선형대수학팀

3팀
박이현
송지현
유종석
김민서
이주형

행렬식과 역행렬

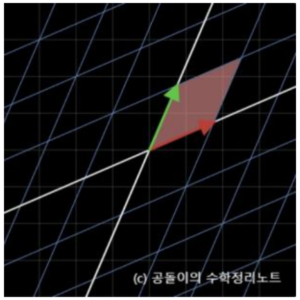
행렬식 (Determinant)

정사각행렬에 수를 대응시키는 함수로,

$det(A)$ 또는 $|A|$로 나타냄

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$에 대해, 행렬식은

$$det(A) = ad - bc$$



선형결합으로 변한

기저벡터들이 이루는

평행사변형의 넓이

역행렬 (Inverse)

행렬에서 역원의 역할

곱해서 항등원이 나오게 하는 행렬

A^{-1}으로 표현하며, 역행렬은 단 **하나**만 존재

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$에 대해,

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

노름과 내적

노름 (Norm)

p = Norm의 차수

$$\|\vec{v}\|_p = L_p = \sqrt[p]{|v_1|^p + |v_2|^p + \cdots + |v_n|^p}$$

벡터의 크기를 의미, 거리와 비슷한 개념으로 이해

일반적으로 p=1인 L1 Norm과 p=2인 L2 Norm이 자주 사용됨

L1 Norm(맨해튼 노름)

각 벡터들의 절댓값을 합한 결과

$$L_1 = |v_1| + |v_2| + \cdots + |v_n|$$

L2 Norm(유클리드 노름)

원점에서 벡터에 연결된 직선거리

$$L_2 = \sqrt{|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2}$$

내적 (Inner Product)

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = |x|\,|y|\,cos\theta$$

두 벡터의 성분끼리 곱한 후 더한 것, 스칼라값 출력

내적의 기하학적 정의를 통해서 두 벡터 사이의 각도 θ를 계산 가능

$$\cos(\theta) = \frac{\vec{x} \cdot \vec{y}}{|x|\,|y|} \qquad \theta = cos^{-1}\left(\frac{\vec{x} \cdot \vec{y}}{|x|\,|y|}\right)$$

3 직교성과 정사영

직교성과 정사영

직교성 (Orthogonality)

두 벡터가 이루는 각도 θ가 90°가 되는 경우 / **두 벡터가 직교한다 = 두 벡터의 내적이 0이다**

$$\cos(90°) = 0 \quad \rightarrow \quad \cos^{-1}(0) = 90° \qquad\qquad 90° = \cos^{-1}\left(\frac{\vec{x}\cdot\vec{y}}{|x||y|}\right) \leftarrow$$

앞에서 구한 공식 사용

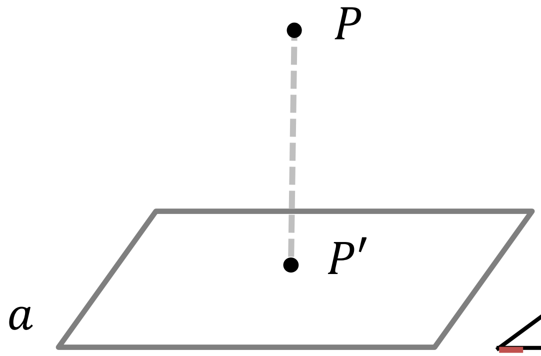
$$\theta = cos^{-1}\left(\frac{\vec{x}\cdot\vec{y}}{|x|\,|y|}\right)$$

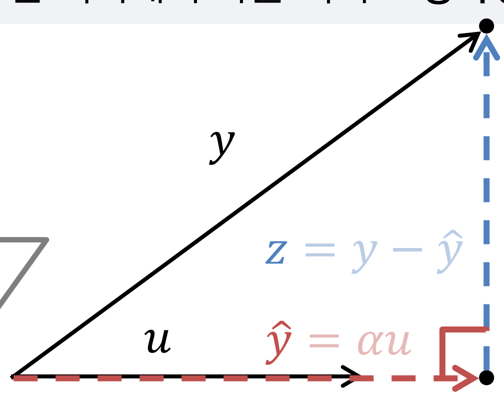
$$x \ and \ y \ is \ orthogonal \ \longrightarrow \ \vec{x}\cdot\vec{y} = 0$$

