactor expansion을 생각해보자. $a_{11} - \lambda$ 를 제외한 나머지 $a_{1i}, i \geq 2$ 에 대해 전개된 항에 포함된 cofactor들은 1행과 i 열에 포함된 $a_{11} - \lambda$ 와 $a_{ii} - \lambda$ 항이 제외된

submatrix에 대한 determinant를 포함하므로 λ 에 대한 최고 차수가 $n-2$ 차가 된다. 따라서 $n-1$ 차항은 $a_{11} - \lambda$ 에 대한 expansion에서만 생겨난다. 이 절차를 반복하면 결국 λ^{n-1} 항은 $(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$ 의 전개에서만 생겨나고 계수는 $(-1)^{n-1}(a_{11} + \cdots + a_{nn})$ 이다. 따라서 $\text{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$ 이 된다.

■

A 와 A^\top 의 행렬식이 같으므로 $\det(A - \lambda I) = \det(A^\top - \lambda I)$ 이고, 따라서 A 와 A^\top 의 eigenvalue는 일치한다.

Fact 9.3 *The eigenvalues of A and A^\top coincide.*

(λ, \mathbf{v}) 가 A 의 eigenpair라면 $A\mathbf{v} = \lambda\mathbf{v}$ 이므로 $A^2\mathbf{v} = A(A\mathbf{v}) = A(\lambda\mathbf{v}) = \lambda A\mathbf{v} = \lambda^2\mathbf{v}$ 이어서 (λ^2, \mathbf{v}) 가 A^2 의 eigenpair가 된다.

Fact 9.4 *If (λ, \mathbf{v}) is an eigenpair of A , then (λ^2, \mathbf{v}) is an eigenpair of A^2 .*

(λ, \mathbf{v}) 가 A 의 eigenpair이고 A 가 가역행렬이면 $\lambda \neq 0$ 이고 $A\mathbf{v} = \lambda\mathbf{v}$ 로부터 $A^{-1}\mathbf{v} = \lambda^{-1}\mathbf{v}$ 이어서 $(\lambda^{-1}, \mathbf{v})$ 가 A^{-1} 의 eigenpair가 된다.

Fact 9.5 *If A is invertible and (λ, \mathbf{v}) is an eigenpair of A , then $\lambda \neq 0$ and $(\lambda^{-1}, \mathbf{v})$ is an eigenpair of A^{-1} .*

그리고 특별한 구조의 행렬에 대한 결과로서 복소행렬이 Hermitian이면 모든 eigenvalue가 실수라는 성질이 있다. 특히 실수로 이루어진 대칭행렬도 이 경우에 해당한다. Appendix C에 Hermitian의 정의를 포함하여 복소수에 관해 간략하게 정리되어 있으니 필요한 독자는 복습하기 바란다.

Lemma 9.1 *Let A be a Hermitian matrix. Then the eigenvalues of A are real. Furthermore, if A is a real symmetric matrix, real eigenvectors exist.*

Proof: Let A be Hermitian and (λ, \mathbf{v}) an eigenpair of A . Recall that $|\mathbf{v}| \neq 0$. From $A\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v}^\text{H}A\mathbf{v} = \lambda\mathbf{v}^\text{H}\mathbf{v} = \lambda|\mathbf{v}|^2$. So,

$$(\lambda|\mathbf{v}|^2)^\text{H} = \lambda^\text{H}|\mathbf{v}|^2 = (\mathbf{v}^\text{H}A\mathbf{v})^\text{H} = \mathbf{v}^\text{H}A^\text{H}(\mathbf{v}^\text{H})^\text{H} = \mathbf{v}^\text{H}A\mathbf{v} = \lambda|\mathbf{v}|^2, \text{ that is, } \lambda^\text{H} = \lambda$$

which implies that λ is real by Fact C.1 in Appendix. In addition, let us assume that A is real and $(\lambda, \mathbf{v} + i\mathbf{w})$ be an eigenpair where λ is real, and \mathbf{v} and \mathbf{w} are real n -vectors. Then, the eigenpair satisfies

$$A\mathbf{v} + iA\mathbf{w} = A(\mathbf{v} + i\mathbf{w}) = \lambda(\mathbf{v} + i\mathbf{w}) = \lambda\mathbf{v} + i\lambda\mathbf{w}$$

which implies (λ, \mathbf{v}) and (λ, \mathbf{w}) are also eigenpairs. ■

추가적으로 행렬이 Hermitian이면 서로 다른 eigenvalue들의 짝인 eigenvector들은 orthogonal하다.

Fact 9.6 *Let A be Hermitian. If eigenvectors \mathbf{v}_1 and \mathbf{v}_2 correspond to different eigenvalues λ_1 and λ_2 , those eigenvectors are orthogonal.*

Proof: By Lemma 9.1, the eigenvalues are real. From $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$, $\mathbf{v}_2^H A \mathbf{v}_1 = \lambda_1 \mathbf{v}_2^H \mathbf{v}_1$ and $\mathbf{v}_1^H A \mathbf{v}_2 = \lambda_2 \mathbf{v}_1^H \mathbf{v}_2$. Since $\lambda_1 \neq \lambda_2$ and

$$0 = (\mathbf{v}_1^H A \mathbf{v}_2)^H - \mathbf{v}_2^H A \mathbf{v}_1 = (\lambda_2 \mathbf{v}_1^H \mathbf{v}_2)^H - \lambda_1 \mathbf{v}_2^H \mathbf{v}_1 = (\lambda_2 - \lambda_1) \mathbf{v}_2^H \mathbf{v}_1,$$

we conclude $\mathbf{v}_2^H \mathbf{v}_1 = 0$. ■

행렬이 orthogonal한 경우, eigenvalue들의 절댓값이 1이며 서로 다른 eigenvalue들의 짝인 eigenvector들은 orthogonal하다.

Fact 9.7 Let Q be a real orthogonal matrix and (λ, \mathbf{v}) an eigenpair of Q . λ and \mathbf{v} may be complex-valued. Then

1. $|\lambda| = 1$.
2. $(\lambda^{-1}, \mathbf{v})$ is an eigenpair of Q^\top .
3. If eigenvectors \mathbf{v}_1 and \mathbf{v}_2 correspond to different eigenvalues λ_1 and λ_2 , then those eigenvectors are orthogonal.

Proof: Note that $Q^\top Q = I$ and $Q\mathbf{v} = \lambda\mathbf{v}$.

1. $|\mathbf{v}|^2 = \mathbf{v}^H \mathbf{v} = \mathbf{v}^H (Q^\top Q) \mathbf{v} = (Q\mathbf{v})^H (Q\mathbf{v}) = (\lambda\mathbf{v})^H (\lambda\mathbf{v}) = |\lambda|^2 |\mathbf{v}|^2$ and $|\lambda|^2 = 1$.
2. $\mathbf{v} = (Q^\top Q)\mathbf{v} = Q^\top(\lambda\mathbf{v}) = \lambda Q^\top \mathbf{v}$. So, $Q^\top \mathbf{v} = \frac{1}{\lambda} \mathbf{v}$.
3. Let $Q\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $Q\mathbf{v}_2 = \lambda_2\mathbf{v}_2$. Then,

$$\mathbf{v}_1^H \mathbf{v}_2 = \mathbf{v}_1^H (Q^\top Q) \mathbf{v}_2 = (Q\mathbf{v}_1)^H (Q\mathbf{v}_2) = \lambda_1^H \lambda_2 \mathbf{v}_1^H \mathbf{v}_2$$

and we have to have $\lambda_1^H \lambda_2 = \overleftrightarrow{\lambda}_1 \lambda_2 = 1$ or $\mathbf{v}_1^H \mathbf{v}_2 = 0$. However, $1 = |\lambda_1|^2 = \lambda_1^H \lambda_1 = \overleftrightarrow{\lambda}_1 \lambda_1$ implies $\overleftrightarrow{\lambda}_1 \lambda_2 = \frac{\lambda_2}{\lambda_1} \neq 1$ since $\lambda_1 \neq \lambda_2$. Therefore, $\mathbf{v}_1^H \mathbf{v}_2 = 0$. ■

Fact 9.8 If A be a real matrix and (λ, \mathbf{v}) an eigenpair of A , $(\overleftrightarrow{\lambda}, \overleftrightarrow{\mathbf{v}})$ is also an eigenpair of A .

Proof: $A\mathbf{v} = \lambda\mathbf{v}$ 에 conjugation을 취하면 $\overleftrightarrow{A}\mathbf{v} = \overleftrightarrow{\lambda}\mathbf{v}$ 인데 복소수 연산 규칙들에 의해 $\overleftrightarrow{A}\overleftrightarrow{\mathbf{v}} = \overleftrightarrow{\lambda}\overleftrightarrow{\mathbf{v}}$ 이다. 그런데 A 가 실수로 구성되었으므로 $\overleftrightarrow{A} = A$ 이어서 $A\overleftrightarrow{\mathbf{v}} = \overleftrightarrow{\lambda}\overleftrightarrow{\mathbf{v}}$ 이고 $(\overleftrightarrow{\lambda}, \overleftrightarrow{\mathbf{v}})$ 도 eigenpair가 된다. ■

Section 7.5의 결과를 이용하면 임의의 n 차원 ellipsoid의 부피를 단위구의 부피와 determinant로 표현할 수 있다.

Fact 9.9 For a positive definite matrix A , an n -dimensional ellipsoid is given as $\mathcal{E} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^\top A^{-1} \mathbf{x} \leq 1\}$. Then,

$$\text{vol}(\mathcal{E}) = \sqrt{\det(A)} \text{vol}(B_n). \quad (9.2)$$

Proof: 이미 (7.5)를 얻었다. Fact 9.2에 의해 $\prod_{i=1}^n \sqrt{\lambda_i(A)} = \sqrt{\det(A)}$ 므로 $\text{vol}(\mathcal{E}) = \sqrt{\det(A)} \times \text{vol}(B_n)$ 가 성립한다. ■

9.3 Similarity and Change of Basis

두 개의 행렬이 (operator로서) 비슷하다(similar)는 것을 적절한 기저에 대해서 두 행렬이 동일한 선형변환을 표현하는 것으로 정의한다.

9.3.1 Change of Basis

Vector space \mathbb{V} 의 기저 \mathcal{B}_1 이 있는데 \mathbb{V} 의 새로운 기저 \mathcal{B}_2 에 대한 표현에 관심이 생겼다고 하자. 기저 \mathcal{B}_1 에 의해 기술되던 여러 성질들을 기저 \mathcal{B}_2 기준으로 설명하기 위해서는 \mathcal{B}_1 과 \mathcal{B}_2 간의 관계를 파악해야 한다. 이 과정에 Section 3.8.1의 방법을 적용할 수 있다. $\mathcal{B}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $\mathcal{B}_2 = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ 이라 하자. \mathcal{B}_2 가 기저이므로 모든 $j = 1, \dots, n$ 에 대해 \mathbf{v}_j 를 \mathcal{B}_2 의 basic vector들의 linear combination으로

$$\mathbf{v}_j = \sum_{i=1}^n p_{ij} \mathbf{w}_i \quad (9.3)$$

로 표현할 수 있다. \mathcal{B}_1 -좌표가 $\mathbf{x} = (x_1, \dots, x_n)^\top$ 인 vector

$$\mathbf{v} = \sum_{j=1}^n x_j \mathbf{v}_j$$

의 \mathcal{B}_2 -좌표 \mathbf{y} 를 계산해보자. (9.3)를 대입하면

$$\mathbf{v} = \sum_{j=1}^n x_j \left(\sum_{i=1}^n p_{ij} \mathbf{w}_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n p_{ij} x_j \right) \mathbf{w}_i \quad (9.4)$$

가 된다. Linear combination 계수를 정사각행렬 $P = (p_{ij})$ 로 표시하면 \mathcal{B}_1 -좌표 \mathbf{x} 와 \mathcal{B}_2 -좌표 \mathbf{y} 간에는

$$y_i = \sum_{j=1}^n p_{ij} x_j \quad \text{or} \quad \mathbf{y} = P\mathbf{x} \quad (9.5)$$

의 관계가 성립한다. p_{ij} 가 (9.3)에서 두 기저 간의 관계를 표현하므로 (9.5) 관계가 벡터에 관계없이 고정된다. $P = (p_{ij})$ 에 대해 성립함에 유의하자.

거꾸로 \mathcal{B}_1 과 \mathcal{B}_2 의 역할을 바꾸고 \mathcal{B}_1 도 기저임을 이용하여 위의 과정을 반복하면 정사각행렬 \bar{P} 가 존재하여 \mathcal{B}_1 -좌표 \mathbf{x} 와 \mathcal{B}_2 -좌표 \mathbf{y} 간에 $\mathbf{x} = \bar{P}\mathbf{y}$ 가 성립한다. 이를 (9.5)와 연결하면

$$\mathbf{x} = \bar{P}P\mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^n \quad \text{and} \quad \mathbf{y} = P\bar{P}\mathbf{y} \quad \text{for all } \mathbf{y} \in \mathbb{R}^n$$

이어서 $\bar{P}P = P\bar{P} = I_n$ 이 된다. 즉 change of basis를 나타내는 행렬은 invertible하고 두 기저에 대한 좌표 간의 변환을 나타내는 두 행렬은 서로 역행렬 관계이다.

역으로 임의의 기저 $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 와 임의의 invertible matrix $P = (p_{ij})$ 가 주어졌을 때, (9.3)를 통해 얻어지는 $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ 가 기저가 되는지, 즉 linearly independent한지 살펴보자. \mathbf{w}_i 는 $\mathbf{y} = \mathbf{e}_i$ 를 (9.5)에 대입하여 구한 \mathcal{B} -좌표 $P^{-1}\mathbf{e}_i$ 에 대응하는 vector이다. $\lambda = (\lambda_1, \dots, \lambda_n)^\top \in \mathbb{R}^n$ 에 대해서 $\lambda_1\mathbf{w}_1 + \dots + \lambda_n\mathbf{w}_n = \mathbf{0}$

라 하자. 그러면, \mathcal{B} -좌표 $\lambda_1 P^{-1}\mathbf{e}_1 + \dots + \lambda_n P^{-1}\mathbf{e}_n = P^{-1}(\lambda_1\mathbf{e}_1 + \dots + \lambda_n\mathbf{e}_n) = P^{-1}\boldsymbol{\lambda}$ 에 대응하는 vector가 $\mathbf{0}$ 이다. \mathcal{B} 가 기저이므로 linearly independent여서 이러한 \mathcal{B} -좌표는 $\mathbf{0}$ 뿐이므로 $P^{-1}\boldsymbol{\lambda} = \mathbf{0}$ 로부터 $\boldsymbol{\lambda} = P\mathbf{0} = \mathbf{0}$ 이고 $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ 가 기저가 된다.

9.3.2 Change of Orthonormal Bases

\mathcal{B}_1 과 \mathcal{B}_2 가 모두 orthonormal한 경우 두 기저 간의 change of basis 행렬 P 가 어떤 추가적인 성질을 가지는지 살펴보자. (9.3)와 orthonormality를 합치면 $\mathbf{w}_i^\top \mathbf{w}_\ell \circ | i = \ell$ 인 경우에만 1이고 나머지 경우에는 0이므로

$$\begin{aligned} \mathbf{v}_j^\top \mathbf{v}_k &= \left(\sum_{i=1}^n p_{ij} \mathbf{w}_i \right)^\top \left(\sum_{\ell=1}^n p_{\ell k} \mathbf{w}_\ell \right) = \left(\sum_{i=1}^n p_{ij} \mathbf{w}_i^\top \right) \left(\sum_{\ell=1}^n p_{\ell k} \mathbf{w}_\ell \right) \\ &= \sum_{i=1}^n \sum_{\ell=1}^n p_{ij} p_{\ell k} \mathbf{w}_i^\top \mathbf{w}_\ell \\ &= \sum_{i=1}^n p_{ij} p_{ik} \end{aligned}$$

인데 $\mathbf{v}_j^\top \mathbf{v}_k$ 역시 $j = k$ 인 경우에만 1이고 나머지 경우에는 0이다. 그런데 $\sum_{i=1}^n p_{ij} p_{ik}$ 는 $P^\top P$ 의 j, k 원소이므로

$$P^\top P = I_n$$

이다. 즉 P 는 orthogonal matrix이다. 결론적으로 orthonormal basis간의 change of basis matrix는 orthogonal하다.

9.3.3 Similarity

Vector space \mathbb{V} 상의 선형변환 $T : \mathbb{V} \rightarrow \mathbb{V}$ 과 \mathbb{V} 의 기저 $\mathcal{B}_1, \mathcal{B}_2$ 를 생각하자. A 를 \mathcal{B}_1 에 대한 T 의 행렬표현이라 하자. 즉 \mathcal{B}_1 -좌표가 \mathbf{x} 인 vector \mathbf{v} 가 선형변환된 $T(\mathbf{v})$ 의 \mathcal{B}_1 -좌표는 $A\mathbf{x}$ 이다. 이 상황에서 기저 \mathcal{B}_2 에 대한 T 의 행렬표현을 찾아보자. T 에 대해서 알려진 것은 \mathcal{B}_1 에 대한 행렬표현 A 뿐이다.

(9.5)에 의하면 invertible인 적절한 change of basis 행렬 P 에 대해 \mathcal{B}_1 -좌표가 \mathbf{x} 인 벡터의 \mathcal{B}_2 -좌표가 $P\mathbf{x}$ 이다. 반대로 \mathcal{B}_2 -좌표가 \mathbf{y} 인 벡터 \mathcal{B}_1 -좌표는 $P^{-1}\mathbf{y}$ 이다. \mathcal{B}_2 -좌표가 \mathbf{y} 인 vector \mathbf{v} 가 T 를 통해 변환된 $T(\mathbf{v})$ 의 \mathcal{B}_1 -좌표는 $AP^{-1}\mathbf{y}$ 이다. 이 \mathcal{B}_1 -좌표를 \mathcal{B}_2 -좌표로 바꾸면 $PAP^{-1}\mathbf{y}$ 가 된다.

$$\begin{array}{ccc} \mathcal{B}_1 : & P^{-1}\mathbf{y} & \xrightarrow{A} AP^{-1}\mathbf{y} \\ & \uparrow P^{-1} & \downarrow P \\ \mathcal{B}_2 : & \mathbf{y} & \xrightarrow{B} B\mathbf{y} = PAP^{-1}\mathbf{y} \end{array}$$

따라서 기저 \mathcal{B}_2 하에서 T 의 표현을 B 로 표시하면 $B\mathbf{y} = PAP^{-1}\mathbf{y}$ 가 모든 $\mathbf{y} \in \mathbb{R}^n$ 에 대해 성립하므로

$$B = PAP^{-1} \quad \text{and} \quad A = P^{-1}BP$$

의 관계를 갖는다. 즉 A 와 PAP^{-1} 는 동일한 선형변환 T 의 두 기저에 대한 표현들이므로 \mathbb{V} 의 vector가 변환되는 측면에서는 동일하다고 생각할 수 있다. 이런 이유로 두 행렬을 similar하다고 부른다.

Definition 9.1 Matrices A and B are **similar** if there is an invertible matrix P such that $A = P^{-1}BP$.

Fact 9.10 If A and B are similar, they share the same eigenvalues.

Proof: $A = P^{-1}BP$ 이고 (λ, \mathbf{v}) 가 A 의 eigenpair라 하자. 즉 $A\mathbf{v} = \lambda\mathbf{v}$ 이다. 그러면,

$$B(P\mathbf{v}) = (BP)\mathbf{v} = (PA)\mathbf{v} = P(A\mathbf{v}) = P(\lambda\mathbf{v}) = \lambda(P\mathbf{v})$$

이므로 $(\lambda, P\mathbf{v})$ 가 B 의 eigenpair가 된다. ■

9.4 Diagonalization

가장 간단한 표현을 갖는 행렬은 대각행렬(diagonal matrix)이다. 따라서 어떤 행렬이 기저변환을 통해서 대각행렬로 나타낼 수 있다면 쉽게 해당 선형변환을 이해할 수 있다. 이렇게 어떤 행렬이 기저변환에 의해서 대각행렬로 표현 가능하면 대각화가능(diagonalizable)이라고 말한다. 즉

Definition 9.2 A is diagonalizable if A is similar to a diagonal matrix.

어떤 행렬이 대각화 가능이기 위한 가장 간단한 필요충분 조건을 살펴보자.

Fact 9.11 $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

Proof: 대각행렬 Λ 를

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

으로 놓자.

- **if part :** $(\lambda_1, \mathbf{v}_1), \dots, (\lambda_n, \mathbf{v}_n)$ 이 n 개의 eigenpair인데 \mathbf{v}_i 들이 선형독립이라 하자. 물론, eigenvalue들은 같을 수도 있다. \mathbf{v}_i 들을 열로 갖는 행렬 $P = [\mathbf{v}_1 | \cdots | \mathbf{v}_n]$ 는 invertible이고 $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ 를 고려하면 $AP = P\Lambda$ 이므로 $P^{-1}AP = \Lambda$ 가 되어 A 가 대각화가능이다.
- **only if part :** A 가 diagonalizable이므로 어떤 가역행렬 P 와 대각행렬 Λ 에 대하여 $A = P^{-1}\Lambda P$ 이다. 그러면 $AP^{-1} = P^{-1}\Lambda$ 이므로 P^{-1} 의 각 열이 eigenvector인데 P^{-1} 가 가역행렬이므로 각 열은 선형독립이다.

■

행렬의 eigenvector와 관련하여 언제 두 행렬의 eigenvector들이 완전히 일치할까라는 질문이 있을수 있다. 일반적인 두 $n \times n$ 행렬에 대해서는 알려진 바가 없지만 두 행렬이 대각화 가능할 때에는 n 개의 eigenvector를 공유하기 위한 필요충분 조건이 두 행렬의 곱이 교환되는 것이다.¹

Theorem 9.1 Let A and B be $n \times n$ diagonalizable matrices. Then, A and B share the same eigenvector matrix if and only if $AB = BA$.

Proof: Let A and B share the same eigenvectors and denote the square matrix of common eigenvectors as S . Then $A = S^{-1}\Lambda_1 S$ and $B = S^{-1}\Lambda_2 S$. So, since diagonal matrices commute,

$$AB = S^{-1}\Lambda_1 S S^{-1}\Lambda_2 S = S^{-1}\Lambda_1 \Lambda_2 S = S^{-1}\Lambda_2 \Lambda_1 S = S^{-1}\Lambda_2 S S^{-1}\Lambda_1 S = BA.$$

Conversely, assume that diagonalizable matrices A and B commute. Let S be an invertible matrix of eigenvectors of A such that $A = S^{-1}\Lambda S$. Set $\bar{B} = S^{-1}BS$. The commutativity of A and B implies the commutativity of Λ and \bar{B} . Assume $\Lambda = \text{diag}(\lambda_1 I_{n_1}, \dots, \lambda_k I_{n_k})$ where each I_{n_i} is an identity matrix of size n_i such that $n = n_1 + \dots + n_k$. Of course, λ_i 's are different from each other. Denote the matrix \bar{B} by a partitioned matrix (\bar{B}_{ij}) where \bar{B}_{ij} is an $n_i \times n_j$ matrix. Then, since

$$\Lambda \bar{B} = \begin{bmatrix} \lambda_1 \bar{B}_{11} & \lambda_1 \bar{B}_{12} & \cdots & \lambda_1 \bar{B}_{1k} \\ \lambda_2 \bar{B}_{21} & \lambda_2 \bar{B}_{22} & \cdots & \lambda_2 \bar{B}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_k \bar{B}_{k1} & \lambda_k \bar{B}_{k2} & \cdots & \lambda_k \bar{B}_{kk} \end{bmatrix} \quad \text{and} \quad \bar{B} \Lambda = \begin{bmatrix} \lambda_1 \bar{B}_{11} & \lambda_2 \bar{B}_{12} & \cdots & \lambda_k \bar{B}_{1k} \\ \lambda_1 \bar{B}_{21} & \lambda_2 \bar{B}_{22} & \cdots & \lambda_k \bar{B}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 \bar{B}_{k1} & \lambda_2 \bar{B}_{k2} & \cdots & \lambda_k \bar{B}_{kk} \end{bmatrix},$$

$\Lambda \bar{B} = \bar{B} \Lambda$ holds if and only if $\bar{B}_{ij} = \mathbf{0}$ if $i \neq j$. That is, $\bar{B} = \text{diag}(\bar{B}_{11}, \dots, \bar{B}_{kk})$. From the diagonalizability of B , \bar{B} is also diagonalizable, and each \bar{B}_{ii} is diagonalizable in turn. Let some \bar{S}_i satisfy $\bar{B}_{ii} = \bar{S}_i^{-1}\bar{\Lambda}_i \bar{S}_i$ where $\bar{\Lambda}_i$ is a diagonal matrix. Denote $\bar{S} = \text{diag}(\bar{S}_1, \dots, \bar{S}_k)$ and $\bar{\Lambda} = \text{diag}(\bar{\Lambda}_1, \dots, \bar{\Lambda}_k)$ such that $\bar{B} = \bar{S}^{-1}\bar{\Lambda}\bar{S}$. Note that $\Lambda = \bar{S}^{-1}\bar{\Lambda}\bar{S}$. Hence,

$$A = S^{-1}\Lambda S = S^{-1}\bar{S}^{-1}\bar{\Lambda}\bar{S}S \quad \text{and} \quad B = S^{-1}\bar{B}S = S^{-1}\bar{S}^{-1}\bar{\Lambda}\bar{S}S$$

holds and A and B share the common matrix $\bar{S}S$ of eigenvectors. ■

Theorem 5.2에 의하면 대칭행렬은 diagonalizable이다. 그리고 A 와 B 가 대칭행렬이면 $(AB)^\top = B^\top A^\top = BA$ 므로 $AB = BA$ 조건이 AB 가 대칭이라는 조건과 일치한다. 따라서 다음 결과를 얻는다.

Corollary 9.1 Let A and B be $n \times n$ symmetric matrices. Then, A and B share the same eigenvector matrix if and only if AB is symmetric.

추가적으로 diagonalizable matrix의 eigenvector들이 orthogonal한 경우에는 normalize하여 orthonormal한 eigenvector들을 찾을 수 있다. 이런 경우에 orthogonally diagonalizable이라고 부른다.

¹ $AB = BA$ 와 같이 두 행렬의 곱이 교환되는 경우 “matrices A and B commute”라고 말한다.

Definition 9.3 A square matrix A is **orthogonally diagonalizable** if there exists an orthogonal matrix Q such that $Q^{-1}AQ = Q^\top AQ$ is diagonal

$n \times n$ 행렬 A 가 orthogonally diagonalizable이면 $Q^{-1} = Q^\top$ 인 orthogonal matrix Q 가 존재하여 $Q^{-1}AQ = \Lambda$ 가 대각행렬이 된다. Orthonormal한 벡터들 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 에 대해서 $Q = [\mathbf{v}_1 | \mathbf{v}_2 | \cdots | \mathbf{v}_n]$ 이고 $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ 이고 λ_i 들은 서로 같을 수도 있다. Corollary 3.1의 (3.6)을 상기하면 orthogonally diagonalizable한 행렬 A 는

$$A = Q\Lambda Q^\top = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top \quad (9.6)$$

와 같이 orthonormal한 벡터들의 rank-one 합으로 표현되고 orthogonally diagonalizable한 행렬은 대칭이라는 것도 알 수 있다. 반면 Theorem 5.2의 spectral decomposition은 대칭행렬이 orthogonally diagonalizable이라는 것을 알려준다. 이 두 결과를 합치면 이러한 성질이 대칭 행렬의 고유한 특성이라는 다음 결과를 얻는다.

Theorem 9.2 (The Fundamental Theorem of Symmetric Matrices) A real matrix A is orthogonally diagonalizable if and only if A is symmetric.

Invertibility and Diagonalizability in the Lens of Eigenpairs

모든 $n \times n$ 행렬은 중근을 포함하여 n 개의 eigenvalue를 복소수 범위에서 갖는데,

- 만약 모든 eigenvalue가 0이 아니라면 행렬은 가역(invertible)이다;
- 만약 n 개의 선형독립인 eigenvector가 존재하면 행렬은 대각화가능(diagonalizable)하다. 중근인 eigenvalue에 대해서는 중근 개수만큼의 선형독립인 eigenvector가 존재하지 않을 수도 있다. Eigenvalue λ 가 characteristic equation $\det(A - xI)$ 의 해로서 k 개의 중근을 갖아서 $(x - \lambda)^k$ 인수를 갖는 경우 algebraic multiplicity가 k 라고 한다. 우리가 증명하지는 않지만 이 경우에 알려진 사실은 λ 에 대응되는 linearly independent한 eigenvector의 개수를 geometric multiplicity라고 부르는데 이 개수가 k 개 이하라는 것이다. 각 eigenvalue들의 algebraic multiplicity의 합은 행렬의 열 또는 행의 개수와 일치한다. 행렬의 diagonalizability는 geometric multiplicity의 합이 열의 개수와 같은지 여부에 달려있다.

9.4.1 An Example of Non-diagonalizable Matrix

행렬 $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ 로 놓자. A 의 characteristic polynomial이 $p_A(x) = (x - 1)^2$ 므로 A 는 한 개의 eigenvalue $\lambda = 1$ 을 갖고 algebraic multiplicity가 2이다. $A - \lambda I = A - I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ 이고 $2 - \text{rank } A = 1$ 이므로 $A - I$ 의 null space의 차원은 1이다. 따라서 A 의 eigenvector는 한 개를 찾을 수 있고 eigenvalue 1의 geometric

multiplicity는 1이다. 앞에서 언급한 바와 같이 geometric multiplicity가 algebraic multiplicity보다 작아서 대각화 가능하지 않다는 것을 알 수 있다. 그런데 multiplicity간의 관계는 아직 증명하지 않았기 때문에 행렬 A 가 diagonalizable하지 않다는 것을 직접 증명해보자.

$P^{-1}AP$ 가 대각행렬이 되는 가역행렬 $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 가 존재한다고 가정하자. $P^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ 이므로 $P^{-1}AP = \frac{1}{ad-bc} \begin{bmatrix} ad+cd-bc & d^2 \\ -c^2 & ad-cd-bc \end{bmatrix}$ 이다. 마지막 행렬이 대각행렬이기 위해서는 $c = d = 0$ 인데 그러면 P 가 singular해지므로 행렬 A 는 diagonalizable하지 않다. 행렬 A 처럼 대각화가 불가능한 모든 $n \times n$ 행렬을 대각행렬과 유사한 Jordan form으로 표현²할 수 있다는 정리를 Section 11.4에서 증명한다.

