선형대수학팀

3팀
김진혁
김보근
노정아
심수현
이상혁

INDEX

1

행렬식

행렬식 (Determinant)

행렬식 (Determinant)

$n \times n$ 정사각행렬에 스칼라를 대응시키는 일종의 함수

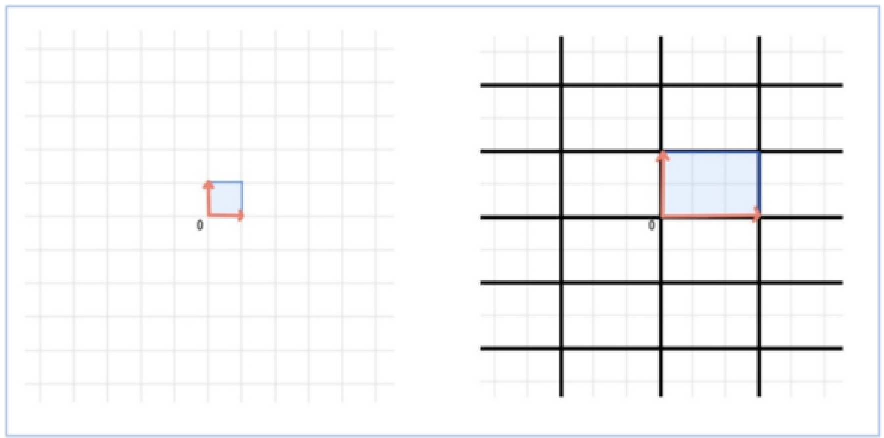
$det(A)$라고 표현

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 일 때, **행렬식은** $det(A) = ad - cb$

3×3 행렬에서는 여인수분해를 통해 행렬식 계산

1 행렬식

행렬식의 기하학적 의미



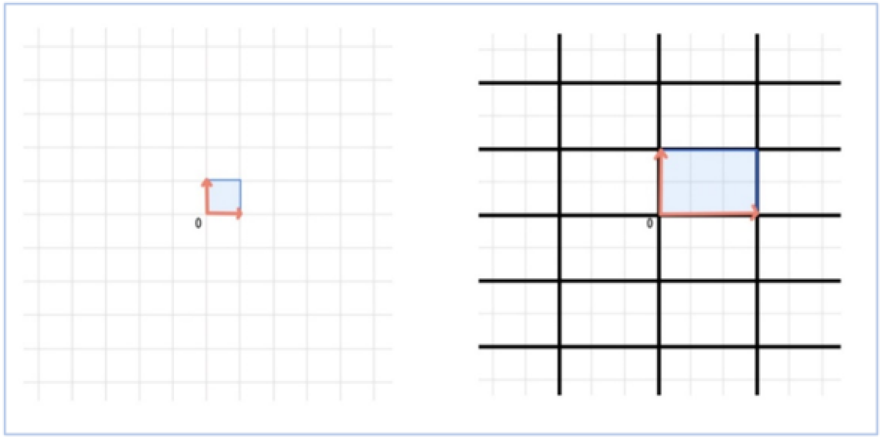
선형변환을 하면 **기저축이 변환**됨

선형변환 후 공간이 얼마나 확장되거나 축소되는지에 대한 정보를

행렬식이 가지고 있음

1 행렬식

행렬식의 기하학적 의미



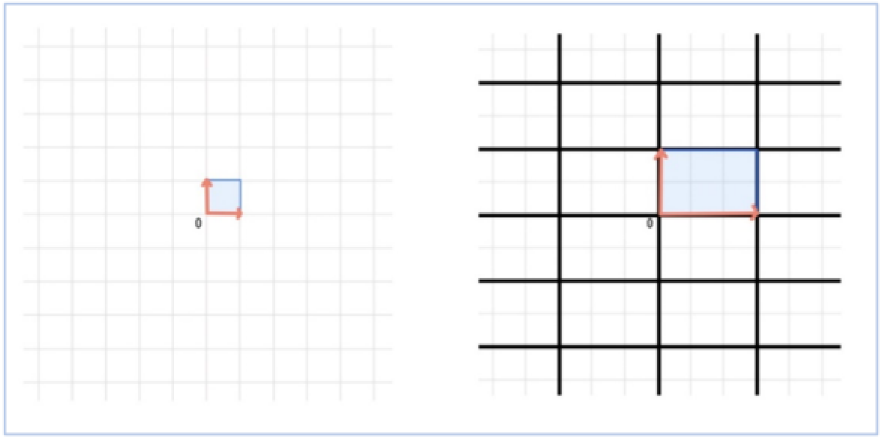
선형변환을 하면 **기저축이 변환**됨

선형변환 후 공간이 얼마나 확장되거나 축소되는지에 대한 정보를

행렬식이 가지고 있음

1 행렬식

행렬식의 기하학적 의미



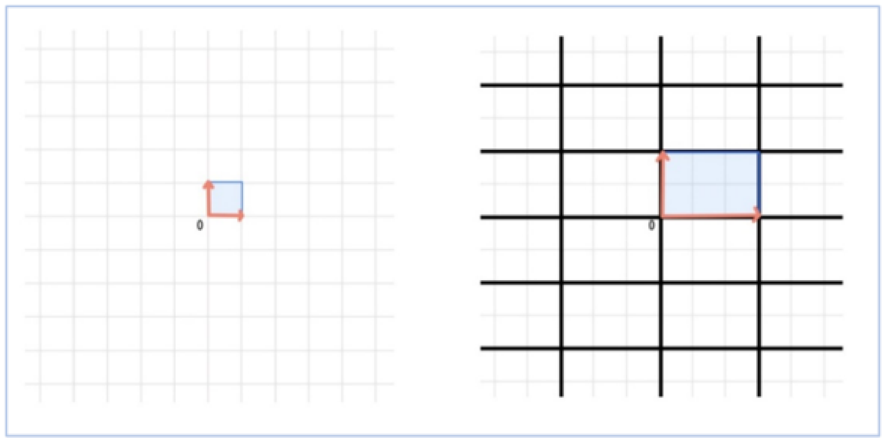
행렬 $\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$을 통해 선형변환

기저벡터 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ / 넓이 = 1 $\longrightarrow$ 기저벡터 $\begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ / 넓이 = 6

$\det\left(\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}\right) = 6 =$ **선형변환 후 평행사변형의 넓이**

1 행렬식

행렬식의 기하학적 의미



행렬 $\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$을 통해 선형변환

행렬식은 R^2 공간에서는 넓이, R^3 공간에서는 부피와 상관 있음

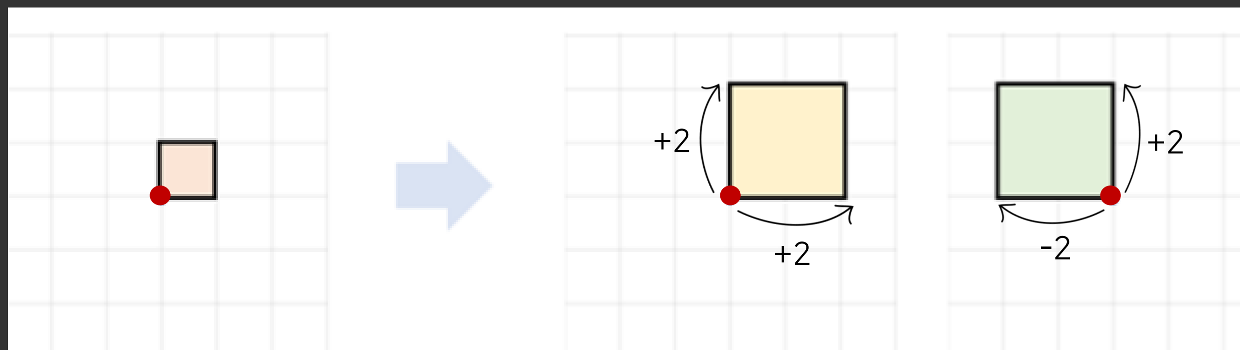
$\det\left(\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}\right) = 6 =$ 선형변환 후 평행사변형의 넓이

1 행렬식

행렬식의 기하학적 의미

행렬식이 음수인 것의 의미

행렬식의 부호 = 공간의 반전 유무



노란색으로의 선형변환

기저벡터를 모두 +2한 상황

$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $\det(A) = 4$

초록색으로의 선형변환

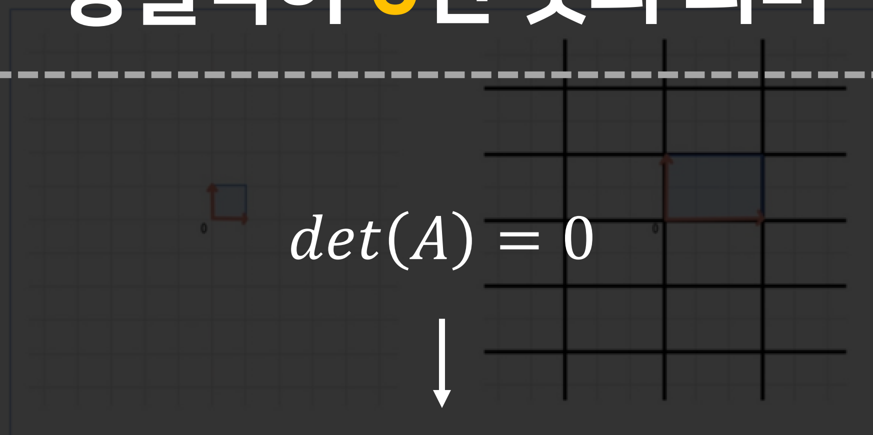
기저벡터를 각각 -2, +2한 상황

$A = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$, $\det(A) = -4$

면적은 같지만 좌우반전됨

행렬식의 기하학적 의미

행렬식이 0인 것의 의미



$$det(A) = 0$$

↓

행렬 공간의 넓이, 부피가 없음

↓

행렬 A가 공간을 압축하는 선형변환임

1 행렬식

행렬식의 쓰임

행렬식을 통해 역행렬의 존재 유무 판단

역행렬이 존재함

= $Ax = b$ 가 유일한 해를 가짐

= x 와 Ax 가 서로 일대일 대응

공간이 압축되는 선형변환의 경우 역행렬이 존재하지 않음

선형대수학팀 클린업 1주차 참고해주세요

$\det(A) = 0 \Leftrightarrow$ 행렬 A가 공간 압축 선형변환

$\det(A) = 0$ 인 경우 A의 역행렬은 존재하지 않음

1 행렬식

행렬식의 쓰임

행렬식을 통해 역행렬의 존재 유무 판단

역행렬이 존재함

= $Ax = b$ 가 유일한 해를 가짐

= x 와 Ax 가 서로 일대일 대응

공간이 압축되는 선형변환의 경우 역행렬이 존재하지 않음

선형대수학팀 클린업 1주차 참고해주세요

$\det(\mathrm{A}) = 0 \Leftrightarrow$ 행렬 A가 공간 압축 선형변환

$\det(\mathrm{A}) = 0$ 인 경우 A의 역행렬은 존재하지 않음

2

노름, 내적, 직교성

노름 (Norm)