정사영 (Orthogonal Projection)

평면 a밖의 점 P에서 a에 그은 수선의 발 P'를 **점 P 의 평면 a 위로의 정사영**이라 함

내적의 기하학적 의미 → 한 벡터에서 다른 벡터로 **정사영** / 두 벡터의 **최소거리**에 위치!





$$y = \hat{y} + z$$

$$y\cdot u = (\hat{y}+z)\cdot u = \hat{y}\cdot u + z\cdot u$$

$$= \hat{y}\cdot u = (au)\cdot u = a(u\cdot u)$$

$$a = \frac{y\cdot u}{u\cdot u}$$

최소제곱법에 적용된 정사영의 원리

두 변수 사이의 관계를 가장 잘 나타내는 직선 추정

회귀분석의 원리

모든 데이터를 통과하는 직선을 구해야 함

= 선형시스템의 해 구하는 것

데이터가 많은 경우

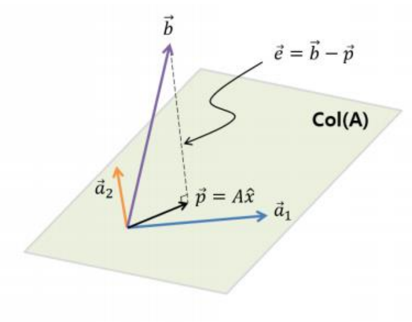
모든 점을 통과하는 직선은 존재하지 않음

선형시스템의 해가 존재하지 않으므로 추정 필요!

즉, 선형시스템 Ax=b는

두 열벡터 $\overrightarrow{a_1}, \overrightarrow{a_2}$ 의 선형결합으로 span한 열공간에 열벡터 $\vec{b}$ 가 존재하는가?

해가 존재하지 않는 경우



열벡터 $\vec{b}$가 열공간 내에 존재하지 않음

열공간으로 최단거리로 **직교하여 투영**시킴

열공간과 수직으로 내린 벡터는 직교함!

$$A^T \cdot (\vec{b} - A\hat{x}) = 0$$
$$A^T\vec{b} - A^T A\hat{x} = 0$$
$$A^T\vec{b} = A^T A\hat{x}$$
$$\hat{x} = (A^T A)^{-1} A^T \vec{b}$$