9.5 Alternative Proof of the Spectral Decomposition Theorem

일반적으로 임의의 $n \times n$ 행렬 A 는 (실수 원소만으로 구성된 경우에도) eigenvalue가 복소수일 수 있다. 하지만 놀랍게도 Lemma 9.1에서 살펴본 것처럼 Hermitian matrix는 모든 eigenvalue가 실수이다. 그런데 Hermitian matrix의 특수한 경우인 대칭인 실수 행렬 (real symmetric matrix)은 eigenvalue가 실수일 뿐만 아니라 실벡터인 eigenvector를 n 개 갖는다. 그리고 이 n 개의 eigenvector들은 서로 orthogonal하다. 이러한 결과는 이미 Theorem 5.2에 내포되어 있었다. 정리를 다시 한번 기술하고 SVD를 이용하지 않는 다른 방식의 증명을 기술한다.

Theorem 5.2 (Real Spectral Decomposition) *Let A be a real symmetric matrix. Then, A is orthogonally diagonalizable, and indeed*

$$A = V\Lambda V^\top = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top,$$

where V is an orthogonal matrix with orthonormal columns $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, $|\mathbf{v}_i| = 1$ and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$.

Proof: $\mathbf{e}_1 = (1, 0, \dots, 0)^\top$ 로 표시하자. 대칭행렬이므로 적어도 한 개 이상의 실수로 이루어진 eigenpair가 존재하고 이를 $|\mathbf{v}| = 1$ 인 벡터에 대해서 (λ, \mathbf{v}) 로 표시하자. 이 eigenpair에 대해서 $Q\mathbf{v} = \mathbf{e}_1$ 이고 $Q^{-1} = Q^\top$ 를 만족하는 real orthogonal matrix Q 가 존재한다.³ $A\mathbf{v} = \lambda\mathbf{v}$ 이므로

$$Q A Q^\top \mathbf{e}_1 = Q A \mathbf{v} = \lambda Q \mathbf{v} = \lambda \mathbf{e}_1$$

²Section 11.4에서 알 수 있듯이 이 예에서는 행렬 A 자체가 Jordan form으로 표현되어 있다.

³Gram-Schmidt process를 이용하여 \mathbf{v} 에서 시작하여 orthonormal한 벡터들 $\mathbf{v}_2, \dots, \mathbf{v}_n$ 을 차례로 만들 수 있다. 이 벡터들을 행으로 하는 행렬 $Q^\top = [\mathbf{v} | \mathbf{v}_2 | \dots | \mathbf{v}_n]$ 를 만들면 $Q\mathbf{v} = \mathbf{e}_1$ 을 만족하고 $Q Q^\top = I$ 가 된다.

가 성립하여 QAQ^\top 의 첫 열이 $\begin{bmatrix} \lambda \\ \mathbf{0} \end{bmatrix}$ 인데 QAQ^\top 가 대칭행렬이므로

$$QAQ^\top = \begin{bmatrix} \lambda & \mathbf{0}^\top \\ \mathbf{0} & A_{n-1} \end{bmatrix}$$

로 표현할 수 있다. 여기서, A_{n-1} 은 적당한 $(n-1) \times (n-1)$ 대칭행렬이다. 수학적 귀납법을 적용하기 위해 $(n-1) \times (n-1)$ 대칭행렬 A_{n-1} 에 대해서 $(n-1) \times (n-1)$ real orthogonal matrix Q_{n-1} 가 존재하여 $Q_{n-1}A_{n-1}Q_{n-1}^\top = \Lambda_{n-1}$ 이 대각행렬이 된다고 가정하자. $Q_n = \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & Q_{n-1} \end{bmatrix}$ 로 놓으면 $n \times n$ real orthogonal matrix로 $Q_n \begin{bmatrix} \lambda & \mathbf{0}^\top \\ \mathbf{0} & A_{n-1} \end{bmatrix} Q_n^\top = \begin{bmatrix} \lambda & \mathbf{0}^\top \\ \mathbf{0} & \Lambda_{n-1} \end{bmatrix}$ 이므로

$$Q_n Q A Q^\top Q_n^\top = (Q_n Q) A (Q_n Q)^\top = \begin{bmatrix} \lambda & \mathbf{0}^\top \\ \mathbf{0} & \Lambda_{n-1} \end{bmatrix} = \Lambda$$

로 대각행렬이 된다. 특히 두 orthogonal matrix의 곱은 orthogonal matrix이므로 $Q_n Q$ 는 orthogonal하다. $V = (Q_n Q)^\top$ 으로 놓으면 $A = V \Lambda V^\top$ 이 증명된다. 그리고 Corollary 3.1의 (3.6)에 의해

$$V \Lambda V^\top = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$$

로 표현된다. ■

9.6 How to obtain Eigenvalues and Eigenvectors

$n \times n$ 정사각행렬 A 의 eigenvalue λ 와 eigenvector vector \mathbf{v} 를 찾는 것은

$$A\mathbf{v} = \lambda\mathbf{v}$$

를 만족하는 $\lambda \in \mathbb{C}$ 와 $\mathbf{v} \in \mathbb{C}^n$ 을 찾는 eigenvalue problem과 동일하다. A 가 실수로 구성되어 있더라도 eigenvalue와 eigenvector는 복소수 범위에서 존재한다. 컴퓨터를 활용할 수 없고 손계산으로 eigenvalue와 eigenvector를 찾아야 하는 경우에는 확실한 방법은 없고 다음과 같은 시도를 해볼 수 있다.

- 먼저 행렬 A 를 잘 관찰하여 eigenvalue와 eigenvector를 찾는 hint를 발견하려고 시도한다.
- 특별한 hint를 못 찾은 경우, n 이 작고 행렬의 원소가 정수로 주어져서 $\det(A - \lambda I)$ 의 전개가 가능한 경우에는 characteristic equation을 풀어본다. 방정식의 해인 eigenvalue λ 를 구했다면 $\text{Null}(A - \lambda I)$ 에 속하는 $n - \text{rank}(A - \lambda I)$ 개의 linearly independent인 vector를 찾는다. 이 과정은 Gaussian elimination 으로 수행할 수 있다.

추가적으로 A 가 symmetric하다면 eigenvalue와 eigenvector를 실수 범위에서 찾을 수 있다. 특히 real spectral decomposition (Theorem 5.2)를 활용하려는 시도를 해본다. 특히 식 (5.12)과 유사하게 행렬이 (구체적인 숫자든 문자든) rank-one 행렬의 합으로 주어진 경우에는 이 구조를 잘 활용하여 eigenvalue와 eigenvector를 찾아 본다. 하지만 어느 경우든 eigenvalue problem을 푸는 확실한 절차는 없다.

컴퓨터를 활용할 수 있는 경우에는 다양한 수치적 방법이 존재한다. Iterative한 power method, QR-decomposition을 활용하는 QR-method를 포함하여 여러 방법이 있으며 관심있는 독자는 관련 서적을 읽어 보기 바란다.

9.7 Application to Data Science: Power Iteration and Google PageRank

어떤 시스템의 상태를 나타내는 벡터 $\mathbf{x}_t = (x_t^1, x_t^2, \dots, x_t^n)^\top \in \mathbb{R}^n$ 의 dynamics가 이산 시간에 대해

$$\mathbf{x}_{t+1} = A\mathbf{x}_t, \quad t = 0, 1, 2, \dots$$

또는 연속시간에 대해

$$\frac{d}{dt}\mathbf{x}_t = A\mathbf{x}_t, \quad t \geq 0$$

으로 주어진 경우에, 임의의 시간 t 에서의 상태벡터는 이산 시간인 경우

$$\mathbf{x}_t = A^t \mathbf{x}_0, \quad t = 1, 2, \dots$$

또는 연속시간인 경우

$$\mathbf{x}_t = e^{tA} \mathbf{x}_0 = \left(I + \frac{t}{1!} A + \frac{t^2}{2!} A^2 + \frac{t^3}{3!} A^3 + \dots \right) \mathbf{x}_0, \quad t \geq 0$$

으로 표현된다.⁴ 어느 경우에나 $A^k, k = 1, 2, \dots$ 의 계산이 필요하다. 만약 A 가 diagonalizable하여 $A = V\Lambda V^{-1}$ 이라면

$$A^k = V\Lambda^k V^{-1}$$

로 간단히 계산되고, matrix exponential은

$$e^{tA} = V \left(I + \frac{t}{1!} \Lambda + \frac{t^2}{2!} \Lambda^2 + \frac{t^3}{3!} \Lambda^3 + \dots \right) V^{-1} = V \text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t}) V^{-1}$$

로 계산된다. A^k 를 계산하는 것은 다양한 모형에서 활용된다. 그런데 행렬의 대각화는 시간에 따른 변화를 찾는 경우 말고도 A 의 특별한 eigenvector를 찾는 방법으로도 활용 가능하다. 특히 행렬의 거듭제곱의 특성을 대각화를 가정하고 분석하지만 대각화된 행렬의 분해된 성분을 모르더라도 원하는 계산을 할 수 있다는 것이 중요한 점이다.

⁴연속시간의 해를 표현할 때는 행렬의 거듭제곱이 무한개 더해진 행렬이 나타나는데 이를 matrix exponential이라 부르며 정사각 행렬 B 에 대해서

$$e^B = I + \frac{1}{1!} B + \frac{1}{2!} B^2 + \frac{1}{3!} B^3 + \dots = \sum_{k=0}^{\infty} \frac{1}{k!} B^k$$

로 정의된다.

행렬 A 가 diagonalizable하더라도 크기가 클 경우에는 eigenpair를 구하는 과정이 어렵거나 불가능할 수 있다. 이런 경우에 절댓값이 가장 큰 eigenvalue와 그 짹인 eigenvector를 근사적으로 구하는 방법을 소개한다. 이 방법은 Google Pagerank의 핵심 아이디어이기도 하다. Google Pagerank는 절댓값이 가장 큰 eigenvalue의 짹인 eigenvector의 성분의 크기를 이용하여 webpage를 ranking한다.

Diagonalizable한 행렬 $A = V\Lambda V^{-1}$, $V = [\mathbf{v}_1 | \mathbf{v}_2 | \cdots | \mathbf{v}_n]$ 가 주어졌다고 가정하자. 그리고 모든 i 에 대해서 $|\mathbf{v}_i| = 1$ 이고 $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots$ 도 가정한다. 먼저 크기가 1인 vector \mathbf{x} 를 Gaussian분포에서 샘플링한다. $\mathbf{w} = V^{-1}\mathbf{x}$ 로 표시하면 \mathbf{x} 가 Gaussian이므로 확률 1로 $w_1 \neq 0$ 이다.⁵ 필요하면 -1 을 곱하여 $w_1 > 0$ 이라고 하자. 이미 알고 있는 $A^k V = V\Lambda^k$ 를 활용하여

$$A^k \mathbf{x} = A^k V \mathbf{w} = V \Lambda^k \mathbf{w} = \sum_{i=1}^n w_i \lambda_i^k \mathbf{v}_i$$

를 얻고, λ_1 과 \mathbf{v}_1 위주로 더 정리하면

$$\begin{aligned} A^k \mathbf{x} &= w_1 \lambda_1^k \mathbf{v}_1 + \sum_{i=2}^n w_i \lambda_i^k \mathbf{v}_i \\ &= w_1 \lambda_1^k \left(\mathbf{v}_1 + \sum_{i=2}^n \frac{w_i}{w_1} \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{v}_i \right) \end{aligned}$$

이다. 팔호속의 합을 \mathbf{z}_k 로 표시하면

$$\begin{aligned} |\mathbf{z}_k| &\leq \sum_{i=2}^n \left| \frac{w_i}{w_1} \right| \left| \frac{\lambda_i}{\lambda_1} \right|^k |\mathbf{v}_i| \\ &= \sum_{i=2}^n \left| \frac{w_i}{w_1} \right| \left| \frac{\lambda_i}{\lambda_1} \right|^k \\ &\leq \left| \frac{\lambda_2}{\lambda_1} \right|^k \frac{\sum_{i=2}^n |w_i|}{|w_1|} \end{aligned}$$

을 얻는다. 여기서 $\frac{1}{|w_1|} \sum_{i=2}^n |w_i|$ 는 양의 상수이고 $\rho = \left| \frac{\lambda_2}{\lambda_1} \right| < 1$ 이므로 $\lim_{k \rightarrow \infty} \mathbf{z}_k = \mathbf{0}$ 은 보장되지만 ρ 가 1에 가까우면 수렴이 느릴 수도 있다. Section 10.6의 Example 10.5에서 Google Pagerank 알고리듬에서 $\rho \leq 0.85$ 가 되도록 적절한 rank-one 행렬을 더하는 아이디어가 설명된다. 실용적인 알고리듬을 생각해 보자.

$$\frac{A^k \mathbf{x}}{|A^k \mathbf{x}|} = \frac{w_1 \lambda_1^k (\mathbf{v}_1 + \mathbf{z}_k)}{w_1 \lambda_1^k |\mathbf{v}_1 + \mathbf{z}_k|} = \frac{\mathbf{v}_1 + \mathbf{z}_k}{|\mathbf{v}_1 + \mathbf{z}_k|}$$

이므로

$$\lim_{k \rightarrow \infty} \frac{A^k \mathbf{x}}{|A^k \mathbf{x}|} = \lim_{k \rightarrow \infty} \frac{\mathbf{v}_1 + \mathbf{z}_k}{|\mathbf{v}_1 + \mathbf{z}_k|} = \mathbf{v}_1$$

이어서 원하는 vector로 수렴하는 것을 알 수 있다.

$$1. \mathbf{x}_0 = \mathbf{x} \sim N(\mathbf{0}, I);$$

$$2. \mathbf{y}_{k+1} = A\mathbf{x}_k, \mathbf{x}_{k+1} = \frac{1}{|\mathbf{y}_{k+1}|} \mathbf{y}_{k+1};$$

⁵ \mathbf{x} 의 샘플링을 통해 얻는 \mathbf{w} 가 가져야 하는 성질이 $w_1 \neq 0$ 이므로 샘플링 분포가 Gaussian이라는 것이 필수적이라기 보다는 연속분포에서 샘플링하는 것이 핵심이다.

$$3. \alpha_{k+1} = \mathbf{x}_{k+1}^\top A \mathbf{x}_{k+1}.$$

2번을 반복하면 $\mathbf{x}_k \rightarrow \mathbf{v}_1$ 을 얻는다. 3번의 경우,

$$\alpha_k = \frac{(\mathbf{v}_1 + \mathbf{z}_k)^\top A(\mathbf{v}_1 + \mathbf{z}_k)}{|\mathbf{v}_1 + \mathbf{z}_k|^2} = \frac{(\mathbf{v}_1^\top + \mathbf{z}_k^\top)(\lambda_1 \mathbf{v}_1 + A \mathbf{z}_k)}{|\mathbf{v}_1 + \mathbf{z}_k|^2} = \frac{\lambda_1 + \mathbf{v}_1^\top A \mathbf{z}_k + \lambda_1 \mathbf{z}_k^\top \mathbf{v}_1 + \mathbf{z}_k^\top A \mathbf{z}_k}{|\mathbf{v}_1 + \mathbf{z}_k|^2}$$

이므로 $\lim_{k \rightarrow \infty} \alpha_k = \lambda_1$ 이다. 따라서 2,3번을 반복하면 원하는 가장 절댓값이 큰 eigenpair로 수렴하게 된다. 이 방법의 장점은 벡터와 행렬의 곱으로 iteration이 수행되어 계산량이 행렬곱보다 상대적으로 적다는 것이다.

Kang & Cho (2024)

Chapter 10

Advanced Results in Linear Algebra

이번 장에서는 아직 다루지 못했던 선형대수학 및 행렬과 관련된 몇 가지 결과를 모아 놓았다. 먼저 벡터공간의 쌍대공간(dual space)을 살펴본다. 유한차원에서는 쌍대공간이 비교적 쉽게 규명되지만 함수를 모아 놓은 벡터공간처럼 무한차원 벡터공간의 경우 단순하지 않으며 고급적분론이나 함수해석학에서 다룬다. 행렬의 transpose를 벡터공간의 operator와 연관지어 해석하기 어려운데 inner product가 정의된 벡터공간에서 선형사상의 adjoint라는 개념으로 설명할 수 있다. 대다수 독자에게 projection matrix가 idempotent해야 된다는 성질은 projection의 정의로부터 자연스럽지만 왜 symmetric해야 하는지 설명하기 쉽지 않다. 그런데 adjoint를 이용하면 projection matrix의 symmetry를 설명할 수 있다. 그리고 positiveness에 기인한 부등식으로부터 positive definite matrix를 모아놓은 집합이 convex cone을 이루는데 이와 관련된 몇 가지 성질도 살펴본다. 또한 특별한 형태의 행렬인 모든 원소가 양수인 행렬은 양수 eigenvalue와 양수 원소로 구성된 eigenvector를 갖는다는 Perron-Frobenius 정리는 확률론에서 중요하다. 이런 특별한 행렬이 아닌 일반적인 행렬이 orthogonal matrix나 unitary matrix로 기술되는 변수변환에 의해 대각원소가 eigenvalue들인 upper triangular matrix가 된다는 Schur triangularization은 활용도가 매우 높다. 마지막으로 rank-one matrix를 더하여 원래 행렬의 eigenvalue들을 조정하는 Brauer의 결과도 유용하다.

10.1 Dual Space

벡터공간 상에서 정의되어 실수 또는 복소수 값을 갖는 선형사상을 linear functional이라 부른다. 그리고 벡터공간 \mathbb{V} 상의 linear functional을 모두 모은 집합

$$\mathbb{V}^* = \{f : f \text{ is a linear functional on } \mathbb{V}\}$$

를 \mathbb{V} 의 쌍대공간(dual space)이라고 부른다. 쌍대공간은 어떤 집합일까? \mathbb{V} 상의 linear functional들을 더하거나 상수배하여도 linear functional이므로 쌍대공간 \mathbb{V}^* 도 벡터공간임을 알 수 있다. 벡터공간을 이해하는데 가장 중요한 성질이 벡터공간의 차원이다. $\dim \mathbb{V} = n$ 일 때 \mathbb{V}^* 의 차원은 얼마일까?

공변역 \mathbb{R} 을 1차원 벡터공간으로 생각하고 scalar 1을 기저벡터로 생각하자. 그러면 linear functional $f : \mathbb{V} \rightarrow \mathbb{R}$ 은 n 차원 벡터공간에서 1차원 벡터공간으로 가는 선형사상이다. \mathcal{B}_v 를 \mathbb{V} 의 기저라 하면 Section

3.8.1의 결과에 의해 $1 \times n$ 행렬 A 가 존재하여 $\mathbf{v} \in \mathbb{V}$ 의 $\mathcal{B}_{\mathbb{V}}$ 에 대한 좌표가 $\mathbf{x} \in \mathbb{R}^n$ 이라면 $f(\mathbf{v}) = A\mathbf{x}$ 이다. 특히 A 의 유일한 행을 $\mathbf{a} \in \mathbb{R}^n$ 로 표시하면 $A = [\mathbf{a}^\top]$ 이고 $f(\mathbf{v}) = \mathbf{a}^\top \mathbf{x}$ 이다. 즉 n 차원 벡터공간의 linear functional은 n 차원 Euclidean vector로 characterize될 수 있다. Section 3.8.1에서 $\mathcal{B}_{\mathbb{V}} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 이라면 $a_i = f(\mathbf{v}_i)$ 임도 밝혔는데 이를 이용하여 \mathbb{V}^* 와 \mathbb{R}^n 간의 대응관계가 injective하다는 것을 보이자. 어떤 두 linear functional f 와 g 가 다르다면 적어도 한개 이상의 basic vector \mathbf{v}_j 에 대해서 $f(\mathbf{v}_j) \neq g(\mathbf{v}_j)$ 이어야 하고 f 와 g 에 대응하는 \mathbb{R}^n 벡터를 각각 \mathbf{a} 와 \mathbf{b} 라면 $a_j \neq b_j$ 이므로 $\mathbf{a} \neq \mathbf{b}$ 이어서 대응관계가 injective하다. 또한 임의의 $\mathbf{a} \in \mathbb{R}^n$ 에 대해서 $h(\mathbf{v}) = \mathbf{a}^\top \mathbf{x}$, 즉

$$h\left(\sum_{i=1}^n x_i \mathbf{v}_i\right) = \sum_{i=1}^n a_i x_i$$

로 정의하면 h 가 \mathbb{V}^* 에서 \mathbb{R} 로 가는 선형사상임이 자명하여 대응관계가 surjective하다. 따라서 linear functional과 \mathbb{R}^n 벡터 간의 대응관계가 bijective하다는 것을 알 수 있어 $\dim \mathbb{V}^* = n$ 이고 본질적으로는 \mathbb{V}^* 는 \mathbb{R}^n 과 동일한 구조라고 생각해도 무방하다.

특별하게 $\mathbf{a} = \mathbf{e}_j$ 일 때 linear functional을 f_j 라 하자. 즉 $f_j\left(\sum_{i=1}^n x_i \mathbf{v}_i\right) = \mathbf{e}_j^\top \mathbf{x} = x_j$ 를 만족한다. 이 f_j 들의 선형독립 여부를 살펴보자. 적당한 scalar α_j 들에 대해서 linear functional 간의 관계로 $\sum_{j=1}^n \alpha_j f_j = 0$ 이 성립한다면¹ 임의의 $\mathbf{v} \in \mathbb{V}$ 에 대해서 $(\sum_{j=1}^n \alpha_j f_j)(\mathbf{v}) = \sum_{j=1}^n \alpha_j f_j(\mathbf{v}) = 0$ 인데 이는 모든 $\mathbf{x} \in \mathbb{R}^n$ 에 대하여 $\sum_{j=1}^n \alpha_j f_j\left(\sum_{i=1}^n x_i \mathbf{v}_i\right) = \sum_{j=1}^n \alpha_j x_j = 0$ 이 성립하는 것으로 $\alpha_1 = \dots = \alpha_n = 0$ 이어야 한다. 따라서 $\{f_1, \dots, f_n\} \subset \mathbb{V}^*$ 은 linearly independent하고 \mathbb{V}^* 의 차원과 같은 개수의 선형독립인 벡터들이므로 쌍대공간 \mathbb{V}^* 의 기저인데 이를 dual basis라고 부른다. \mathbb{V}^* 가 벡터공간이므로 이 공간의 쌍대 공간 $(\mathbb{V}^*)^* = \mathbb{V}^{**}$ 도 당연히 생각할 수 있으나 유한차원 공간의 경우 본질적으로 원래 벡터공간의 구조를 유지하는 것으로 생각할 수 있고 자세한 내용은 advanced linear algebra textbook에서 참조하기 바란다.

최적화 문제들 중 linearly constrained problem들은 linear functional로 정의된 등식이나 부등식들을 만족하는 벡터들 중에 objective function값을 최대 또는 최소로 하는 벡터를 찾는 문제이다. 원 문제의 제약식을 구성하는 linear functional 각각에 쌍대 변수 (dual variable)라고 부르는 새로운 변수를 정의하여 새로운 관점에서 원 최적화 문제를 살펴보는데 이를 원래 최적화 문제의 쌍대 문제(dual problem)라 부르고 최적화이론에서 중요한 역할을 한다.

10.2 Transpose of Matrices and Adjoint of Linear Transformations

Inner product가 정의된 두 vector space \mathbb{V} 와 \mathbb{W} 를 생각하자. 두 vector space의 inner product는 각각 $\langle \cdot, \cdot \rangle_{\mathbb{V}}$, $\langle \cdot, \cdot \rangle_{\mathbb{W}}$ 로 표시하자. 행렬의 transpose를 함수 관점으로 해석할 수 있는 adjoint의 개념은 다음과 같다.

¹우변의 0은 모든 벡터에 대해서 값이 0이 되는 상수 함수를 말한다.

Definition 10.1 A function $f : \mathbb{W} \rightarrow \mathbb{V}$ is an **adjoint** of a linear transformation $T : \mathbb{V} \rightarrow \mathbb{W}$ if

$$\langle T(\mathbf{v}), \mathbf{w} \rangle_{\mathbb{W}} = \langle \mathbf{v}, f(\mathbf{w}) \rangle_{\mathbb{V}} \quad (10.1)$$

for all $\mathbf{v} \in \mathbb{V}$ and $\mathbf{w} \in \mathbb{W}$.

Adjoint의 유일성을 보이는데 유용한 다음 결과를 inner product의 성질로부터 얻을 수 있다.

Fact 10.1 Let f and g be functions from \mathbb{W} to \mathbb{V} . If $\langle \mathbf{v}, f(\mathbf{w}) \rangle_{\mathbb{V}} = \langle \mathbf{v}, g(\mathbf{w}) \rangle_{\mathbb{V}}$ for all $\mathbf{v} \in \mathbb{V}$ and all $\mathbf{w} \in \mathbb{W}$, then $f = g$.

Proof: For $\mathbf{v}^* = f(\mathbf{w}) - g(\mathbf{w}) \in \mathbb{V}$,

$$0 = \langle \mathbf{v}^*, f(\mathbf{w}) \rangle_{\mathbb{V}} - \langle \mathbf{v}^*, g(\mathbf{w}) \rangle_{\mathbb{V}} = \langle \mathbf{v}^*, f(\mathbf{w}) - g(\mathbf{w}) \rangle_{\mathbb{V}} = \langle f(\mathbf{w}) - g(\mathbf{w}), f(\mathbf{w}) - g(\mathbf{w}) \rangle_{\mathbb{V}}$$

implies $f(\mathbf{w}) - g(\mathbf{w}) = 0$. ■

이 결과를 활용하면 Fact 10.2에서 adjoint가 유일함을 알 수 있다. 하지만, 무한차원 공간에서는 아래 Example 10.2에서 볼 수 있듯이 linear transformation의 adjoint가 존재하지 않을 수도 있음에 유의하자.

Fact 10.2 If a linear transformation has an adjoint, then it is unique.

Proof: f 와 g 가 linear transformation T 의 adjoint라면 (10.1)에 의해서 $\langle \mathbf{v}, f(\mathbf{w}) \rangle_{\mathbb{V}} = \langle \mathbf{v}, g(\mathbf{w}) \rangle_{\mathbb{V}}$ for all $\mathbf{v} \in \mathbb{V}$ 가 성립한다. Fact 10.1에 의해 $f = g$ 임을 알 수 있다. ■

Linear transformation T 의 adjoint가 존재하면 이 유일한 adjoint를 T^* 로 표시한다. 만약 $T^* = T$ 라면 T 를 self-adjoint라고 한다. 그리고 다음 Fact에서 T^* 도 linear transformation임을 알 수 있다.

Fact 10.3 If a linear transformation has an adjoint, then it is also linear.

Proof: Linear transformation T 의 adjoint인 T^* 가 존재한다고 하자. 그러면, 임의의 $\mathbf{v} \in \mathbb{V}$, $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{W}$ 에 대해서

$$\begin{aligned} \langle \mathbf{v}, T^*(\alpha \mathbf{w}_1 + \mathbf{w}_2) \rangle_{\mathbb{V}} &= \langle T(\mathbf{v}), \alpha \mathbf{w}_1 + \mathbf{w}_2 \rangle_{\mathbb{W}} \\ &= \langle T(\mathbf{v}), \alpha \mathbf{w}_1 \rangle_{\mathbb{W}} + \langle T(\mathbf{v}), \mathbf{w}_2 \rangle_{\mathbb{W}} \\ &= \alpha \langle T(\mathbf{v}), \mathbf{w}_1 \rangle_{\mathbb{W}} + \langle T(\mathbf{v}), \mathbf{w}_2 \rangle_{\mathbb{W}} \\ &= \alpha \langle \mathbf{v}, T^*(\mathbf{w}_1) \rangle_{\mathbb{V}} + \langle \mathbf{v}, T^*(\mathbf{w}_2) \rangle_{\mathbb{V}} \\ &= \langle \mathbf{v}, \alpha T^*(\mathbf{w}_1) \rangle_{\mathbb{V}} + \langle \mathbf{v}, T^*(\mathbf{w}_2) \rangle_{\mathbb{V}} \\ &= \langle \mathbf{v}, \alpha T^*(\mathbf{w}_1) + T^*(\mathbf{w}_2) \rangle_{\mathbb{V}} \end{aligned}$$

인 관계가 임의의 $\mathbf{v} \in \mathbb{V}$ 에 대해 성립하므로 Fact 10.1와 유사하게 $T^*(\alpha \mathbf{w}_1 + \mathbf{w}_2) = \alpha T^*(\mathbf{w}_1) + T^*(\mathbf{w}_2)$ 이어서 T^* 가 linear transformation임을 알 수 있다. ■

Example 10.1 $\mathbb{V} = \mathbb{W} = \mathbb{P} = \{a_0 + a_1x + \cdots + a_nx^n : a_i \in \mathbb{R} \text{ for } 0 \leq i \leq n, n = 0, 1, 2, \dots\}$ 인 다항식으로 이루어진 벡터공간을 생각해 보자. Inner product는 $f, g \in \mathbb{P}$ 에 대해서

$$\langle f, g \rangle_{\mathbb{P}} = \int_0^1 f(x)g(x)dx$$

로 정의하자. 다항식 $p \in \mathbb{P}$ 를 고정하고 이 다항식 p 의 곱으로 정의된 linear transformation $T_p(f) = pf$ 의 adjoint를 찾아보자.

$$\begin{aligned}\langle T_p(f), g \rangle_{\mathbb{P}} &= \langle pf, g \rangle_{\mathbb{P}} \\ &= \int_0^1 (p(x)f(x))g(x)dx \\ &= \int_0^1 f(x)(p(x)g(x))dx \\ &= \langle f, pg \rangle_{\mathbb{P}} \\ &= \langle f, T_p(g) \rangle_{\mathbb{P}}\end{aligned}$$

로부터 $T_p^* = T_p$ 임을 알 수 있다. 즉 T_p 는 self-adjoint라는 것을 알 수 있다. ■

Example 10.2 [Non-existence of Adjoint] Example 10.1의 공간을 다시 고려하자. Linear transformation으로 미분을 표현하는 $T(f) = f'$ 을 고려하자. 이 경우의 adjoint T^* 가 존재한다고 가정하자. \mathbb{P} 에 속하는 $g(x) \equiv 1$ 인 상수함수에 대해 $T^*(g) = h \in \mathbb{P}$ 로 놓자. $f \in \mathbb{P}$ 에 대해서 미적분학의 기본정리에 의해

$$\langle T(f), g \rangle_{\mathbb{P}} = \int_0^1 f'(x)dx = f(1) - f(0)$$

이고

$$\langle f, T^*(g) \rangle_{\mathbb{P}} = \int_0^1 f(x)h(x)dx$$

이므로 임의의 $f \in \mathbb{P}$ 에 대해서

$$\int_0^1 f'(x)dx = f(1) - f(0) = \int_0^1 f(x)h(x)dx$$

이 성립하여야 한다. 특별히 $f(x) = x^2(x-1)^2h(x)$ 로 놓아서 $f(1) = f(0) = 0$ 이라면

$$0 = \int_0^1 f(x)h(x)dx = \int_0^1 x^2(x-1)^2h(x)^2dx$$

이다. 특히 $q(x) = x(x-1)h(x) \in \mathbb{P}$ 에 대하여 $\langle q, q \rangle_{\mathbb{P}} = 0$ 이므로 $q \equiv 0$ 이고, $T^*(g) = h \equiv 0$ 이라는 것을 알 수 있다.² 따라서 임의의 $f \in \mathbb{P}$ 에 대해서

$$f(1) - f(0) = \langle T(f), g \rangle_{\mathbb{P}} = \langle f, T^*(g) \rangle_{\mathbb{P}} = \langle f, 0 \rangle_{\mathbb{P}} = 0,$$

²정적분으로 정의된 식이 inner product라는 사실을 생각할 때도 같은 논증을 해야하는데, 미적분학 지식으로부터 만약 $[0, 1]$ 에서 $h(x)$ 가 0이 아닌 함수값을 갖는 구간의 길이가 0보다 크면 위의 적분이 0이 될 수 없고, $h(x)$ 가 다항식이므로 $h(x) \equiv 0$ 이어야 한다는 것을 알 수 있다.