노름 (Norm)

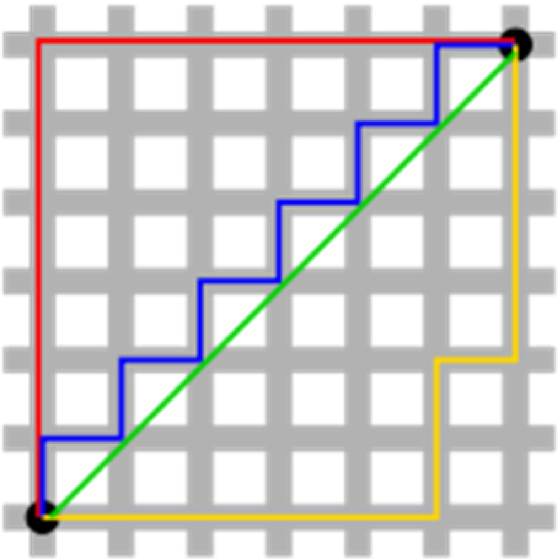
$$\|\vec{v}\|_p = L_p = \sqrt[p]{|v_1|^p + |v_2|^p + \cdots + |v_n|^p}$$

벡터의 **크기**를 의미

거리와 비슷한 개념으로 이해

일반적으로 P=1인 L1 Norm과 p=2인 L2 Norm을 주로 사용

노름 (Norm)



L1 Norm(맨해튼 노름)

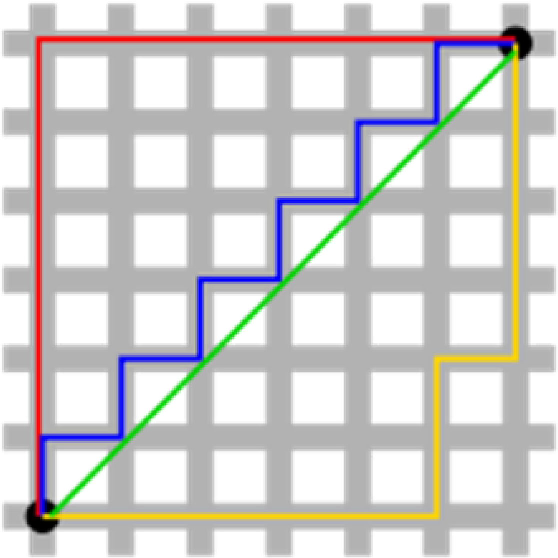
각 벡터들의 **절댓값**을 합한 결과

좌표축을 따라 움직이는 거리

Lasso Regression에서 사용됨

$$\mathrm{L1} = |v_1| + |v_2| + \cdots + |v_n|$$

노름 (Norm)



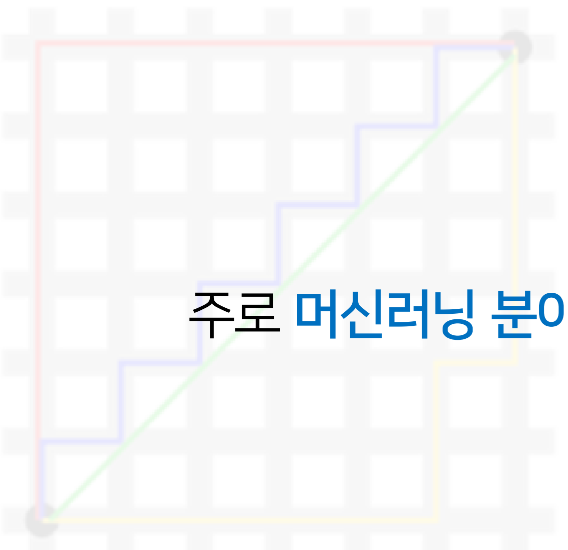
L2 Norm(유클리드 노름)

원점에서 벡터에 연결된 **직선거리**

Ridge Regression에서 사용됨

$$\mathrm{L2} = \sqrt{|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2}$$

노름 (Norm)



L2 Norm(유클리드 노름)

원점에서 벡터에 연결된 직선거리

Ridge Regression에서 사용됨

$$L2 = \sqrt{|v_1| + |v_2| + \cdots + |v_n|}$$

L1, L2 Norm은

주로 **머신러닝 분야에서 제약조건을 설정**할 때 많이 사용

내적 (Inner Product)

내적 (Inner Product)

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

두 벡터의 성분끼리 **곱한 후 더한 것**

일반적으로 두 벡터의 방향이 얼마나 일치하는지를 알기 위한 용도로 사용

$$x^T y = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

내적의 결과는 **상수**

내적 (Inner Product)

내적 (Inner Product)

$$\vec{x} \cdot \vec{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

$$= \|x\|\|y\| \cos\theta$$

: 두 벡터의 길이 × 두 벡터가 이루는 사잇각의 코사인 값

내적의 기하학적 정의를 통해 두 벡터 사이의 각도 θ를 구할 수 있음

$$x \cdot y = \|x\|\|y\| \cos\theta \xrightarrow{\text{이항}} \cos\theta = \frac{x \cdot y}{\|x\|\|y\|}$$

$$\xrightarrow{\text{양변에 } \cos^{-1}} \theta = \cos^{-1}\left(\frac{x \cdot y}{\|x\|\|y\|}\right)$$

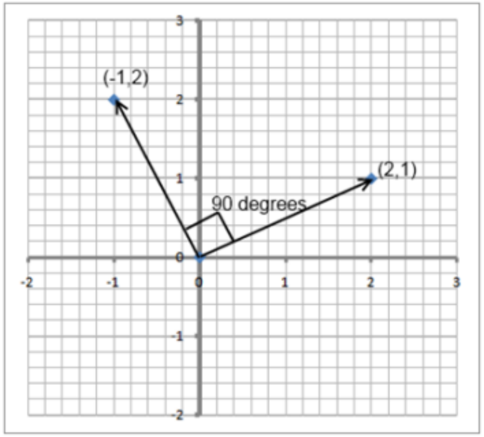
직교성 (Orthogonality)

직교성 (Orthogonality)

두 벡터가 90°를 이루는 것

두 벡터가 직교함 → $cos(90^{\circ}) = 0 \rightarrow \vec{x} \cdot \vec{y} = 0$

두 벡터의 내적값이 0임



Ex)

$$v = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, u = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$v^T u = \begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = -2 + 2 = 0$$

직교성 (Orthogonality)

직교성 (Orthogonality)

임의의 벡터 u와 영벡터는 orthogonal 할까?

두 벡터가 90°를 이루는 것

두 벡터가 직교함 $\rightarrow cos(90^\circ) = 0 \rightarrow \vec{x} \cdot \vec{y} = 0$

두 벡터의 내적값이 0임

두 벡터의 내적값이 0이면 두 벡터는 직교함

\+

어떤 벡터든 영벡터와 내적하면 0이 됨

$$v = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, u = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$v^T u = \begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = -2 + 2 = 0$$

영벡터와 임의의 벡터는 직교함

직교성 (Orthogonality)

직교하는 벡터들의 집합 = 직교집합(orthogonal set)

직교집합의 벡터

① 서로 **선형 독립**

② 직교집합의 벡터들의 span으로 **만들어지는**

벡터공간의 basis가 **됨**

정규직교성 (Orthonomality)

정규직교성 (Orthonomality)

두 벡터의 노름이 모두 1이면서 서로 직교하는 벡터

노름이 1이기 때문에 크기가 아닌 오로지 방향 성분만을 나타냄

벡터 x, y가 orthonomal하다

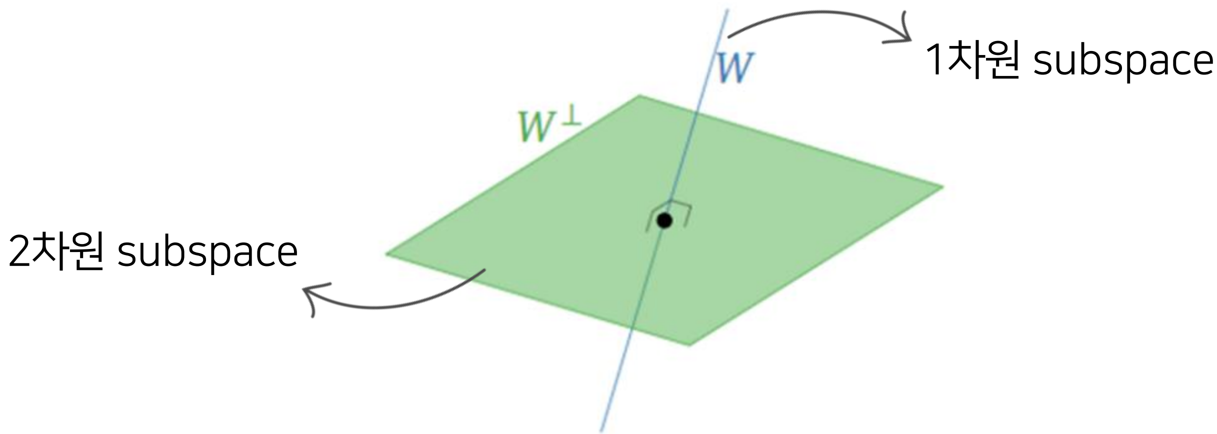
$\Leftrightarrow$

내적 값 $x \cdot y$이 0 이며, $\|x\| = \|y\| = 1$ 임

부분 공간의 직교 (Orthogonality of Subspace)