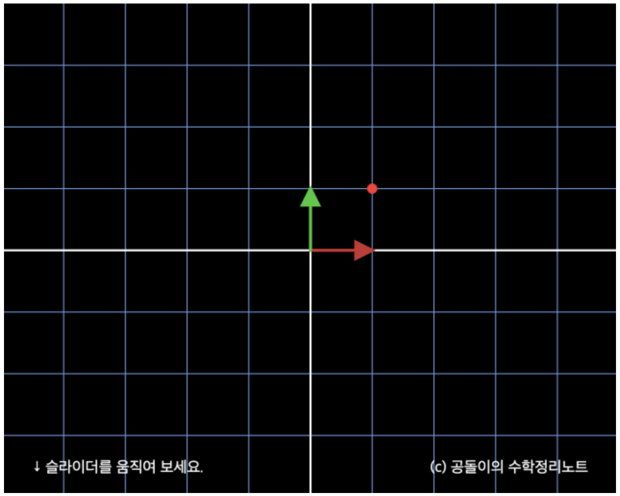
다중선형회귀의 최적화 식과 동일함을 확인 가능

고유값과 고유벡터

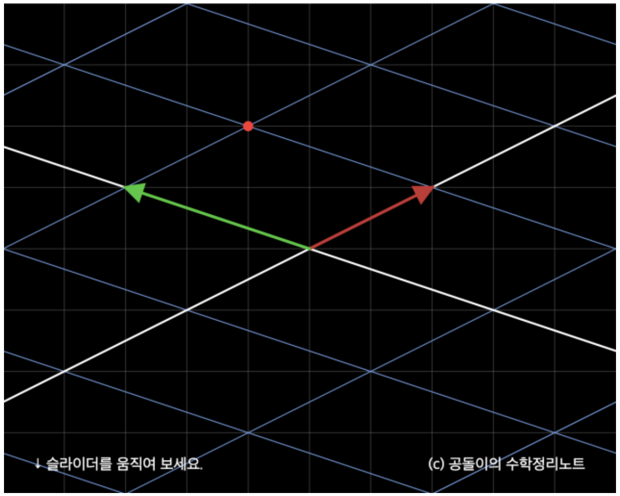
벡터 x를 행렬 A를 통해 선형 변환시

기저가 변환되어 새로운 벡터 공간에서 벡터가 변환

→ 대부분의 **벡터는 크기와 방향**이 기존과 달라짐



→



고유값과 고유벡터

벡터 x를 행렬 A를 통해 선형 변환시 기저가 변환되어 새로운 벡터 공간에서 벡터가 변환
→ 대부분의 벡터는 크기와 방향이 기존과 달라짐

하지만, 선형변환 후에도 **벡터의 방향은 변하지 않으면서 크기만 변하는 벡터** 존재

⇔ $A\boldsymbol{x} = \lambda\boldsymbol{x}$ 를 만족하는 벡터 x 존재

고유값과 고유벡터

벡터 x와 스칼라 λ, 행렬 A에 대해 $A\boldsymbol{x} = \lambda\boldsymbol{x}$을 만족하는 경우 해당하는 x를 **고유벡터**, λ를 **고유값**이라 함

고유벡터 (Eigenvector)

선형변환 시 방향은 유지한 채 길이만 변하는
영벡터가 아닌 벡터

벡터 $\boldsymbol{x}$를 행렬 A를 통해 선형변환 시
방향은 유지된 채 길이만 변화
벡터 $\boldsymbol{x}$는 고유벡터

고유값 (Eigenvalue)

선형변환시 고유벡터의 길이가 변하는 정도
고유벡터에 대한 상수배(λ)

벡터 $\boldsymbol{x}$를 행렬 A를 통해 선형변환 시
λ배 만큼 고유벡터의 길이가 변화
λ는 벡터 $\boldsymbol{x}$의 고유값

고유벡터 찾기

이항 및 정리

$$Ax = \lambda x \Leftrightarrow (A - \lambda I)x = 0$$

위 방정식의 해를 구하면 고유벡터를 찾을 수 있음

①$(A - \lambda I)$의 역행렬이 존재할 경우

양변에 $(A - \lambda I)^{-1}$을 곱해주면 해결 가능

$$(A - \lambda I)\boldsymbol{x} = 0 \Leftrightarrow \boldsymbol{x} = 0$$

이 경우,

$\boldsymbol{x} = \mathbf{0}$: trivial solution

그러나,

정의에 의해 영벡터는 고유벡터에서 제외

원하는 형태의 고유벡터를 구할 수 없음

②$(A - \lambda I)$의 역행렬이 존재하지 않을 경우

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

① $\det(A - \lambda I) = 0$ 를 풀어서 고유값(λ) 구하기

② 구한 고유값(λ)을 $A\boldsymbol{x} = \lambda \boldsymbol{x}$ 에 대입해 고유벡터($\boldsymbol{x}$) 찾기

전치 / 대칭행렬 / 대각행렬

전치 (Transpose)

원래 원소의 행과 열의 위치를 바꾼 행렬 (A^T으로 표현) $A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

① 전치를 행렬의 곱에 적용하면 곱하는 순서가 변함 $(XYZ)^T = Z^T Y^T X^T$

② 역행렬에 전치를 적용시키면 역행렬과 전치의 순서를 바꿀 수 있음 $(A^{-1})^T = (A^T)^{-1}$

대칭행렬 (Symmetric Matrix)

전치를 해도 원래 행렬과 동일한 모습을 보이는 행렬 $A = A^T$ $ex)$ 단위행렬 $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

