Chap13

고유벡터 - EigenVector

13.1 ~ 13.2 생략

13.3 고유값과 고유벡터

Definition : 정방행렬 A에 대하여, 스칼라(scalar)인 λ와 영이 아닌 벡터 v에 대해 $Av = \lambda v$가 만족하는 경우, λ는 A의 *고유값(eigenvalue)*, v 는 대응하는 *고유벡터(eigenvector)* 라고 한다.

만약, λ가 행렬 A의 고유값이면, 대응하는 고유벡터는 무수히 많다. 집합 $\{v : Av = \lambda v\}$ 는 벡터공간이며 고유값 λ에 대응하는 *고유공간(eigenspace)* 이라 한다. 따라서, 고유공간에 있는 임의의 영이 아닌 벡터는 고유벡터로 간주된다. 일반적으로 고유벡터의 크기($norm$)가 1이라는 제한을 두는 것이 다루기에 편리하다.

Example 13.3.3 : 행렬 A를 아래와 같은 대각행렬이라 하자.

$$A = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

행렬 A에 대한 고유벡터와 고유값은 무엇일까? 표준 기저 벡터 $e_1, \ldots, e_n$ 에 대해, $Ae_1 = \lambda_1, \ldots, Ae_n = \lambda_n e_n$ 이므로, $e_1, \ldots, e_n$ 은 고유벡터이고 대각원소인 $\lambda_1, \ldots, \lambda_n$ 은 고유값이다.

Example 13.3.5 : 행렬 A의 한 고유값은 0이라고 하자. 이 고유값에 대응하는 고유벡터는 $Av = 0v$를 만족하는 영이 아닌 벡터 v 이다. 즉, 벡터 v 는 Av가 영벡터가 되게 하는 영이 아닌 벡터이며, v 는 영공간(null space)에 속한다. 역으로, 만약 A의 영공간이 자명하지 않으면 0은 A의 고유값이다.

위의 Example 13.3.5는 고유값 0에 대응하는 고유벡터(즉, 영공간에 속하는 영이 아닌 벡터)를 찾는 방법에 대한 설명이다.
행렬 A 의 고유값을 λ, 대응하는 고유벡터를 v라고 하면, $Av = \lambda v$ 이다. 따라서, $Av - \lambda v = 0$ 이다.
$Av - \lambda v = (A - \lambda I)v$ 이므로 $(A - \lambda I)v$ 는 영벡터이다. 이것은 벡터 v가 $A - \lambda I$의 영공간에 속하는 영이 아닌 벡터임을 의미한다. 따라서, $A - \lambda I$는 비가역적이다.

Corollary : 만약 λ 가 행렬 A의 고유값일 경우 λ 는 또한 A^T 의 고유값이다.

- Proof : λ를 행렬 A의 고유값이라 하면, $A - \lambda I$는 비가역적이다. 7.4.7에 의하면 $(A - \lambda I)^T$ 또한 비가역적이다. $(A - \lambda I)^T = A^T - \lambda I$ 이므로, λ는 A^T의 고유값이다.

python에서 `numpy` 모듈의 `numpy.linalg.eig()` 를 이용하여 고유값과 고유벡터를 구할 수 있다.

```
import numpy as np

A = np.matrix([[1, 2],
               [3, 4]])

w, v = np.linalg.eig(A)

print('eigenvalue =', w)
print('eigenvector =\n', v)

'''출력결과
eigenvalue = [-0.37228132  5.37228132]
eigenvector =
 [[-0.82456484 -0.41597356]
 [ 0.56576746 -0.90937671]]
'''
```

13.3.1 유사성과 대각화 가능성 - Diagonalizability

Definition : 가역행렬 S에 대해 $S^{-1}AS = B$ 가 만족되면 두 정방행렬 A와 B는 '유사' 또는 '닮은(similar)' 행렬이라고 한다.

Proposition : 유사행렬(similar matrix)들은 동일한 고유값을 가진다.

- Proof : λ 를 행렬 A의 고유값이라 하고, v 를 고유벡터라고 하면, $Av = \lambda v$ 가 성립한다. $S^{-1}AS = B$라 하고, $w = S^{-1}v$라 하면 다음이 성립한다.

$$\begin{aligned} Bw &= S^{-1}ASw \\ &= S^{-1}ASS^{-1}v \\ &= S^{-1}Av \\ &= S^{-1}\lambda v \\ &= \lambda S^{-1}v \\ &= \lambda w \end{aligned}$$

- 따라서, λ 는 행렬 B의 고유값이다.