subspace S와 subspace T가 직교함

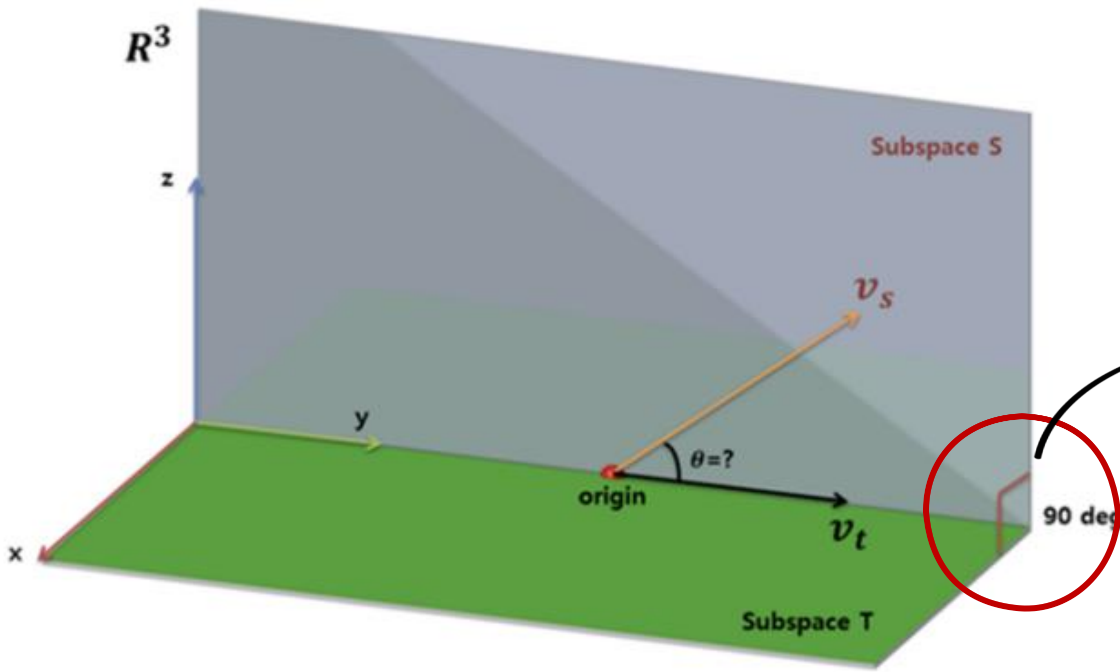
= S 안에 있는 모든 벡터가 T 안에 있는 모든 벡터와 직교함



두 space가 직교함

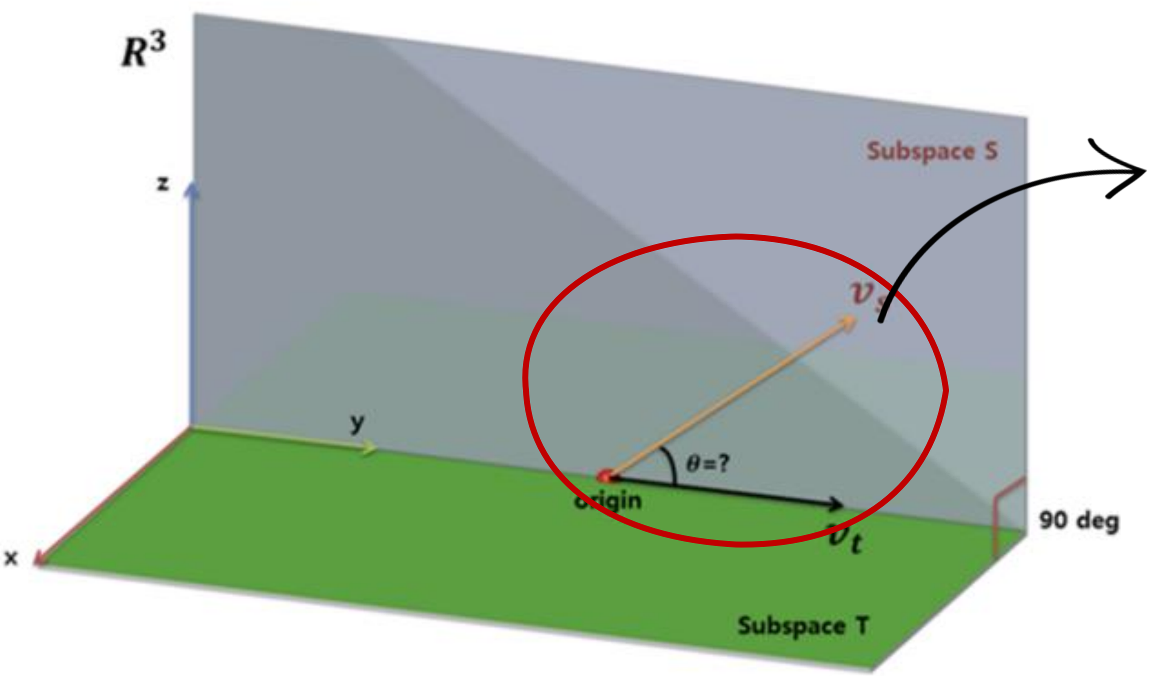
= 평면과 직선이 수직으로 만남

부분 공간의 직교 (Orthogonality of Subspace)



평면 자체는 직교

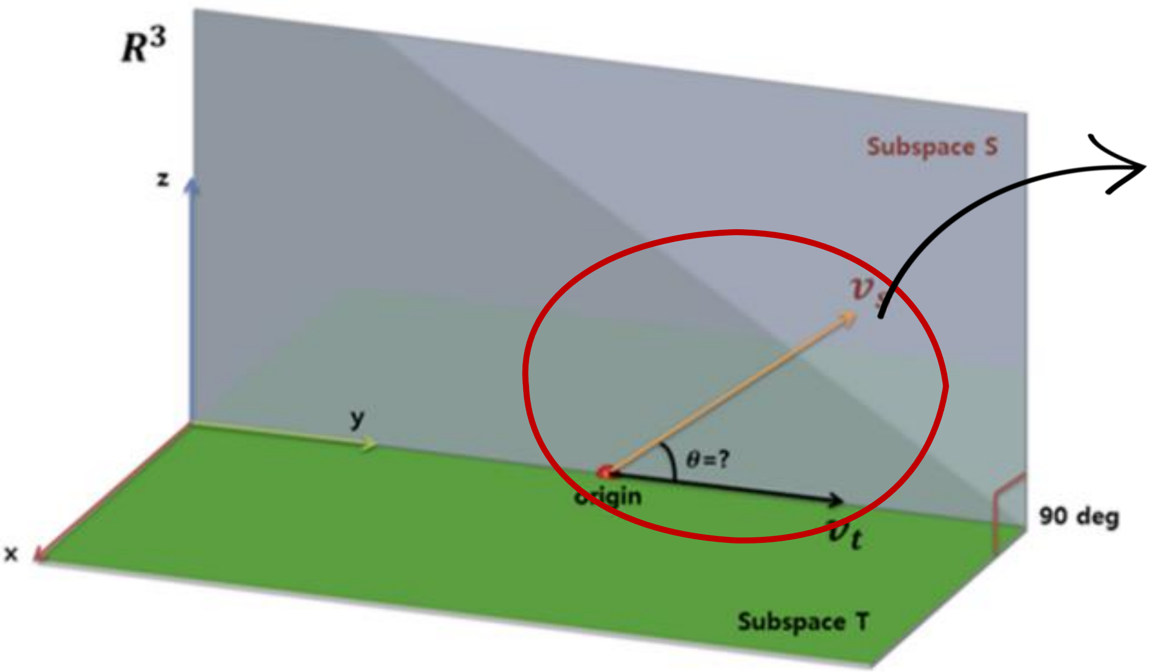
부분 공간의 직교 (Orthogonality of Subspace)



Subspace T에 존재하는 v_t와
subspace S에 존재하는 v_s가
직교하지 않음

직교하지 않는 벡터가
존재하므로
두 부분 공간은 직교하지 않음

부분 공간의 직교 (Orthogonality of Subspace)



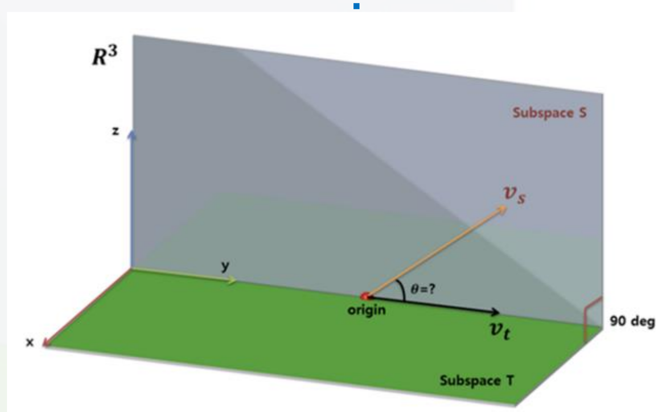
Subspace T에 존재하는 v_t와 subspace S에 존재하는 v_s가 직교하지 않음

직교하지 않는 벡터가 존재하므로 두 부분 공간은 직교하지 않음

부분 공간의 직교 (Orthogonality of Subspace)

두 부분공간이 직교하기 위해서는

각 부분 공간에 존재하는 모든 벡터가 서로 직교해야함



Subspace T에 존재하는 v_t와

Subspace S에 존재하는 v_s가

직교하지 않음

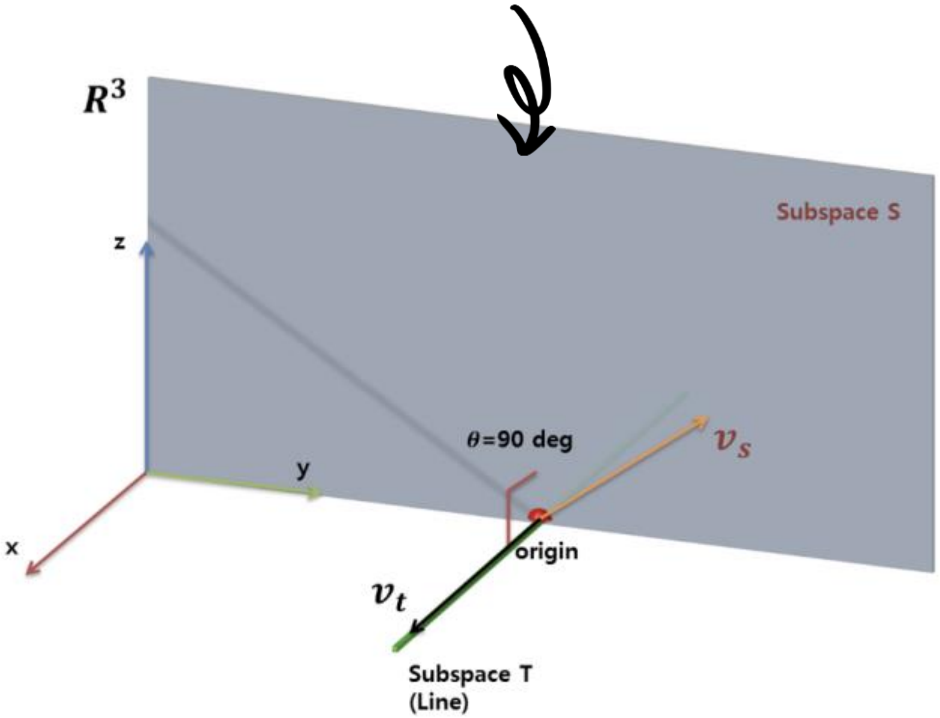
영벡터가 아닌 다른 벡터에서 만나게 되면

직교하지 않는 벡터가 존재하게 되어 두 부분 공간은 직교하지 않음

부분 공간의 직교 (Orthogonality of Subspace)

두 부분 공간이 **직교**하기 위해서는

오직 원점에서만 두 공간이 만나야 함



2 노름, 내적, 직교성

Orthogonality of 4 Fundamental Subspaces

(선형대수학팀 1주차 클린업 참고해 주세요!)