즉 $f(1) = f(0)$ 이 성립해야 한다. 하지만 $f(x) = x \in \mathbb{P}$ 는 $f(1) = 1 \neq 0 = f(0)$ 이어서 모순이다. 즉 미분을 나타내는 linear transformation은 adjoint를 갖지 않는다. ■

다음 Fact에서 adjoint와 matrix transpose 간의 관계를 알 수 있다.

Fact 10.4 Consider vector spaces $\mathbb{V} = \mathbb{R}^n$ and $\mathbb{W} = \mathbb{R}^m$ with their standard inner products. Let A be an $m \times n$ matrix. Then, the adjoint of $T(\mathbf{v}) = A\mathbf{v}$ is $T^*(\mathbf{w}) = A^\top \mathbf{w}$.

Proof: Fact 10.2의 유일성 때문에 어떤 방법으로든 adjoint를 찾으면 된다. 그런데

$$\langle T(\mathbf{v}), \mathbf{w} \rangle_{\mathbb{W}} = (A\mathbf{v})^\top \mathbf{w} = \mathbf{v}^\top A^\top \mathbf{w} = \langle \mathbf{v}, A^\top \mathbf{w} \rangle_{\mathbb{V}}$$

이므로 $T^*(\mathbf{w}) = A^\top \mathbf{w}$ 임을 알 수 있다. ■

이러한 연유로 transpose와 adjoint를 혼용하여 사용하기도 한다. 또한, 유한차원 공간의 self-adjoint operator를 나타내는 행렬은 symmetric함을 알 수 있고 self-adjoint와 symmetric을 혼용하기도 한다.

끝으로 finite dimensional vector space간의 linear transformation의 adjoint를 생각해 보자. 유한차원 공간에서는 linear transformation을 언제나 유한한 크기의 행렬로 나타낼 수 있고 Fact 10.4은 이러한 경우에 adjoint가 행렬의 transpose로 정의됨을 알 수 있다. 즉 유한차원 공간 간에서는 항상 adjoint가 존재한다.

10.2.1 Adjoint and Projection

Fact 4.16에서 orthogonal projection matrix P 는 symmetric하고 $P^2 = P$ 를 만족해야 함을 알 수 있었다. $P^2 = P$ 와 같이 어떤 행렬 또는 함수의 제곱이 원래 행렬 또는 함수와 일치하는 성질을 idempotent라고 부른다. Projection matrix P 의 경우 P 를 통해 부분공간의 vector를 일단 대응시키면 이미 부분공간에 위치하므로 P 를 추가로 적용하여도 더 이상 변화하지 않아서 idempotent가 자연스럽다. 하지만 symmetry는 어떻게 projection과 연관지을 수 있을까? Projection을 꼭 orthogonal한 방향으로 할 필요는 없다. 2차원에서 1차원 부분공간에 45° 사선 방향으로 projection을 수행해도 1차원 부분공간 상의 vector를 대응시킬 수는 있다. 하지만, 우리가 흔히 projection이라고 할 때는 orthogonal한 방향으로 대응시키는 orthogonal projection을 의미한다. 그러면 이 orthogonal projection을 어떻게 기술할 수 있을까? Inner product $\langle \cdot, \cdot \rangle$ 가 정의된 vector space \mathbb{V} 에서 orthogonal projection을 다음과 같이 생각해 보자. $P(\cdot)$ 가 부분공간 \mathbb{W} 로의 orthogonal projection을 나타내는 linear transformation이라 하자. 그러면 임의의 $\mathbf{v}, \mathbf{w} \in \mathbb{V}$ 에 대해 $\mathbf{v} - P(\mathbf{v}) \perp \mathbb{W}$ 이고 $P(\mathbf{w}) \in \mathbb{W}$ 이므로

$$0 = \langle \mathbf{v} - P(\mathbf{v}), P(\mathbf{w}) \rangle = \langle \mathbf{v}, P(\mathbf{w}) \rangle - \langle P(\mathbf{v}), P(\mathbf{w}) \rangle$$

이고 마찬가지로 \mathbf{v} 와 \mathbf{w} 의 역할을 바꾸면 $\langle P(\mathbf{v}), \mathbf{w} \rangle = \langle P(\mathbf{v}), P(\mathbf{w}) \rangle$ 도 얻는다. 따라서

$$\langle P(\mathbf{v}), \mathbf{w} \rangle = \langle P(\mathbf{v}), P(\mathbf{w}) \rangle = \langle \mathbf{v}, P(\mathbf{w}) \rangle,$$

즉 linear transformation으로서 orthogonal projection은 self-adjoint임을 알 수 있다. Fact 10.4를 통해 유한차원에서 self-adjoint operator를 나타내는 행렬은 symmetric함을 보였으므로, 이로부터 projection을 나타내는 행렬이 symmetric함도 자연스럽게 알 수 있다.

반대로 idempotent P 가 self-adjoint하다고 가정하자. 즉 임의의 $\mathbf{v}, \mathbf{w} \in \mathbb{V}$ 에 대해 $\langle P(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, P(\mathbf{w}) \rangle$ 가 성립한다. \mathbf{v} 대신에 $P(\mathbf{v})$, \mathbf{w} 대신에 \mathbf{v} 를 넣으면 $\langle P(P(\mathbf{v})), \mathbf{v} \rangle = \langle P(\mathbf{v}), P(\mathbf{v}) \rangle$ 인데 P 가 idempotent이므로 $P^2 = P$ 이어서 $\langle P(\mathbf{v}), \mathbf{v} \rangle = \langle P(\mathbf{v}), P(\mathbf{v}) \rangle$ 를 얻게 된다. 우변을 이항하면

$$\langle P(\mathbf{v}), \mathbf{v} - P(\mathbf{v}) \rangle = 0$$

이므로 P 가 orthogonal projection임을 알 수 있다. 이를 정리하면 다음 Lemma와 같이 기술할 수 있다.

Lemma 10.1 *An idempotent linear transformation represents a projection. The projection is orthogonal if and only if the transformation is self-adjoint.*

10.3 Further Theory of Positive Definite Matrices

덧셈과 곱셈이 잘 정의되는 정사각행렬은 곱셈의 비교환성 등 제약이 있지만 실수와 유사한 면도 많다. 정사각행렬 중에서 특히 positive definite matrix는 양의 실수에 해당하는 행렬로서 양수의 성질과 유사한 성질을 많이 가지고 있다. 변수가 많은 복잡한 시스템을 방정식으로 모형화할 때 정사각행렬이 변수인 경우가 종종 생기는데 변수행렬이 positive definite인 경우 문제를 해결하는 데 positive definite matrix에 관한 다양한 결과들이 적용 가능하다. Positive definite matrix에 관한 여러 이야기를 할 때 다음과 같은 기호들을 사용하면 여려모로 편리하다.

- $\mathbb{R}^{m,n}$: set of $m \times n$ matrices with real entries;
- \mathbb{S}^n : set of $n \times n$ symmetric matrices;
- \mathbb{S}_+^n : set of $n \times n$ positive semi-definite matrices. $A \succeq \mathbf{0}$ if $A \in \mathbb{S}_+^n$;
- \mathbb{S}_{++}^n : set of $n \times n$ positive definite matrices. $A \succ \mathbf{0}$ if $A \in \mathbb{S}_{++}^n$;
- $A \succ B$ (resp., $A \succeq B$) if and only if $A - B \succ \mathbf{0}$ (resp., $A - B \succeq \mathbf{0}$).

10.3.1 Congruence Transformations

앞에서 행렬의 similarity에 관해서 Definition 9.1에서 살펴보았다. 이 개념은 동일한 linear transformation이 다른 기저에 대해서 표현되는 상황을 말하는 것으로 transformation이 동일하므로 두 행렬이 similar하다고 말한다. 이에 반해서 다음 congruence 개념은 이차함수의 계수로 행렬이 사용되는 경우에 원래 변수를 linear transform한 이차함수의 계수 행렬을 표현하고 이 경우 행렬의 양수 성질에 해당하는 positive (semi-)definiteness가 잘 보존됨을 설명한다.

Theorem 10.1 *Let $A \in \mathbb{S}^n$ and $B \in \mathbb{R}^{n,m}$, and consider the product*

$$C = B^\top AB \in \mathbb{S}^m. \quad (10.2)$$

1. If $A \succeq \mathbf{0}$, then $C \succeq \mathbf{0}$;

2. If $A \succ \mathbf{0}$, then $C \succ \mathbf{0}$ if and only if $\text{rank } B = m$;
3. If B is square and invertible, then $A \succ \mathbf{0}$ (resp. $A \succeq \mathbf{0}$) if and only if $C \succ \mathbf{0}$ (resp. $C \succeq \mathbf{0}$).

Proof: For $\mathbf{x} \in \mathbb{R}^m$, set $\mathbf{y} = B\mathbf{x} \in \mathbb{R}^n$.

1. $\mathbf{x}^\top C \mathbf{x} = \mathbf{y}^\top A \mathbf{y} \geq 0$ since $A \succeq \mathbf{0}$;
2. If $\text{rank } B = m$, $\mathbf{y} = \mathbf{0} \Leftrightarrow \mathbf{x} = \mathbf{0}$. So, $\mathbf{x}^\top C \mathbf{x} = \mathbf{y}^\top A \mathbf{y} > 0$ if $\mathbf{x} \neq \mathbf{0}$. That is, $C \succ \mathbf{0}$. Conversely, if $C \succ \mathbf{0}$, then $\mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{y} = \mathbf{0}$, which implies B has a full-column rank, i.e., $\text{rank } B = m$;
3. If B is square and invertible, $\text{rank } B = m$. By applying 1 and 2 to both (10.2) and $A = B^{-\top} C B^{-1}$, we obtain the results.

■

(10.2) is called a **congruence transformation**. Compare this congruence transformation with the similarity transformation in Definition 9.1.

Corollary 10.1 For any matrix $A \in \mathbb{R}^{m,n}$, it holds that:

1. $A^\top A \succeq \mathbf{0}$ and $AA^\top \succeq \mathbf{0}$;
2. $A^\top A \succ \mathbf{0}$ if and only if A has full-column rank, i.e., $\text{rank } A = n$;
3. $AA^\top \succ \mathbf{0}$ if and only if A has full-row rank, i.e., $\text{rank } A = m$.

Proof: By setting $A = I_n$ in Theorem 10.1, we obtain all results. ■

행렬이 포함된 수리 모형에서 그 행렬들의 대각화가 가능하면 문제의 난이도를 낮출 수 있는 경우가 종종 있어 행렬의 대각화는 중요한 문제이다. 다음 결과들은 특히 두 개의 행렬이 동시에 대각화 가능하거나 곱이 대각화 가능한 경우에 대한 이야기이다.

Theorem 10.2 (Joint diagonalization by similarity transformation) Let $A_1, A_2 \in \mathbb{S}^n$ be such that

$$A = \alpha_1 A_1 + \alpha_2 A_2 \succ \mathbf{0}$$

for some scalars α_1, α_2 . Then, there exists a nonsingular matrix $B \in \mathbb{R}^{n,n}$ such that $B^\top A_1 B$ and $B^\top A_2 B$ are diagonal.

Proof: If both α_1 and α_2 vanish, $A \not\succ \mathbf{0}$. Hence, at least one of them should be non-zero, and we assume that $\alpha_2 \neq 0$. Since $A \succ \mathbf{0}$, $A = C^\top C$ for some invertible matrix C . If we plug it into the original equation, we get $C^\top C = \alpha_1 A_1 + \alpha_2 A_2$. We modify it as

$$I_n = \alpha_1 C^{-\top} A_1 C^{-1} + \alpha_2 C^{-\top} A_2 C^{-1}.$$

Here, the point is that $C^{-\top} A_i C^{-1}$ is still symmetric. So the symmetric spectral decomposition guarantees $C^{-\top} A_1 C^{-1} = Q \Lambda Q^\top$ where Q is orthogonal and Λ is a diagonal matrix with eigenvalues as its diagonal entries. We multiply Q^\top and Q on both sides of the equation once again and obtain

$$I_n = Q^\top I_n Q = \alpha_1 Q^\top C^{-\top} A_1 C^{-1} Q + \alpha_2 Q^\top C^{-\top} A_2 C^{-1} Q = \alpha_1 \Lambda + \alpha_2 Q^\top C^{-\top} A_2 C^{-1} Q.$$

Since $\alpha_2 \neq 0$, we can modify the above equation as

$$\frac{1}{\alpha_2} (I_n - \alpha_1 \Lambda) = Q^\top C^{-\top} A_2 C^{-1} Q$$

where the right side is diagonal because the left side is diagonal. Hence, if we set $B = C^{-1}Q$, then B is invertible and diagonalizes both A_1 and A_2 . ■

Corollary 10.2 *Let $A \succ \mathbf{0}$ and $C \in \mathbb{S}^n$. Then there exists a nonsingular matrix B such that $B^\top C B$ is diagonal and $B^\top A B = I_n$.*

Proof: If we apply Theorem 10.2 to $A = 1A + 0C \succ \mathbf{0}$, then there exists an invertible \bar{B} such that both $\bar{B}^\top A \bar{B}$ and $\bar{B}^\top C \bar{B}$ are diagonal. By 3 of Theorem 10.1, $\bar{B}^\top A \bar{B} \succ \mathbf{0}$. Since $\bar{B}^\top A \bar{B} = \text{diag}(\lambda_i)$, all $\lambda_i > 0$. Define $D = \text{diag}(\sqrt{\lambda_i})$. Set $B = \bar{B}D^{-1}$. Then, $B^\top A B = D^{-1}\bar{B}^\top A \bar{B} D^{-1} = D^{-1}D^2 D^{-1} = I_n$, and $B^\top C B = D^{-1}\bar{B}^\top C \bar{B} D^{-1}$ is diagonal since both $\bar{B}^\top C \bar{B}$ and D^{-1} are diagonal. ■

Corollary 10.3 *Let $A, B \in \mathbb{S}^n$ with $A \succ \mathbf{0}$. Then, the matrix AB is diagonalizable and has purely real eigenvalues.*

Proof: Let $A^{1/2}$ be the square root of positive definite A . Then,

$$A^{-1/2}ABA^{1/2} = A^{1/2}BA^{1/2}.$$

The matrix on the right side is symmetric, diagonalizable, and has real eigenvalues. Since AB and $A^{1/2}BA^{1/2}$ are similar (as in Definition 9.1), both share the same eigenvalues as well as their diagonalizability. ■

10.3.2 Positive Semi-definite Cone and Partial Order

많은 대상을 수학적으로 살펴볼 때 대상을 기술하는 집합의 중요한 구조 중 하나가 집합의 볼록성(convexity)이다. 최적화나 확률적 분석 등에서 대상이 가지고 있는 볼록성을 이해하면 수학적 분석과 문제 해결이 우아해지는 경우가 종종 있다. 양의 실수 집합이나 실수평면에서 1사분면 집합과 같이 한쪽 방향으로 한없이 뻗어나갈 수 있는 구조를 가지고 있는 집합을 cone이라 부르는데 이 cone의 볼록성을 살펴본다.

Definition 10.2 Let \mathbb{V} be a vector space.

A subset K of \mathbb{V} is **convex** if

$$\lambda\mathbf{v}_1 + (1 - \lambda)\mathbf{v}_2 \in K \text{ for all } \mathbf{v}_1, \mathbf{v}_2 \in K \text{ and all } \lambda \in (0, 1).$$

A subset K of \mathbb{V} is a **cone** if

$$\lambda\mathbf{v} \in K \text{ for all } \mathbf{v} \in K \text{ and all } \lambda \geq 0.$$

A subset of \mathbb{V} is a **convex cone** if it is convex and a cone.

행렬과 연관되어 positive (semi-)definite matrix들이 convex cone 구조를 갖는다. 앞에서 언급한 것처럼 Positive definiteness가 양의 실수와 유사하다는 것을 생각하면 자연스러운 결과로 생각할 수 있다.

Fact 10.5 \mathbb{S}_+^n and \mathbb{S}_{++}^n are convex cones.

Proof: It is clear that \mathbb{S}_+^n is a cone. If $A, B \in \mathbb{S}_+^n$, then $\mathbf{x}^\top(\lambda A + (1 - \lambda)B)\mathbf{x} = \lambda\mathbf{x}^\top A\mathbf{x} + (1 - \lambda)\mathbf{x}^\top B\mathbf{x} \geq 0$ since $\mathbf{x}^\top A\mathbf{x} \geq 0$ and $\mathbf{x}^\top B\mathbf{x} \geq 0$. Hence, $\lambda A + (1 - \lambda)B \in \mathbb{S}_+^n$. It is parallel to show that \mathbb{S}_{++}^n is a convex cone. ■

하지만 \mathbb{S}_+^n 나 \mathbb{S}_{++}^n 와 양의 실수 간의 중요한 차이는 임의의 두 대상 간에 대소관계가 늘 존재하는지 여부이다. 행렬의 경우 대소 비교가 불가능한 예가 존재한다.

Example 10.3 \mathbb{S}_+^n 에 포함된 positive semi-definite matrix간에 크기 순서(order)를 정할 수 있을까? 자연스러운 생각은 앞에서 정의된 $A \succeq B$ iff $A - B \succeq \mathbf{0}$ 를 이용하는 것이다. \succeq 에 대해서 $A \succeq B$ 이고 $B \succeq C$ 이면 $A \succeq C$ 인 transitivity는 성립한다. 그런데 $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ 간에는 $A \succeq B$, $C \succeq B$ 는 성립하지만 A 와 C 간에는 $A \succeq C$ 나 $C \succeq A$ 가 성립하지 않느다. 즉 임의의 두 positive semi-definite matrix간에 \succeq 에 기반한 관계는 존재하지 않을 수 있는 것이다. 따라서 \mathbb{S}_+^n 상에서 \succeq 는 partial order가 된다. ■

두 양의 실수 a, b 가 $a \geq b$ 라면 $ba^{-1} \leq 1$ 인 것과 유사한 성질을 행렬에 대해 기술한 것이 다음 결과이다.

Theorem 10.3 Let $A \succ \mathbf{0}$ and $B \succeq \mathbf{0}$, and denote by $\rho(\cdot)$ the spectral radius of a matrix (that is, the maximum absolute value of the eigenvalues of a matrix). Then,

$$A \succeq B \Leftrightarrow \rho(BA^{-1}) \leq 1, \quad (10.3)$$

$$A \succ B \Leftrightarrow \rho(BA^{-1}) < 1. \quad (10.4)$$

Proof: By Corollary 10.2, there exists a nonsingular matrix M such that $M^{-\top}AM^{-1} = I_n$ and $M^{-\top}BM^{-1} = D = \text{diag}(d_i)$ is a diagonal matrix. Note that $d_i \geq 0$ since $B, D \in \mathbb{S}_+^n$. We apply Theorem 10.1 to

$$A - B = M^\top I_n M - M^\top DM = M^\top (I_n - D)M \succeq \mathbf{0}$$

and obtain that $I_n - D \succeq \mathbf{0}$. Therefore, $d_i \leq 1$ for all i and $\rho(D) \leq 1$.

Since $B = M^\top DM$ and $A^{-1} = M^{-1}M^{-\top}$, $BA^{-1} = M^\top DMM^{-1}M^{-\top} = M^\top DM^{-\top}$, that is, D and BA^{-1} are similar to each other. By Fact 9.10, D and BA^{-1} share the same eigenvalues and $\rho(BA^{-1}) \leq 1$. (10.4) can be proved parallel to the proof of (10.3). ■

두 정사각행렬 A, B 에 대해서 AB 와 BA 는 동일한 eigenvalue들을 가진다. 따라서 $\rho(AB) = \rho(BA)$ 이다. 만약 $A \succ \mathbf{0}$ 이고 $B \succ \mathbf{0}$ 라면 Theorem 10.3에 의해

$$A \succeq B \Leftrightarrow \rho(BA^{-1}) = \rho(A^{-1}B) \leq 1 \Leftrightarrow B^{-1} \succeq A^{-1}$$

이 성립한다. 양의 실수와 역수 사이의 대소관계와 매우 비슷하다. 행렬간의 대소관계인 $A \succeq B$ 가 다른 대소관계로 자연스럽게 해석되는 경우를 살펴보자. $A, B \in \mathbb{S}_+^n$ 라고 하자. Fact 7.8에 의해 $A \succeq B$ 이면 모든 i 에 대하여

$$\lambda_i(A) = \lambda_i(B + (A - B)) \geq \lambda_i(B)$$

가 성립한다. 따라서 다음과 같은 부등식을 얻는다.

$$\begin{aligned} \det A &= \prod_{i=1}^n \lambda_i(A) \geq \prod_{i=1}^n \lambda_i(B) = \det B \\ \text{trace } A &= \sum_{i=1}^n \lambda_i(A) \geq \sum_{i=1}^n \lambda_i(B) = \text{trace } B. \end{aligned}$$

\mathbb{S}_+^n 상에서 정의된 함수로 생각할 수 있는 determinant와 trace가 \succeq 관계에 대해서 monotone함을 알 수 있다.

약간 생소한 형태이지만 system control theory 등에서 자주 나타나는 문제와 연관된 symmetric sum에 관한 결과도 종종 유용하다.

Theorem 10.4 (Symmetric Sum) Let $A \succ \mathbf{0}$ and $B \in \mathbb{S}^n$, and consider the symmetric sum

$$S = AB + BA.$$

Then, $S \succeq \mathbf{0}$ (resp., $S \succ \mathbf{0}$) implies that $B \succeq \mathbf{0}$ (resp., $B \succ \mathbf{0}$).

Proof: B 가 symmetric하므로 real spectral decomposition에 의해 적당한 orthogonal matrix $Q = [\mathbf{q}_1 | \mathbf{q}_2 | \cdots | \mathbf{q}_n]$ 와 B 의 eigenvalue로 이루어진 대각행렬 $\Lambda = \text{diag}(\lambda_i(B))$ 에 대해서 $B = Q^\top \Lambda Q$ 으로 표현된다. $S \succeq \mathbf{0}$ 이므로 Theorem 10.1에 의해 $Q^\top SQ \succeq \mathbf{0}$ 이고, 특히 대각원소들도 $(Q^\top SQ)_{ii} \geq 0$ 을 만족한다.

$$\begin{aligned} Q^\top SQ &= Q^\top ABQ + Q^\top BAQ \\ &= Q^\top AQQ^\top BQ^\top + Q^\top BQQ^\top AQ \\ &= Q^\top AQL + \Lambda Q^\top AQ \end{aligned}$$

이므로 $(Q^\top SQ)_{ii} = 2\lambda_i(B)(QAQ^\top)_{ii} = 2\lambda_i(B)\mathbf{q}_i^\top A\mathbf{q}_i \geq 0$ 이다. $A \succ \mathbf{0}$ 이므로 $\mathbf{q}_i^\top A\mathbf{q}_i > 0$ 이어서 $\lambda_i(B) \geq 0$ 이어야 한다. 즉 $B \succeq \mathbf{0}$ 이어야 한다. $S \succ \mathbf{0}$ 인 경우도 유사하게 증명할 수 있다. ■

Example 10.4 [Matrix square-root preserves the PSD ordering] $A \succ \mathbf{0}^\circ$ 이고 $B \succ \mathbf{0}$ 라 하자. 그러면,

$$2(A - B) = (A^{1/2} + B^{1/2})(A^{1/2} - B^{1/2}) + (A^{1/2} - B^{1/2})(A^{1/2} + B^{1/2})$$

인데 $A^{1/2} + B^{1/2} \succ \mathbf{0}^\circ$ 이고 $A^{1/2} - B^{1/2} \in \mathbb{S}_+^n$ 이므로 Theorem 10.4를 적용할 수 있다. 즉 $A \succ B$ 이면 $2(A - B) \succ \mathbf{0}^\circ$ 에서 $A^{1/2} - B^{1/2} \succ \mathbf{0}$, 즉 $A^{1/2} \succ B^{1/2}$ 가 성립한다. 양의 실수에서 square-root가 대소관계를 유지하는 것과 유사하다. 하지만, 역은 일반적으로 성립하지 않는다. $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1.2 & 1 \\ 1 & 0.9 \end{bmatrix}$ 의 예에서 살펴보면 $A \succ \mathbf{0}^\circ$ 이고 $B \succ \mathbf{0}$, $A \succ B^\circ$ 지만 $A^2 \not\succ B^2$ 이다. ■

앞에서 여러 결과들을 block matrix에 대해 살펴보면 새로운 표현이나 조건을 얻는 경우가 있었다. Positive definiteness에 관해서도 block diagonal 관점에서 살펴보자. 다음과 같은 block diagonal matrix의 경우 행렬 전체의 positive definiteness가 diagonal block 각각의 positive definiteness에 의해 결정된다. 즉 $M = \begin{bmatrix} A & \mathbf{0}_{n,m} \\ \mathbf{0}_{m,n} & B \end{bmatrix}$ 에 대해서

$$M \succeq \mathbf{0} \text{ (resp., } M \succ \mathbf{0}) \Leftrightarrow A \succeq \mathbf{0}, B \succeq \mathbf{0} \text{ (resp., } A \succ \mathbf{0}, B \succ \mathbf{0})$$

가 성립한다.³ 이 사실을 이용하면 symmetric block diagonal matrix에 대해서 확장된 다음 결과를 얻을 수 있다.

Fact 10.6 (Schur Complement) Let $A \in \mathbb{S}^n$, $B \in \mathbb{S}^m$, $C \in \mathbb{R}^{n,m}$ with $B \succ \mathbf{0}$. Consider the symmetric block matrix

$$M = \begin{bmatrix} A & C \\ C^\top & B \end{bmatrix},$$

and consider the Schur complement $S = A - CB^{-1}C^\top$ of A with respect to M . Then,

$$M \succeq \mathbf{0} \text{ (resp., } M \succ \mathbf{0}) \Leftrightarrow S \succeq \mathbf{0} \text{ (resp., } S \succ \mathbf{0}).$$

Proof: Block matrix에 대한 Gaussian elimination을 진행하면 다음과 같이 block diagonal matrix를 얻는다. $L = \begin{bmatrix} I_n & \mathbf{0} \\ -B^{-1}C^\top & I_m \end{bmatrix}$ 으로 놓으면

$$\begin{aligned} L^\top M L &= \begin{bmatrix} I_n & -CB^{-1} \\ \mathbf{0} & I_m \end{bmatrix} \begin{bmatrix} A & C \\ C^\top & B \end{bmatrix} \begin{bmatrix} I_n & \mathbf{0} \\ -B^{-1}C^\top & I_m \end{bmatrix} \\ &= \begin{bmatrix} S & \mathbf{0} \\ C^\top & B \end{bmatrix} \begin{bmatrix} I_n & \mathbf{0} \\ -B^{-1}C^\top & I_m \end{bmatrix} \\ &= \begin{bmatrix} S & \mathbf{0}_{n,m} \\ \mathbf{0}_{m,n} & B \end{bmatrix} =: D \end{aligned}$$

³이 사실을 증명하는 것은 각자 해보기 바란다.

이다. $B > \mathbf{0}$ 이므로 D 의 positive definiteness는 S 의 positive definiteness와 동일하다. 그런데 L 은 대각원소가 모두 1인 lower triangular matrix로 invertible하므로 Theorem 10.1 3에 의해 M 의 positive definiteness는 D 의 positive definiteness와 동일하여 원하는 결과를 얻는다. ■

10.4 Schur triangularization

이후에 다룰 중요한 정리들의 증명에 Schur triangularization이 자주 활용된다. Theorem 10.5은 그 자체도 매우 중요하고 여러 분야에서 활용이 많이 된다. 행렬의 원소가 복소수일 수도 있으면 실수 행렬에서 orthogonal matrix의 복소수 버전인 행렬을 도입할 필요가 생기는데 이를 unitary matrix라고 부른다. 행렬 Q 가 unitary라 함은 $Q^{-1} = Q^H$ 인 것이다. 즉 $QQ^H = Q^HQ = I$ 이다. 어떤 행렬 A 가 $A = QBQ^H$ 이면 $A^n = QB^nQ^H$ 인데 거듭제곱 계산에서 유용하다.