Example 13.3.11 : 뒤에서 다루겠지만, 행렬 U는 상삼각행렬로, 행렬 $A = \begin{bmatrix} 6 & 3 & -9 \\ 0 & 9 & 15 \\ 0 & 0 & 15 \end{bmatrix}$ 의 고유값은 대각원소들인 $6, 9, 15$ 이다. 행렬 $B = \begin{bmatrix} 92 & -32 & -15 \\ -64 & 34 & 39 \\ 176 & -68 & -99 \end{bmatrix}$ 는 $B = S^{-1}AS$ 인 성질을 가진다. 여기서, $S = \begin{bmatrix} -2 & 1 & 4 \\ 1 & -2 & 1 \\ -4 & 3 & 5 \end{bmatrix}$ 이다. 따라서, B의 고유값 또한 $6, 9, 15$ 이다.

Definition : 만약 어떤 정방행렬 A가 대각행렬과 유사행렬이면, 즉 대각행렬 Λ에 대해 $S^{-1}AS = \Lambda$를 만족하는 가역행렬 S가 있으면, A는 *'대각화가 가능하다(diagonalizable)'* 라고 한다.

만약 Λ 가 대각행렬 $\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ 이면, 이 행렬의 고유값은 그 대각원소들인 $\lambda_1, \ldots, \lambda_n$ 이다. 만약 행렬 A와 Λ가 유사행렬이면, 위의 Proposition에 의해 A의 고유값은 Λ의 고유값 즉, Λ의 대각원소들이다.

Lemma : 만약 $\Lambda = S^{-1}AS$ 가 대각행렬이면, Λ 의 대각원소들은 고유값들이고, S의 열들은 선형독립인 고유벡터들이다.

$n \times n$ 행렬 A가 n개의 선형독립인 고유벡터 $v_1, \ldots, v_n$을 가진다고 하고, $\lambda_1, \ldots, \lambda_n$은 대응하는 고유값들이라 하자. 행렬 S를 $\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}$ 로 나타내고, Λ를 행렬 $\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ 라고 하면, $AS = S\Lambda$이다. S는 정방행렬이고, 그 열들은 선형독립이므로 가역행렬이다. 위의 식에서 오른쪽에 S^{-1} 을 곱하면 $A = S\Lambda S^{-1}$이 구해진다. 이것은 A가 대각화가 가능하다는 것을 보여주며, 아래 lemma로 정리할 수 있다.

Lemma : 만약 $n \times n$ 행렬 A가 n개의 선형독립인 고유벡터를 가지면 A는 대각화가 가능하다.

Theorem : $n \times n$ 행렬이 대각화 가능할 필요충분조건은 이 행렬이 n개의 선형독립인 고유벡터를 가지는 것이다.

13.4 고유벡터에 대한 좌표표현

A를 $n \times n$ 행렬이라 하고, $t = 1, 2, \ldots$에 대해, $x^{(t)} = A^t x^{(0)}$ 라 하자. 또한, 행렬 A 는 대각화 가능하다고 가정하자. 즉, $S^{-1}AS = \Lambda$를 만족하는 가역행렬 S와 대각행렬 Λ가 존재한다. $\lambda_1, \ldots, \lambda_n$ 은 Λ의 대각원소라 하자. 이 대각원소들은 A의 고유값이다. 그리고 $v_1, \ldots, v_n$은 고유벡터들이며 행렬 S의 열이다. 고유벡터들에 대한 $x^{(t)}$ 의 좌표표현을 $u^{(t)}$라고 하면, $x^{(t)} = A^t x^{(0)}$ 는 훨씬 단순한 형태로 표현된다.

$$\begin{bmatrix} \\ u^{(t)} \end{bmatrix} = \begin{bmatrix} \lambda_1^t & & \\ & \ddots & \\ & & \lambda_n^t \end{bmatrix} \begin{bmatrix} \\ u^{(0)} \end{bmatrix}$$

위 식이 단순한 이유는 $u^{(0)}$ 의 해당 원소에 대응하는 고유값의 t 제곱을 곱하면 $u^{(t)}$의 각 원소가 구해지기 때문이다.

이것을 다른 각도에서 한 번 살펴보자.

고유벡터들은 $\mathbb{R}^n$에 대한 기저를 형성한다. 따라서, 임의의 벡터 x는 고유벡터들의 선형결합으로 나타낼 수 있다.