행렬의 부분공간 : 행렬 A로부터 만들 수 있는 벡터들이 존재할 수 있는 영역

행공간 : 행렬에서 행들의 선형결합으로 만들 수 있는 선형결합

열공간 : 행렬에서 열들의 선형결합으로 만들 수 있는 선형결합

영공간 : 선형변환 후에 모두 0을 출력하게 만들어주는 벡터들의 집합

Left null space : A^T라는 선형변환 후에 모두 0을 출력하게 만들어주는 벡터들의 집합

공간의 직교를 바탕으로 $Ax = b$ 를 다시 한 번 생각해보자!

Orthogonality of 4 Fundamental Subspaces

(선형대수학팀 1주차 클린업 참고해 주세요!)

행렬의 부분공간 : 행렬 A로부터 만들 수 있는 벡터들이 존재할 수 있는 영역

행공간 : 행렬에서 행들의 선형결합으로 만들 수 있는 선형결합

열공간 : 행렬에서 열들의 선형결합으로 만들 수 있는 선형결합

영공간 : 선형변환 후에 모두 0을 출력하게 만들어주는 벡터들의 집합

Left null space : A^T라는 선형변환 후에 모두 0을 출력하게 만들어주는 벡터들의 집합

공간의 직교를 바탕으로 $Ax = b$ 를 다시 한 번 생각해보자!

Orthogonality of 4 Fundamental Subspaces

1. 행공간과 영공간의 관계

$$Ax = 0$$

$$Ax = \begin{bmatrix} - & a_1 & - \\ - & a_2 & - \\ & \vdots & \\ - & a_m & - \end{bmatrix} x = \begin{bmatrix} - & a_1 x & - \\ - & a_2 x & - \\ & \vdots & \\ - & a_m x & - \end{bmatrix} = 0$$

행공간에 있는 벡터와 영공간에 있는 벡터의 내적이 모두 0

내적이 0 = 두 벡터가 직교함

→ 행공간과 영공간이 직교함

Orthogonality of 4 Fundamental Subspaces

1. 행공간과 영공간의 관계

$$Ax = 0$$

$$Ax = \begin{bmatrix} - & a_1 & - \\ - & a_2 & - \\ & \vdots & \\ - & a_m & - \end{bmatrix} x = \begin{bmatrix} - & a_1 x & - \\ - & a_2 x & - \\ & \vdots & \\ - & a_m x & - \end{bmatrix} = 0$$

행공간에 있는 벡터와 영공간에 있는 벡터의 내적이 모두 0

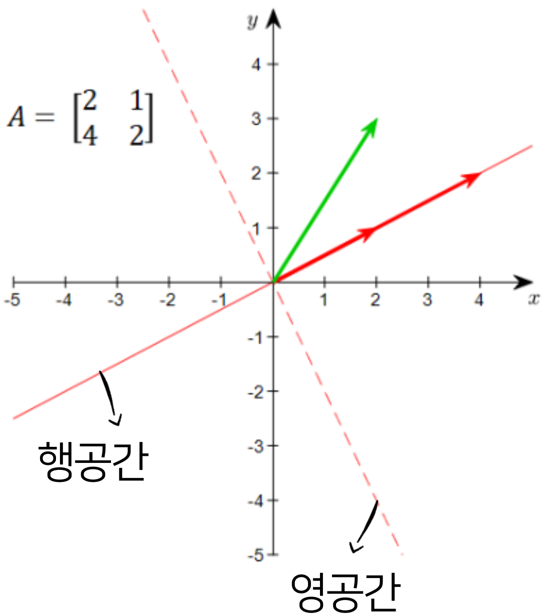
내적이 0 = 두 벡터가 직교함

→ **행공간과 영공간이 직교함**

Orthogonality of 4 Fundamental Subspaces

1. 행공간과 영공간의 관계

Ex)



행공간과 영공간이 직교한다는 사실을

이용해 $(2,3)$이라는 벡터를

선형결합으로 표현할 수 있음

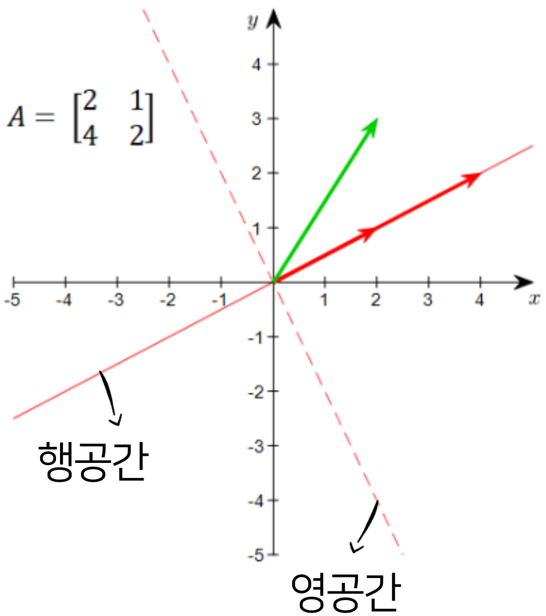
행공간과 영공간이

선형변환의 정의역이 됨

Orthogonality of 4 Fundamental Subspaces

1. 행공간과 영공간의 관계

Ex)



행공간과 영공간이 직교한다는 사실을

이용해 $(2, 3)$이라는 벡터를

선형결합으로 표현할 수 있음

행공간과 영공간이

선형변환의 정의역이 됨

Orthogonality of 4 Fundamental Subspaces

2. 열공간과 Left Null Space의 관계

$$Ax = b$$

$$\downarrow$$

$$Ax = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= x_1 \begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} + x_2 \begin{bmatrix} | \\ a_2 \\ | \end{bmatrix} + \cdots + x_n \begin{bmatrix} | \\ a_n \\ | \end{bmatrix} = b$$

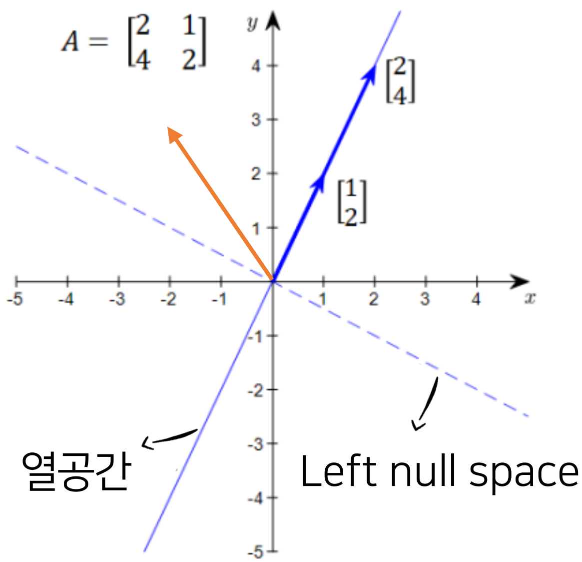
$Ax = b$를 만족하는 벡터 b가 존재한다면 b는 **행렬 A의 열공간 안에 존재**

하지만 **항상 해가 존재하는 것은 아님**

Orthogonality of 4 Fundamental Subspaces

2. 열공간과 Left Null Space의 관계

Ex)



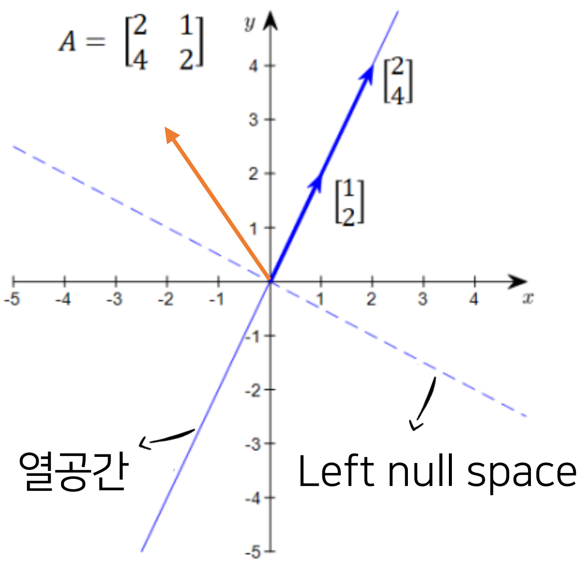
열공간과 left null space가 직교한다는 사실을 이용해 $(-2, 3)$이라는 벡터를 **선형결합으로 표현**할 수 있음