Theorem 10.5 (Schur Triangularization) *Let the eigenvalues of $n \times n$ matrix A be arranged in any given order $\lambda_1, \lambda_2, \dots, \lambda_n$ (including multiplicities), and let (λ_1, \mathbf{x}) be an eigenpair of A , in which \mathbf{x} is a unit vector. Then,*

- (a) *There is an $n \times n$ unitary $Q = [\mathbf{x}|Q_2]$ such that $A = QUQ^H$, in which $U = (u_{ij})$ is upper triangular and has diagonal entries $u_{ii} = \lambda_i$ for $i = 1, 2, \dots, n$.*
- (b) *If A is real, each eigenvalue is real, and \mathbf{x} is real, then there is an $n \times n$ real orthogonal $Q = [\mathbf{x}|Q_2]$ such that $A = QUQ^T$, in which $U = (u_{ij})$ is upper triangular and has diagonal entries $u_{ii} = \lambda_i$ for $i = 1, 2, \dots, n$.*

Proof: n 에 대한 수학적 귀납법을 사용하려고 한다. 먼저 $n = 1$ 인 경우는 별도의 증명없이 (a)가 성립함을 알 수 있다. $n \geq 2$ 인 경우를 생각해 보자. $n - 1$ 까지는 성립한다고 가정한다. \mathbf{x} 를 포함하는 unitary한 n 개의 vector를 (Gram-Schmidt 등을 활용하여) 구해서 $n \times n$ unitary matrix $\hat{Q} = [\mathbf{x}|V]$ 를 구성한다. 여기서 V 는 $n \times (n - 1)$ 행렬로 $V^H \mathbf{x} = \mathbf{0}$ 이다.

$$A\hat{Q} = [A\mathbf{x}|AV] = [\lambda_1 \mathbf{x}|AV]$$

이므로

$$\hat{Q}^H A \hat{Q} = \hat{Q}^H [\lambda_1 \mathbf{x} | AV] = \begin{bmatrix} \mathbf{x}^H \\ V^H \end{bmatrix} [\lambda_1 \mathbf{x} | AV] = \begin{bmatrix} \lambda_1 \mathbf{x}^H \mathbf{x} & \mathbf{x}^H AV \\ \lambda_1 V^H \mathbf{x} & V^H AV \end{bmatrix} = \begin{bmatrix} \lambda_1 & \mathbf{x}^H AV \\ \mathbf{0} & V^H AV \end{bmatrix}$$

이다. 마지막 upper triangular block matrix의 eigenvalue는 각 block의 eigenvalue들로 구성된다. 즉 λ_1 과 $V^H AV$ 의 eigenvalue들이 마지막 block matrix의 eigenvalue들이다. 그런데 A 와 $\hat{Q}^H A \hat{Q}$ 는 similar하므로 동일한 eigenvalue를 갖는다. 따라서 $V^H AV$ 의 eigenvalue들이 $\lambda_2, \dots, \lambda_n$ 임을 알 수 있다. $V^H AV$ 가 $(n-1) \times (n-1)$ 이므로 수학적 귀납법 가정에 의해 적당한 $(n-1) \times (n-1)$ unitary matrix Q_{n-1} 와 대각원소가 $\lambda_2, \dots, \lambda_n$ 인 upper triangular matrix U_{n-1} 이 존재하여 $V^H AV = Q_{n-1} U_{n-1} Q_{n-1}^H$ 가 성립한다. 적절한 $(n-1)$ 차원

벡터 \mathbf{a} 에 대해서 block matrix로 표현하면

$$\hat{Q}^H A \hat{Q} = \begin{bmatrix} \lambda_1 & \mathbf{x}^H A V \\ \mathbf{0} & Q_{n-1} U_{n-1} Q_{n-1}^H \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q_{n-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & \mathbf{a}^H \\ \mathbf{0} & U_{n-1} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q_{n-1} \end{bmatrix}^H$$

이 된다. Q_{n-1} 가 unitary이므로 $\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q_{n-1} \end{bmatrix}$ 도 unitary이고 $Q = \hat{Q} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q_{n-1} \end{bmatrix}$ 도 unitary임을 알 수

있다. 또한, $U = \begin{bmatrix} \lambda_1 & \mathbf{a}^H \\ \mathbf{0} & U_{n-1} \end{bmatrix}$ 는 대각원소가 $\lambda_1, \lambda_2, \dots, \lambda_n$ 인 upper triangular matrix이다. 그리고 이 Q 와 U 에 대해서 $A = Q U Q^H$ 가 성립하여 (a)가 증명된다. (b)는 (a)에서 conjugation이 transpose로 바뀌어 자동적으로 성립한다. ■

Schur triangularization의 간단한 응용으로 trace와 determinant를 eigenvalue로 표현할 수 있다.

Corollary 10.4 Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of $n \times n$ matrix A . Then,

$$\text{trace } A = \lambda_1 + \lambda_2 + \dots + \lambda_n \quad \text{and} \quad \det A = \lambda_1 \lambda_2 \dots \lambda_n.$$

Proof: Theorem 10.5에 의해 $A = Q U Q^H$ 로 표현될 수 있다. 그러면,

$$\text{trace } A = \text{trace}(Q U Q^H) = \text{trace}(U Q^H Q) = \text{trace } U = \lambda_1 + \lambda_2 + \dots + \lambda_n$$

이고

$$\det A = \det(Q U Q^H) = \det Q \det U \det Q^H = \det(Q Q^H) \det U = \det U = \lambda_1 \lambda_2 \dots \lambda_n$$

이어서 증명이 완료된다. ■

10.5 Perron-Frobenius Theorem

$n \times n$ 행렬 $A = (a_{ij})$ 가 $a_{ij} \geq 0$ 이고 $\sum_{j=1}^n a_{ij} = 1$, 즉 각 행의 합이 1이고 원소들이 0보다 크거나 같다고 하자. 이런 특성을 만족하는 행렬을 **Markov matrix**라고 부르며 Markov chain이라는 중요한 확률 모형에서 시스템 상태의 확률적 전이를 기술하는데 주로 사용된다. 이 경우, a_{ij} 는 시스템이 상태 i 에서 상태 j 로 바뀔 확률을 나타낸다. 모든 행의 합이 1이므로 Markov matrix는 $(1, 1)$ 를 eigenpair로 갖는다. 그리고 다른 eigenvalue λ_i 들은 $|\lambda_i| \leq 1$ 임을 보일 수 있다.

Fact 10.7 $\rho(A) = 1$ if A is a Markov matrix.

Proof: (λ, \mathbf{v}) 가 A 의 임의의 eigenpair라 하자. $k = \text{argmax}_{1 \leq i \leq n} |v_i|$ 라면 $0 < |v_k|$ 이다. $A\mathbf{v} = \lambda\mathbf{v}$ 의 k 번째 등식 $\sum_{j=1}^n a_{kj} v_j = \lambda v_k$ 을 살펴보자. A 가 Markov matrix이므로 $a_{kj} \geq 0$ 이고 $\sum_{j=1}^n a_{kj} = 1$ 이다. 따라서

$$|\lambda| |v_k| = |\lambda v_k| = \left| \sum_{j=1}^n a_{kj} v_j \right| \leq \sum_{j=1}^n a_{kj} |v_j| \leq \sum_{j=1}^n a_{kj} |v_k| = |v_k|$$

이므로 $|\lambda| \leq 1$ 이다. 그런데 $(1, 1)$ 가 eigenpair이므로 $\rho(A) = 1$ 이다. ■

어떤 vector \mathbf{p} 가 probability vector라 함은 $\mathbf{p} \geq \mathbf{0}$ 이고 $\mathbf{p}^\top \mathbf{1} = 1$ 을 만족한다는 의미이다. Probability vector \mathbf{p} 가 어느 시점에서 시스템 상태의 분포라 할 때, $\mathbf{p}^\top A$ 는 전이확률 A 에 따라 한번 전이된 다음 시점에서 시스템의 상태 분포를 의미한다. 따라서 어떤 probability vector \mathbf{p} 가 $\mathbf{p}^\top A = \mathbf{p}^\top$ 을 만족하는 경우, 시스템 상태가 \mathbf{p} 분포를 따르면 시스템이 A 에 따라 전이하더라도 시스템 상태 분포가 변하지 않음을 의미한다. 이러한 probability vector를 stationary distribution 혹은 equilibrium distribution이라 부른다.

다음은 모든 원소가 양수인 행렬에 대한 일반적인 결과로 Perron-Frobenius theorem으로 불리는 여러 결과의 일부이다.

Theorem 10.6 (Perron-Frobenius) *Let $A = (a_{ij})$ be an $n \times n$ positive matrix, i.e., $a_{ij} > 0$ for all i and j . Then, there exists a positive eigenvalue of the spectral radius of A , and its associated eigenvector with at least one positive element is unique up to scaling by a positive constant and has only positive components.*

Proof: $\lambda_0 = \max\{\lambda : \text{there exists } \mathbf{x} \neq \mathbf{0}, \mathbf{x} \geq \mathbf{0} \text{ such that } A\mathbf{x} \geq \lambda\mathbf{x}\}$ 을 생각해 보자. $A > \mathbf{0}$ 이므로 $A\mathbf{1} > \mathbf{0}$ 이고 따라서 vector $A\mathbf{1}$ 의 제일 작은 원소를 λ' 이라 놓으면 $A\mathbf{1} \geq \lambda'\mathbf{1}$ 를 만족하므로 $\lambda_0 \geq \lambda' > 0$ 이다. sup 대신에 max로 해도 괜찮은 이유는 각자 생각해 보자. λ_0 에 대해서 $A\mathbf{x}_0 \geq \lambda_0\mathbf{x}_0$ 인 vector \mathbf{x}_0 를 생각하자. 만약 $A\mathbf{x}_0 \neq \lambda_0\mathbf{x}_0$ 라면 $\mathbf{y} = A\mathbf{x}_0 - \lambda_0\mathbf{x}_0 \geq \mathbf{0}$ 이면서 $\mathbf{y} \neq \mathbf{0}$ 이다. 다시 한번 $A > \mathbf{0}$ 를 이용하면 $A\mathbf{y} > \mathbf{0}$ 인데 $A\mathbf{y} = AA\mathbf{x}_0 - \lambda_0 A\mathbf{x}_0$ 이어서 $A(A\mathbf{x}_0) > \lambda_0(A\mathbf{x}_0)$ 를 얻는다. 그런데 $A\mathbf{x}_0$ 에 대해서는 λ_0 보다 약간 더 큰 λ 에 대해서 $A(A\mathbf{x}_0) \geq \lambda(A\mathbf{x}_0)$ 가 성립될 수 있어서 λ_0 의 선택과정에 모순된다. 즉 $A\mathbf{x}_0 = \lambda_0\mathbf{x}_0$ 이고 $\mathbf{x}_0 > \mathbf{0}$ 이다.

$\lambda \neq \lambda_0$ 에 대해서 (λ, \mathbf{y}) 가 eigenpair라 하자. $|y_j|$ 를 원소로 하는 vector를 $\hat{\mathbf{y}}$ 으로 표시하자. $A\mathbf{y} = \lambda\mathbf{y}$ 의 i 번째 행인 $\sum_{j=1}^n a_{ij}y_j = \lambda y_i$ 의 양변에 절댓값을 취하면

$$|\lambda|\hat{y}_i = |\lambda||y_i| = |\lambda y_i| = \left| \sum_{j=1}^n a_{ij}y_j \right| \leq \sum_{j=1}^n a_{ij}|y_j| = \sum_{j=1}^n a_{ij}\hat{y}_j$$

가 성립한다. 이를 vector식으로 표현하면 $A\hat{\mathbf{y}} \geq |\lambda|\hat{\mathbf{y}}$ 이므로 λ_0 의 정의때문에 $|\lambda| \leq \lambda_0$ 가 성립한다.

마지막으로 $A\mathbf{x}_1 = \lambda_0\mathbf{x}_1$ 이면서 \mathbf{x}_0 에 linearly independent한 vector $\mathbf{x}_1 > \mathbf{0}$ 를 생각하자. $\mathbf{w} = \alpha\mathbf{x}_0 + \mathbf{x}_1$ 은 λ_0 에 대한 eigenvector인데 α 를 적당한 음수로 잡으면 $\mathbf{w} \geq \mathbf{0}$ 이지만 $\mathbf{w} \not> \mathbf{0}$ 가 된다. 이 vector도 λ_0 에 대한 eigenvector인데 앞에서 positive vector임을 보였기 때문에 모순이다. 따라서 λ_0 와 짹이 되는 eigenvector는 유일하다. ■

만약 Markov matrix A 가 positive하면 Theorem 10.6에 의해 eigenvalue 1에 대한 positive left-eigenvector 가 존재하는데 이 vector를 normalize하면 위에서 언급한 stationary distribution이 된다. 원소들이 non-negative한 Markov matrix에 대해서는 Theorem 10.6를 적용할 수 없고 이후에 증명할 Theorem 10.8에 의해 stationary distribution을 갖는다는 것을 보일 수 있다.

10.6 Eigenvalue Adjustments and Applications

정방행렬 A 의 eigenpair (λ, \mathbf{v}) 에 대해 임의의 벡터 \mathbf{w} 와 eigenvector \mathbf{v} 로 만든 rank-one 행렬 \mathbf{vw}^H 를 원래 행렬에 더한 $A + \mathbf{vw}^H$ 는 원래 eigenvalue λ 만 $\lambda + \mathbf{w}^H \mathbf{v}$ 로 변하고 다른 eigenvalue는 변화가 없다는 결과가 다음 Bauer정리이다. 특정한 eigenvalue를 조정하고 싶을 때 유용한 결과로 Google의 초창기 search engine에 사용되었다.

Theorem 10.7 (Brauer) Let (λ, \mathbf{v}) be an eigenpair of $n \times n$ matrix A and let $\lambda, \lambda_2, \dots, \lambda_n$ be its eigenvalues. For any $\mathbf{w} \in \mathbb{C}^n$, the eigenvalues of $A + \mathbf{vw}^H$ are $\lambda + \mathbf{w}^H \mathbf{v}, \lambda_2, \dots, \lambda_n$, and $(\lambda + \mathbf{w}^H \mathbf{v}, \mathbf{v})$ is an eigenpair of $A + \mathbf{vw}^H$.

Proof: $\mathbf{u} = \frac{1}{|\mathbf{v}|}\mathbf{v}$ 인 unit vector로 놓는다. 그러면, (λ, \mathbf{u}) 는 A 의 eigenpair이다. Schur triangularization (Theorem 10.5)을 적용하면 $Q = [\mathbf{u} | Q_2]$ 인 unitary matrix가 존재하여 $Q^H A Q = U = \begin{bmatrix} \lambda & \mathbf{a}^H \\ \mathbf{0} & U_{n-1} \end{bmatrix}$ 가 성립한다. 여기서 U_{n-1} 은 $\lambda_2, \dots, \lambda_n$ 을 대각원소로 갖는 upper triangular matrix이다. 그런데

$$\begin{aligned} Q^H \mathbf{vw}^H Q &= (Q^H \mathbf{v})(\mathbf{w}^H Q) = \begin{bmatrix} \mathbf{u}^H \mathbf{v} \\ Q_2^H \mathbf{v} \end{bmatrix} [\mathbf{w}^H \mathbf{u} | \mathbf{w}^H Q_2] \\ &= \begin{bmatrix} |\mathbf{v}| \\ \mathbf{0} \end{bmatrix} [\mathbf{w}^H \mathbf{u} | \mathbf{w}^H Q_2] = \begin{bmatrix} \mathbf{w}^H \mathbf{v} & \mathbf{b}^H \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \end{aligned}$$

이다. 따라서

$$Q^H (A + \mathbf{vw}^H) Q = \begin{bmatrix} \lambda & \mathbf{a}^H \\ \mathbf{0} & U_{n-1} \end{bmatrix} + \begin{bmatrix} \mathbf{w}^H \mathbf{v} & \mathbf{b}^H \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \lambda + \mathbf{w}^H \mathbf{v} & (\mathbf{a} + \mathbf{b})^H \\ \mathbf{0} & U_{n-1} \end{bmatrix}$$

인데 마지막 행렬은 $\lambda + \mathbf{w}^H \mathbf{v}, \lambda_2, \dots, \lambda_n$ 이 eigenvalue이므로 similar한 $A + \mathbf{vw}^H$ 의 eigenvalue도 동일하다. 그리고 $(A + \mathbf{vw}^H)\mathbf{v} = A\mathbf{v} + \mathbf{vw}^H \mathbf{v} = (\lambda + \mathbf{w}^H \mathbf{v})\mathbf{v}$ 가 되어 증명이 완료된다. ■

Theorem 10.7를 적절한 \mathbf{w} 에 적용하면 다음과 같이 행렬의 eigenvalue를 조정할 수 있다.

Corollary 10.5 Let (λ, \mathbf{v}) be an eigenpair of $n \times n$ matrix A and let $\lambda, \lambda_2, \dots, \lambda_n$ be its eigenvalues. Let $\mathbf{w} \in \mathbb{C}^n$ be such that $\mathbf{w}^H \mathbf{v} = 1$ and let $\tau \in \mathbb{C}$. Then the eigenvalues of $A_\tau = \tau A + (1 - \tau)\lambda \mathbf{vw}^H$ are $\lambda, \tau\lambda_2, \dots, \tau\lambda_n$.

Proof: τA 의 eigenvalue는 $\tau\lambda_1, \tau\lambda_2, \dots, \tau\lambda_n$ 이다. Theorem 10.7를 적용하면 $A_\tau = \tau A + \mathbf{v}((1 - \tau)\lambda \mathbf{w})^H$ 의 eigenvalue는 $\tau\lambda + ((1 - \tau)\lambda \mathbf{w})^H \mathbf{v} = \tau\lambda + (1 - \tau)\lambda \mathbf{w}^H \mathbf{v} = \tau\lambda + (1 - \tau)\lambda = \lambda$ 와 $\tau\lambda_2, \dots, \tau\lambda_n$ 이다. ■

다음은 Theorem 10.6을 확장한 non-negative matrix에 대한 Perron-Frobenius theorem의 일부이다.

Theorem 10.8 (Perron-Frobenius) Any Markov matrix has a stationary distribution.

Proof: Markov matrix를 A 로 나타내자. Fact 10.7에 의해 eigenvalue $1, \lambda_2, \dots, \lambda_n$ 들이 $|\lambda_i| \leq 1$ 이다. 만약 $A > \mathbf{0}$ 이면 Theorem 10.6 때문에 stationary distribution이 유일하게 존재한다. A 에 0인 원소가 있는 경우를 생각하자. 이 경우에는 앞에서 얻어진 Perron-Frobenius 결과를 A 에 적용할 수 없다. 그래서, Theorem 10.7을 이용하여 행렬 A 를 positive matrix 바꾸어서 생각하려고 한다. $(1, 1)$ 은 A 의 eigenpair이므로 작은 $\epsilon > 0$ 에 대해서 $\mathbf{v} = \mathbf{1}$ 과 $\mathbf{w} = \epsilon\mathbf{1}$ 로 놓고 Theorem 10.7을 적용하면 positive matrix $A + \epsilon\mathbf{1}\mathbf{1}^\top$ 은 $1 + n\epsilon, \lambda_2, \dots, \lambda_n$ 을 eigenvalue로 갖는다. 특히 Theorem 10.6에 의해 eigenvalue $1 + n\epsilon$ 에 대응되는 $A + \epsilon\mathbf{1}\mathbf{1}^\top$ 의 unit left-eigenvector \mathbf{u}_ϵ 은 각 ϵ 에 대해 유일하게 존재하고 positive이다. 또한, $\{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| = 1\}$ 이 compact 하므로 이 sphere상의 sequence인 \mathbf{u}_ϵ 의 subsequence \mathbf{u}_{ϵ_k} 가 $k \rightarrow \infty$ 일 때 sphere상의 vector로 수렴하여야 하고, 이를 \mathbf{u} 로 표시하자. 그러면,

$$\mathbf{u}_{\epsilon_k}^\top (A + \epsilon_k \mathbf{1}\mathbf{1}^\top) = (1 + n\epsilon_k) \mathbf{u}_{\epsilon_k}^\top$$

에서 $k \rightarrow \infty$ 를 양 변에 취하면

$$\mathbf{u}^\top A = \mathbf{u}^\top$$

이고 $\mathbf{u} = \lim_{k \rightarrow \infty} \mathbf{u}_{\epsilon_k} \geq \mathbf{0}$ 이어서 증명이 완료된다. ■

Example 10.5 [Google Matrix] 인터넷 상의 webpage들 간의 연결성을 나타내는 커다란 Markov matrix A 를 가정하자. 이 행렬의 left-eigenvector로부터 webpage들의 ranking을 정하는 것이 Google search engine의 중요한 아이디어이다. 따라서 Markov matrix의 left-eigenvector를 구하는 것이 중요하다. 그런데 Markov matrix가 크면 보통 power iteration과 같은 iterative algorithm을 사용하는데 이 경우, 가장 큰 eigenvalue와 두 번째로 큰 eigenvalue가 차이가 충분하여야 빨리 수렴한다. 즉 계산 상의 편의를 위해서 절댓값이 1인 A 의 eigenvalue는 1 뿐이고 나머지 eigenvalue의 절댓값은 1보다 작아지도록 A 를 수정하되 Markov matrix는 유지하도록 해보자. 이를 위해 $\mathbf{v} = \mathbf{1}, \mathbf{w} = \frac{1}{n}\mathbf{1}, E = \mathbf{1}\mathbf{1}^\top, 0 < \tau < 1$ 로 놓자. $\mathbf{w}^\top \mathbf{v} = 1$ 이다. 행렬 A 에 rank-one matrix인 \mathbf{vw}^\top 을 τ 와 $1 - \tau$ 로 가중하여 합하자. 이 가중합 행렬

$$A_\tau = \tau A + (1 - \tau)\mathbf{vw}^\top = \tau A + \frac{1 - \tau}{n} E \quad (10.5)$$

의 eigenvalue들은 Corollary 10.5에 의해 $1, \tau\lambda_2, \dots, \tau\lambda_n$ 이 되어 1을 제외한 나머지 eigenvalue들은 τ 배로 작아졌다. 또한 $A_\tau \mathbf{1} = \tau A \mathbf{1} + (1 - \tau)\mathbf{1} \frac{1}{n} \mathbf{1}^\top \mathbf{1} = \tau \mathbf{1} + (1 - \tau)\mathbf{1} = \mathbf{1}$ 로 행의 합이 1로 유지되고 원소들이 0보다 크다는 성질도 바뀌지 않았다. (10.5)식의 행렬 A_τ 를 Google matrix라고 부르며 Google이 초기에 web page ranking에 활용하였다. Google 창업자들은 초기에 $\tau = 0.85$ 를 사용하였다. 연관된 power iteration에 관해서는 Section 9.7을 참조하기 바란다. ■

Chapter 11

Big Theorems in Linear Algebra

$p(x) = ax^n$ 같이 한 개의 항으로 이루어진 다항식의 변수 x 대신에 정사각행렬 A 를 대입한 aA^n 은 잘 정의되어 $p(A)$ 를 aA^n 으로 생각해도 아무런 문제가 없다. 일반적으로 $p(x) = z^x + c_{n-1}x^{n-1} + \dots + c_1x + c_0$ 와 같은 다항식의 변수 x 대신에 행렬 A 를 대입하는 것을 생각해 보자. 이 경우 마지막 상수항 c_0 을 c_0x^0 로 생각하여 $c_0A^0 = c_0I$ 로 계산하는 걸로 약속하자. 그러면 자연스럽게 $p(A) = A^n + c_{n-1}A^{n-1} + \dots + c_1A + c_0I$ 로 계산된다. 복소수 원소도 가능한 일반적인 행렬 A 의 characteristic polynomial을 $p_A(x)$ 로 표시하자. 행렬 Q 가 unitary이고 $A = QBQ^H$ 이면 $A^n = QB^nQ^H$ 이고 $p(A) = Qp(B)Q^H$ 가 되는 것을 쉽게 확인할 수 있다. 선형대수학에서 가장 추상적이면서 일반적인 결과를 서술하는 두 정리가 Cayley-Hamilton theorem과 Jordan normal form theorem이다. Cayley-Hamilton theorem은 임의의 행렬 A 에 대해서 $p_A(A) = \mathbf{0}$ 이라는 놀라운 결과를 기술한다. 그리고 모든 행렬은 대각행렬 또는 대각행렬과 매우 유사한 Jordan form과 similar하다는 결과도 증명할 수 있다.

11.1 The First Big Theorem: Cayley-Hamilton Theorem

본 절에서는 대다수 독자의 상상을 뛰어넘는 결과로서 정사각행렬 A 의 characteristic polynomial $p_A(x)$ 로 주어질 때 $p_A(A) = \mathbf{0}$ 이 항상 성립한다는 결과인 Cayley-Hamilton 정리를 보이고자 한다. Schur triangulation (Theorem 10.5)을 이용하여 Cayley-Hamilton theorem을 증명해보자.

Theorem 11.1 (Cayley-Hamilton) *Let*

$$p_A(x) = \det(A - xI) = (-1)^n x^n + c_{n-1}x^{n-1} + \dots + c_1x + c_0$$

be the characteristic polynomial of an $n \times n$ matrix A . Then

$$p_A(A) = (-1)^n A^n + c_{n-1}A^{n-1} + \dots + c_1A + c_0I_n = \mathbf{0}.$$

Proof: $\lambda_1, \lambda_2, \dots, \lambda_n$ 이 행렬 A 의 eigenvalue라고 하자. 그러면,

$$p_A(x) = (\lambda_1 - x)(\lambda_2 - x) \cdots (\lambda_n - x).$$

으로 표현할 수 있다. Schur triangularization (Theorem 10.5)에 의해서 적당한 unitary matrix Q 와 $\lambda_1, \lambda_2, \dots, \lambda_n$ 을 대각원소로 갖는 upper triangular matrix U 가 존재하여 $A = QUQ^H$ 가 된다. 그리고 $p_A(A) = Qp_A(U)Q^H$ 이므로 $p_A(U) = (-1)^n(U - \lambda_1 I)(U - \lambda_2 I) \cdots (U - \lambda_n I) = \mathbf{0}$ 임을 보이면 된다. 우선 $j = 1, 2, \dots, n$ 에 대해서

$$U_j = (U - \lambda_1 I)(U - \lambda_2 I) \cdots (U - \lambda_j I)$$

로 놓자. 그런데 $U - \lambda_i I$ 는 upper triangular이면서 $(U - \lambda_i I)_{ii} = 0$ 이다. 특히 $U_1 = U - \lambda_1 I$ 는 첫 번째 열이 영벡터이다. 곰곰이 생각해보면 $U_2 = (U - \lambda_1 I)(U - \lambda_2 I)$ 도 첫 번째와 두 번째 열이 모두 영벡터가 됨을 알 수 있다.

행렬 U_{j-1} 의 처음 $(j-1)$ 개 열이 모두 영벡터라고 가정하자. $U_j = U_{j-1}(U - \lambda_j I)$ 에서 U_j 의 j 번째 열은 U_{j-1} 의 각 행과 $U - \lambda_j I$ 의 j 번째 열의 내적으로 원소들이 결정된다. 그런데 $U - \lambda_j I$ 의 j 번째 열은 j 번째 원소가 0이어서 처음 $(j-1)$ 개만 0이 아닐 수 있다. 그리고 U_{j-1} 의 각 행은 처음 $(j-1)$ 개가 모두 0이므로 계산할 내적은 모두 0이 된다. 즉 U_j 의 j 번째 열은 영벡터이다. 마찬가지 논리로 U_j 의 $(j-1)$ 번째까지 열들 모두 영벡터가 된다. 따라서 U_j 의 j 번째까지 열은 모두 영벡터이고 $p_A(U) = (-1)^n U_n = \mathbf{0}$ 이다. ■

Cayley-Hamilton theorem은 이론 전개나 사고실험에 주로 활용된다. 이 정리를 실제 계산에 활용하려면 characteristic equation의 계수를 알아야 하는데 이 계수를 얻는 계산양이 많아 부담이 된다. 구체적 계산 측면에서 실용적이지는 않지만 재미있는 결과의 예로 역행렬의 다향식 표현을 생각해보자. 가역행렬 A 의 characteristic polynomial이 $(-1)^n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0$ 라고 가정하자. 그러면 Cayley-Hamilton 정리에 의해 $(-1)^n A^n + c_{n-1} A^{n-1} + \cdots + c_1 A + c_0 I_n = \mathbf{0}$ 이다. 양변에 A^{-1} 를 곱하고 정리하면

$$A^{-1} = -\frac{c_1}{c_0} I - \frac{c_2}{c_0} A - \cdots - \frac{c_{n-1}}{c_0} A^{n-2} - \frac{(-1)^n}{c_0} A^{n-1}.$$

를 얻는데 역행렬을 A 의 거듭제곱으로 나타낸다는 측면에서 매력적이나 c_i 들을 알아야 한다는 제약이 있다.