대각행렬 (Diagonal Matrix)

대각성분을 제외한 **나머지가 모두** 0인 행렬 $ex)\ A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \cdots \quad A^n = \begin{bmatrix} 3^n & 0 & 0 \\ 0 & 7^n & 0 \\ 0 & 0 & 1^n \end{bmatrix}$

거듭제곱 시 **대각성분에만 연산**을 하면 되기 때문에 계산 용이

5 고유값 분해와 대각화

역행렬의 존재성과 선형독립의 관계 / 고유값 분해

선형독립

n개의 벡터 $x_1, x_2, \ldots, x_n$ 에 대한 벡터의 **선형결합**

$c_1x_1 + c_2x_2 + \cdots + c_nx_n = 0$ 에서 상수 c가 모두 0인 경우를 의미

역행렬의 존재성

A의 **역행렬이 존재**한다면 양변에 A^{-1}을 곱할 수 있음

$\rightarrow \boldsymbol{c_1} = \boldsymbol{c_2} = \cdots = \boldsymbol{c_n} = 0$ (trivial solution)

따라서, 선형독립인 벡터들을 모아놓은 행렬은 **역행렬이 존재!**

고유값 분해 (Eigenvalue Decomposition)

$$\boldsymbol{A = V^{-1}\Lambda V}$$

A : n x n 정방행렬 V : A의 고유벡터들로만 구성, 역행렬이 존재하는 행렬(선형독립)

Λ : 고유값을 대각성분의 원소로 갖는 **대각행렬**

인수분해를 통해 복잡한 방정식을 쉽게 해결하는 것과 유사한 원리

$(x^2 - 3x + 2)^4 = 0$의 해를 구하기 → $[(x-1)(x-2)]^4 = 0$의 해를 구하기

고유값 분해의 수식적 이해

고유벡터   고유값 대각행렬

$$\boldsymbol{AV} = \begin{bmatrix} | & | & & | \\ \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \boldsymbol{V\Lambda}$$

$\boldsymbol{\Lambda}$: 고유값을 대각성분의 원소로 갖는 대각행렬

$$\boldsymbol{AV} = \boldsymbol{V\Lambda} \qquad \boldsymbol{AVV^{-1}} = \boldsymbol{V\Lambda\, V^{-1}} \qquad \boldsymbol{A} = \boldsymbol{V\Lambda\, V^{-1}}$$

거듭제곱 연산에 적용된 고유값분해

$$\boldsymbol{A^2 = (V\Lambda\, V^{-1})(V\Lambda\, V^{-1}) = V\Lambda\, V^{-1}V\Lambda\, V^{-1} = V\Lambda^2\, V^{-1}}$$

$$\boldsymbol{A^n = V\Lambda^n\, V^{-1}}$$

대각행렬이기 때문에 거듭제곱 연산이 용이!

$\boldsymbol{A} = \boldsymbol{V\Lambda\, V^{-1}}$의 기하학적 이해

Step 1

$\boldsymbol{V}$: 축 방향 회전

Step 2

$\boldsymbol{\Lambda}$: 길이 변화

Step 3

$\boldsymbol{V^{-1}}$: 축 방향 회전

대칭행렬의 고유값 분해

대칭행렬에 대해 고유값 분해를 해줄 경우 특수한 상황이 발생

대칭행렬의 고유값 분해

A가 대칭행렬일 때,

앞과 같이 고유값 분해 진행

$$\begin{aligned} A = V\Lambda V^{-1} &= (V\Lambda V^{-1})^T \\ &= (V^{-1})^T \Lambda^T V^T \\ &= (V^T)^{-1} \Lambda V^T \end{aligned}$$

$\Lambda = \Lambda^T$ Λ가 대칭행렬이기 때문

역행렬과 전치 순서 교환

$V^{-1} = V^T \longrightarrow V^T V = I$ 를 만족

$$V^T V = \begin{bmatrix} - & v_1 & - \\ - & v_2 & - \\ \vdots & \vdots & \vdots \\ - & v_n & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} {v_1}^2 & v_1 v_2 & \cdots & v_1 v_n \\ v_2 v_1 & {v_2}^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ v_n \cdot v_1 & \cdots & \cdots & {v_n}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix} = I$$