열공간과 left null space가 선형변환의 공역이 됨

Orthogonality of 4 Fundamental Subspaces

2. 열공간과 Left Null Space의 관계

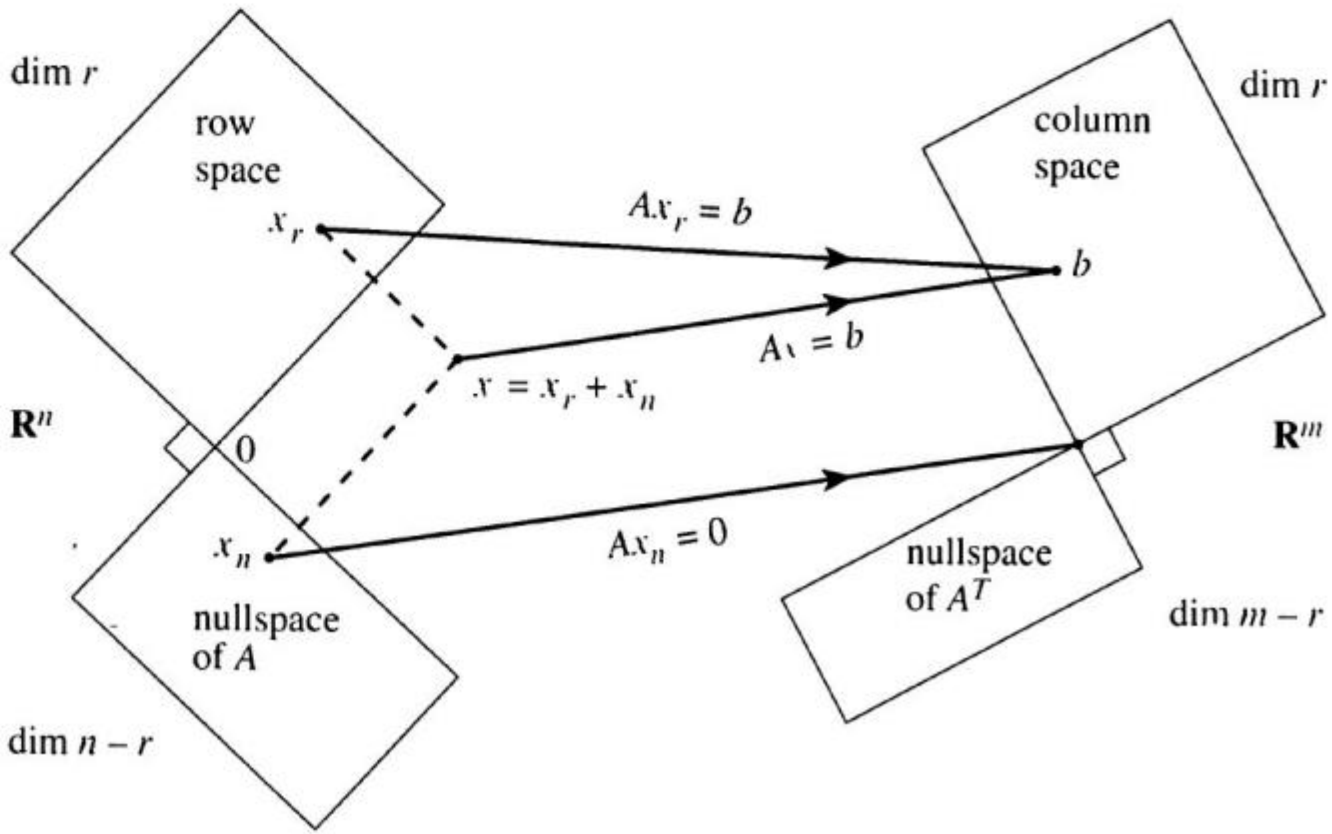
Ex)



열공간과 left null space가 직교한다는 사실을 이용해 $(-2, 3)$이라는 벡터를 선형결합으로 표현할 수 있음

열공간과 left null space가 선형변환의 공역이 됨

Orthogonality of 4 Fundamental Subspaces



Orthogonality of 4 Fundamental Subspaces

선형변환

행공간 + 영공간 = **정의역**

열공간 + left null space = **공역**

열공간 = **치역**

행렬 = **함수**

3

정사영과 회귀적용

정사영의 회귀적용

연립선형방정식 $Ax = b$의 문자를 바꾸면 **회귀문제**로 생각 가능

$$Ax = b \rightarrow X\beta = y$$

행렬 A : 독립변수들의 집합 X

벡터 b : 종속변수 y

벡터 x : 회귀계수 β

정사영의 회귀적용

X로 y를 예측하고자 할 때 **오차** 발생

이때 **오차를 최소화**하는 방향으로 분석 진행

실제로 어떤 관계가 있는지 모르는 X로 y를

정사영 (Orthogonal Projection)

정사영(Orthogonal Projection)

투영(Projection)

어떤 벡터를 다른 벡터의 공간에 옮겨 표현하는 것

어떤 벡터 b가 만드는 공간으로 **공간을 압축시키는 선형변환**과 같은 말

Projection은 선형변환의 일종

주황색 벡터들이 모두 Projection

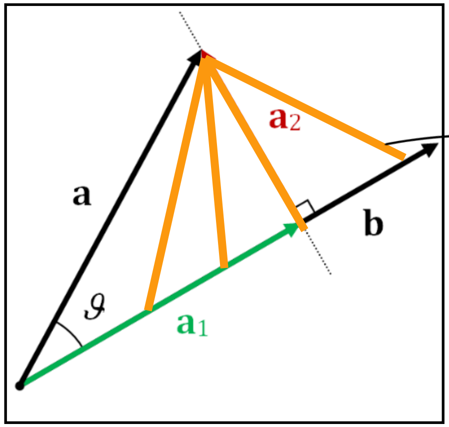
정사영(Orthogonal Projection)

투영(Projection)

어떤 벡터를 다른 벡터의 공간에 옮겨 표현하는 것

어떤 벡터 b가 만드는 공간으로 **공간을 압축시키는 선형변환**과 같은 말

Projection은 선형변환의 일종



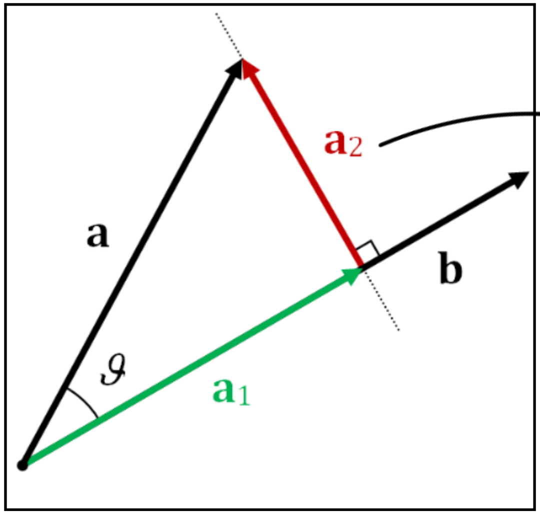
주황색 벡터들이 모두 Projection

정사영(Orthogonal Projection)

정사영(Orthogonal Projection)

벡터를 다른 벡터로 투영시킬 때 여러 각도로 투영 가능

정사영은 **직각을 이루어 투영시키는 것**을 의미



벡터 a_2는 벡터 b에 대한

Orthogonal Projection

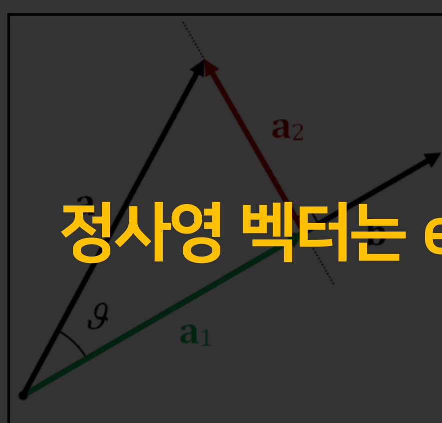
정사영(Orthogonal Projection)

정사영(Orthogonal Projection)

벡터를 다른 벡터로 투영시킬 때 여러 각도로 투영이 가능

어떤 벡터를 투영시키면 반드시 원래 벡터와 차이 발생

이를 오차(Error)라고 하면

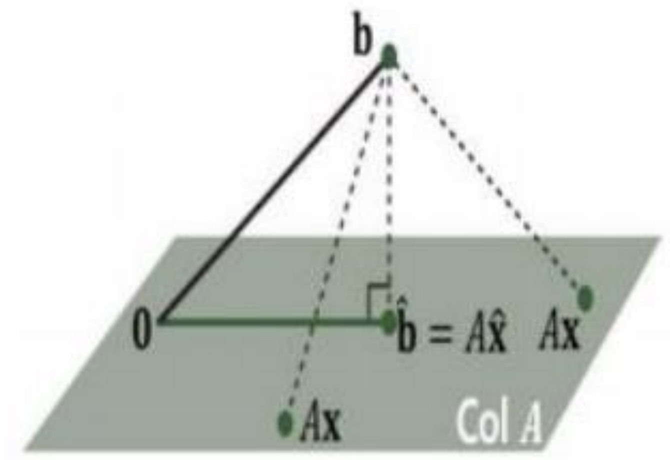


벡터 a_2는 벡터 b에 대한

정사영 벡터는 error가 최소가 되도록 함

정사영(Orthogonal Projection)

정사영은 높은 차원의 부분공간에 속한 벡터를 저차원으로 낮춤



3차원 공간에서 선형변환으로

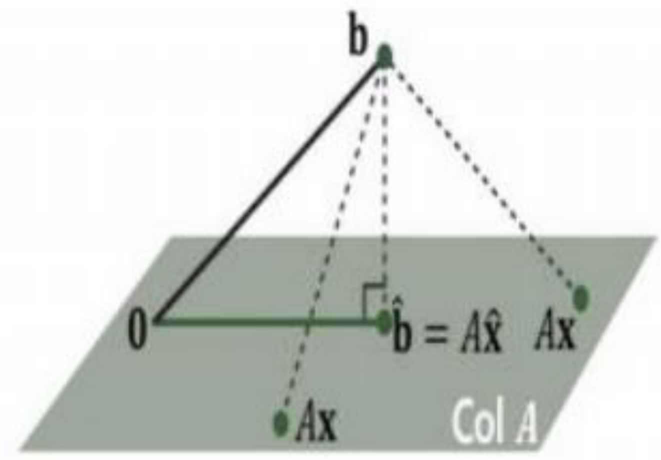
방정식의 해가 2차원 공간만을 구성

이는 회귀분석의 최소제곱법(Least Square Method)에서 사용

회귀분석의 관점에서 정사영을 이해할 수 있음

정사영(Orthogonal Projection)