11.2 Decomposition of Nilpotency into Cyclic Subspaces

Nilpotency of Matrices

정사각행렬 A 가 자연수 r 에 대해서 $A^{r-1} \neq \mathbf{0}$ 이면서 $A^r = \mathbf{0}$ 이면 nilpotent of degree r 라고 한다. 좀 더 일반적으로 어떤 부분공간 \mathbb{W} 에 대해서 $A^r \mathbf{v} = \mathbf{0}$ for all $\mathbf{v} \in \mathbb{W}$ 이지만 어떤 벡터 $\mathbf{w} \in \mathbb{W}$ 에 대해서 $A^{r-1} \mathbf{w} \neq \mathbf{0}$ 이면 “ A is nilpotent of degree r on \mathbb{W} ”라고 말한다. 행렬 A 에 대해서 자연스러운 \mathbb{W} 의 예를 생각해보자.

n 차원 벡터공간 \mathbb{V} 와 $n \times n$ 행렬 A 에 대해서

$$\mathbb{V}_A = \{\mathbf{v} \in \mathbb{V} : A^j \mathbf{v} = \mathbf{0} \text{ for some } j\} \quad (11.1)$$

를 정의하자. $A^{j_1} \mathbf{v} = \mathbf{0}$ 이면 $j \geq j_1$ 에 대해서 $A^j \mathbf{v} = \mathbf{0}$ 이라는 사실로부터 $A^{j_1} \mathbf{v}_1 = \mathbf{0}$ 이고 $A^{j_2} \mathbf{v}_2 = \mathbf{0}$ 이면 $A^{\max\{j_1, j_2\}}(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{0}$ 이 된다. 이로부터 \mathbb{V}_A 가 \mathbb{V} 의 부분공간이라는 것을 알 수 있고 \mathbb{V}_A 도 유한차원 벡터공간이므로 벡터 $\mathbf{v} \in \mathbb{V}_A$ 들에 대해서 $A^j \mathbf{v} = \mathbf{0}$ 가 되는 최소의 j 값들이 bounded일 것이라는

추측을 할 수 있다. 즉 부분공간 \mathbb{V}_A 의 basis가 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 라 하면 $A^{j_i} \mathbf{v}_i = \mathbf{0}$ 이 되는 j_1, \dots, j_k 가 존재하고 $r \geq r_0 = \max\{j_1, \dots, j_k\}$ 으로 놓으면 모든 \mathbf{v}_i 에 대해서 $A^r \mathbf{v}_i = \mathbf{0}$ 이 된다. 따라서 임의의 \mathbb{V}_A 의 벡터 $\mathbf{v} = x_1 \mathbf{v}_1 + \dots + x_k \mathbf{v}_k$ 도 $A^r \mathbf{v} = x_1 A^r \mathbf{v}_1 + \dots + x_k A^r \mathbf{v}_k = \mathbf{0}$ 를 만족한다. 그러므로 $\mathbb{V}_A \subset \text{Null}(A^r)$ 이 성립한다. (11.1)에서 $\mathbb{V}_A \supset \text{Null}(A^r)$ 은 자명하여

$$\mathbb{V}_A = \text{Null}(A^r) \quad \text{for some } r$$

를 얻는다.

$\mathbf{v} \in \text{Null}(A^r) \cap \text{Col}(A^r)$ 를 생각하자. $\mathbf{v} \in \text{Col}(A^r)$ 이므로 적당한 $\mathbf{w} \in \mathbb{V}$ 에 대해서 $\mathbf{v} = A^r \mathbf{w}$ 가 되는 동시에 $\mathbf{v} \in \text{Null}(A^r)$ 이므로 $A^r \mathbf{v} = A^{2r} \mathbf{w} = \mathbf{0}$ 이다. 그러면 (11.1)에 의해서 $\mathbf{w} \in \mathbb{V}_A$ 이고 $\mathbf{v} = A^r \mathbf{w} = \mathbf{0}$ 가 되어 $\text{Null}(A^r) \cap \text{Col}(A^r) = \{\mathbf{0}\}$ 임을 알게 된다. 이로부터 \mathbb{V} 의 두 부분공간 $\text{Null}(A^r)$ 과 $\text{Col}(A^r)$ 은 서로 독립이 된다. 즉 $\mathbf{v}_1 \in \text{Null}(A^r)$ 이고 $\mathbf{v}_2 \in \text{Col}(A^r)$ 이면 $x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 = \mathbf{0}$ 일 때 $x_1 \mathbf{v}_1 = -x_2 \mathbf{v}_2 \in \text{Null}(A^r) \cap \text{Col}(A^r) = \{\mathbf{0}\}$ 이어서 $x_1 = x_2 = 0$ 이 된다. 그러면 두 부분공간의 기저들이 서로 독립이므로 $\dim(\text{Null}(A^r) + \text{Col}(A^r)) = \dim \text{Null}(A^r) + \dim \text{Col}(A^r)$ 이 되고 rank-nullity theorem (3.4)에 의해 $\dim(\text{Null}(A^r) + \text{Col}(A^r)) = n = \dim \mathbb{V}$ 를 얻는다. 따라서 임의의 정사각행렬 A 에 대해서 적당한 자연수 r 이 존재하여

$$\mathbb{V} = \text{Null}(A^r) \oplus \text{Col}(A^r) \tag{11.2}$$

가 성립한다. 추가적인 성질로 두 부분공간 $\text{Null}(A^r)$ 과 $\text{Col}(A^r)$ 은 A 에 대해 invariant하다. 즉 $\mathbf{v} \in \text{Null}(A^r)$ 이면 $A^r(A\mathbf{v}) = A^{r+1}\mathbf{v} = A(A^r\mathbf{v}) = \mathbf{0}$ 이어서 $A\mathbf{v} \in \text{Null}(A^r)$ 이고, $\mathbf{v} \in \text{Col}(A^r)$ 이면 적당한 $\mathbf{w} \in \mathbb{V}$ 에 대해서 $\mathbf{v} = A^r \mathbf{w}$ 인데 $A\mathbf{v} = A^r(A\mathbf{w})$ 이므로 $A\mathbf{v} \in \text{Col}(A^r)$ 이다. 또한 α 가 임의의 scalar일 때 αI 에 대해서 두 부분공간이 invariant하다는 것은 자명하다. 따라서 $A + \alpha I$ 에 대해서도 두 부분공간이 invariant하다. 이상의 결과를 정리하면 다음과 같다.

Theorem 11.2 Let A be an $n \times n$ matrix and \mathbb{V} be an n -dimensional vector space. Let

$$\mathbb{V}_A = \{\mathbf{v} \in \mathbb{V} : A^j \mathbf{v} = \mathbf{0} \quad \text{for some } j\}.$$

Then there exists r_0 such that for any $r \geq r_0$

$$\mathbb{V}_A = \text{Null}(A^r)$$

and

$$\mathbb{V} = \text{Null}(A^r) \oplus \text{Col}(A^r)$$

where $\text{Null}(A^r)$ and $\text{Col}(A^r)$ are invariant under $A + \alpha I$ for any scalar α .

다음으로 \mathbb{V}_A 의 선형독립인 벡터들을 찾는 방법을 생각해 보자. 어떤 벡터 $\mathbf{v}_1 \in \mathbb{V}_A$ 에 대해서 $\mathbf{v}_2 = A\mathbf{v}_1 \neq \mathbf{0}, \dots, \mathbf{v}_k = A\mathbf{v}_{k-1} = A^{k-1}\mathbf{v}_1 \neq \mathbf{0}$ 이면서 $A^k \mathbf{v}_1 = A\mathbf{v}_k = \mathbf{0}$ 이라 가정하자. 선형독립인지 확인하기 위해서 $x_1 \mathbf{v}_1 + \dots + x_k \mathbf{v}_k = \mathbf{0}$ 이라고 놓자. 양변에 A^{k-1} 을 곱하면 $i \geq 2$ 일 때 $A^{k-1}\mathbf{v}_i = A^{i-2}(A^k\mathbf{v}_1) = \mathbf{0}$ 이어서 $x_1 A^{k-1}\mathbf{v}_1 = x_1 \mathbf{v}_k = \mathbf{0}$ 으로부터 $x_1 = 0$ 을 얻는다. $x_2 \mathbf{v}_2 + \dots + x_k \mathbf{v}_k = \mathbf{0}$ 양변에 A^{k-2} 를 곱하면 비슷하게 $x_2 = 0$ 을 얻는다. 이 과정을 반복하면 $x_1 = \dots = x_k = 0$ 을 얻게되어 $\mathcal{B} = \{\mathbf{v}_1, A\mathbf{v}_1, \dots, A^{k-1}\mathbf{v}_1\}$ 이 선형독립이라는 것을 알 수 있다.

Lemma 11.1 Assume that a nonzero vector $\mathbf{v} \in \mathbb{V}_A$ satisfies $A\mathbf{v} \neq \mathbf{0}, \dots, A^{k-1}\mathbf{v} \neq \mathbf{0}$ and $A^k\mathbf{v} = \mathbf{0}$. Then, $\{\mathbf{v}, A\mathbf{v}, \dots, A^{k-1}\mathbf{v}\}$ is linearly independent.

Lemma 11.1의 basic vector들이 갖는 특별한 형태를 cyclic structure라고 부르는데 임의의 행렬을 Jordan form으로 분해하는데 핵심적인 역할을 한다. 또한 이 lemma는 임의의 행렬 A 의 nilpotent degree r 이 벡터 공간의 차원 n 보다 클 수 없다는 것을 말해주므로 Theorem 11.2에서 $r_0 \leq n$ 로 놓을 수 있다.

Direct Sum Decomposition of the Null Space of Nilpotents

Nilpotent of degree r on $\mathbb{W} = \text{Null } A^r$ 인 행렬 A 를 분석해보자. 일반적으로는 $\mathbb{W} \subset \text{Null } A^r$ 이다. $A^k\mathbf{v} = \mathbf{0}$ 이면 $A^{k+1}\mathbf{v} = \mathbf{0}$ 이므로 $\text{Null } A^k \subset \text{Null } A^{k+1}$ 이어서

$$\text{Null } A \subset \text{Null } A^2 \subset \cdots \subset \text{Null } A^{r-1} \subset \text{Null } A^r = \mathbb{W} \quad (11.3)$$

이 성립한다. 그리고 적어도 하나의 벡터 $\mathbf{w} \in \mathbb{W}$ 에 대해서 $A^{r-1}\mathbf{w} \neq \mathbf{0}$ 이므로 $A^{r-k}\mathbf{w} \in \text{Null } A^k \setminus \text{Null } A^{k-1}$ 이고 (11.3)의 포함관계는 strict하다. 행렬 A 의 nilpotency structure를 이용하여 부분공간 $\mathbb{W} = \text{Null } A^r$ 를 direct sum으로 분해하자. 먼저 linearly independent의 개념의 확장이 필요하다.

Definition 11.1 Vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent of a subspace \mathbb{W} if $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k \in \mathbb{W}$ implies $a_1 = \cdots = a_k = 0$.

Definition 11.1에서 $\mathbf{v}_1, \dots, \mathbf{v}_k \not\in$ linearly independent of \mathbb{W} 이면 $\mathbf{v}_1, \dots, \mathbf{v}_k \not\in$ linearly independent하다는 것을 알 수 있다. 첫 단계로 이 개념을 이용하여 linearly independent of $\text{Null } A^{k-1}$ 이면서 $\text{Null } A^k \setminus \text{Null } A^{k-1}$ 에 포함된 벡터를 최대로 얻어서 \mathbb{W} 의 기저를 구성할 수 있다.

Theorem 11.3 Let a square matrix A be nilpotent of degree r on $\text{Null } A^r$. Then there exists a number m and vectors $\mathbf{v}_1, \dots, \mathbf{v}_m \in \text{Null } A^r$ such that the non-zero vectors of the form $A^j\mathbf{v}_\ell$, for $j \geq 0$ and $1 \leq \ell \leq m$, are a basis for $\text{Null } A^r$. Any vectors linearly independent of $\text{Null } A^{r-1}$ can be included among $\mathbf{v}_1, \dots, \mathbf{v}_m$.

Proof: 행렬 A 가 nilpotent of degree r on $\text{Null } A^r$ 이라고 하자. 그러면 어떤 벡터 \mathbf{v} 가 $A^r\mathbf{v} = \mathbf{0}$ 과 $A^j\mathbf{v} \neq \mathbf{0}$ for $j < r$ 을 만족한다. 만약 새로운 벡터를 $\mathbf{w} = A\mathbf{v}$ 로 놓으면 $A^{r-1}\mathbf{w} = A^r\mathbf{v} = \mathbf{0}$ 이면서 $A^j\mathbf{w} = A^{j+1}\mathbf{v} \neq \mathbf{0}$ for $j < r - 1$ 이다. 따라서 A 는 nilpotent of degree $r - 1$ on $\text{Null } A^{r-1}$ 이기도 하자. 이 사실을 이용하면 수학적 귀납법으로 정리를 증명할 수 있다: 수학적 귀납법을 적용하자:

- $r = 1$ 인 경우에는 A 가 nilpotent of degree 1이므로 $\dim \text{Null } A \geq 1$ 이고 따라서 $\text{Null } A$ 의 기저를 찾으면 정리를 만족한다.
- $r - 1$ 이하의 degree에 대해서 nilpotent인 정사각행렬에 대해서 정리가 성립한다고 가정하자.
- A 가 nilpotent of degree r on $\text{Null } A^r$ 이라고 하자. Linearly independent of $\text{Null } A^{r-1}$ 인 $\text{Null } A^r$ 의 non-zero vector $\mathbf{w}_1, \dots, \mathbf{w}_k$ 를 생각하자. 이때 $\text{Null } A^r \setminus \text{Null } A^{r-1}$ 에 속하는 다른 벡터를 이 k 개의

벡터들에 추가하면 선형종속이 되도록 $\mathbf{w}_1, \dots, \mathbf{w}_k$ 를 최대로 선택한다.¹ 이 벡터들은 $\mathbf{w}_i \in \text{Null } A^r$ 이므로 $A\mathbf{w}_i \in \text{Null } A^{r-1}$ 이다. 만약 $a_1A\mathbf{w}_1 + \dots + a_kA\mathbf{w}_k = A(a_1\mathbf{w}_1 + \dots + a_k\mathbf{w}_k) \in \text{Null } A^{r-2}$ 이라면 $a_1\mathbf{w}_1 + \dots + a_k\mathbf{w}_k \in \text{Null } A^{r-1}$ 인데 Null A^{r-1} 와 선형독립이 되도록 선택했으므로 $a_1 = \dots = a_k = 0$ 이다. 따라서 $A\mathbf{w}_1, \dots, A\mathbf{w}_k$ 는 linearly independent of Null A^{r-2} 이다. A 가 nilpotent of degree $r-1$ on Null A^{r-1} 이므로 induction hypothesis에 따라서 $A\mathbf{w}_1, \dots, A\mathbf{w}_k$ 를 포함한 Null A^{r-1} 의 벡터들 $\mathbf{v}_1, \dots, \mathbf{v}_{m-k}$ 가 존재하여 $A^j\mathbf{v}_\ell$ 들이 Null A^{r-1} 의 기저가 된다. 이 Null A^{r-1} 의 기저에 $\mathbf{w}_1, \dots, \mathbf{w}_k$ 를 합치면 정리의 내용을 만족하는 Null A^r 의 기저를 얻는다.

따라서 모든 r 에 대해서 정리가 증명된다. ■

Theorem 11.3에서 얻게되는 기저를 자세하게 살펴보자. 각 \mathbf{v}_ℓ 에 대해서 $r_\ell \mid A^{r_\ell}\mathbf{v}_\ell = \mathbf{0} \circ$ 이고 $A^{r_\ell-1}\mathbf{v}_\ell \neq \mathbf{0}$ 라고 하자. $\{A^j\mathbf{v}_\ell : j \geq 0, 1 \leq \ell \leq m\} = \cup_{\ell=1}^m \{\mathbf{v}_\ell, \dots, A^{r_\ell-1}\mathbf{v}_\ell\}$ 이고 $\mathbb{V}_\ell = \text{span}\{\mathbf{v}_\ell, \dots, A^{r_\ell-1}\mathbf{v}_\ell\}$ 로 표시하면

$$\text{Null } A^r = \mathbb{V}_1 \oplus \dots \oplus \mathbb{V}_m, \quad r_\ell = \dim \mathbb{V}_\ell \quad (11.4)$$

으로 분해된다. 또한 $\mathbf{v} \in \mathbb{V}_\ell$ 이면 $A\mathbf{v} \in \mathbb{V}_\ell$ 이어서 \mathbb{V}_ℓ 들은 A 에 invariant하다. 그리고 $\mathbf{v}_1, \dots, \mathbf{v}_m$ 이 유일하지 않기 때문에 \mathbb{V}_ℓ 들이 유일하게 결정되지는 않는다 (Table 11.1 참조). 하지만 \mathbb{V}_ℓ 의 차원은 direct sum을 구성하는 \mathbb{V}_ℓ 들의 순서를 무시하면 유일하게 결정된다.

$k = 1, \dots, r$ 에 대해서

$$d_k = \dim \text{Null } A^k - \dim \text{Null } A^{k-1}$$

로 놓자. $k \leq r$ 일 때 $\text{Null } A^k \setminus \text{Null } A^{k-1}$ 가 공집합일 수는 없으므로 $d_k \geq 1$ 이다. Null A^r 의 기저 $\{A^j\mathbf{v}_\ell : j \geq 0, 1 \leq \ell \leq m\}$ 에 속한 basic vector 중 $\text{Null } A^k \setminus \text{Null } A^{k-1}$ 에 포함된 d_k 개의 $A^j\mathbf{v}_\ell$ 들은 $A^{k-1+j}\mathbf{v}_\ell$ 이 \mathbb{V}_ℓ 에 포함되므로 해당 $r_\ell \geq k$ 이어야 하고 \mathbb{V}_ℓ 들의 차원은 최소 k 이다. 따라서 (11.4) 분해에는 최소 k 차원 이상인 \mathbb{V}_ℓ 이 d_k 개 있음을 알 수 있다. 또한 $d_k \leq d_{k-1}$ 임을 알 수 있고 $d_{r+1} = 0$ 으로 놓으면 차원이 k 인 \mathbb{V}_ℓ 은 $d_k - d_{k+1}$ 개 존재한다. 행렬 A 에 대해서 d_k 는 일의적으로 결정되므로 분해 (11.4)를 구성하는 부분공간의 개수와 각 부분공간의 차원은 일의적으로 결정된다는 것을 알 수 있다. $\text{span}\{\mathbf{v}_\ell, \dots, A^{r_\ell-1}\mathbf{v}_\ell\}$ 에서 basic vector \mathbf{v}_ℓ 의 cyclic structure를 강조하여 부분공간 \mathbb{V}_ℓ 을 cyclic subspace라고 부르기도 한다. 이상의 내용을 다음과 같이 정리할 수 있다.

Theorem 11.4 Let a square matrix A be nilpotent of degree r on Null A^r . Then, there exists a unique numbers m and r_ℓ 's such that

$$\text{Null } A^r = \mathbb{V}_1 \oplus \dots \oplus \mathbb{V}_m$$

where $\mathbb{V}_\ell = \text{span}\{\mathbf{v}_\ell, \dots, A^{r_\ell-1}\mathbf{v}_\ell\}$. Furthermore, the number of summands \mathbb{V}_ℓ of dimension at least k is $\dim \text{Null } A^k - \dim \text{Null } A^{k-1}$. Therefore, the decomposition is unique up to the number of summands and their dimensions.

이해를 돋기 위해 nilpotent of degree 5인 행렬 A 의 예를 들어보면 Table 11.1과 같다. 각 열은 \mathbb{V}_ℓ 의 기저를 나타내고 아래 쪽 k 개의 행들을 합치면 Null A^k 의 기저를 나타낸다. 즉 k 가 1, 2, 3인 행들을 합치면 Null A^3 의 기저가 된다. 특히 $k = 1$ 인 행은 행렬 A 의 eigenvalue 0에 대한 eigenvector를 나타낸다.

¹구체적인 예로 Null A^{r-1} 의 기저로부터 Gram-Schmidt를 통해서 얻어진 Null A^r 의 벡터를 $\mathbf{w}_1, \dots, \mathbf{w}_k$ 로 사용해도 된다.

Table 11.1: Demonstration of a Decomposition of Nilpotent Matrix

k	\mathbb{V}_1	\mathbb{V}_2	\mathbb{V}_3	\mathbb{V}_4	\mathbb{V}_5	\mathbb{V}_6	\mathbb{V}_7	$\dim \text{Null } A^k$	d_k
5	\mathbf{v}_1	\mathbf{v}_2						25	2
4	$A\mathbf{v}_1$	$A\mathbf{v}_2$	\mathbf{v}_3	\mathbf{v}_4	\mathbf{v}_5			23	5
3	$A^2\mathbf{v}_1$	$A^2\mathbf{v}_2$	$A\mathbf{v}_3$	$A\mathbf{v}_4$	$A\mathbf{v}_5$			18	5
2	$A^3\mathbf{v}_1$	$A^3\mathbf{v}_2$	$A^2\mathbf{v}_3$	$A^2\mathbf{v}_4$	$A^2\mathbf{v}_5$	\mathbf{v}_6		13	6
1	$A^4\mathbf{v}_1$	$A^4\mathbf{v}_2$	$A^3\mathbf{v}_3$	$A^3\mathbf{v}_4$	$A^3\mathbf{v}_5$	$A\mathbf{v}_6$	\mathbf{v}_7	7	7

이 절에서 고민한 nilpotency structure는 eigenvalue가 λ 인 행렬 A 의 경우 변경된 행렬 $A - \lambda I$ 에 적용할 수 있다. 다음 절에서 $A - \lambda I$ 가 적당한 $r \leq n$ 에 대해서 nilpotent of degree r on Null $(A - \lambda I)^r$ 임을 보이고 Theorem 11.3과 11.4를 적용하여 Jordan block을 얻게된다.

11.3 Nilpotency of $A - \lambda I$

Generalized Eigenvectors

$n \times n$ 행렬 A 의 한 eigenvalue λ 에 대해서 Null $(A - \lambda I)$ 에 속한 벡터들ㅇ이 eigenvector 들이고 이 null space의 차원을 geometric multiplicity라고 부른다. 이 geometric multiplicity가 characteristic equation의 근으로서 λ 의 중근 정도인 algebraic multiplicity보다 작으면 대각화 과정에서 문제가 생긴다. 이러한 경우에 $A - \lambda I$ 를 더 분석해보기 위해 먼저 Theorem 11.2을 행렬 $A - \lambda I$ 에 적용하여 다음의 corollary를 얻는다.

Corollary 11.1 *Let A be an $n \times n$ complex matrix with an eigenvalue λ . Then, there exists r_λ such that for any $r \geq r_\lambda$*

$$\{\mathbf{v} \in \mathbb{V} : (A - \lambda I)^j \mathbf{v} = \mathbf{0} \text{ for some } j\} = \text{Null } (A - \lambda I)^r$$

and

$$\mathbb{C}^n = \text{Null } (A - \lambda I)^r \oplus \text{Col } (A - \lambda I)^r,$$

where Null $(A - \lambda I)^r$ and Col $(A - \lambda I)^r$ are invariant under $A - cI$ for any scalar c .

벡터공간 Null $(A - \lambda I)^{r_\lambda}$ 의 기저벡터들이 최대 n 개이어서 $r_\lambda \leq n$ 으로 잡을 수 있어서 Null $(A - \lambda I)^{r_\lambda} = \text{Null } (A - \lambda I)^n$ 이므로 이후의 전개에서는 Null $(A - \lambda I)^n$ 를 통해서 $A - \lambda I$ 의 nilpotent structure를 살펴본다. Geometric multiplicity가 algebraic multiplicity보다 작은 경우에 수학자들이 찾아낸 우회로가 Null $(A - \lambda I)^n$ 에 속한 벡터들을 **generalized eigenvector**로 정의하고 이 벡터들로 이루어진 부분공간들로 \mathbb{C}^n 을 분해하여 행렬의 표준형태를 얻는 것이다.

직관적으로 당연해 보이는 결과 한가지가 Null $(A - \lambda I)^n$ 에 속한 벡터들에 대한 행렬 A 의 eigenvalue는 λ 가 유일하다는 것이다. $\mathbf{v} \in \text{Null } (A - \lambda I)^n$ 일 때 (μ, \mathbf{v}) 가 A 의 eigenpair라 하자. 그러면 $(A - \lambda I)\mathbf{v} = A\mathbf{v} - \lambda\mathbf{v} = (\mu - \lambda)\mathbf{v}$ 이어서 $\mathbf{0} = (A - \lambda I)^n \mathbf{v} = (\mu - \lambda)^n \mathbf{v}$ 이므로 $\mu = \lambda$ 이다. 게다가 서로 다른 eigenvalue에

대응되는 eigenvector들이 선형독립이었던 것처럼 서로 다른 eigenvalue에 대응되는 generalized eigenvector들도 선형독립이다. 즉 서로 다른 두 eigenvalue $\lambda_1 \neq \lambda_2$ 에 대해서

$$\text{Null}(A - \lambda_1 I)^n \cap \text{Null}(A - \lambda_2 I)^n = \{\mathbf{0}\} \quad (11.5)$$

이 성립한다. 이유를 살펴보기 위해서 $\mathbf{0} \neq \mathbf{v} \in \text{Null}(A - \lambda_1 I)^n \cap \text{Null}(A - \lambda_2 I)^n$ 인 벡터를 생각하자. $\mathbf{v} \in \text{Null}(A - \lambda_2 I)^n$ 이므로 적당한 $k \geq 0$ 에 대해서 $(A - \lambda_2 I)^{k+1}\mathbf{v} = \mathbf{0}$ 와 $(A - \lambda_2 I)^k\mathbf{v} \neq \mathbf{0}$ 가 동시에 성립한다. 새로운 벡터를 $\mathbf{w} = (A - \lambda_2 I)^k\mathbf{v} \neq \mathbf{0}$ 로 놓으면 $A\mathbf{w} = \lambda_2\mathbf{w}$ 이 성립한다. 그런데 $A - \lambda_1 I$ 와 $A - \lambda_2 I$ 의 행렬곱이 교환되므로

$$(A - \lambda_1 I)^n \mathbf{w} = (A - \lambda_1 I)^n (A - \lambda_2 I)^k \mathbf{v} = (A - \lambda_2 I)^k (A - \lambda_1 I)^n \mathbf{v} = \mathbf{0}$$

이 성립하여 $\mathbf{w} \in \text{Null}(A - \lambda_1 I)^n$ 이다. 즉 (λ_2, \mathbf{w}) 가 A 의 $\text{Null}(A - \lambda_1 I)^n$ 상에서의 eigenpair가 되는데 앞에서 살펴본 generalized eigenvector들에 대한 eigenvalue의 유일성에 모순되어 (11.5)이 성립한다.

임의의 $n \times n$ 복소행렬 A 는 중근을 포함하여 n 개의 eigenvalue를 갖고 Theorem 10.5에 의해 대각원소가 eigenvalue들인 upper triangular matrix와 similar하다. 즉 upper triangular matrix U 와 unitary matrix Q 가 존재하여 $A = Q U Q^H$ 이다. 그리고 서로 다른 eigenvalue 값이 $\lambda_1, \dots, \lambda_m$ 이고 같은 eigenvalue들은 U 의 대각원소로서 인접하다고 가정하자. λ_i 의 algebraic multiplicity가 k_i 라면 U 의 i 번째 diagonal block의 k_i 개 대각원소들이 λ_i 이다. 따라서

$$A - \lambda_i I = Q U Q^H - \lambda_i Q Q^H = Q(U - \lambda_i I)Q^H$$

이고 $U - \lambda_i I$ 의 i 번째 $k_i \times k_i$ diagonal block은 대각원소가 0인 upper triangular matrix가 된다. 임의의 k 에 대해서 $(A - \lambda_i I)^k = Q(U - \lambda_i I)^k Q^H$ 인데 Fact 2.1에 의해 $(U - \lambda_i I)^{k_i}$ 의 i 번째 block인 $k_i \times k_i$ 행렬은 영행렬이고 나머지 대각원소들은 $(\lambda_j - \lambda_i)^{k_i} \neq 0$ 이므로 $(U - \lambda_i I)^{k_i}$ 의 rank는 $n - k_i$ 이고 Q 가 가역이므로 $(A - \lambda_i I)^{k_i}$ 의 rank도 동일하게 $n - k_i$ 이다. 따라서 $\dim \text{Null}(A - \lambda_i I)^{k_i} = \dim \text{Null}(U - \lambda_i I)^{k_i} = k_i$ 이다. k_i 들은 n 차 방정식인 characteristic equation의 중근의 multiplicity이고 $\text{Null}(A - \lambda_i I)^{k_i} \subset \text{Null}(A - \lambda_i I)^n$ 이므로

$$\begin{aligned} \dim(\mathbb{C}^n) &= n = k_1 + \cdots + k_m \\ &= \dim \text{Null}(A - \lambda_1 I)^{k_1} + \cdots + \dim \text{Null}(A - \lambda_m I)^{k_m} \\ &\leq \dim \text{Null}(A - \lambda_1 I)^n + \cdots + \dim \text{Null}(A - \lambda_m I)^n \end{aligned}$$

이 성립한다. 동시에 null space들이 부분공간이므로 당연한 포함관계인

$$\mathbb{C}^n \supset \text{Null}(A - \lambda_1 I)^n + \cdots + \text{Null}(A - \lambda_m I)^n$$

와 (11.5)를 합치면

$$\mathbb{C}^n \supset \text{Null}(A - \lambda_1 I)^n \oplus \cdots \oplus \text{Null}(A - \lambda_m I)^n$$

를 얻고 이 포함관계의 차원을 생각하면 $n \geq \dim \text{Null}(A - \lambda_1 I)^n + \cdots + \dim \text{Null}(A - \lambda_m I)^n$ 를 얻는다. 따라서 직전의 결과와 합치면 $k_i \leq \dim \text{Null}(A - \lambda_i I)^n$ 이면서

$$n = k_1 + \cdots + k_m \leq \dim \text{Null}(A - \lambda_1 I)^n + \cdots + \dim \text{Null}(A - \lambda_m I)^n \leq n$$

이 성립하여 $k_i = \dim \text{Null}(A - \lambda_i I)^n$ 이고

$$\mathbb{C}^n = \text{Null}(A - \lambda_1 I)^n \oplus \cdots \oplus \text{Null}(A - \lambda_m I)^n$$

으로 분해된다. 이를 정리하면 다음과 같다.

Theorem 11.5 Let A be an $n \times n$ matrix and $\lambda_1, \dots, \lambda_m$ be its eigenvalues. Then

$$\mathbb{C}^n = \text{Null}(A - \lambda_1 I)^n \oplus \cdots \oplus \text{Null}(A - \lambda_m I)^n. \quad (11.6)$$

특히 $\text{Null}(A - \lambda_i I)^n$ 부분공간 각각의 기저를 얻을 수 있으면 그 부분공간의 기저들을 합한 것이 \mathbb{C}^n 의 기저가 된다.