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

위 식에서 양변의 왼쪽에 A를 곱해보자.

$$\begin{aligned} Ax &= A(\alpha_1 v_1) + \cdots + A(\alpha_n v_n) \\ &= \alpha_1 A v_1 + \cdots + \alpha_n A v_n \\ &= \alpha_1 \lambda_1 v_1 + \cdots + \alpha_n \lambda_n v_n \end{aligned}$$

같은 방식으로 $A(Ax)$를 계산하면, 다음과 같이 된다.

$$A^2x = \alpha_1\lambda_1^2 v_1 + \cdots + \alpha_n\lambda_n^2 v_n$$

이를 좀더 일반적인, 임의의 음이 아닌 정수 t에 대해 다음과 같이 쓸 수 있다.

$$A^t x = \alpha_1\lambda_1^t v_1 + \cdots + \alpha_n\lambda_n^t v_n$$

이제, 어떤 고유값의 절대값이 다른 것들보다 약간이라도 큰 경우를 생각해 보자. 이때, t가 충분히 클 경우, 위의 식에서 우변은 절대값이 큰 고유값이 포함된 항에 의해 결정되고 다른 항들은 상대적으로 작은 값이 될 것이다.

특히, λ_1의 절대값이 다른 모든 고유값보다 크다고 가정해 보자. 이 경우, t가 충분히 크다면 $A^t x \approx \alpha_1\lambda_1^t v_1$ 이 될 것이다. 실제로, 절대값이 1보다 작은 고유값에 대응하는 항은 t가 증가함에 따라 그 값이 점점 작아지게 된다.

13.5 생략

13.6 고유값의 존재

어떠한 상황에서 정방행렬이 고유값을 가지는지 알 수 있을까? 또한 대각화가 가능할까?

13.6.1 양의 정부호(Positive-Definite) 행렬과 양의 준정부호(Positive-Semidefinite) 행렬

A를 임의의 가역행렬이라고 하면, 이 행렬에 12장에서 배운 특이값 분해(SVD) 를 적용하면 다음과 같다.

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

위의 식에서 양변 왼쪽에 $A^T = \left(U\Sigma V^T\right)^T = V\Sigma U^T$ 를 곱하면 다음 식이 얻어진다.

$$\begin{aligned} A^T A &= V\Sigma U^T U\Sigma V^T \\ &= V\Sigma\Sigma V^T \\ &= V\Sigma^2 V^T \end{aligned}$$

위의 식에서 왼쪽에 V^T를 곱하고 오른쪽에 V를 곱하면, 다음 식이 얻어진다.

$$V^T\left(A^T A\right)V = \Sigma^2$$

여기서, $A^T A$는 대각화가 가능하고 고유값은 A의 특이값(singular value)의 제곱이다. 이 고유값들은 모두 양의 실수이다.

또한, $A^T A$는 아래의 식에서 알 수 있듯이 대칭행렬이다.

$$\left(A^T A\right)^T = A^T\left(A^T\right)^T = A^T A$$

Definition : 고유값이 모두 양의 실수인 대칭행렬은 *양의 정부호행렬* 이라 한다.

13.6.2 고유값이 모두 다른 행렬

정방행렬을 대각화가 가능하게 하는 또 다른 조건에 대해 알아보자.

Lemma : 행렬 A 의 모두 다른(distinct) 고유값으로 이루어진 임의의 집합 T에 대해, 대응하는 *고유벡터들은 선형독립* 이다.

- Proof : 고유벡터들이 선형종속이라고 가정해 보자.

$$0 = \alpha_1 v_1 + \cdots + \alpha_r v_r \quad \rightarrow (1)$$

- 위 식은 T에 속하는 고유값들에 대응하는 고유벡터들의 부분집합으로 구성된 선형결합이며, 특히 최소 크기의 부분집합으로 된 선형결합이라고 하자. $\lambda_1, \ldots, \lambda_r$은 대응하는 고유값이라 하면, 다음이 성립한다.

$$\begin{aligned} 0 &= A(0) \\ &= A(\alpha_1 v_1 + \cdots + \alpha_r v_r) \\ &= \alpha_1 A v_1 + \cdots + \alpha_r A v_r \\ &= \alpha_1 \lambda_1 v_1 + \cdots + \alpha_r \lambda_r v_r \end{aligned} \quad \rightarrow (2)$$

- 위의 식처럼 $v_1, \ldots, v_r$ 사이에 새로운 선형종속이 얻어진다. λ_1 을 식(1)에 곱한 뒤 식(2)에서 빼면 다음 식을 얻는다.