정사영은 높은 차원의 부분공간에 속한 벡터를 저차원으로 낮춤



정사영은
3차원 공간에 있는 벡터를 2차원 벡터로
바꾸어서 표현하는 방식

이는 회귀분석의 최소제곱법(Least Square Method)에서 사용

회귀분석의 관점에서 정사영을 이해할 수 있음

최소제곱법에 적용된 정사영

모든 데이터를 통과하는 직선을 구하는 것이 이상적인 목표

= 연립방정식의 해를 구하는 것이 선형대수의 목표

높은 확률로 모든 점을 통과하는 직선은 없음

이는 $Ax = b$에서 b가

행렬 A의 열공간에 존재하지 않는다는 의미

최소제곱법에 적용된 정사영

모든 데이터를 통과하는 직선을 구하는 것이 이상적인 목표

= 연립방정식의 해를 구하는 것이 선형대수의 목표

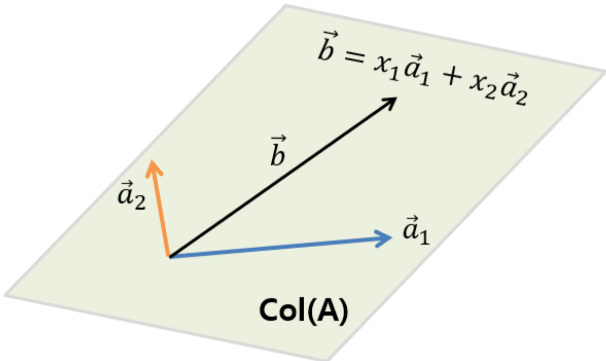
높은 확률로 모든 점을 통과하는 직선은 없음

이는 $Ax = b$에서 b가

행렬 A의 열공간에 존재하지 않는다는 의미

최소제곱법에 적용된 정사영

해가 존재하는 경우



열공간 안에 벡터 b 가 존재하는 간단한 경우

↓

Gausss-Jordan Elimination 해 구하기

해가 없는 경우

열공간 안에 벡터 b 가 존재하지 않는 경우

↓

벡터를 열공간으로 최단거리로 직교하여 투영

최소제곱법에 적용된 정사영

해가 존재하는 경우

x 대신 추정값 $\hat{x}$을 사용

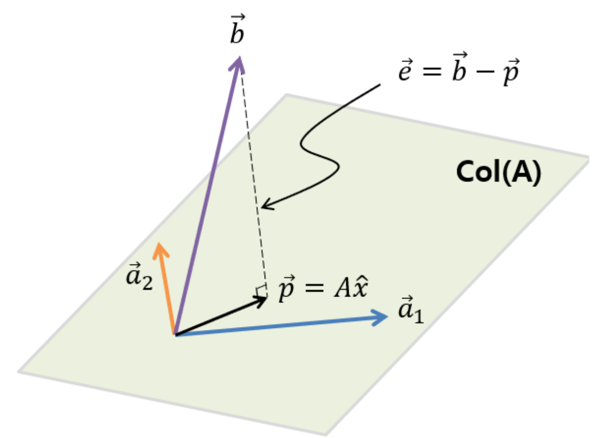
투영된 벡터를 $\hat{p} = A\hat{x}$라고 표현

$(b - A\hat{x})$가 열공간과 직교함

따라서 둘의 내적값이 0임을 이용하면

$\hat{x}$을 구할 수 있음

해가 없는 경우



열공간 안에 벡터 b 가 존재하지 않는 경우

↓

벡터를 열공간으로

최단거리로 직교하여 투영

3 정사영과 회귀적용

최소제곱법에 적용된 정사영

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

$$\rightarrow \hat{x} = (A^T A)^{-1} A^T b$$

다중선형회귀에서 **최적의 $\hat{\beta}$**을 구하는 과정과 동일

$$\hat{\beta} \ = \ (X^T X)^{-1} X^T y$$

근사한 해를 찾기 위해 **최단거리로 열공간에 투영**시켜

차원을 낮추는 선형대수의 원리가 적용됨

최소제곱법에 적용된 정사영

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

다중선형회귀에서 **최적의 $\hat{\beta}$**을 구하는 과정과 동일

$$\hat{\beta} \ = \ (X^T X)^{-1} X^T y$$

근사한 해를 찾기 위해 **최단거리로 열공간에 투영**시켜
차원을 낮추는 선형대수의 원리가 적용됨

4

고유값과 고유벡터

고유값과 고유벡터

기저는 유일(unique)하지 않음

선형변환을 하게 되면 기저가 변환되어 새로운

벡터공간에서 새롭게 벡터가 생성

고유값과 고유벡터

선형변환 이후에 방향이 변하지 않는 벡터가 있다면,

변환의 기준 축이 될 수 있지 않을까?

기저는 유일(unique)하지 않음

선형변환을 하게 되면 기저가 변환되어 새로운

벡터공간에서 새롭게 벡터가 생성

고유값과 고유벡터

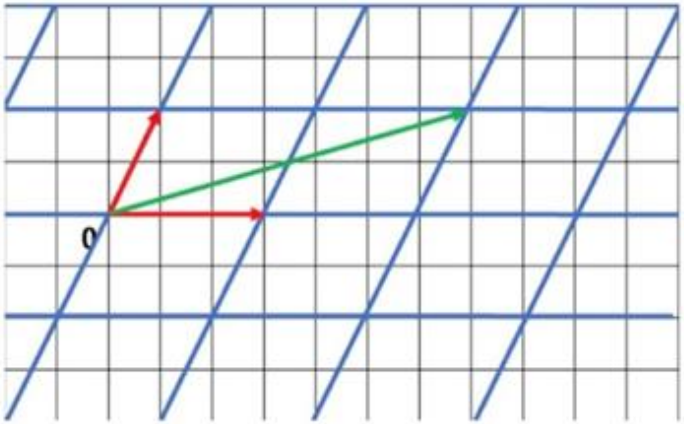
고유벡터(Eigenvector)

방향은 변하지 않으면서 크기만 변하는 **벡터**

고유값(Eigenvalue)

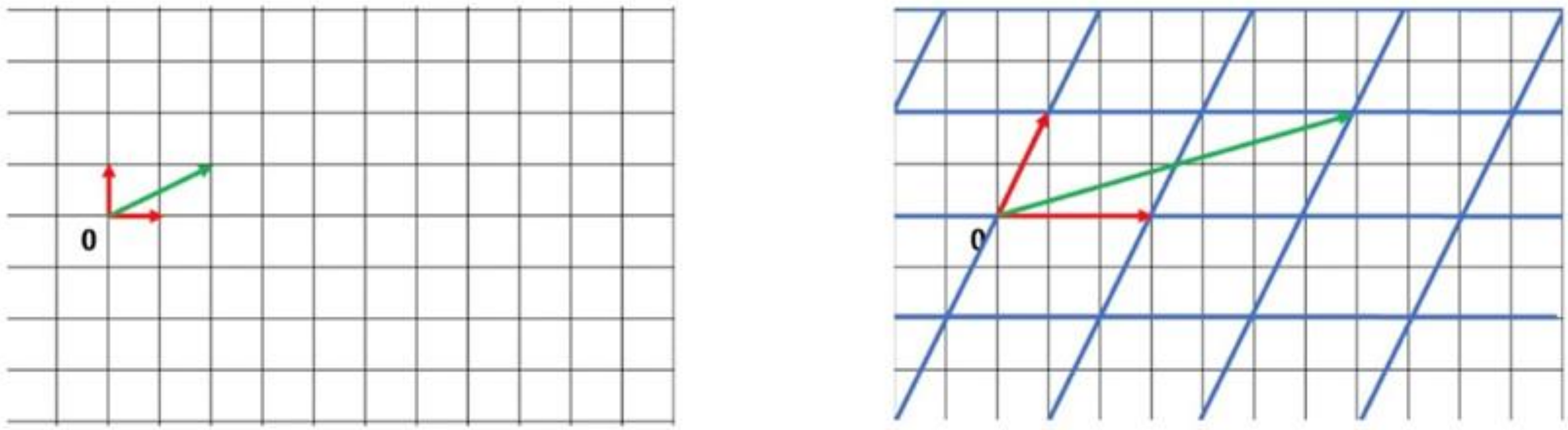
고유 벡터의 **길이가 변하는 정도**

고유값과 고유벡터



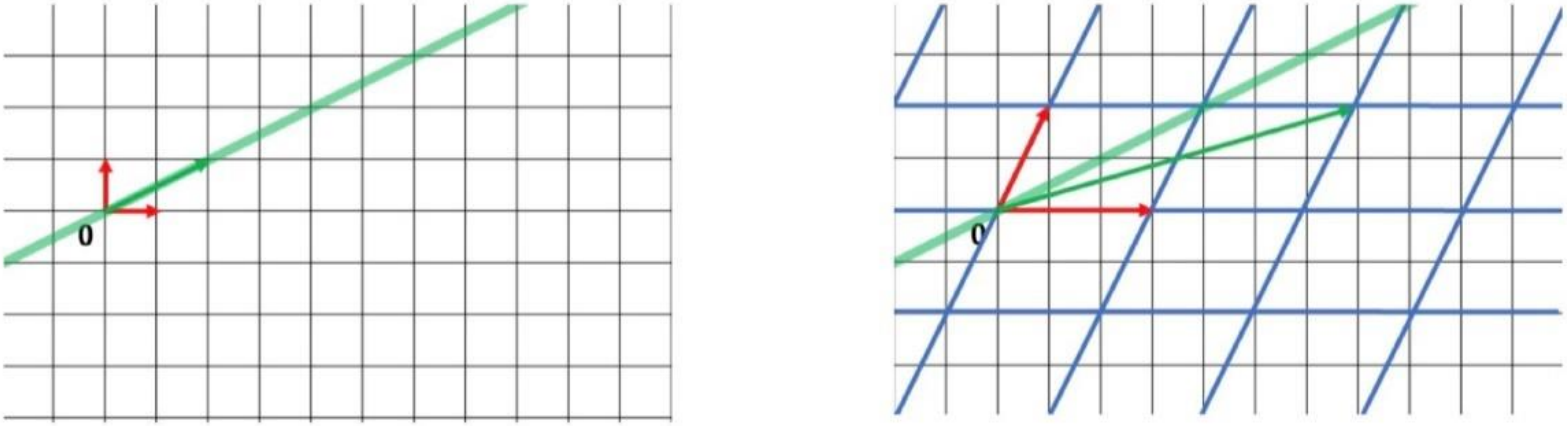
$\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}$벡터에서 $\begin{bmatrix}3\\0\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}$벡터로의 선형 변환

고유값과 고유벡터



초록색 벡터 $\begin{bmatrix}2\\1\end{bmatrix}$는 $\begin{bmatrix}3 & 1\\0 & 2\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}7\\2\end{bmatrix}$로 매핑

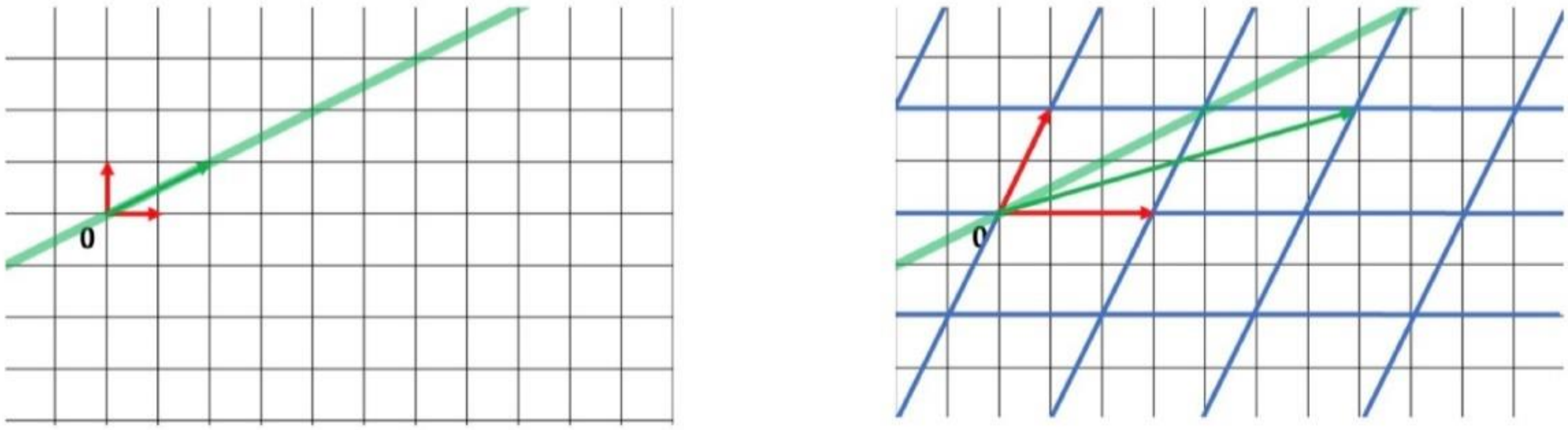
고유값과 고유벡터



초록색 벡터 $\begin{bmatrix}2\\1\end{bmatrix}$ 은 새로운 초록색 벡터 $\begin{bmatrix}7\\2\end{bmatrix}$ 와 서로 다른 직선 위에 있음