11.4 The Second Big Theorem: Jordan Normal Form Theorem

$n \times n$ 복소행렬 A 가 algebraic multiplicity가 k_i 인 eigenvalue λ_i 를 갖는다고 가정하자. Corollary 11.1에 의해 $A - \lambda_i I$ 가 nilpotent on $\text{Null}(A - \lambda_i I)^n$ 이라는 것을 알 수 있다. 따라서 Theorem 11.3과 11.4를 nilpotent $A - \lambda_i I$ 에 적용할 수 있다. Theorem 11.3에 의해 $\{\mathbf{v}_\ell, \dots, (A - \lambda_i I)^{r_\ell-1}\mathbf{v}_\ell\}$ 의 기저를 갖는 부분공간들의 direct sum으로 $\text{Null}(A - \lambda_i I)^n$ 이 분해된다. 특히 $\mathbf{w}_k = (A - \lambda_i I)^{k-1}\mathbf{v}_\ell$ 로 표시하면 기저가 $\{\mathbf{w}_1, \dots, \mathbf{w}_{r_\ell}\}$ 이고 기저들 간에

$$\mathbf{w}_{k+1} = (A - \lambda_i I)\mathbf{w}_k \quad \text{for } k = 1, \dots, r_\ell - 1, \quad (A - \lambda_i I)\mathbf{w}_{r_\ell} = \mathbf{0}$$

의 관계가 있다. 이를 다시 쓰면

$$A\mathbf{w}_k = \lambda_i \mathbf{w}_k + \mathbf{w}_{k+1} \quad \text{for } k = 1, \dots, r_\ell - 1, \quad A\mathbf{w}_{r_\ell} = \lambda_i \mathbf{w}_{r_\ell}$$

이다. Jordan block이라고 불리는 $r_\ell \times r_\ell$ 행렬 J_ℓ 를 다음과 같이 정의하자:

$$J_\ell = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & & \lambda_i & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{bmatrix} = \lambda_i I + [\mathbf{0} | \mathbf{e}_1 | \cdots | \mathbf{e}_{r_\ell-1}].$$

그러면 $W_\ell = [\mathbf{w}_{r_\ell} | \cdots | \mathbf{w}_1]$ 일 때

$$AW_\ell = W_\ell J_\ell \quad (11.7)$$

이 성립한다.² 즉 $\{\mathbf{w}_1, \dots, \mathbf{w}_{r_\ell}\}$ 를 기저로 생성된 부분공간 상에서는 행렬 A 로 정의되는 linear transformation이 Jordan block J_ℓ 로 표현된다는 것이다. 각 \mathbf{v}_ℓ 에 대해서 Jordan block J_ℓ 이 존재하여 (11.7)를 만족하므로 $W_{\lambda_i} = [W_1 | \cdots | W_{r_\ell}]$ 와 $J_{\lambda_i} = \text{diag}(J_1, \dots, J_{r_\ell})$ 에 대해서

$$AW_{\lambda_i} = W_{\lambda_i} J_{\lambda_i} \quad (11.8)$$

² W 의 열벡터 index가 역순임에 주의하자. 관습적으로 Jordan block J 는 superdiagonal이 1인 upper triangular 행렬로 정의한다.

가 성립하는 것을 쉽게 확인할 수 있다. W_{λ_i} 는 $n \times k_i$ 행렬이고 J_{λ_i} 는 $k_i \times k_i$ 행렬이다. Theorem 11.4에 의해서 행렬 J_{λ_i} 를 구성하는 Jordan block들의 크기와 개수는 유일하게 결정된다. 행렬 A 의 eigenvalue λ_i 들 각각에 대해서 W_{λ_i} 와 J_{λ_i} 가 존재하여 (11.8)을 만족하므로 $W = [W_{\lambda_1} | \cdots | W_{\lambda_m}]$ 와 $J = \text{diag}(J_{\lambda_1}, \dots, J_{\lambda_m})$ 으로 정의하면

$$AW = WJ \quad (11.9)$$

를 만족한다. 이 J 를 Jordan (normal) form이라고 부른다. 특히 basic vector를 열로 갖는 W 가 가역이므로 $J = W^{-1}AW$ 이어서 A 와 J 는 similar하다. 이 결과를 요약하면 다음과 같다.

Theorem 11.6 (Jordan Normal Form Theorem) *Any $n \times n$ matrix is similar to a Jordan normal form. The Jordan form is unique up to the number and size of Jordan blocks.*

Jordan form J 는 대각행렬은 아니지만 대각행렬과 상당히 유사하게 생겼다. k 가 열의 개수보다 크면 J^k 은 일반적으로 upper triangular 행렬이다. 하지만 Jordan block의 특별한 구조를 활용하여 J^k 를 A^k 보다는 효율적으로 계산할 수 있다. 따라서 대각화가 불가능한 일반적인 행렬들에 대해서 Section 9.7의 대각행렬을 이용한 power iteration의 효과적인 계산 방법들에 Jordan block을 대치할 수 있다. 이러한 방법으로 대각행렬보다는 복잡하고 느리지만 원래 행렬보다는 효과적으로 여러 계산을 수행할 수 있다.

Kang & Cho (2024)

Chapter 12

Homework Assignments

Chapter 2

1. Consider the following simultaneous linear equations in 3 unknowns, $\mathbf{x} = (u, v, w)^\top$:

$$\begin{cases} u - 4v + 7w = -9 \\ 2u - 6v + 9w = -10 \\ u - 2v + 5w = -7 \end{cases}$$

- (a) Convert the equations in matrix-vector form as $A\mathbf{x} = \mathbf{b}$. What are A and \mathbf{b} ?
- (b) Decompose A in the form of LDU where L and U are lower and upper triangular with unit diagonals, and D is a diagonal matrix. What are L , D , and U ?
- (c) Find \mathbf{x} by solving $DU\mathbf{x} = L^{-1}\mathbf{b}$.
- (d) Compute A^{-1} .
2. Repeat 1 with randomly generated 10×10 matrix A and $\mathbf{b} \in \mathbb{R}^{10}$.
3. Prove that matrix multiplication is associative $(AB)C = A(BC)$ and distributive $A(B + C) = AB + AC$, $(B + C)D = BD + CD$.
4. Find an example showing that matrix multiplication is not commutative, that is, find two matrices A and B such that $AB \neq BA$.
5. Show that $AI_n = A = I_mA$ for any $m \times n$ matrix A .
6. Show that $A^\top DA$ is symmetric where A is $m \times n$ and D is an $m \times m$ diagonal matrix.
7. Show the following multiplication rule for two block matrices:

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

8. Show that the product of two lower triangular matrices is also lower triangular.

9. Compute A^2 , A^3 , and A^4 where

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

Chapter 3

1. Show that $\{\mathbf{0}\}$, $\{c\mathbf{x} : c \in \mathbb{R}\}$ for $\mathbf{x} \in \mathbb{V}$, and $\{c_1\mathbf{x}_1 + \dots + c_n\mathbf{x}_n : c_1, \dots, c_n \in \mathbb{R}\}$ for $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{V}$ are subspaces of a vector space \mathbb{V} .

2. Show that $\{(x, y) : x \geq 0, y \geq 0\}$ is not a subspace of \mathbb{R}^2 .

3. Show that the set of all $n \times n$ lower triangular matrices is a vector space and a subspace of the vector space consisting of all $n \times n$ matrices. Do the same work for the set of all $n \times n$ symmetric matrices.

4. Let U be the row echelon form of A . Show that the vectors, in the null space of U , obtained by setting 1 for a single free variable and other free variables as zeros constitute a basis for the null space of A .

5. Consider the following 4×5 matrix A :

$$A = \begin{bmatrix} 2 & 2 & 1 & -6 & 4 \\ 4 & 4 & 1 & 10 & 13 \\ 6 & 6 & 0 & 20 & 19 \\ 8 & 8 & 1 & 14 & 23 \end{bmatrix}.$$

(a) Compute the row echelon form and reduced row echelon form of A .

(b) What is the rank of A ?

(c) Characterize the null space of A .

(d) What is the dimension of $\text{Null}(A)$?

(e) Find a maximally independent set of column vectors of A .

(f) What is the dimension of $\text{Col}(A)$?

6. Let a set \mathbb{V} include all polynomials of degree n or less, $\mathbb{V} = \{a_0 + a_1t + \dots + a_nt^n : a_0, a_1, \dots, a_n \in \mathbb{R}\}$. Let T be a transform that maps a polynomial $f(t)$ to the polynomial $\int_0^t f(s)ds$. Note that the n is a fixed integer in this question.

(a) Show that \mathbb{V} is a vector space over the multiplication scalar field \mathbb{R} .

- (b) What is the dimension of \mathbb{V} ?
- (c) To make T be a map from \mathbb{V} into \mathbb{W} , what should be \mathbb{W} if \mathbb{W} is a vector space?
- (d) Is T a linear map?
- (e) Find bases $\mathcal{B}_{\mathbb{V}}$ and $\mathcal{B}_{\mathbb{W}}$ of \mathbb{V} and \mathbb{W} .
- (f) Find the transform matrix A representing T with respect to the bases $\mathcal{B}_{\mathbb{V}}$ and $\mathcal{B}_{\mathbb{W}}$ chosen above.
7. In a finite dimensional vector space, show that any linearly independent set of vectors can be extended to a basis.
8. Let \mathbb{V} be a vector space of dimension n . \mathbb{W}_1 and \mathbb{W}_2 are two subspaces of \mathbb{V} . Suppose that $\dim \mathbb{W}_1 = n_1$, $\dim \mathbb{W}_2 = n_2$, and $n_1 + n_2 > n$. Show that $\dim(\mathbb{W}_1 \cap \mathbb{W}_2) \geq n_1 + n_2 - n$.

Chapter 4

1. Let a vector space \mathbb{V} include all polynomials of degrees 2 or less, $\mathbb{V} = \{a_0 + a_1 t + a_2 t^2 : a_0, a_1, a_2 \in \mathbb{R}\}$. For two polynomials $f(t)$ and $g(t)$ in \mathbb{V} , we define an inner product
- $$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$
- (a) Compute $|f|$, $|g|$, and $|f - g|$ for $f(t) = t$ and $g(t) = t^2$.
- (b) Show that $\{1, t, t^2\}$ is linearly independent.
- (c) For $\mathbb{W} = \text{span}\{1, t^2\}$, compute $\mathbf{P}_{\mathbb{W}}(t)$.
- (d) Show that $\mathcal{B} = \{1, t, t^2\}$ is a basis for \mathbb{V} .
- (e) Find a matrix representation of the inner product $\langle f, g \rangle$ with respect to the basis \mathcal{B} .
2. Consider the Euclidean vector space \mathbb{R}^4 . For two vectors $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$ and $\mathbf{y} = (y_1, y_2, y_3, y_4)^T$ in \mathbb{R}^4 , we define a *non-standard* inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = 2x_1y_1 + \frac{2}{3}(x_1y_3 + x_2y_2 + x_3y_1) + \frac{2}{5}(x_2y_4 + x_3y_3 + x_4y_2) + \frac{2}{7}x_4y_4,$$

or, in matrix form as

$$\langle \mathbf{x}, \mathbf{y} \rangle = [x_1, x_2, x_3, x_4] \begin{bmatrix} 2 & 0 & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{2}{5} \\ \frac{2}{3} & 0 & \frac{2}{5} & 0 \\ 0 & \frac{2}{5} & 0 & \frac{2}{7} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}.$$

- (a) Compute $|\mathbf{x}|$, $|\mathbf{y}|$, and $|\mathbf{x} - \mathbf{y}|$ for $\mathbf{x} = (1, 0, 0, 0)^T$ and $\mathbf{y} = (0, 0, 1, 0)^T$.
- (b) Find an orthonormal basis of $\text{span}\{(1, 0, 0, 0)^T, (0, 0, 1, 0)^T\}$.

- (c) Let $\mathbb{W} = \text{span}\{(1, 0, 0, 0)^\top, (0, 0, 1, 0)^\top\}$. Find an orthonormal basis of \mathbb{W}^\perp .
3. Let a vector space \mathbb{V} include all polynomials of degree 3 or less, $\mathbb{V} = \{a_0 + a_1t + a_2t^2 + a_3t^3 : a_0, a_1, a_2, a_3 \in \mathbb{R}\}$. For two polynomials $f(t), g(t) \in \mathbb{V}$, we define an inner product
- $$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$
- (a) Compute $|f|$, $|g|$, and $|f - g|$ for $f(t) = 1$ and $g(t) = t^2$.
- (b) Find an orthonormal basis of $\text{span}\{1, t^2\}$.
- (c) Let $\mathbb{W} = \text{span}\{1, t^2\}$. Find an orthonormal basis of \mathbb{W}^\perp .
- (d) Find the matrix representation of the inner product $\langle f, g \rangle$ with respect to the basis $\{1, t, t^2, t^3\}$ of \mathbb{V} .
4. Compare 2 and 3.

Chapter 5

1. Find singular values and singular vectors of the following **rotation matrix**:

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

2. Let a 4×4 matrix A is given as

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

- (a) Find singular values and right-/left-singular vectors of A .
- (b) Find $\|A\|_2$ and $\|A\|_F$.
- (c) Find the pseudoinverse A^+ of A .
- (d) Find a 4×4 matrix B of rank 2 that minimizes $\|A - B\|_2$.
- (e) Can you find eigenvalues and eigenvectors of A by hand calculation?

3. Let a 4×4 matrix A is given as

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

- (a) Find singular values and right-/left-singular vectors of A .
- (b) Can you find eigenvalues and eigenvectors of A by hand calculation?
- (c) Find rank A .

4. Suppose A is a 3×4 matrix given as $A = \begin{bmatrix} 2 & 1 & -1 & 1 \\ 6 & 3 & -3 & 3 \\ -2 & -1 & 1 & -1 \end{bmatrix}$.

- (a) Find eigenvalues and eigenvectors of $A^\top A$.
- (b) Find eigenvalues and eigenvectors of AA^\top .
- (c) Is there a 3×4 matrix B of rank 2 that minimizes $\|A - B\|_2$?

Chapter 6

1. Let A and B be two $n \times n$ positive definite matrices. Consider the following sum of two quadratic forms

$$f(\mathbf{x}) = (\mathbf{x} - \mathbf{a})^\top A(\mathbf{x} - \mathbf{a}) + (\mathbf{x} - \mathbf{b})^\top B(\mathbf{x} - \mathbf{b})$$

where $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. Re-formulate the sum as a single quadratic form: that is, find C , \mathbf{d} , and r such that

$$f(\mathbf{x}) = (\mathbf{x} - \mathbf{d})^\top C(\mathbf{x} - \mathbf{d}) + r.$$

2. Let A be an $n \times n$ symmetric positive definite matrix. Denote its eigenvalues as $\lambda_1, \dots, \lambda_n$. Let B be the square root of A , that is, $A = B^2$.

- (a) Let $A = U^\top U$ be the Cholesky factorization of A . Find $|\det U|$.
 - (b) Find the eigenvalues of B .
 - (c) Let $B = QR$ be the QR -decomposition of B . Find $|\det R|$.
 - (d) Does there exist the square root C of B ? If yes, what are the eigenvalues of C ?
3. Let A be an $n \times n$ symmetric positive definite matrix whose spectral decomposition is described as $A = V\Lambda V^\top$ where V is an orthogonal matrix and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ for $\lambda_i > 0$. We also define $\Lambda^{1/k} = \text{diag}(\lambda_1^{1/k}, \dots, \lambda_n^{1/k})$.
- (a) Characterize a symmetric positive definite matrix B_2 satisfying $A = B_2^2$ in terms of V and Λ .
 - (b) Characterize a symmetric positive definite matrix B_k satisfying $A = B_k^k$ in terms of V and Λ for every positive integer k .
 - (c) What would be $\lim_{k \rightarrow \infty} B_k$? Validate your answer as reasonably as possible.

Chapter 7

1. Compute the determinants of A , U , U^\top , U^{-1} , and M where

$$A = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix} \begin{bmatrix} 3 & -2 & 2 \end{bmatrix}, \quad U = \begin{bmatrix} 4 & 4 & 8 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix}, \quad M = \begin{bmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 1 & 2 & 2 \\ 4 & 4 & 8 & 7 \end{bmatrix}.$$

2. Let A be an $n \times n$ tridiagonal matrix of

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

Find $\det A$.

3. Find the determinant of the following **Vandermonde matrix**:

$$V_3 = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}.$$

4. Find the determinant of the following **rotation matrix**:

$$\begin{bmatrix} \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \\ \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \\ -\sin \phi & \cos \phi & 0 \end{bmatrix}.$$

Chapter 8

1. Show that the set of eigenvectors associated with a single eigenvalue is a subspace without the origin.

2. Find all eigenvalues of A , B , U , U^{-1} , and T where

$$A = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 2 \end{bmatrix}, \quad U = \begin{bmatrix} 4 & 4 & 8 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}, \quad T = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}.$$

3. Find all eigenvectors of A in 2.

4. Find all eigenvalues and eigenvectors of the following **rotation matrix**:

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

Chapter 10

1. Show that \mathbb{V}_λ in (11.1) is a subspace of \mathbb{V} .

Kang & Cho (2024)

Chapter 13

Problems

13.1 Problems for Chapter 1 ~ 4

Problem Set 1

1. Suppose A is a 3×4 matrix and \mathbf{b} is a vector in \mathbb{R}^3 .
 - (a) Give an example of the matrix A with rank 2. No component of A is allowed to be zero and PA admits an LU decomposition with $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ whereas A itself does not allow LU decomposition.
 - (b) For the matrix A in (a), give two examples of \mathbf{b} , one of which allows a solution of $A\mathbf{x} = \mathbf{b}$ and the other one doesn't.
 - (c) Find two vectors in $\text{Col}(A)$ closest to the two vectors in (b), respectively.
2. Suppose A and B are $n \times m$ and $m \times n$ matrices, respectively.
 - (a) Recalling that each column of AB is a linear combination of the columns of A , prove that $\text{rank}(AB) \leq \text{rank}(A)$. Can we conclude $\text{rank}(AB) \leq \text{rank}(B)$ by straightforwardly applying the above statement? If yes, why?
 - (b) Assume $AB = I_n$ where I_n is the identity matrix. What is the rank of A ? If $m = n$, what is BA ?
 - (c) Assume $AB = I_n$ where I is the identity matrix and $m > n$. Does $A\mathbf{x} = \mathbf{b}$ have a solution for any $\mathbf{b} \in \mathbb{R}^n$? Does $B\mathbf{y} = \mathbf{c}$ have a solution for any $\mathbf{c} \in \mathbb{R}^m$?
3. Let a vector space \mathbb{V} include all polynomials of degree 3 or less, $\mathbb{V} = \{a_0 + a_1t + a_2t^2 + a_3t^3 : a_0, a_1, a_2, a_3 \in \mathbb{R}\}$. Let T be a transformation that maps a polynomial $f(t)$ to the derivative $f'(t)$.

- (a) To make T a map from \mathbb{V} into \mathbb{W} , what should be \mathbb{W} if \mathbb{W} is a vector space?
- (b) Find bases $\mathcal{B}_{\mathbb{V}}$ and $\mathcal{B}_{\mathbb{W}}$ of \mathbb{V} and \mathbb{W} .
- (c) Find the matrix A representing the transform T with respect to the bases $\mathcal{B}_{\mathbb{V}}$ and $\mathcal{B}_{\mathbb{W}}$ chosen above.
4. Let a vector space \mathbb{V} include all polynomials of degree 3 or less, $\mathbb{V} = \{a_0 + a_1t + a_2t^2 + a_3t^3 : a_0, a_1, a_2, a_3 \in \mathbb{R}\}$. For two polynomials $f(t), g(t) \in \mathbb{V}$, we define an inner product
- $$\langle f, g \rangle = \int_{-1}^1 f(t)g(t)dt.$$
- (a) Compute $|f|$, $|g|$, and $|f - g|$ for $f(t) = t$ and $g(t) = t^2$.
- (b) Find an orthonormal basis of $\text{span}\{1, t^2\}$.
- (c) Let $\mathbb{W} = \text{span}\{t^2\}$. Find a basis of \mathbb{W}^\perp .
5. Consider the Euclidean vector space \mathbb{R}^4 . For two vectors $\mathbf{x} = (x_1, x_2, x_3, x_4)^\top$ and $\mathbf{y} = (y_1, y_2, y_3, y_4)^\top$ in \mathbb{R}^4 , we define a *non-standard* inner product
- $$\langle \mathbf{x}, \mathbf{y} \rangle = 2x_1y_1 + \frac{2}{3}(x_1y_3 + x_2y_2 + x_3y_1) + \frac{2}{5}(x_2y_4 + x_3y_3 + x_4y_2) + \frac{2}{7}x_4y_4,$$
- or, in matrix form as
- $$\langle \mathbf{x}, \mathbf{y} \rangle = [\mathbf{x}] \begin{bmatrix} 2 & 0 & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{2}{5} \\ \frac{2}{3} & 0 & \frac{2}{5} & 0 \\ 0 & \frac{2}{5} & 0 & \frac{2}{7} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}.$$
- (a) Compute $|\mathbf{x}|$, $|\mathbf{y}|$, and $|\mathbf{x} - \mathbf{y}|$ for $\mathbf{x} = (0, 1, 0, 0)^\top$ and $\mathbf{y} = (0, 0, 1, 0)^\top$.
- (b) Find an orthonormal basis of $\text{span}\{(1, 0, 0, 0)^\top, (0, 0, 1, 0)^\top\}$.
- (c) Let $\mathbb{W} = \text{span}\{(0, 0, 1, 0)^\top\}$. Find a basis of \mathbb{W}^\perp .
6. (a) Let a subspace \mathbb{W} of \mathbb{R}^3 be spanned by two vectors $(-1, 0, 1)^\top$ and $(1, 1, 0)^\top$. Find the projection of $(1, 1, 1)^\top$.
- (b) Consider a linear model $z = ax + by$. We have three observations of (x, y, z) : $(-1, 1, 1)$, $(0, 1, 1)$, $(1, 0, 1)$. Compute the least square estimates of a and b .
- (c) Compare the above two questions.
7. True or False. No need to explain your guesses.
- (a) For an $m \times n$ ($m > n$) matrix, the number of linearly independent rows equals the number of linearly independent columns.

- (b) For an $n \times n$ matrix A , a map T from the vector space of $n \times n$ matrix onto itself, defined as $T(X) = AX - XA$ is a linear transformation.
- (c) Suppose that a square matrix A is invertible. Then, the following block matrix B is invertible, and the inverse is given as

$$B = \begin{bmatrix} A & \mathbf{0} \\ C & I \end{bmatrix} \quad \text{and} \quad B^{-1} = \begin{bmatrix} A^{-1} & \mathbf{0} \\ -CA^{-1} & I \end{bmatrix}$$

where $\mathbf{0}$ is a matrix with 0's and I is an identity matrix in appropriate sizes, respectively.

- (d) A set of linearly independent vectors is orthogonal.
- (e) For a square matrix A , $\dim \text{Null}(A) = \dim \text{Null}(A^\top)$.

Problem Set 2

1. Suppose A is a 3×4 matrix and \mathbf{b} is a vector in \mathbb{R}^3 .
 - (a) Give an example of the matrix A with rank 2. No component of A is allowed to be zero and PA admits an LU decomposition with $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ whereas A itself does not allow LU decomposition.
 - (b) For the matrix A in (a), give two examples of \mathbf{b} , one of which allows a solution of $A\mathbf{x} = \mathbf{b}$ and the other one doesn't.
 - (c) Find vectors in $\text{Col}(A)$ which are closest to the two vectors in (b), respectively.
2. Assume that $\mathbf{u}_1, \dots, \mathbf{u}_k$ and $\mathbf{v}_1, \dots, \mathbf{v}_k$ are sets of orthonormal vectors in \mathbb{R}^n , respectively. Find the rank of $n \times n$ matrix

$$\sum_{j=1}^k j \mathbf{u}_j \mathbf{v}_j^\top.$$
3. Let a vector space \mathbb{V} include all functions in the form of $f(t) = a_1 e^{-t} + a_2 e^{-2t} + a_3 e^{-3t}$ where $a_1, a_2, a_3 \in \mathbb{R}$, that is, $\mathbb{V} = \{a_1 e^{-t} + a_2 e^{-2t} + a_3 e^{-3t} : a_1, a_2, a_3 \in \mathbb{R}\}$. For two functions $f(t)$ and $g(t)$ in \mathbb{V} , we define an inner product

$$\langle f, g \rangle = \int_0^\infty f(t)g(t)dt.$$
 - (a) Check that \mathbb{V} is a vector space over the scalar \mathbb{R} .
 - (b) Compute $|f|$, $|g|$, and $|f - g|$ for $f(t) = e^{-t}$ and $g(t) = e^{-2t}$.
 - (c) Show that $\mathcal{B}_1 = \{e^{-t}, e^{-2t}, e^{-3t}\}$ is linearly independent.
 - (d) Find the matrix representation of the inner product $\langle \cdot, \cdot \rangle$ with respect to the basis \mathcal{B}_1 .

- (e) For $\mathbb{W} = \text{span}\{e^{-t}, e^{-3t}\}$, compute $\mathbf{P}_{\mathbb{W}}(e^{-2t})$.
- (f) Find an orthonormal basis \mathcal{B}_2 for \mathbb{V} .
- (g) Find the matrix representation of the inner product $\langle \cdot, \cdot \rangle$ with respect to the basis \mathcal{B}_2 .
- (h) The density level of some chemical at time t is described by $a_1 e^{-t} + a_2 e^{-2t} + a_3 e^{-3t} + \varepsilon$ for some fixed $a_1, a_2, a_3 \in \mathbb{R}$. At n time points $0 < t_1 < t_2 < \dots < t_n$, the observed density levels are y_1, y_2, \dots, y_n . For $\varepsilon_i = y_i - (a_1 e^{-t_i} + a_2 e^{-2t_i} + a_3 e^{-3t_i})$, characterize $\hat{a}_1, \hat{a}_2, \hat{a}_3$ that minimizes $\sum_{i=1}^n \varepsilon_i^2$.
4. Let \mathbb{V} be a vector space and T be a linear transform representing a projection.
- Show that $\mathbb{W} = \{T(\mathbf{v}) : \mathbf{v} \in \mathbb{V}\}$ be a subspace of \mathbb{V} .
 - Show that $I - T$ is also a projection where $(I - T)(\mathbf{v}) = \mathbf{v} - T(\mathbf{v})$.
 - What is the linear transform $T \circ (I - T)$?
 - Describe \mathbb{W}^\perp in terms of T .
5. True or False. No need to explain your guesses.
- For an $n \times n$ matrix A , $I - A$ has rank $n - k$ if $\text{rank } A = k$.
 - For an $n \times n$ projection matrix P , $I - P$ has rank $n - k$ if $\text{rank } P = k$.
 - For an $n \times n$ matrix A , a map T from the vector space of $n \times n$ matrix onto itself, defined as $T(X) = A^\top X - X^\top A$ is a linear transformation.
 - Suppose that U and V are $n \times k$ matrices. Then
- $$\begin{bmatrix} I_n & \mathbf{0} \\ V^\top & I_k \end{bmatrix} \begin{bmatrix} I_n + UV^\top & U \\ \mathbf{0} & I_k \end{bmatrix} \begin{bmatrix} I_n & \mathbf{0} \\ -V^\top & I_k \end{bmatrix} = \begin{bmatrix} I_n & U \\ \mathbf{0} & I_k - V^\top U \end{bmatrix}.$$
- (e) A set of orthogonal vectors is linearly independent.

13.2 Problems for Chapter 5 ~ 8

Problem Set 1

1. A 3×3 square matrix has a QR -type decomposition (not an exact QR -decomposition) as $A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1/2 & -1 \\ 1 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.
- Find the Q and R factors in the QR -decomposition of A .
 - Compute the volume of a parallelepiped(평행육면체) whose six faces are parallelograms formed by the column vectors of A .

- (c) Compute A^{-1} . You may describe the inverse in a decomposed form.
2. Suppose A is a 3×4 matrix. We computed its singular value decomposition (SVD) using a computer program. Unfortunately, there was a problem with the screen, and we could not recognize some figures of the SVD results. We have U , V , and Σ such that $A = U\Sigma V^\top$ where
- $$U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & * & * \\ -1/\sqrt{2} & * & * \end{bmatrix}, V = \begin{bmatrix} 1/2 & 1/\sqrt{2} & 0 \\ 1/2 & 0 & * \\ 1/2 & * & * \\ 1/2 & 0 & * \end{bmatrix}, \Sigma = \begin{bmatrix} 3 & 0 & * \\ 0 & 2 & * \\ 0 & * & 1 \end{bmatrix}, \text{ and } * \text{ means missing figure on the screen.}$$
- (a) Fill out V .
- (b) Find the largest eigenvalue and corresponding eigenvectors of $A^\top A$.
- (c) Find a rank 2 matrix B that minimizes $\|A - B\|_2$.
3. Suppose A is a 3×4 matrix given as $A = \begin{bmatrix} 2 & 1 & -1 & 1 \\ 4 & 2 & -2 & 2 \\ -2 & -1 & 1 & -1 \end{bmatrix}$.
- (a) Find a singular value decomposition of A .
- (b) Find the pseudoinverse A^+ of A .
- (c) Find $\|A\|_2$.
4. Assume an $m \times n$ matrix $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ has non-zero columns $\mathbf{a}_i \in \mathbb{R}^m$, $i = 1, \dots, n$ that are orthogonal to each other: $\mathbf{a}_i^\top \mathbf{a}_j = 0$ for $i \neq j$. Find an SVD for A , in terms of \mathbf{a}_i 's. Be as explicit as you can.
5. Assume an $m \times n$ matrix A with $m \geq n$, have singular values $\sigma_1, \dots, \sigma_n$. Set an $(m+n) \times n$ matrix $\tilde{A} = \begin{bmatrix} A \\ I_n \end{bmatrix}$. (10 pts each)
- (a) Find the singular values of \tilde{A} .
- (b) Find an SVD of the matrix \tilde{A} in terms of the SVD for A .
6. $A = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$ where \mathbf{v}_i 's are orthonormal, $\lambda_1 > 0$, and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.
- (a) Compute $\det A$.
- (b) Find all eigenvalues and corresponding eigenvectors of A .
- (c) For A to be positive semidefinite, what conditions on λ_i do we need?
- (d) If A is positive semidefinite, characterize an SVD of A .
- (e) If A is positive semidefinite, characterize the pseudoinverse A^+ of A .