$\rightarrow v_1 \cdot v_2 = v_1 \cdot v_3 = \cdots = 0$

각 고유벡터의 내적이 모두 0

대칭행렬에서 고유벡터끼리는 모두 서로 직교함! (orthogonal)

특이값 분해

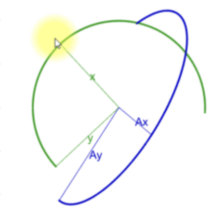
특이값 분해 (Singular Value Decomposition)

일반적인 m×n 행렬 A에 대해서 행렬을 분해하는 과정, 고유값 분해의 **일반화된 형태**

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

$A : m \times n \quad rectangular\ matrix$ $\quad U : m \times m \quad orthogonal\ matrix$

$\Sigma : m \times n \quad diagonal\ matrix$ $\quad V : n \times n \quad orthogonal\ matrix$



직교하는 열벡터($\vec{x}, \vec{y}$)를 선형변환했을 때,

형태는 변하지만 여전히 직교하는 벡터들($\mathrm{A}\vec{x}, \mathrm{A}\vec{y}$)을 찾는 것

$$AV = (U\Sigma V^T)V = (U\Sigma)(V^TV) = U\Sigma$$

원래 서로 직교하던 벡터들의 집합 V에 $\boldsymbol{A}$라는 선형변환을 시켜주었지만,

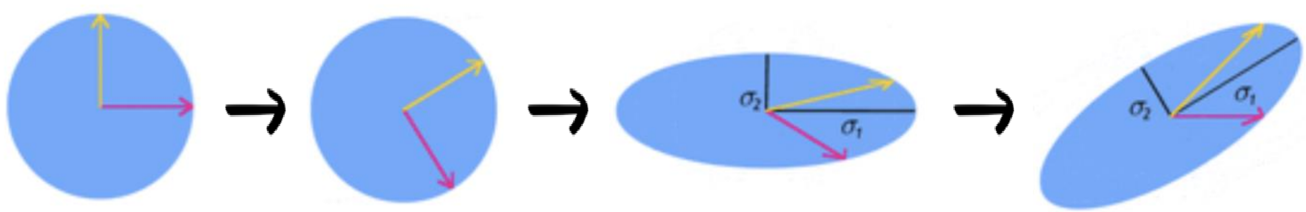
여전히 직교하는 벡터들의 집합 U에 대해서 $\boldsymbol{\Sigma}$만큼 크기만 변화함

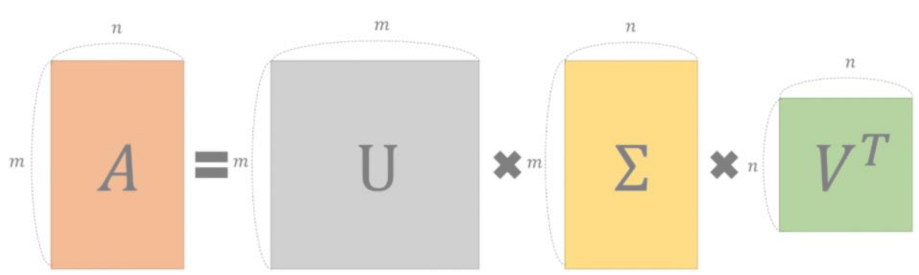
특이값 분해의 기하학적 이해

A라는 선형변환은 V를 통해 회전시키고,

크기를 조정해준 후, 다시 회전을 시켜주는 방식

순서대로도 살펴봅시다





특이값 분해의 활용

 → 

고화질 사진에서 저화질 사진으로 축소됨!

정보량을 줄여서 **데이터를 압축**하는 데 쓰임

기존 사진이 갖고 있는 형태는 유지하면서 (U, V^T)

계속 직교하므로!

데이터 자체의 scale(Σ)만 줄여주는 것

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