= $\begin{bmatrix}2\\1\end{bmatrix}$ 벡터가 span하는 공간에 $\begin{bmatrix}7\\2\end{bmatrix}$ 벡터가 존재하지 않음

고유벡터의 기하학적 의미

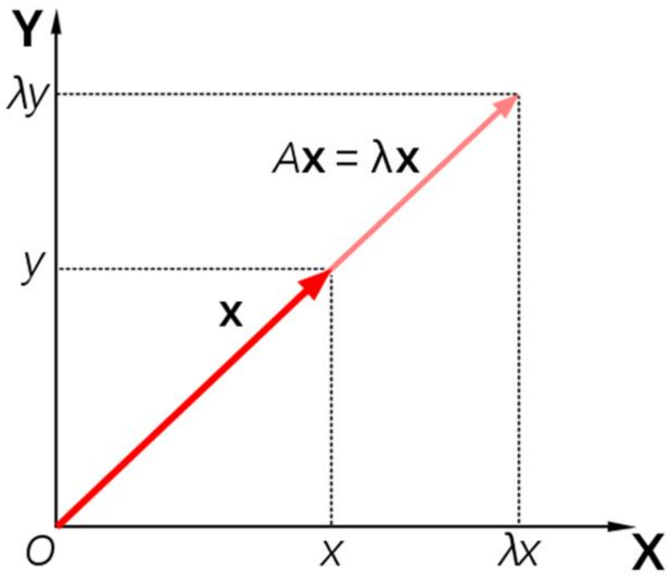


초록색 사선은 변환 후에도 고유한 span 공간에 남아있음

행렬 변환을 벡터의 길이를 늘리고 줄이는 **스칼라**처럼 여기게 해 줌

4 고유값과 고유벡터

고유벡터의 기하학적 의미



벡터 x 에 행렬 A 를 곱해도

방향은 유지한 채 **길이만 변함**

벡터 x 가 행렬 A 의 고유벡터

고유값, 고유벡터 계산

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

위 조건을 만족하는 벡터 x = **고유벡터**, λ= **고유값**

① $(A - \lambda I)$ **역행렬이 존재하는 경우** ② 역행렬이 존재하지 않는 경우

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

But, x = 0 : trivial solution

고유벡터의 정의에 의해 영벡터는 고유벡터에서 제외됨

고유값, 고유벡터 계산

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

위 조건을 만족하는 벡터 x = **고유벡터**, λ= **고유값**

① $(A - \lambda I)$ 역행렬이 존재하는 경우

② 역행렬이 존재하지 않는 경우

자명하지 않은 해(nontrivial solution)를 구할 수 있음

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

① **고유값을 구한다**

② **고유값을 $Ax = \lambda x$ 에 대입해 고유벡터를 찾는다**

고유값, 고유벡터 찾기

Ex) $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ 의 고유값, 고유벡터 찾기

① $\det(A - \lambda I) = 0$ 을 풀어 고유값을 구함

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

$$\lambda^2 - 4\lambda + 3 = 0 \rightarrow (\lambda - 1)(\lambda - 3) = 0, \qquad \boldsymbol{\lambda = 1, 3}$$

$\lambda_1 = 1$ → 어떤 고유벡터는 A라는 선형변환 후에 **자기 자신과 같은 벡터**가 됨

$\lambda_2 = 3$ → 또다른 고유벡터는 A라는 선형변환 후에 **자기 자신의 3배**가 됨

고유값, 고유벡터 찾기

Ex) $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ 의 고유값, **고유벡터** 찾기

② 고유값을 $Ax = \lambda x$ 에 대입해 고유벡터를 찾음

$$if\ \lambda_1 = 1,$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

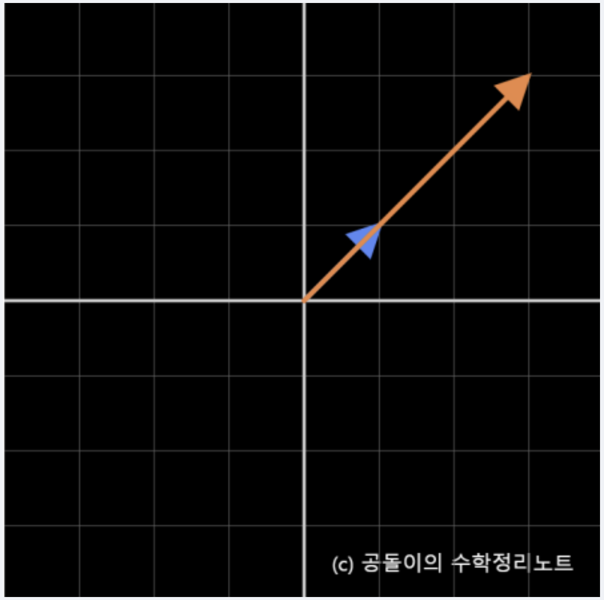
$$if\ \lambda_2 = 3,$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

고유값이 각각 1,3일 때 **고유벡터=** $\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

고유값, 고유벡터 찾기

Ex) $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ 의 고유값, 고유벡터 찾기



고유벡터 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 이 선형변환 후,

고유값 (3)만큼 길이가 늘어남

고유공간(Eigenspace)

고유공간 (Eigenspace)

각 **고유값** λ에 대해서 **고유벡터들을 모아놓은 공간**

앞의 예시에서 λ_1 의 고유공간은

선형변환 A 적용해도 길이가 1배가 되는 벡터의 span으로 형성됨

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ -5 \end{bmatrix} = \begin{bmatrix} 5 \\ -5 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} n \\ -n \end{bmatrix} = \begin{bmatrix} n \\ -n \end{bmatrix}$$

고유벡터 $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$에 상수배 한 후 선형변환하면, 자기 자신이 나옴

고유공간 = span($\begin{bmatrix} 1 \\ -1 \end{bmatrix}$)

고유공간(Eigenspace)

고유공간 (Eigenspace)

각 **고유값** λ에 대해서 **고유벡터들을 모아놓은 공간**

앞의 예시에서 λ_1 의 고유공간은

선형변환 A 적용해도 길이가 1배가 되는 벡터의 span으로 형성됨

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ -5 \end{bmatrix} = \begin{bmatrix} 5 \\ -5 \end{bmatrix}, \qquad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} n \\ -n \end{bmatrix} = \begin{bmatrix} n \\ -n \end{bmatrix}$$

고유벡터 $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$에 상수배 한 후 선형변환하면, 자기 자신이 나옴

고유공간 = span($\begin{bmatrix} 1 \\ -1 \end{bmatrix}$)

고유값 분해(Eigen Decomposition)

$$(x^2-4x+3)^4=0$$

$$[(x-1)(x-3)]^4=(x-1)^4(x-3)^4=0$$

$$x=1\ or\ 3$$

복잡한 방정식의 해를 구할 때 사용하는 인수분해라는 도구처럼

어떤 복잡한 행렬 계산에서 간단히 해를 구하기 위해 활용

대각화

대각행렬

비대각 성분이 모두 0인 행렬

행렬식, 거듭제곱, 역행렬 계산 등에서 이점을 가짐

$$D = \begin{bmatrix} C_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & C_2 \end{bmatrix}$$

$$A = \begin{matrix} 3 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{matrix} \rightarrow A^2 = \begin{matrix} 3^2 & 0 & 0 \\ 0 & 7^2 & 0 \\ 0 & 0 & 1^2 \end{matrix} \rightarrow A^n = \begin{matrix} 3^n & 0 & 0 \\ 0 & 7^n & 0 \\ 0 & 0 & 1^n \end{matrix}$$

대각화

대각행렬

비대각 성분이 모두 0인 행렬

행렬식, 거듭제곱, 역행렬 계산 등에서 이점을 가짐

어떤 행렬 $A \in R^{nxn}$ **이 대각행렬** $D \in R^{nxn}$ **로 대각화** 가능하다면,

행렬 A의 행렬식, 거듭제곱 등을 쉽게 계산할 수 있음

$$D = P^{-1}AP, P \in R^{nxn}$$

고유값 분해는 대각화의 한 종류

고유값 분해(Eigen Decomposition)

$$Av = \lambda v$$

v= 고유벡터, λ=고유값

행렬 A가 n x n **정방행렬**이고, **n개의 고유벡터, 고유값**을 가진다고 가정

$$AV = A[v_1\ v_2 \cdots v_n] = [\lambda_1 v_1\ \lambda_2 v_2\ \cdots \lambda_n v_n]$$

행렬 A의 고유벡터들로만 구성된 역행렬이 존재하는 행렬

$$= [v_1\ v_2\ \cdots v_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ & & \ddots & \vdots \\ 0 & \vdots & & \lambda_n \end{bmatrix} = \mathrm{V}\Lambda$$

행렬 A의 고유값을 대각성분의 원소로 갖는 대각행렬

고유값 분해(Eigen Decomposition)

즉, **고유값 분해**는 선형변환 이후에

방향은 변하지 않고 크기만 λ만큼 변하는

고유벡터를 찾는 것에서 시작

고유값 분해(Eigen Decomposition)

고유값 분해

행렬 A를 **고유벡터들을 열벡터**로 하는 행렬과

고유값을 대각원소로 하는 행렬의 곱으로 **대각분해** 하는 것

행렬 A의 고유벡터들로만 구성된

역행렬이 존재하는 행렬

$$AV{\color{red}V^{-1}} = V\Lambda{\color{red}V^{-1}}$$

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

행렬 A의 고유값을

대각원소로 갖는 대각행렬

고유값 분해를 이용한 다양한 계산

행렬식

$$\det(A) = \det(\,V\Lambda V^{-1}) = \det(V)\det(\Lambda)\det(V^{-1}) = \det(\,\Lambda) = \lambda_1\lambda_2\cdots\lambda_n$$