- (f) If A is positive definite, characterize the inverse A^{-1} of A .
7. Let A be an $n \times n$ symmetric positive definite matrix whose spectral decomposition is described as $A = V\Lambda V^\top$ where V is an orthogonal matrix and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ for $\lambda_i > 0$. We also define $\Lambda^{1/k} = \text{diag}(\lambda_1^{1/k}, \dots, \lambda_n^{1/k})$.
- Characterize a symmetric positive definite matrix B_2 satisfying $A = B_2^2$ in terms of V and Λ .
 - Characterize a symmetric positive definite matrix B_k satisfying $A = B_k^k$ in terms of V and Λ for every positive integer k .
 - What would be $\lim_{k \rightarrow \infty} B_k$? Validate your answer as reasonably as possible.
8. True or False. No need to explain your guesses.
- For an orthogonal matrix Q , $\det Q = 1$.
 - For $n \times n$ symmetric matrices A and B , $\lambda_k(A + B) > \lambda_k(A)$ if B is positive definite.
 - Suppose that B and C are square matrices. Then, the following block matrix $A = \begin{bmatrix} B & \mathbf{0} \\ \mathbf{0} & C \end{bmatrix}$ is positive definite if and only if both B and C are positive definite.
 - For an $m \times n$ matrix A , $A^+ = (A^\top A)^{-1}A$ if $\text{rank}(A) = m$.
 - Every projection matrix has 0 as its determinant.

Problem Set 2

1. Let a matrix A be $m \times n$ and a matrix B be $n \times m$. Consider a block matrix

$$M = \begin{bmatrix} \mathbf{0} & A \\ -B & I_n \end{bmatrix}.$$

Find $\det M$.

2. Assume that a matrix B has eigenvalues 1, 2, 3, a matrix C has eigenvalues 4, 0, -4, and a matrix D has eigenvalues -1, -2, -3. Set a 6×6 matrix $A = \begin{bmatrix} B & C \\ \mathbf{0} & D \end{bmatrix}$ where B, C , and D are 3×3 matrices.

We also define a linear transformation $\ell : \mathbb{R}^6 \rightarrow \mathbb{R}^6$ as $\ell(\mathbf{x}) = A\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^6$.

- What are the eigenvalues of the 6×6 matrix A ?
- Consider the unit cube Q in \mathbb{R}^6 , that is, $Q = \{(x_1, \dots, x_6)^\top : 0 \leq x_i \leq 1, i = 1, \dots, 6\} \subset \mathbb{R}^6$. Find the volume of $\ell(Q) = \{\ell(\mathbf{x}) : \mathbf{x} \in Q\}$.

3. Let a 4×4 matrix A is given as

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

- (a) Find eigenvalues and eigenvectors of A .
- (b) Find singular values and right-/left-singular vectors of A .
- (c) Find $\|A\|_2$ and $\|A\|_F$.
- (d) Find the pseudoinverse A^+ of A .
- (e) Find a 4×4 matrix B of rank 2 that minimizes $\|A - B\|_2$.

4. Suppose A is a 3×4 matrix given as $A = \begin{bmatrix} 2 & 1 & -1 & 1 \\ 4 & 2 & -2 & 2 \\ -2 & -1 & 1 & -1 \end{bmatrix}$.

- (a) Find eigenvalues and eigenvectors of $A^\top A$.
- (b) Find eigenvalues and eigenvectors of AA^\top .
- (c) Is there a 3×4 matrix B of rank 2 that minimizes $\|A - B\|_2$?

5. Let A be an $m \times n$ matrix. Define its symmetrization $s(A)$ as

$$s(A) = \begin{bmatrix} \mathbf{0} & A \\ A^\top & \mathbf{0} \end{bmatrix},$$

which is an $(m+n) \times (m+n)$ **symmetric** matrix as the name suggests.

- (a) Let $(\sigma, \mathbf{v}, \mathbf{u})$ be a singular triplet of A . Then, σ and $-\sigma$ are eigenvalues of $s(A)$. Find the eigenvectors of $s(A)$ associated with the eigenvalues σ and $-\sigma$, respectively.
 - (b) Let (λ, \mathbf{w}) be an eigenpair of $s(A)$ where $\lambda \neq 0$. Then, $-\lambda$ is also an eigenvalue of $s(A)$. Find the eigenvector of $s(A)$ associated with the eigenvalue $-\lambda$.
 - (c) Let $\lambda < 0$ and (λ, \mathbf{w}) be an eigenpair of $s(A)$. Find right-/left-singular vectors of A associated with the singular value $-\lambda$.
6. Let A be an $m \times n$ matrix with at least one non-zero entry. What are the projection $\mathbf{P}_{\text{Col}(A)}$ onto $\text{Col}(A)$ and the pseudoinverse A^+ of A if
- (a) A consists of a single column \mathbf{v} , that is, $n = 1$;
 - (b) the columns of A are orthonormal;
 - (c) the columns of A are linearly independent (that is $\text{rank } A = n$);

- (d) the columns of A are possibly linearly dependent, and one of its compact SVD is $U\Sigma V^\top$.
7. Let A be an $n \times n$ symmetric positive definite matrix. Denote its eigenvalues as $\lambda_1, \dots, \lambda_n$. Let B be the square root of A , that is, $A = B^2$.
- Let $A = U^\top U$ be the Cholesky factorization of A . Find $|\det U|$.
 - Find the eigenvalues of B .
 - Let $B = QR$ be the QR -decomposition of B . Find $|\det R|$.
 - Does there exist the square root C of B ? If yes, what are the eigenvalues of C ?

Bibliography

- [1] Blum, A., Hopcroft, J., and Kannan, R. (2020) *Foundations of Data Science*, Cambridge University Press, UK.
- [2] Garcia, S. R., and Horn, R. A. (2017) *A Second Course in Linear Algebra*, Cambridge University Press, UK.
- [3] Hofmann, T., Schölkopf, B., and Smola, A. J. (2008) Kernel Methods in Machine Learning. *The Annals of Statistics* 36(3), 1171–1220.
- [4] Langville, A. N., and Meyer, C. D. (2006) *Google's Page Rank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, USA.
- [5] Strang, G. (2006) *Linear Algebra and Its Applications*, Fourth Edition, Brooks/Cole, Cengage Learning, USA.

Kang & Cho (2024)

Appendix A

Permutation and its Matrix Representation

A permutation σ is simply a bijective function from and onto $\{1, \dots, n\}$, that is, $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ such that $\sigma(i) \neq \sigma(j)$ if $i \neq j$. You may think that a permutation shuffles the order of $1, \dots, n$. Therefore, its inverse function σ^{-1} always exists, and the inverse function is a permutation, too. The size n of the permuted set is usually clear in the context. We can also define a $n \times n$ matrix associated with a permutation σ so that $(i, \sigma(i))$ -element of the matrix is 1 and all other elements are zero in the i -th row. Since $(i, \sigma(i))$ equals $(\sigma^{-1}(j), j)$ if we set $j = \sigma(i)$, each column of the matrix also has only one non-zero element. This matrix is called a permutation matrix.

Let us consider an example of size 4, $\sigma(1) = 3$, $\sigma(2) = 2$, $\sigma(3) = 4$, and $\sigma(4) = 1$. Then, its permutation matrix P is

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

You may regard a permutation matrix as a row-shuffled identity matrix. Its inverse permutation is $\sigma^{-1}(1) = 4$, $\sigma^{-1}(2) = 2$, $\sigma^{-1}(3) = 1$, and $\sigma^{-1}(4) = 3$, whose permutation matrix Q is

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

If you know the transpose of a matrix, you can see that $Q = P^\top$ from this example. That is, the matrix of inverse permutation is the transpose of the original permutation matrix. Furthermore, it is easy to see $PQ = I$ and $QP = I$, which implies that Q is the inverse P^{-1} of P . That is, the matrix of inverse

permutation is the inverse of the original permutation matrix. Combining these two relations, every permutation matrix is invertible and its transpose is its inverse matrix, that is, every permutation matrix is orthogonal.

This conclusion holds not only for this small example. To generalize these results, it is convenient to borrow a summation representation of rank-one matrices if you are exposed to the rank-one matrix. For a permutation σ , its permutation matrix P can be compactly expressed as

$$P = \sum_{i=1}^n \mathbf{e}_i \mathbf{e}_{\sigma(i)}^\top \quad (\text{A.1})$$

and

$$P^\top = \left(\sum_{i=1}^n \mathbf{e}_i \mathbf{e}_{\sigma(i)}^\top \right)^\top = \sum_{i=1}^n (\mathbf{e}_i \mathbf{e}_{\sigma(i)}^\top)^\top = \sum_{i=1}^n \mathbf{e}_{\sigma(i)} \mathbf{e}_i^\top = \sum_{j=1}^n \mathbf{e}_j \mathbf{e}_{\sigma^{-1}(j)}^\top \quad (\text{A.2})$$

where the last equality is obtained by replacing $j = \sigma(i)$ and $\sigma^{-1}(j) = \sigma^{-1}(\sigma(i)) = i$. Furthermore,

$$\begin{aligned} PP^\top &= \left(\sum_{i=1}^n \mathbf{e}_i \mathbf{e}_{\sigma(i)}^\top \right) \left(\sum_{j=1}^n \mathbf{e}_j \mathbf{e}_{\sigma^{-1}(j)}^\top \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \mathbf{e}_i \mathbf{e}_{\sigma(i)}^\top \mathbf{e}_j \mathbf{e}_{\sigma^{-1}(j)}^\top \\ &= \sum_{i=1}^n \sum_{j=\sigma(i)} \mathbf{e}_i (\mathbf{e}_{\sigma(i)}^\top \mathbf{e}_j) \mathbf{e}_{\sigma^{-1}(j)}^\top \\ &= \sum_{i=1}^n \mathbf{e}_i (\mathbf{e}_{\sigma(i)}^\top \mathbf{e}_{\sigma(i)}) \mathbf{e}_{\sigma^{-1}(\sigma(i))}^\top \\ &= \sum_{i=1}^n \mathbf{e}_i \mathbf{e}_i^\top \\ &= I. \end{aligned}$$

That is, for a permutation matrix P ,

$$PP^\top = I, \quad (\text{A.3})$$

which also implies

$$P^{-1} = P^\top. \quad (\text{A.4})$$

Appendix B

Existence of Optimizers

정의구역이 X 인 함수 $f : X \rightarrow \mathbb{R}$ 에 대해서 어떤 실수 α 가 upper bound라는 것은 $f(\mathbf{x}) \leq \alpha$ for all $\mathbf{x} \in X$ 를 만족하는 것이다. d -차원 유클리드 공간의 단위 구를 $B = \{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| \leq 1\}$ 라고 놓자.

함수 $f : B \rightarrow \mathbb{R}$ 에 대해서 최대화 문제

$$\max_{\mathbf{x} \in B} f(\mathbf{x})$$

를 생각하자. 함수 f 가 연속이고 upper bound를 가지면 적어도 한 개의 $\mathbf{x}^* \in B$ 에 대해 $f(\mathbf{x}) \leq f(\mathbf{x}^*)$ for all $\mathbf{x} \in B$ 가 성립한다는 것을 보일 수 있다. 즉 함수의 최댓값을 갖는 B 의 원소가 존재한다는 사실이다.¹ Advanced calculus나 topology로부터 “continuous image of a compact set is compact”와 “a compact set in \mathbb{R} is closed and bounded”라는 명제를 아는 경우 자명한 결과이나 모르는 경우를 위하여 기초적인 calculus를 이용한 증명을 소개한다. 증명을 위한 첫번째 단계로 다음 결과부터 살펴보자.

Lemma B.1 어떤 실수 α 가 연속함수 $f : B \rightarrow \mathbb{R}$ 의 upper bound라면

- $\mathbf{x}^* \in B$ 가 존재하여 $\alpha = f(\mathbf{x}^*)$ 거나;
- 임의의 $\bar{\mathbf{x}} \in B$ 에 대해서

1. $f(\bar{\mathbf{x}}) < f(\tilde{\mathbf{x}}) < \alpha$ 면서 $\alpha - f(\tilde{\mathbf{x}}) = \frac{1}{2}(\alpha - f(\bar{\mathbf{x}}))$ 를 만족하는 $\tilde{\mathbf{x}} \in B$ 가 존재하거나;
2. $\beta - f(\bar{\mathbf{x}}) = \frac{1}{2}(\alpha - f(\bar{\mathbf{x}}))$ 를 만족하는 upper bound β 가 존재한다.

Proof:

- 방정식 $f(\mathbf{x}) = \alpha$ 가 해를 가지면 이 해를 \mathbf{x}^* 로 놓는다. 만약 방정식의 해가 없다면 함수 $f(\mathbf{x}) < \alpha$ for all $\mathbf{x} \in B$ 를 의미한다.
- $f(\mathbf{x}) = \alpha$ 의 해가 없는 경우 $\bar{\mathbf{x}} \in B$ 를 생각하면 $f(\bar{\mathbf{x}}) < \alpha$ 이다. f 의 최댓값보다 큰 α 와 $f(\bar{\mathbf{x}})$ 의 중간값인 $\beta = f(\bar{\mathbf{x}}) + \frac{1}{2}(\alpha - f(\bar{\mathbf{x}}))$ 에 대해 방정식 $f(\mathbf{x}) = \beta$ 를 고려하자.

¹ 예를 들면 집합 $(0, 1) = \{x \in \mathbb{R} : 0 < x < 1\}$ 상에서 함수 $f(x) = x$ 는 최댓값을 갖는 원소를 $(0, 1)$ 에서 찾을 수 없다.

1. 해가 존재하는 경우 그 해를 $\tilde{\mathbf{x}}$ 로 놓으면 $f(\tilde{\mathbf{x}}) = \beta$ 이므로

$$\alpha - f(\tilde{\mathbf{x}}) = \alpha - \beta = \alpha - f(\bar{\mathbf{x}}) - \frac{1}{2}(\alpha - f(\bar{\mathbf{x}})) = \frac{1}{2}(\alpha - f(\bar{\mathbf{x}}))$$

이다. $f(\tilde{\mathbf{x}}) > f(\bar{\mathbf{x}})$ 라는 것도 확인할 수 있다.

2. 해가 존재하지 않는 경우 함수 f 가 연속함수이기 때문에 β 가 f 의 upper bound이다. 그리고

$$\beta - f(\bar{\mathbf{x}}) = \frac{1}{2}(\alpha - f(\bar{\mathbf{x}}))$$

이다. ■

B 상의 연속함수 f 가 upper bound α 를 갖는다고 가정하자. $\mathbf{x}_0 = \mathbf{0} \in B$ 와 upper bound $\alpha_0 = \alpha$ 에 대해서 Lemma B.1을 반복적으로 적용하여

$$\alpha_k - f(\mathbf{x}_k) = \frac{1}{2}(\alpha_{k-1} - f(\mathbf{x}_{k-1})) \quad (\text{B.1})$$

를 만족하는 $\mathbf{x}_k \in B$ 와 upper bound α_k , $k = 1, 2, \dots$ 를 얻자. 만약 어떤 k 에 대해 $f(\mathbf{x}^*) = \alpha_k$, $\mathbf{x}^* \in B$ 를 찾았다면 \mathbf{x}^* 가 우리가 원하던 $f(\mathbf{x}^*) = \max_{\mathbf{x} \in B} f(\mathbf{x})$ 를 만족한다. 따라서 모든 k 에 대해서 $f(\mathbf{x}) = \alpha_k$, $\mathbf{x} \in B$ 의 해가 존재하지 않는다고 가정하자. Lemma B.1에서 $f(\mathbf{x}) = \alpha$ 의 해가 없는 경우 중

- 1번의 경우 upper bound는 $\alpha_k = \alpha_{k-1}$ 로 이전 값을 유지하고 \mathbf{x}_k 는 새로운 값으로 바꾼다;
- 2번의 경우 argument는 $\mathbf{x}_k = \mathbf{x}_{k-1}$ 로 이전 값을 유지하고 upper bound α_k 는 작은 값을 바꾼다.

(B.1)에 의해

$$0 \leq \alpha_{k-1} - \alpha_k \leq \frac{1}{2^k}(\alpha - f(\mathbf{0}))$$

이므로 calculus의 comparison test에 의해 $\lim_{k \rightarrow \infty} \alpha_k = \alpha + \sum_{k=1}^{\infty} (\alpha_k - \alpha_{k-1}) = \ell$ 이 존재한다. 임의의 $\mathbf{x} \in B$ 에 대해 $f(\mathbf{x}) \leq \alpha_k$ 이므로 $f(\mathbf{x}) \leq \lim_{k \rightarrow \infty} \alpha_k = \ell$ 이고 ℓ 도 upper bound이다. 그리고 (B.1)에 의해 $\lim_{k \rightarrow \infty} (\alpha_k - f(\mathbf{x}_k)) = 0$ 이므로 $\lim_{k \rightarrow \infty} f(\mathbf{x}_k) = \ell$ 이 된다. 즉 $\mathbf{x}_k \in B$ 면서 $f(\mathbf{x}_k) \rightarrow \ell$ 인 sequence를 얻는다.

$\{\mathbf{x}_k : k = 1, 2, \dots\}$ 를 분석하기 위해서 B 를 포함하는 d -dimensional cube $C_0 = [-1, 1]^d$ 를 생각하자. C_0 의 각 모서리를 같은 길이의 두 선분으로 나누어 모서리 길이가 모두 1인 2^d 개의 작은 cube들로 C_0 를 나누자. 2^d 개의 작은 cube들에 무한개의 \mathbf{x}_k 가 담겨있으므로 적어도 한 cube (C_1 이라 부르자)에는 무한개의 \mathbf{x}_k 가 담겨있다. C_1 에 담긴 \mathbf{x}_k 들을 $\mathbf{x}_i^{(1)}$, $i = 1, 2, \dots$ 로 표시하자. C_1 의 각 모서리의 길이가 $1/2$ 이 되도록 2^d 개의 작은 cube로 나누면 이전 단계와 유사하게 C_2 로 부를 작은 cube 한개에 무한개의 $\mathbf{x}_i^{(1)}$ 들이 담겨있다. C_2 에 담긴 $\mathbf{x}_i^{(1)}$ 들을 $\mathbf{x}_i^{(2)}$, $i = 1, 2, \dots$ 로 표시하자. 이 과정을 무한히 반복하여 모서리의 길이가 절반으로 짧아지는 C_j 들과 그 안에 담긴 $\mathbf{x}_i^{(j)}$ 를 얻자. 얻는 과정으로부터 $C_j \subset C_{j-1}$ 이고 $\{\mathbf{x}_i^{(j)}\}_{i=1}^{\infty} \subset \{\mathbf{x}_i^{(j-1)}\}_{i=1}^{\infty}$ 임을 알 수 있다.

Cantor의 diagonal argument를 적용하기 위해서 벡터열 $\{\mathbf{x}_i^{(i)}\}_{i=1}^{\infty}$ 를 생각하자. Diagonal element로 이루어진 벡터열의 $\mathbf{x}_i^{(i)} \in C_j$ for all $i \geq j$ 라는 성질이 중요하다. 이 벡터열의 첫번째 좌표들로 이루어진

실수 수열을 $\{y_i\}$ 로 나타내자. $\mathbf{x}_{j+1}^{(j+1)}$ 와 $\mathbf{x}_j^{(j)}$ 가 모두 C_j 에 포함되고 C_j 의 각 모서리의 길이가 $(\frac{1}{2})^{j-1}$ 이므로 $|y_{j+1} - y_j| \leq (\frac{1}{2})^{j-1}$ 이다. Calculus의 comparison test에 의해 $y_j = y_1 + \sum_{i=1}^{j-1} (y_{i+1} - y_i)$ 은 수렴한다. 나머지 좌표에 대해서도 같은 논리를 적용하면 모든 좌표가 수렴하고 이 수렴점을 \mathbf{x}^* 로 표시하면 $|\mathbf{x}_j^{(j)} - \mathbf{x}^*| \rightarrow 0$ 이다. $f(\mathbf{x}_j^{(j)})$ 가 수렴하는 실수 수열인 $f(\mathbf{x}_k)$ 의 subsequence이므로 $\ell = \lim_{k \rightarrow \infty} f(\mathbf{x}_k) = \lim_{j \rightarrow \infty} f(\mathbf{x}_j^{(j)})$ 이고 f 가 연속함수라는 사실로부터 $\lim_{j \rightarrow \infty} f(\mathbf{x}_j^{(j)}) = f(\mathbf{x}^*)$ 를 알 수 있어서 $\ell = f(\mathbf{x}^*)$ 가 성립한다. 또한 $|\mathbf{x}_j^{(j)}| \leq 1$ 이므로 모든 j 에 대해 $|\mathbf{x}^*| \leq |\mathbf{x}^* - \mathbf{x}_j^{(j)}| + |\mathbf{x}_j^{(j)}| \leq |\mathbf{x}^* - \mathbf{x}_j^{(j)}| + 1$ 이 성립한다. 양변에 limit을 취하면 $|\mathbf{x}^*| \leq 1$ 을 얻어서 $\mathbf{x}^* \in B$ 임을 알 수 있다.

Lemma B.2 *If a continuous function $f : B \rightarrow \mathbb{R}$ has an upper bound, there exists an $\mathbf{x}^* \in B$ such that*

$$f(\mathbf{x}^*) = \max_{\mathbf{x} \in B} f(\mathbf{x}).$$

Lemma B.2를 $-f$ 에 적용하면 lower bound를 갖는 연속함수의 최솟값에 대해서도 동일한 결과를 얻는다. 그리고 최적화 문제를 다룰 때 임의의 집합 B 와 C 에 대하여

$$\text{"if } \mathbf{x}^* \in B \subset C, f(\mathbf{x}^*) \leq \max_{\mathbf{x} \in B} f(\mathbf{x}) \leq \max_{\mathbf{x} \in C} f(\mathbf{x}) \text{ holds."}$$

라는 사실은 너무 간단하여 인용없이 여러 곳에서 자주 활용된다. 다음은 함수가 선형변환된 벡터의 크기로 주어지는 경우로 유용한 결과이다.

Lemma B.3 *Let A be an $n \times d$ matrix. Then there exists a unit vector \mathbf{x}^* such that*

$$|A\mathbf{x}^*| = \max_{|\mathbf{x}| \leq 1} |A\mathbf{x}| = \max_{|\mathbf{x}|=1} |A\mathbf{x}|.$$

Proof: If $A = \mathbf{0}$, the equalities hold trivially. Therefore, let us assume that $A \neq \mathbf{0}$. Then the maximum value is positive.

Consider the first equality. Since the function $|A\mathbf{x}|$ is continuous, we can apply Lemma B.2 once we show that the function $|A\mathbf{x}|$ has an upper bound on B . Let $|\mathbf{x}| \leq 1$ and denote i -th row of A as \mathbf{a}_i^\top . Then, by the Cauchy-Schwartz inequality,

$$|A\mathbf{x}|^2 = \sum_{i=1}^n (\mathbf{a}_i^\top \mathbf{x})^2 \leq \sum_{i=1}^n |\mathbf{a}_i|^2 |\mathbf{x}|^2 \leq \sum_{i=1}^n |\mathbf{a}_i|^2 = \|A\|_F^2$$

and we can conclude that the function $|A\mathbf{x}|$ has an upper bound $\|A\|_F$.

For the second equality, let \mathbf{x}^* be the vector satisfying the first equality. If $|\mathbf{x}^*| = 0$, $|A\mathbf{x}^*| = 0$ contradicts to $A \neq \mathbf{0}$. If $|\mathbf{x}^*| < 1$, then the unit vector $\mathbf{y} = \frac{1}{|\mathbf{x}^*|} \mathbf{x}^*$ provides a bigger value $|A\mathbf{y}| > |A\mathbf{x}^*|$, which is a contradiction. Hence $|\mathbf{x}^*| = 1$ and this implies the second equality. ■

Lemma B.3를 더 확장하여 orthonormal한 vector들에 수직하다는 조건이 추가된 경우를 생각해 보자.

Lemma B.4 *Let A be an $n \times d$ matrix and $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be orthonormal. Then there exists a feasible \mathbf{x}^* satisfying $|\mathbf{x}^*| = 1, \mathbf{v}_1 \perp \mathbf{x}^*, \dots, \mathbf{v}_k \perp \mathbf{x}^*$ such that*

$$|A\mathbf{x}^*| = \max\{|A\mathbf{x}| : |\mathbf{x}| \leq 1, \mathbf{v}_1 \perp \mathbf{x}, \dots, \mathbf{v}_k \perp \mathbf{x}\} = \max\{|A\mathbf{x}| : |\mathbf{x}| = 1, \mathbf{v}_1 \perp \mathbf{x}, \dots, \mathbf{v}_k \perp \mathbf{x}\}.$$

Proof: Assume $k < d$. Expand the k orthonormal vectors such that $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_d\}$ to be an orthonormal basis by Gram-Schmidt procedure. Denote $V = [\mathbf{v}_1 | \dots, | \mathbf{v}_d]$, the matrix whose columns are \mathbf{v}_k 's. If \mathbf{y} is the coordinate vector with respect to the new basis, $\mathbf{x} = V\mathbf{y}$ holds. Note that $\mathbf{v}_i \perp \mathbf{x}$ is equivalent to $y_i = 0$ since $\mathbf{v}_i^\top \mathbf{x} = \mathbf{v}_i^\top V\mathbf{y} = y_i$. Furthermore, $|\mathbf{x}| = \sqrt{\mathbf{x}^\top \mathbf{x}} = \sqrt{\mathbf{y}^\top V^\top V\mathbf{y}} = \sqrt{\mathbf{y}^\top \mathbf{y}} = |\mathbf{y}|$. Set \tilde{A} to be the last $(d - k)$ columns of AV and $\tilde{\mathbf{y}} = (y_{k+1}, \dots, y_d)^\top$. Then, the following set equalities

$$\{|A\mathbf{x}| : |\mathbf{x}| \leq 1, \mathbf{v}_1 \perp \mathbf{x}, \dots, \mathbf{v}_k \perp \mathbf{x}\} = \{|AV\mathbf{y}| : |\mathbf{y}| \leq 1, y_1 = 0, \dots, y_k = 0\} = \{|\tilde{A}\tilde{\mathbf{y}}| : |\tilde{\mathbf{y}}| \leq 1\}$$

imply

$$\max\{|A\mathbf{x}| : |\mathbf{x}| \leq 1, \mathbf{v}_1 \perp \mathbf{x}, \dots, \mathbf{v}_k \perp \mathbf{x}\} = \max\{|\tilde{A}\tilde{\mathbf{y}}| : |\tilde{\mathbf{y}}| \leq 1\}.$$

By applying Lemma B.3 to $\max\{|\tilde{A}\tilde{\mathbf{y}}| : |\tilde{\mathbf{y}}| \leq 1\}$, we get a unit vector $\tilde{\mathbf{y}}^*$ and $\tilde{\mathbf{y}}^{**} = \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{y}}^* \end{bmatrix}$ satisfying

$$\max\{|\tilde{A}\tilde{\mathbf{y}}| : |\tilde{\mathbf{y}}| \leq 1\} = |\tilde{A}\tilde{\mathbf{y}}^*| = |AV\tilde{\mathbf{y}}^{**}| = |A\mathbf{x}^*|$$

where $\mathbf{x}^* = V\tilde{\mathbf{y}}^*$ with $|\mathbf{x}^*| = |\tilde{\mathbf{y}}^*| = 1$ and $\mathbf{v}_1 \perp \mathbf{x}^*, \dots, \mathbf{v}_k \perp \mathbf{x}^*$. This proves the first equality. Since $|\mathbf{x}^*| = 1$, the second equality holds. ■

Appendix C

Complex Numbers and Matrices

복소수(complex number)는 $i^2 = -1$ 로 정의되는 허수기호 i 와 두 실수 a, b 에 대해서 $a + ib$ 로 정의되는 수이다. 복소수를 모아 놓은 집합을 \mathbb{C} 로 표시한다. 그리고 복소수를 원소로 갖는 n -벡터를 모아 놓은 벡터공간은 \mathbb{C}^n 으로, 복소수를 원소로 갖는 $n \times n$ 행렬을 모아 놓은 집합을 $\mathbb{C}^{n \times n}$ 으로 표현한다. $a + i(-b) = a - ib$ 로 표현하며 실수처럼 복소수도 사칙연산이 잘 정의되고 크기도 쟀 수 있다. 또한, 복소수 $a + ib$ 의 콤팩트복소수(complex conjugate)는 $a - ib$ 인데 이를 $\overleftrightarrow{a+ib}$ 로 표시한다. $b = 0$ 인 경우에는 $a + ib = a$ 로서 복소수는 실수를 특별한 경우로 포함한다.

- $(a_1 + ib_1) \pm (a_2 + ib_2) = (a_1 \pm a_2) + i(b_1 \pm b_2)$;
- $(a_1 + ib_1) \times (a_2 + ib_2) = (a_1a_2 - b_1b_2) + i(a_1b_2 + b_1a_2)$;
- $|a_1 + ib_1|^2 = (a_1 + ib_1) \times (\overleftrightarrow{a_1 + ib_1}) = (a_1 + ib_1) \times (a_1 - ib_1) = a_1^2 + b_1^2 \geq 0$ 이므로 $|a_1 + ib_1| = \sqrt{a_1^2 + b_1^2}$;
- $(a_1 + ib_1)^{-1} = \frac{1}{(a_1 + ib_1) \times (a_1 - ib_1)} (a_1 - ib_1) = \frac{a_1}{a_1^2 + b_1^2} - i \frac{b_1}{a_1^2 + b_1^2}$ if $|a + ib| \neq 0$.

또한, 사칙연산과 콤팩트연산은 호환적 자연스럽게 이루어진다. $z_1, z_2 \in \mathbb{C}$ 일 때,

- $\overleftrightarrow{(a_1 + ib_1) \pm (a_2 + ib_2)} = (\overleftrightarrow{a_1 + ib_1}) \pm (\overleftrightarrow{a_2 + ib_2})$, 즉 $\overleftrightarrow{z_1 \pm z_2} = \overleftrightarrow{z_1} \pm \overleftrightarrow{z_2}$;
- $\overleftrightarrow{(a_1 + ib_1) \times (a_2 + ib_2)} = (\overleftrightarrow{a_1 + ib_1}) \times (\overleftrightarrow{a_2 + ib_2})$, 즉 $\overleftrightarrow{z_1 \cdot z_2} = \overleftrightarrow{z_1} \cdot \overleftrightarrow{z_2}$.