$$0 = (\lambda_1 - \lambda_1)\alpha_1 v_1 + (\lambda_2 - \lambda_1)\alpha_2 v_2 + \cdots + (\lambda_r - \lambda_1)\alpha_r v_r$$

- 위의 식에서 첫 번째 계수는 0 이므로 다음과 같이 쓸 수 있다.

$$0 = (\lambda_2 - \lambda_1)\alpha_2 v_2 + \cdots + (\lambda_r - \lambda_1)\alpha_r v_r$$

- 이 식은 식(1)보다 더 작은 수의 벡터를 가지는데, 이것은 우리가 처음에 가정한 최소 크기의 선형결합이라는 가정에 모순된다.

Theorem : n 개의 모두 다른 고유값을 가지는 $n \times n$ 행렬은 대각화 가능하다.

랜덤한 원소를 가지는 $n \times n$ 행렬은 n 개의 모두 다른 고유값을 가질 가능성이 높다. 따라서, 위의 Theorem은 *대부분*의 정방행렬은 대각화 가능하다는 것을 의미한다. 또한, $n \times n$ 행렬 중 n 개의 모두 다른 고유값을 가지지는 않지만 대각화가 가능한 것이 있다. 바로 $n \times n$ 단위행렬이다. 단위행렬의 고유값은 모두 1 이지만 대각화 가능하다.

13.6.3 대칭행렬 - Symmetric matrices

Theorem(대칭행렬의 대각화) : A를 $\mathbb{R}$ 상의 대칭행렬이라 하면, $Q^T A Q = \Lambda$ 를 만족하는 직교행렬 Q와 실수값(real-valued) 대각행렬 Λ 가 존재한다.

13.6.4 상삼각(Upper-triangular) 행렬

모든 정방행렬이 대각화 되는것은 아니다. 간단한 예로, 행렬 $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$은 대각화 되지 않는다.

Lemma : 상삼각 행렬 U의 대각원소는 U의 고유값이다.

- Proof : 어떤 수 λ 가 행렬 U의 고유값이 될 필요충분 조건은 $U - \lambda I$가 비가역적인 경우이다. $U - \lambda I$는 상삼각행렬이므로 5.13.3에 의해 $U - \lambda I$ 가 비가역적일 필요충분조건은 대각원소 중 적어도 하나가 영인 경우이다. 따라서, $U - \lambda I$의 대각원소가 영일 필요충분조건은 λ가 U의 대각원소 중 하나인 경우이다.

Example 13.6.6 : 행렬 $U = \begin{bmatrix} 5 & 9 & 9 \\ 0 & 4 & 7 \\ 0 & 0 & 3 \end{bmatrix}$이 행렬의 대각원소는 $5, 4, 3$ 이고 이 대각원소들은 이 행렬의 고유값이다.

U의 대각원소로 하나의 값이 여러 번 사용될 수 있다. 예를 들어, 행렬 $U = \begin{bmatrix} 5 & 9 & 9 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{bmatrix}$ 에서 5는 대각원소로서 2번 사용된다.

Definition : 상삼각행렬 U의 스펙트럼(spectrum, 고유값들의 집합)은 대각원소들의 중복집합(multiset)이다. 중복집합에 포함되는 각각의 수는 U의 대각원소에 나타나는 횟수만큼 나타난다.

Example 13.6.8 : $\begin{bmatrix} 5 & 9 & 9 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{bmatrix}$의 스펙트럼은 중복집합 $\{5, 5, 4\}$ 이다.

13.7 ~ 13.9 생략

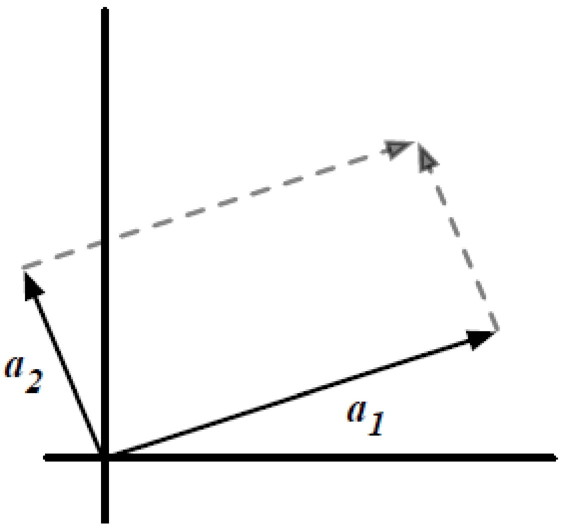
13.10 행렬식 - Determinant

행렬식을 이용하여 2×2 행렬의 행렬식에 기반을 둔 계산 기법을 사용하여 다각형의 면적을 계산하는 예를 살펴보자.