복소행렬은 각 원소가 복소수인 행렬을 말하는데, 실수 행렬의 전치(transpose)에 대응되는 복소행렬의 연산은 conjugate transpose로서 복소행렬 $A = (a_{ij})$ 에 대해서 $A^H = (\overleftrightarrow{a_{ji}})$ 로 정의된다. 실수 행렬의 대칭 $A = A^\top$ 에 대응되는 개념은 $A = A^H$ 로서 이를 만족하는 행렬을 Hermitian matrix라고 부른다. 복소수 z 를 1×1 복소행렬로 생각하면 $z^H = \overleftrightarrow{z}$ 이다. $z \in \mathbb{C}$, $\mathbf{z} \in \mathbb{C}^n$, A, B 복소수 행렬이고 여러 합과 곱들이 잘 정의될 때,

- $\overleftrightarrow{\overleftrightarrow{z}} = z$, $\overleftrightarrow{\overleftrightarrow{\mathbf{z}}} = \mathbf{z}$, $\overleftrightarrow{\overleftrightarrow{A}} = A$;
- $\overleftrightarrow{A + B} = \overleftrightarrow{A} + \overleftrightarrow{B}$, $\overleftrightarrow{AB} = \overleftrightarrow{A} \overleftrightarrow{B}$;

- $(A + B)^H = A^H + B^H$, $(AB)^H = B^H A^H$

이다. 단순하지만 이후에 유용한 사실 한 가지는 다음과 같다.

Fact C.1 *A complex number z is a real number if and only if $z = z^H$.*

Proof: $z = a + ib$ 라 하면 $\overleftrightarrow{z} = a - ib$ 이고 $b = 0$ 이라는 것은 $z = \overleftrightarrow{z}$ 와 동치이다. ■

Fact C.2 *If $A = A^H$, then for all complex vectors \mathbf{x} , $\mathbf{x}^H A \mathbf{x}$ is real.*

Proof: Using Fact C.1, $(\mathbf{x}^H A \mathbf{x})^H = \mathbf{x}^H A^H (\mathbf{x}^H)^H = \mathbf{x}^H A \mathbf{x}$ implies that $\mathbf{x}^H A \mathbf{x}$ is real. ■

그리고 \mathbb{C}^n 의 표준내적은 $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ 에 대해서 $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^H \mathbf{v}$ 로 정의한다. \mathbb{R}^n 의 표준내적과 제일 큰 차이는 $\langle \mathbf{u}, \mathbf{v} \rangle = \overleftrightarrow{\langle \mathbf{v}, \mathbf{u} \rangle}$ 라는 것이고 나머지 bilinearity 등은 유사하다.

Appendix D

Covariance Matrices

모든 성분이 확률변수(random variable)인 벡터를 확률벡터(random vector)라 부른다. 실수를 성분으로 갖는 벡터처럼 확률변수 X_i 가 i 번째 성분이면 $\mathbf{X} = (X_1, \dots, X_d)^\top$ 로 d 차원 확률벡터를 표시한다. 확률벡터의 기댓값은 성분 확률변수의 기댓값이 성분인 \mathbb{R}^d -벡터를 의미한다. 즉

$$\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}, \mathbb{E}[\mathbf{X}] = \mathbb{E} \left[\begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix} \right] = \begin{pmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_d] \end{pmatrix} = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_d])^\top \in \mathbb{R}^d$$

이다. 일반적인 규칙은 각 성분이 확률변수인 벡터나 행렬의 기댓값은 성분 확률변수의 기댓값으로 구성된 동일한 크기의 벡터나 행렬을 의미한다. 불확실성에 노출된 관측값들을 기술하는 기본적인 통계량인 평균과 분산은 분포와 무관하게 중요하다. 평균과의 차이의 제곱을 평균하는 확률변수의 분산과 동일하게 확률벡터의 분산은 평균벡터와의 차이의 제곱을 평균한다. 확률벡터의 경우 제곱을 하기 위해서 곱해지는 두 성분 확률변수의 조합이 많아 표현이 복잡해지는 차이만 있다. 벡터로 표현하면 기댓값을 취하기 전 단계인 확률벡터의 제곱에 해당하는 양은 $d \times d$ random matrix인

$$(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top = ((X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j]))$$

로서 i, j 원소가 $(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])$ 이다. 이 random matrix는 기호적으로나 실현된 값이나 모두 rank-one matrix라는 특징이 있다. 그리고 한 개의 벡터로 만들어진 rank-one matrix이므로 $((\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top)^\top = (\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top$ 이어서 symmetric하고 당연히 기댓값을 취하더라도 symmetry는 유지된다. 기댓값이 취해지고 나면 성분 확률변수 간의 공분산효과로 rank가 높아진다. 이 random matrix의 기댓값이 공분산(covariance)을 나타내는 $d \times d$ 행렬이다. 공분산 행렬(covariance matrix)이라고 부르고

$$\Sigma := \text{Cov}(\mathbf{X}, \mathbf{X}) = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top]$$

로 표시한다. 이 symmetric matrix의 i, j 원소는 $\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])]$ 이다.

평균벡터가 영벡터인 경우 공분산을 포함한 많은 계산에서 여러모로 편리하다. 평균벡터가 $\mathbf{0}$ 이 아닌 경우에는 $\mathbf{X}' = \mathbf{X} - \mathbb{E}[\mathbf{X}]$ 로 놓으면 \mathbf{X}' 은 평균이 $\mathbf{0}$ 이고 \mathbf{X} 와 같은 공분산을 갖는 확률벡터가 된다. 이 경우, 공분산 행렬은 $\Sigma = \mathbb{E}[\mathbf{X}\mathbf{X}^\top]$ 으로 간략하게 표현된다.

D.1 Covariance Matrix of Random Vector

확률벡터 $\mathbf{X} = (X_1, \dots, X_d)^\top$ 의 공분산에 대해서 좀 더 자세하게 살펴보자. 표현을 간단하게 하기 위해 $\mu_i = \mathbb{E}[X_i]$ 로 표시하자. 그러면, 평균벡터로부터 변동의 제곱은 rank-one matrix^{o)}인

$$\begin{aligned} & (\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top \\ &= \begin{bmatrix} (X_1 - \mu_1)^2 & (X_1 - \mu_1)(X_2 - \mu_2) & \cdots & (X_1 - \mu_1)(X_d - \mu_d) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)^2 & \cdots & (X_2 - \mu_2)(X_d - \mu_d) \\ \vdots & \vdots & \ddots & \vdots \\ (X_d - \mu_d)(X_1 - \mu_1) & (X_d - \mu_d)(X_2 - \mu_2) & \cdots & (X_d - \mu_d)^2 \end{bmatrix} \end{aligned}$$

이 된다. 이 random matrix의 기댓값을 취하면

$$\begin{aligned} \Sigma &= \text{Cov}(\mathbf{X}, \mathbf{X}) = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top] \\ &= \mathbb{E} \left[\begin{bmatrix} (X_1 - \mu_1)^2 & (X_1 - \mu_1)(X_2 - \mu_2) & \cdots & (X_1 - \mu_1)(X_d - \mu_d) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)^2 & \cdots & (X_2 - \mu_2)(X_d - \mu_d) \\ \vdots & \vdots & \ddots & \vdots \\ (X_d - \mu_d)(X_1 - \mu_1) & (X_d - \mu_d)(X_2 - \mu_2) & \cdots & (X_d - \mu_d)^2 \end{bmatrix} \right] \\ &= \left[\begin{array}{cccc} \mathbb{E}[(X_1 - \mu_1)^2] & \mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)] & \cdots & \mathbb{E}[(X_1 - \mu_1)(X_d - \mu_d)] \\ \mathbb{E}[(X_2 - \mu_2)(X_1 - \mu_1)] & \mathbb{E}[(X_2 - \mu_2)^2] & \cdots & \mathbb{E}[(X_2 - \mu_2)(X_d - \mu_d)] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}[(X_d - \mu_d)(X_1 - \mu_1)] & \mathbb{E}[(X_d - \mu_d)(X_2 - \mu_2)] & \cdots & \mathbb{E}[(X_d - \mu_d)^2] \end{array} \right] \\ &= \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_d) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \cdots & \text{Cov}(X_2, X_d) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_d, X_1) & \text{Cov}(X_d, X_2) & \cdots & \text{Var}(X_d) \end{bmatrix} \end{aligned}$$

이 된다. 특히 $\mathbb{E}[\mathbf{X}] = \mathbf{0}$ 으로 평균이 $\mathbf{0}$ 인 경우에는 다음과 같이 공분산을

$$\begin{aligned} \Sigma &= \text{Cov}(\mathbf{X}, \mathbf{X}) = \mathbb{E}[\mathbf{X}\mathbf{X}^\top] \\ &= \mathbb{E} \left[\begin{bmatrix} X_1^2 & X_1X_2 & \cdots & X_1X_d \\ X_2X_1 & X_2^2 & \cdots & X_2X_d \\ \vdots & \vdots & \ddots & \vdots \\ X_dX_1 & X_dX_2 & \cdots & X_d^2 \end{bmatrix} \right] \\ &= \left[\begin{array}{cccc} \mathbb{E}[X_1^2] & \mathbb{E}[X_1X_2] & \cdots & \mathbb{E}[X_1X_d] \\ \mathbb{E}[X_2X_1] & \mathbb{E}[X_2^2] & \cdots & \mathbb{E}[X_2X_d] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}[X_dX_1] & \mathbb{E}[X_dX_2] & \cdots & \mathbb{E}[X_d^2] \end{array} \right] \end{aligned}$$

간략하게 표현할 수 있다.

D.1.1 Positive Definiteness of Covariance Matrices

d 차원 확률벡터 \mathbf{X} 와 임의의 $\mathbf{y} \in \mathbb{R}^d$ 에 대해서 $\mathbf{y}^\top \mathbf{X} = \mathbf{X}^\top \mathbf{y}$ 는 확률변수이다. 따라서

$$\mathbf{y}^\top \mathbf{X} \mathbf{X}^\top \mathbf{y} = (\mathbf{y}^\top \mathbf{X})(\mathbf{X}^\top \mathbf{y}) = (\mathbf{y}^\top \mathbf{X})^2$$

이다. 이를 이용하면

$$\mathbf{y}^\top \Sigma \mathbf{y} = \mathbf{y}^\top \mathbb{E}[\mathbf{X} \mathbf{X}^\top] \mathbf{y} = \mathbb{E}[\mathbf{y}^\top \mathbf{X} \mathbf{X}^\top \mathbf{y}] = \mathbb{E}[(\mathbf{y}^\top \mathbf{X})^2] \geq 0$$

이 된다. 따라서 공분산 행렬은 symmetric하고 positive semi-definite하다. 만약 \mathbf{X} 의 공분산 Σ 가 positive definite하지 않다면 어떤 \mathbf{y} 에 대해 $\mathbb{E}[(\mathbf{y}^\top \mathbf{X})^2] = 0$ 이라는 것이므로 $\mathbf{y}^\top \mathbf{X} = 0$ with probability 1이다. 즉 확률벡터 \mathbf{X} 가 linear dependence가 있는 것이고, 이 경우 \mathbf{X} 는 d 차원 확률밀도함수를 갖지 못한다.

D.1.2 A Useful Quadratic Identity

$n_1 \times n_1$, $n_1 \times n_2$, $n_2 \times n_2$ 행렬 A_{11} , A_{12} , A_{22} 로 이루어진 block matrix

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^\top & A_{22} \end{bmatrix}$$

과 n_1 차원 벡터 \mathbf{u}_1 , n_2 차원 벡터 \mathbf{u}_2 를 생각하자. 행렬 A , A_{11} , A_{22} 는 가역이며 대칭이라고 가정하자. $\mathbf{u} =$

$$\begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

$$\mathbf{u}^\top A^{-1} \mathbf{u} - \mathbf{u}_2 A_{22}^{-1} \mathbf{u}_2$$

를 계산해보자. (2.7)과 같이 block matrix의 역행렬을 구할 때 추가적으로 Schur complement

$$S_{11} = A_{11} - A_{12} A_{22}^{-1} A_{12}^\top$$

이 가역이라고 가정하면

$$A^{-1} = \begin{bmatrix} S_{11}^{-1} & -S_{11}^{-1} A_{12} A_{22}^{-1} \\ -A_{22}^{-1} A_{12}^\top S_{11}^{-1} & A_{22}^{-1} + A_{22}^{-1} A_{12}^\top S_{11}^{-1} A_{12} A_{22}^{-1} \end{bmatrix}$$

를 얻는다. 이 역행렬 표현을 이용하면

$$\begin{aligned} \mathbf{u}^\top A^{-1} \mathbf{u} &= \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}^\top \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^\top & A_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} \\ &= \mathbf{u}_1^\top S_{11}^{-1} \mathbf{u}_1 - 2\mathbf{u}_1^\top S_{11}^{-1} A_{12} A_{22}^{-1} \mathbf{u}_2 + \mathbf{u}_2^\top (A_{22}^{-1} + A_{22}^{-1} A_{12}^\top S_{11}^{-1} A_{12} A_{22}^{-1}) \mathbf{u}_2 \\ &= \mathbf{u}_1^\top S_{11}^{-1} \mathbf{u}_1 - 2\mathbf{u}_1^\top S_{11}^{-1} A_{12} A_{22}^{-1} \mathbf{u}_2 + \mathbf{u}_2^\top A_{22}^{-1} \mathbf{u}_2 + \mathbf{u}_2^\top A_{22}^{-1} A_{12}^\top S_{11}^{-1} A_{12} A_{22}^{-1} \mathbf{u}_2 \end{aligned}$$

이므로 구하려던 quadratic difference는

$$\mathbf{u}^\top A^{-1} \mathbf{u} - \mathbf{u}_2^\top A_{22}^{-1} \mathbf{u}_2$$

$$\begin{aligned}
&= \mathbf{u}_1^\top S_{11}^{-1} \mathbf{u}_1 - 2\mathbf{u}_1^\top S_{11}^{-1} A_{12} A_{22}^{-1} \mathbf{u}_2 + \mathbf{u}_2^\top A_{22}^{-1} A_{12}^\top S_{11}^{-1} A_{12} A_{22}^{-1} \mathbf{u}_2 \\
&= (\mathbf{u}_1 - A_{12} A_{22}^{-1} \mathbf{u}_2)^\top S_{11}^{-1} (\mathbf{u}_1 - A_{12} A_{22}^{-1} \mathbf{u}_2)
\end{aligned}$$

가 된다. 마지막 등식을 정리하면 다음과 같은 quadratic identity를 얻게된다.

$$\mathbf{u}^\top A^{-1} \mathbf{u} = \mathbf{u}_2^\top {A_{22}}^{-1} \mathbf{u}_2 + (\mathbf{u}_1 - A_{12} {A_{22}}^{-1} \mathbf{u}_2)^\top S_{11}^{-1} (\mathbf{u}_1 - A_{12} {A_{22}}^{-1} \mathbf{u}_2). \quad (\text{D.1})$$

이 quadratic identity는 multivariate Gaussian distribution의 conditional distribution을 얻을 때 이용된다.
(Appendix D.3 참조)

D.2 Multivariate Gaussian Distribution

\mathbf{X} 의 평균벡터를 $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}]$, 공분산 행렬을 $\boldsymbol{\Sigma} = \mathbb{E}[\mathbf{XX}^\top]$ 로 표시하자. 확률벡터 \mathbf{X} 를 기술하는 다변량 분포의 밀도함수 $f(\mathbf{x})$ 가

$$(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})$$

의 함수로¹ 표시되는 경우를 생각하자. 즉 양의 실수에 대해 정의된 함수 $g(\cdot)$ 에 대해

$$f(\mathbf{x}) = g((\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}))$$

로 표시되는 경우 이러한 다변량 분포들을 **elliptical distribution**이라 부르며 대표적인 예로 multivariate Gaussian distribution, multivariate t -distribution 등이 있다. g 가 감소함수이면 확률밀도함수의 level set인 $\{x \in \mathbb{R}^d : f(x) \geq a\}$ 는 상수항을 무시하면 적절한 α 에 대해서

$$\{\mathbf{x} \in \mathbb{R}^d : (\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \leq \alpha\}$$

이 되어 중심이 평균 벡터 μ 이고 covariance matrix Σ 에 의해 비율이 정해지는 ellipsoid가 된다 (Section 7.5 참조). Level set이 고차원 ellipsoid라는 사실은 data science에서 풍부한 아이디어를 제공하며 elliptical distribution이라는 이름의 연유다.

Σ 가 $d \times d$ positive definite matrix라고 가정하자. 이 경우, 다음과 같은 다중적분 결과가 알려져 있다.

$$\int_{\mathbb{R}^d} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})} d\mathbf{x} = \sqrt{(2\pi)^d \det \Sigma}.$$

따라서 \mathbb{R}^d 상에서 함수 $f_{\mu, \Sigma}$ 를 다음과 같이

$$f_{\boldsymbol{\mu}, \Sigma}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma}} e^{-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})}.$$

로 정의하면

$$f_{\mu, \Sigma}(\mathbf{x}) > 0 \text{ } \circ] \text{과 } \int_{\mathbb{R}^d} f_{\mu, \Sigma}(\mathbf{x}) d\mathbf{x} = 1$$

이어서 f 가 \mathbb{R}^d 에서 값을 갖는 어떤 다변량분포의 확률밀도함수가 되며 이 분포를 **multivariate Gaussian distribution**이라고 부른다. 이 분포의 평균벡터는 μ 이고, 공분산 행렬은 Σ 이어서 $N(\mu, \Sigma)$ 로 표시한다.

¹Positive definite matrix Σ 와 벡터 μ 에 대해 $\sqrt{(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})}$ 을 Section 7.5에서 정의된 것처럼 \mathbf{x} 의 $\boldsymbol{\mu}$ 로부터의 Mahalanobis 거리라 부른다. Mahalanobis 거리는 표준편차로 normalize된 통계적 거리로 이해하면 된다.

다면량분포이므로 일부 확률변수에 대한 조건분 확률분포를 밀도함수가 종종 필요한데 Appendix D.3에서 본문의 여러 결과를 활용하여 얻는다. 또한 여러 기대값 계산을 위해 multivariate Gaussian분포의 random sample^{o]} 사용되는데 Appendix D.4에 서술된 Cholesky분해에서 얻어지는 상삼각행렬을 이용하는 방법이 대표적인 sampling 방법이다.

D.3 Conditional Multivariate Gaussian Distribution

\mathbf{X}_1 은 n_1 차원 확률벡터이고 \mathbf{X}_2 는 n_2 차원 확률벡터라고 하고 평균벡터 $\boldsymbol{\mu}_1 \in \mathbb{R}^{n_1}$ 과 $\boldsymbol{\mu}_2 \in \mathbb{R}^{n_2}$ 를 생각하자. 각각의 covariance는 $\Sigma_{11} = \mathbb{E}[(\mathbf{X}_1 - \boldsymbol{\mu}_1)(\mathbf{X}_1 - \boldsymbol{\mu}_1)^\top]$, $\Sigma_{12} = \mathbb{E}[(\mathbf{X}_1 - \boldsymbol{\mu}_1)(\mathbf{X}_2 - \boldsymbol{\mu}_2)^\top]$, $\Sigma_{22} = \mathbb{E}[(\mathbf{X}_2 - \boldsymbol{\mu}_2)(\mathbf{X}_2 - \boldsymbol{\mu}_2)^\top]$ 로 표시한다. 그러면 $\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$, $\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$ 일 때, \mathbf{X} 의 covariance는

$$\Sigma = \mathbb{E}[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^\top] = \mathbb{E} \left[\begin{bmatrix} \mathbf{X}_1 - \boldsymbol{\mu}_1 \\ \mathbf{X}_2 - \boldsymbol{\mu}_2 \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 - \boldsymbol{\mu}_1 \\ \mathbf{X}_2 - \boldsymbol{\mu}_2 \end{bmatrix}^\top \right] = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

인데 $\Sigma_{21} = \Sigma_{12}^\top$ 이다. $\mathbf{u}_1 = \mathbf{X}_1 - \boldsymbol{\mu}_1$, $\mathbf{u}_2 = \mathbf{X}_2 - \boldsymbol{\mu}_2$, $A_{ij} = \Sigma_{ij}$ 로 놓고 (D.1)를 적용하면

$$\bar{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^\top$$

일 때

$$(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) = (\mathbf{x}_2 - \boldsymbol{\mu}_2)^\top \Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) + (\mathbf{x}_1 - (\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)))^\top \bar{\Sigma}^{-1}(\mathbf{x}_1 - (\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)))$$

를 얻는다. 다변량 정규분포의 conditional density를 구하기 위해서는 $\frac{f_{\boldsymbol{\mu}, \Sigma}(\mathbf{x})}{f_{\boldsymbol{\mu}_2, \Sigma_{22}}(\mathbf{x}_2)}$ 를 계산하면 된다. 먼저 상수 항 부분의 정리를 위해서 (8.3)을 이용하면 $\det \Sigma = \det \Sigma_{22} \det \bar{\Sigma}$ 를 얻는다. 따라서 conditional multivariate Gaussian density는

$$\begin{aligned} \frac{f_{\boldsymbol{\mu}, \Sigma}(\mathbf{x})}{f_{\boldsymbol{\mu}_2, \Sigma_{22}}(\mathbf{x}_2)} &= \frac{\sqrt{(2\pi)^{n_2} \det \Sigma_{22}}}{\sqrt{(2\pi)^{n_1+n_2} \det \Sigma}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}) + \frac{1}{2}(\mathbf{x}_2-\boldsymbol{\mu}_2)^\top \Sigma_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)} \\ &= \frac{1}{\sqrt{(2\pi)^{n_1} \det \bar{\Sigma}}} e^{-\frac{1}{2}(\mathbf{x}_1 - (\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)))^\top \bar{\Sigma}^{-1}(\mathbf{x}_1 - (\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)))} \end{aligned}$$

인데 이 분포는 평균^{o]} $\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$ 과 covariance^{o]} $\bar{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^\top$ 인 multivariate Gaussian distribution의 density이다. 보통 $\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$ 를 conditional mean, $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^\top$ 를 conditional covariance라고 부른다.

D.4 Multivariate Gaussian Sampling using Cholesky Decomposition

확률변수를 sampling하는 방법은 오랫동안 연구된 분야로 다양한 방법이 알려져 있다. 큰 소수 p 에 대해서 $\{1, 2, \dots, p-1, p\}$ 의 p 개 자연수를 동일한 빈도로 sampling하는 (discrete uniform sampling) 방법이 가장 근본적인 알고리듬이다. 이렇게 얻어진 샘플을 p 로 나누면 $[0, 1]$ 에서의 uniform sample^{o]} 된다. 이 [0, 1]-uniform

sample을 누적화률함수(cumulative distribution function)의 역함수 F^{-1} 와 합성하면 분포함수가 F 인 확률변수의 샘플이 된다. 이런 방법으로 쉽게 샘플이 얻어지는 분포의 대표적인 예가 exponential distribution이다. 하지만, F 가 다루기 어렵거나 correlation이 존재하는 확률벡터에서의 sampling은 어려운 경우가 많다. Gaussian distribution의 경우 분포함수가 다루기 어렵지만 워낙 중요한 분포여서 많은 라이브러리가 random number를 생성하는 기능을 제공한다.

본 절에서는 $N(0, 1)$ Gaussian 분포를 따르는 독립적인 random number를 원하는 개수만큼 얻을 수 있다고 가정할 때 $N(\mu, \Sigma)$ 의 샘플을 어떻게 얻을 수 있는지 살펴본다. 정규분포를 따르는 확률변수들은 이 변수들의 선형조합이 다시 정규분포를 따른다는 유용한 성질을 가지고 있다. 먼저 공분산 행렬 Σ 에 대해 $\Sigma = R^\top R$ 으로 Cholesky decomposition을 얻는다. 여기서 R 은 upper triangular matrix이다. $N(0, 1)$ 확률변수는 독립적으로 sampling할 수 있다고 가정했으므로 이렇게 얻어진 d 개의 샘플을 성분으로 하는 벡터는 $N(\mathbf{0}, I)$ 에서 독립적으로 sampling된 sample vector이다. 이 d 차원 random sample vector를 $\mathbf{Z} \sim N(\mathbf{0}, I)$ 로 표시하자. 다음 단계로 d 차원 확률벡터

$$\mathbf{X} = R^\top \mathbf{Z}$$

를 계산한다. 이 확률벡터 \mathbf{X} 는

- 앞서 언급한대로 Gaussian random vector이고;
- 평균은 $\mathbb{E}[\mathbf{X}] = R^\top \mathbb{E}[\mathbf{Z}] = R^\top \mathbf{0} = \mathbf{0}$ 이고;
- 공분산은 $\mathbb{E}[\mathbf{Z}\mathbf{Z}^\top] = I$ 로부터

$$\mathbb{E}[\mathbf{X}\mathbf{X}^\top] = \mathbb{E}[R^\top \mathbf{Z}\mathbf{Z}^\top R] = R^\top \mathbb{E}[\mathbf{Z}\mathbf{Z}^\top] R = R^\top I R = \Sigma \text{이다.}$$

따라서 생성된 \mathbf{X} 가 Gaussian distribution $N(\mathbf{0}, \Sigma)$ 로부터 독립적으로 얻어진 샘플이다.

D.5 Ill-conditioned Sample Covariance Matrices

통계학이나 기계학습에서 관측된 샘플로부터 추정된 공분산 행렬의 역행렬을 이용해야 하는 경우가 많이 있다. 즉 관심있는 확률벡터의 샘플 벡터들이 $\mathbf{x}_1, \dots, \mathbf{x}_n$ 으로 주어지면 covariance matrix의 불편추정량(unbiased estimator)은

$$\hat{\Sigma} = \frac{1}{n-1} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top$$

인데, 실제 데이터로 계산된 경우에는 $\hat{\Sigma}$ 의 역행렬이 존재하지 않는 경우가 종종 발생한다. 이 경우, 적절하게 선택된 작은 양수 ϵ 를 사용하여 공분산 행렬을 $\hat{\Sigma} + \epsilon I$ 로 perturb하여² 사용한다. Fact 7.7을 활용하거나, 또는 eigenpair에 대해 간단하게 계산해보면 이 변형에 의해 $\hat{\Sigma}$ 의 모든 eigenvalue가 정확하게 ϵ 만큼 증가하는 것을 알 수 있다.

² 대부분의 경우에 0.01정도의 ϵ 으로 perturb된 행렬이 역행렬을 갖는다.

Index

- adjoint, 175
- algebraic multiplicity, 166
- argmax, argmin, 80
- back-substitution, 11
- basic vector, 34
- basis, 34
- block matrix, 11, 20
- Brauer theorem, 187
- Cauchy-Schwarz inequality, 56
- Cayley-Hamilton theorem, 189
- change of basis, 162
- Cholesky decomposition, 126
- cofactor expansion, 146
- column space, 26
- condition number, 93
- conditional covariance, 229
- conditional mean, 229
- cone, 181
- congruence transformation, 179
- convex, 181
- convex cone, 181
- cosine similarity, 61
- covariance, 225
- Cramer's rule, 152
- cyclic subspace, 193
- determinant, 140, 185
- diagonal, 6
- diagonalizable, 164
- dimension, 35
- direct sum, 25
- dual basis, 174
- dual problem, 174
- dual space, 173
- dual variable, 174
- echelon form, 28
- eigenvalue, 93, 157
- eigenvalue adjustment, 187
- eigenvalue interlacing, 132
- eigenvector, 93, 157
- elementary matrix, 10
- ellipsoid, 133
- elliptical distribution, 228
- free variable, 29
- Frobenius norm, 77
- fundamental theorem of symmetric matrices, 166
- Gaussian elimination, 10
- Gaussian kernel, 136
- generalized eigenvector, 194
- generalized projection, 103
- geometric multiplicity, 166
- Google matrix, 188
- Gram-Schmidt Procedure, 70, 71
- Hermitian, 160, 223
- Householder matrix, 67
- idempotent, 177
- identity, 6
- inner product, 56

- inverse, 13, 20
- isometry, 76
- joint diagonalization, 179
- Jordan (normal) form, 197
- Jordan block, 196
- Jordan normal form theorem, 197
- kernel trick, 134
- latent space, 112
- least square, 80
- left singular vector, 88
- linear combination, 25
- linear functional, 173
- linear transformation, 44
- linearly dependent, 32
- linearly independent, 32
- low rank approximation, 98
- lower triangular, 10
- LU-factorization, 18
- Mahalanobis distance, 134
- Markov matrix, 185
- matrix, 1, 2
- matrix determinant lemma, 149
- matrix exponential, 169
- matrix norm, 77
- minimax principle, 129
- MNIST, 114
- multivariate Gaussian distribution, 228
- nilpotent, 190
- norm, 56
- null space, 26
- one-hot vector, 3
- orthogonal, 62, 64
- orthogonal complement, 63
- orthogonal matrix, 75
- orthogonally diagonalizable, 166
- orthonormal, 62
- PageRank, 170
- partial order, 181
- Penrose identities, 101
- Perron-Frobenius theorem, 186, 187
- pivot, 28
- pivot variable, 29
- polynomial kernel, 135
- positive definite, 59, 123
- positive semi-definite, 123
- principal component, 106
- principal components analysis (PCA), 106
- projection, 48, 68, 69, 103, 104
- pseudoinverse, 101
- QR-decomposition, 76
- quasi-Newton method, 154
- random vector, 225
- rank, 36
- rank-one, 42
- Rayleigh quotient, 127
- real spectral decomposition, 94, 167
- reduced row echelon form, 28
- reflection matrix, 49
- right singular vector, 88
- rotation matrix, 47
- row echelon form, 28
- Schur complement, 19, 149, 183, 227
- Schur triangularization, 184
- self-adjoint, 175
- Sherman-Morrison formula, 153
- similar, 164
- singular value, 88
- singular value decomposition (SVD), 90
- singular vector, 88

- span, 33
- spectral decomposition, 167
- spectral norm, 78
- spectral radius, 181
- subspace, 25
- symmetric, 7
- symmetric positive definite kernel, 135
- symmetric rank-one update, 153
- symmetric sum, 182
- symmetrization of matrix, 97
- trace, 77, 185
- transpose, 6
- triangularization, 184
- unit vector, 56
- unitary matrix, 184
- upper triangular, 10
- vector, 2
- vector space, 24
- volume of ellipsoid, 133, 161
- volume of parallelogram, 150
- Weyl's inequality, 130
- Woodbury formula, 153
- word embedding, 118