13.10.1 평행사변형의 면적

A를 2×2 행렬이라하고, 이 행렬의 열 a_1, a_2는 직교한다고 하자. 이때 다음 직사각형의 면적을 구해보자.

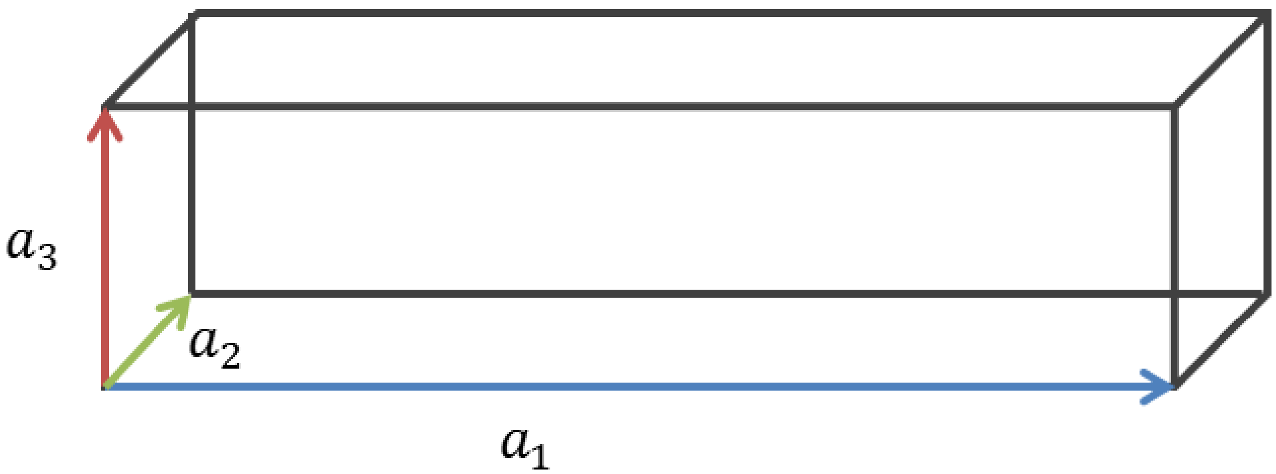
$$\{\alpha_1 a_1 + \alpha_2 a_2 : 0 \leq \alpha_1, \alpha_2 \leq 1\}$$



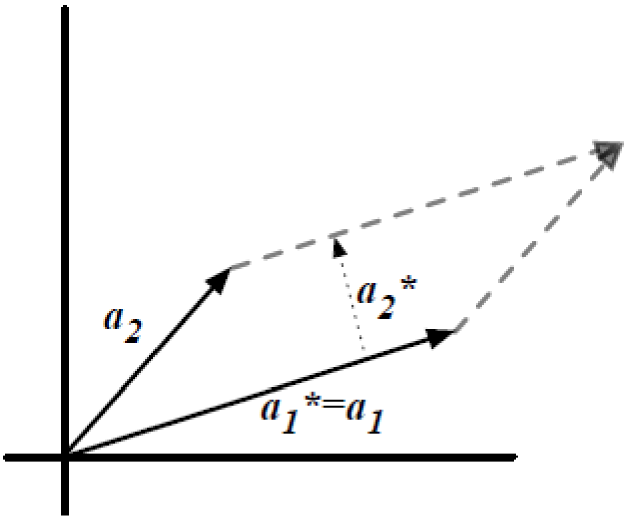
직사각형의 면적은 두 변의 길이의 곱이므로 $\|a_1\| \times \|a_2\|$ 이다.

좀 더 일반적으로, A는 $n \times n$ 행렬이고, 이 행렬의 열 $a_1, \ldots, a_n$은 직교한다고 하자. 이 경우, hyperrectangle의 부피는 n개 변의 길이의 곱, 즉 $\|a_1\| \times \|a_2\| \times \cdots \times \|a_n\|$ 이다.

$$\{\alpha_1 a_1 + \cdots + \alpha_n a_n : 0 \leq \alpha_1, \ldots, \alpha_n \leq 1\}$$



그렇다면, 이제 직교한다는 가정을 없애보자. a_1, a_2는 직교하지 않으므로 아래와 같이 평행사변형이 된다. 이 평행사변형의 면적은 밑변의 길이와 높이의 곱이다.

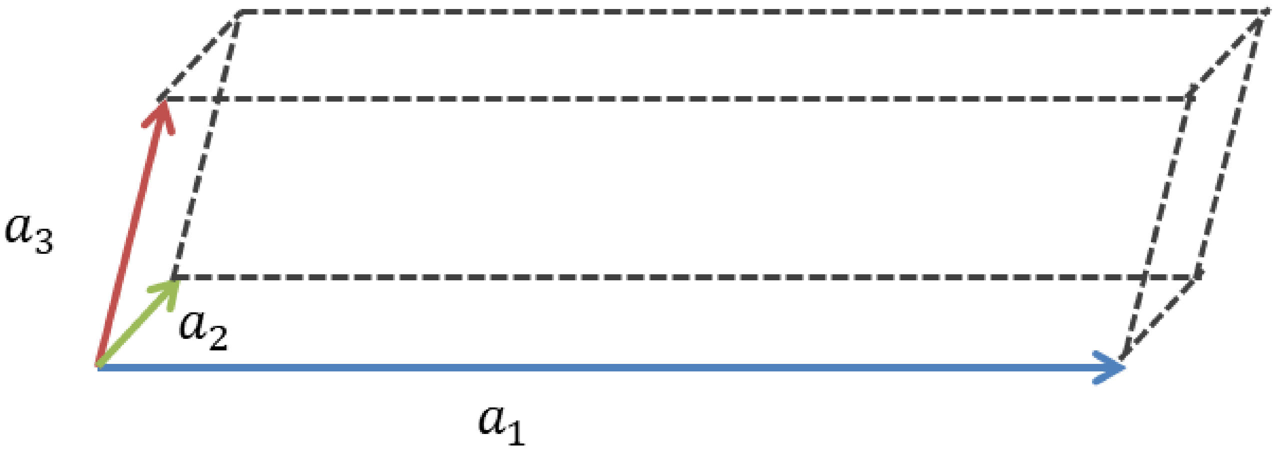


$a_1^* = a_1$ 이라 하고, a_1 을 평행사변형의 밑변으로 간주하자. 높이 a_2^* 는 a_2 의 투영이며 a_1^*과 직교한다. 따라서 면적은 $\|a_1^*\| \times \|a_2^*\|$ 가 된다.

13.10.2 평행육면체(Parallelepiped)의 부피

13.10.1 과 동일한 방식으로 생각해보면, $a_1, \ldots, a_n$ 을 n개의 벡터라 했을 때 아래의 집합은 평행육면체(parallelepiped) 형태가 된다.

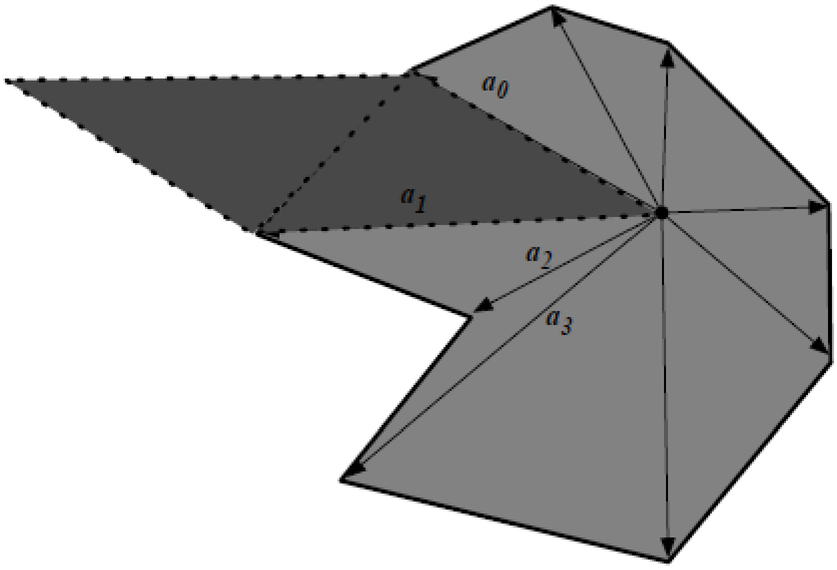
$$\{\alpha_1 a_1 + \cdots + \alpha_n a_n : 0 \leq \alpha_1, \cdots, \alpha_n \leq 1\}$$



이 평행육면체의 부피는 열벡터들을 직교화해 $a_1^*, \ldots, a_n^*$ 를 구하고, 이들의 길이를 곱하면 면적을 구할 수 있다.

13.10.3 평행사변형 면적을 이용한 다각형의 면적 표현

아래의 다각형의 면적을 계산해보자.

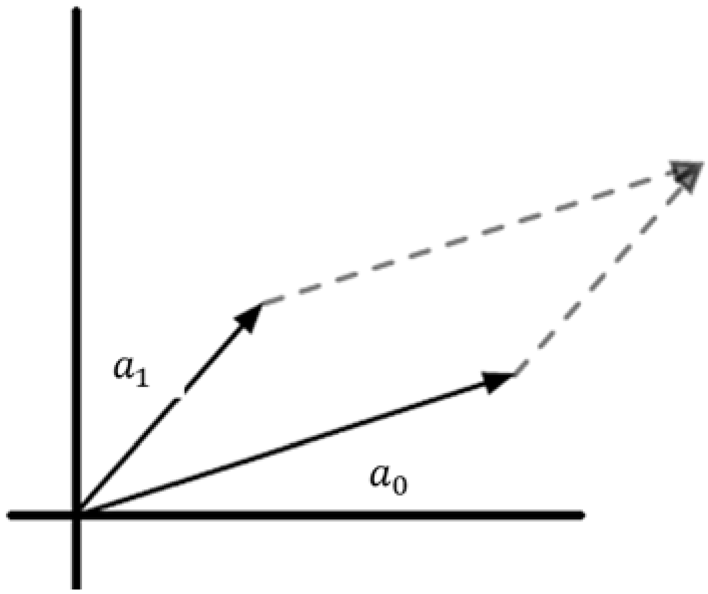


$a_0, \ldots, a_{n-1}$ 을 (x, y) 쌍으로 표현한 다각형의 꼭지점이라 하자. 위의 그림에서 점은 원점의 위치를 나타낸다.

다각형의 면적은 n개의 삼각형의 면적으로 나타낼 수 있다.

- a_0와 a_1로 형성된 삼각형
- a_1과 a_2로 형성된 삼각형
- $\vdots$
- a_{n-2}와 a_{n-1}로 형성된 삼각형
- a_{n-1}와 a_0로 형성된 삼각형

a_0와 a_1로 형성된 삼각형을 복사하여 원래 삼각형에 붙이면 아래와 같이 편행사변형 $\{\alpha_0 a_0 + \alpha_1 a_1 : 0 \le \alpha_0, \alpha_1 \le 1\}$ 이 된다.

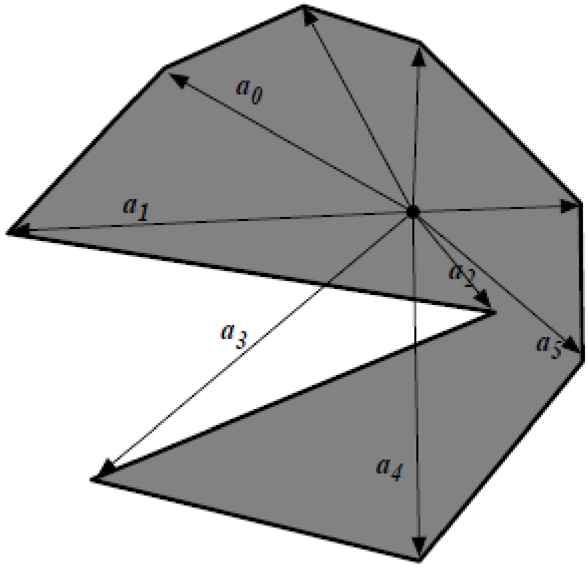


따라서, 삼각형의 면적은 평행사변형 면적의 절반이다. 따라서, 삼각형 면적을 모두 더하면 다각형의 면적을 얻는다.

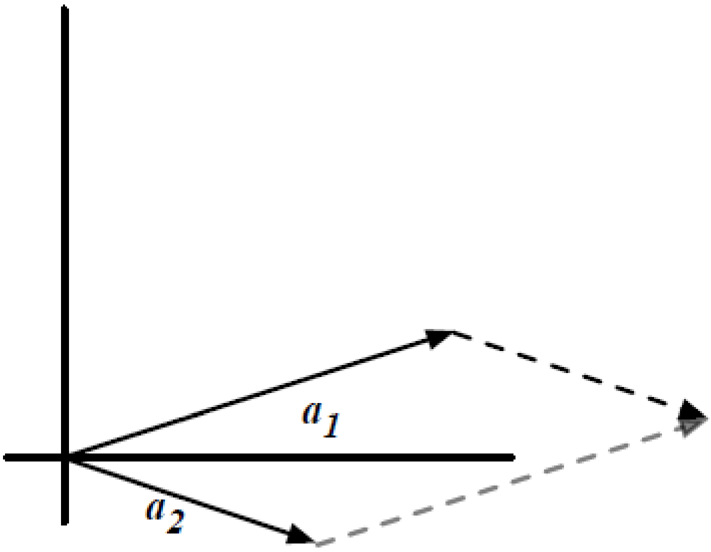
$$\frac{1}{2}(area(a_0, a_1) + \cdots + area(a_{n-1}, a_0))$$

이번에는 위의 방식으로 면적을 구할 수 없는 다각형을 보자.

예를 들어, a_i와 a_{i+1}로 형성된 삼각형이 일부 겹쳐져 있고 다각형 내에 포함되지 않는 경우가 있는 다각형이 있다고 하자.



이러한 이유로, 부호를 가지는 면적($signed$ area)을 고려해준다. 벡터 a_i와 a_{i+1}로 형성된 평행사변형의 부호를 가지는 면적의 부호는 이들 벡터가 어떻게 위치 되어있는지에 따라 결정된다. 아래의 그림과 같은 경우는 면적이 음수가 된다.



부호를 가지는 면적을 이용하여 위의 다각형의 면적을 구할 수 있다.

$$\frac{1}{2}(signedarea(a_0, a_1) + \cdots + signedarea(a_{n-1}, a_0))$$

13.10.4 행렬식 - Determinant

행렬식(Determinant)은 일종의 함수로 볼 수 있다.

$$\det : \text{실수의정방행렬} \longrightarrow \mathbb{R}$$

열 $a_1, \ldots, a_n$을 가지는 $n \times n$ 행렬 A에 대해, $\det A$의 값은 벡터 $a_1, \ldots, a_n$에 의해 정의되는 평행육면체의 *부호*를 가지는 부피이다.

- $a_1, \dots, a_n$은 표준 기저벡터 $e_1, \dots, e_n$이라하면, A는 단위행렬이다. 이 경우 평행육면체는 n차원의 단위 (하이퍼)큐브 (hyper cube)이고 $\det A$는 1이다.
- 여러가지 양수로 벡터들을 확대/축소(scale)하자. 평행육면체는 cube가 아니라 n 차원 (hyper)rectangle이며, A는 양수의 대각원소를 가지는 대각행렬이 되고 $\det A$는 이 대각원소들의 곱이다.

행렬식의 성질: A를 $n \times n$ 행렬이라 하고, $A = \begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix}$로 표시하자.

- 만약 $a_1, \dots, a_n$이 직교하면, $|\det A| = \|a_1\| \, \|a_2\| \cdots \|a_n\|$ 이다.
- 일반적으로는 $|\det A| = \|a_1^*\| \, \|a_2^*\| \cdots \|a_n^*\|$ 이다.
- 열 a_i에 α를 곱하는 것은 행렬식에 α를 곱하는 것과 같다.

$$\det \begin{bmatrix} | & & | & & | \\ a_1 & \cdots & \alpha a_i & \cdots & a_n \\ | & & | & & | \end{bmatrix} = \alpha \det \begin{bmatrix} | & & | & & | \\ a_1 & \cdots & a_i & \cdots & a_n \\ | & & | & & | \end{bmatrix}$$

Proposition: 정방행렬 A가 가역적이 될 필요충분조건은 A의 행렬식이 영이 아니어야 한다.

- *Proof*: $a_1, \dots, a_n$ 을 행렬 A의 열이라 하고, $a_1^*, \dots, a_n^*$ 을 Gram-Shmidt 직교화에 의해 얻은 직교벡터라고 하면, A가 가역적이지 않을 필요충분조건은 $a_1, \dots, a_n$이 선형종속인 경우, $a_1^*, \dots, a_n^*$의 적어도 하나는 영벡터인 경우, $\|a_1^*\| \, \|a_2^*\| \cdots \|a_n^*\|$ 이 영인경우, 행렬식이 영인 경우이다.

`nump.linalg.det()` 를 이용해 detAdetA를 구할 수 있다.

```
A = np.matrix([[1, 2],
               [3, 4]])

A_det = np.linalg.det(A)
print('det A =', A_det)
'''출력결과
det A = -2.0000000000000004
'''
```