거듭제곱

$$A^k = (\,V\Lambda V^{-1})^k = (\,V\Lambda V^{-1})(\,V\Lambda V^{-1})\cdots(\,V\Lambda V^{-1}) \quad = V\Lambda^k V^{-1}$$

대각합

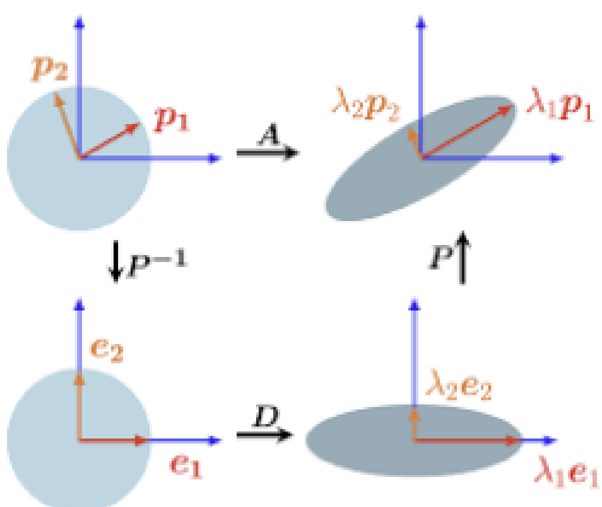
$$tr(A) = tr(V\Lambda V^{-1}) = tr(V^{-1}V\Lambda) = tr(\Lambda) \ = \lambda_1 + \cdots + \lambda_n$$

고유값 분해의 기하학적 이해

고유값 분해

행렬 A를 3가지 행렬, 3가지 선형변환으로 분해하는 것

*선형변환 = 기저벡터(축)를 바꾸는 것



P=고유벡터로 구성된 행렬 V

D=고유값으로 구성된 대각행렬 Λ

행렬 V의 선형변환은 회전과 유사한 형태를 보임

- 행렬 V: 축의 방향을 돌려줌
- 대각행렬 Λ : 몇 배만큼 늘려주는지 결정

고유값 분해의 기하학적 이해

그렇다면 V 선형변환은 회전 변환인가?

일반적으로 방향을 회전시키면

기저벡터의 크기가 약간 변하고, 뒤집어지며 변환할 가능성이 있음

또한 변환 이후 기저벡터들의 각도가 90도가 아닐 수 있음

완전히 일치하진 않고, 유사한 형태를 보임

다만, 행렬 A가 대칭행렬인 경우에는

회전을 시키는 단계에서 기저벡터의 크기가 변하지 않음

대칭행렬에 대한 고유값 분해

$A = A^T$에 대한 고유값 분해

$$(V\Lambda \mathrm{V}^{-1}) = (\mathrm{V}\Lambda\mathrm{V}^{-1})^{\mathrm{T}} \rightarrow \mathrm{V}\Lambda\mathrm{V}^{-1} = (\mathrm{V}^{-1})^{\mathrm{T}}\Lambda^{\mathrm{T}}\mathrm{V}^{\mathrm{T}}$$

Λ는 대각행렬이므로 대칭행렬 $\Lambda=\Lambda^{\mathrm{T}}$

따라서 식을 정리하면

$$V\Lambda\mathrm{V}^{-1} = (\mathrm{V}^{-1})^{\mathrm{T}}\Lambda\mathrm{V}^{\mathrm{T}} \rightarrow \mathrm{V}^{-1} = \mathrm{V}^{\mathrm{T}},\quad \mathrm{V}^{\mathrm{T}}\mathrm{V} = I$$

대칭행렬에 대한 고유값 분해

$$\text{let V}=[v_1\ v_2 \cdots v_n],\ V^TV = \begin{bmatrix} -v_1- \\ -v_2- \\ \vdots \\ -v_n- \end{bmatrix}[v_1\ v_2 \cdots v_n]$$

$$= \begin{bmatrix} v_1^2 & \cdots & v_1v_n \\ \vdots & \ddots & \vdots \\ v_nv_1 & \cdots & v_n^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix} = I \quad (\because v_1v_2 = v_2v_3 = \cdots = 0)$$

내적이 0이므로 직교함

대칭행렬에 대한 고유값 분해

행렬 A가 대칭행렬이면,

각 고유벡터끼리 서로 직교한다

(모든 고유벡터 간의 **내적이 0**이기 때문)

$$\text{let } V=[v_1\ v_2 \cdots v_n],\quad V^T V = \begin{bmatrix} -V_1- \\ -V_2- \\ \vdots \\ -V_n- \end{bmatrix} [v_1\ v_2 \cdots v_n]$$

$$= \begin{bmatrix} V_1^2 & \cdots & V_1V_n \\ \vdots & \ddots & \vdots \\ V_nV_1 & \cdots & V_n^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix} = I \quad (\because V_1V_2 = V_2V_3 = \cdots = 0)$$

내적이 0이므로 직교함

항상 고유값 대각화가 가능

직교행렬로 대각화가 가능

이 특징을 공분산 행렬에 적용한 것: 주성분 분석(PCA)

3주차에서 계속…

고유값 분해

고유값 분해가 가능하려면,

선형독립이 아닌 벡터들 사이에서 nxn 행렬 A의

고유벡터 n개를 찾는 것이 불가능하기 때문에

정방행렬 A의 열벡터가 모두 선형독립

고유값 분해

그러나, **정방행렬이 아니거나**

열벡터가 선형독립이 아닌 행렬도

고유값 분해와 비슷한 방식으로 행렬을 분해할 수 있음

특이값 분해

5

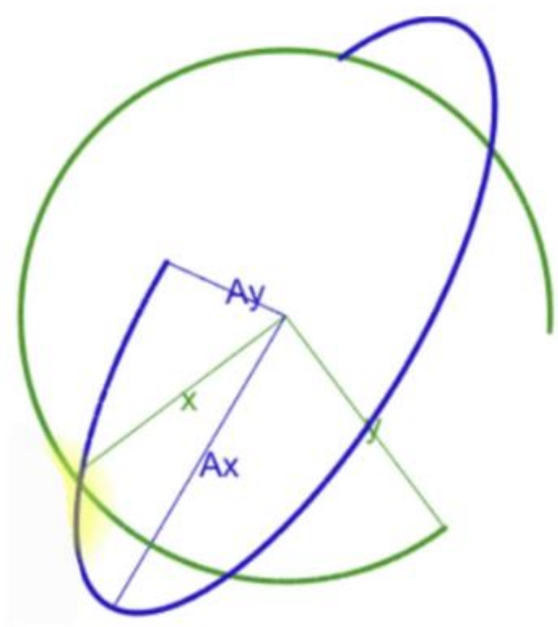
특이값 분해

특이값 분해(SVD)의 목적

특이값 분해 (Singular Value Decomposition)

직교하는 벡터 집합에 대하여, 선형변환 후에 그 크기는 변하지만

여전히 직교하는 직교집합을 찾는 것



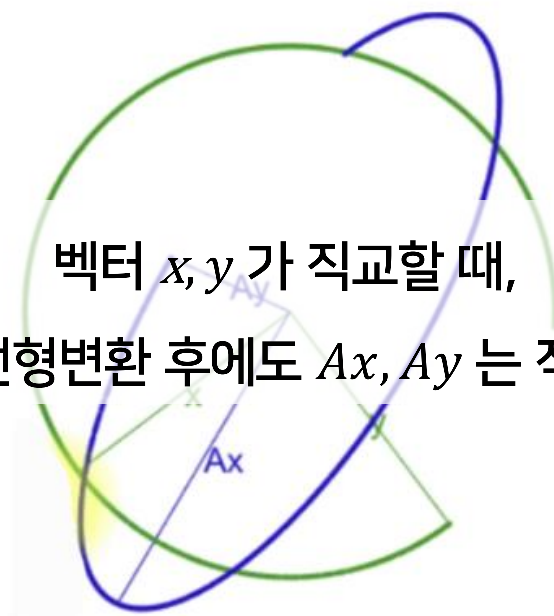
5 특이값 분해(SVD)

특이값 분해(SVD)의 목적

특이값 분해 (Singular Value Decomposition)

직교하는 벡터 집합에 대하여, 선형변환 후에 그 크기는 변하지만

여전히 직교하는 직교집합을 찾는 것

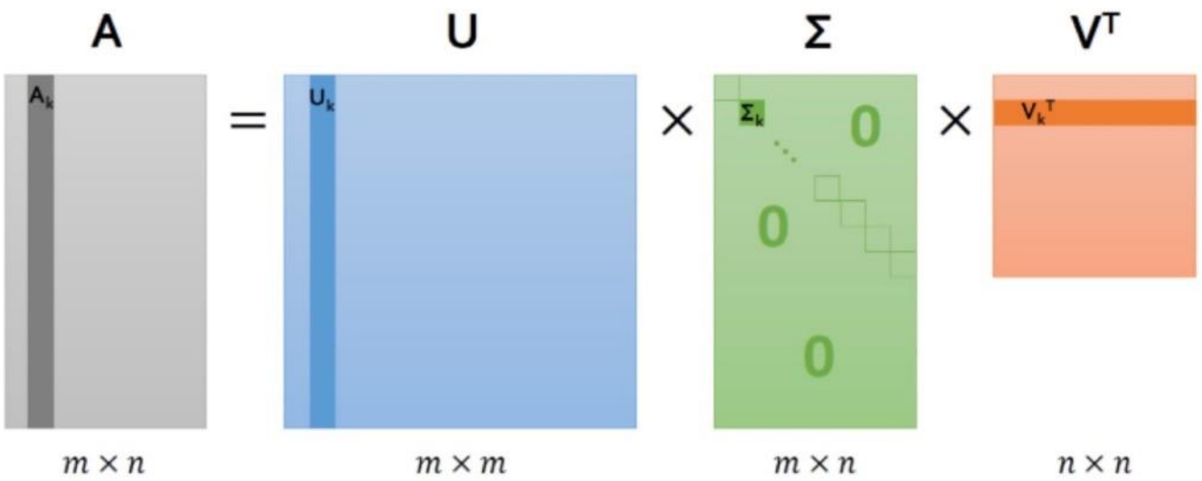


벡터 x, y 가 직교할 때,

A라는 선형변환 후에도 Ax, Ay 는 직교할까?

5 특이값 분해(SVD)

특이값 분해(SVD)의 개념



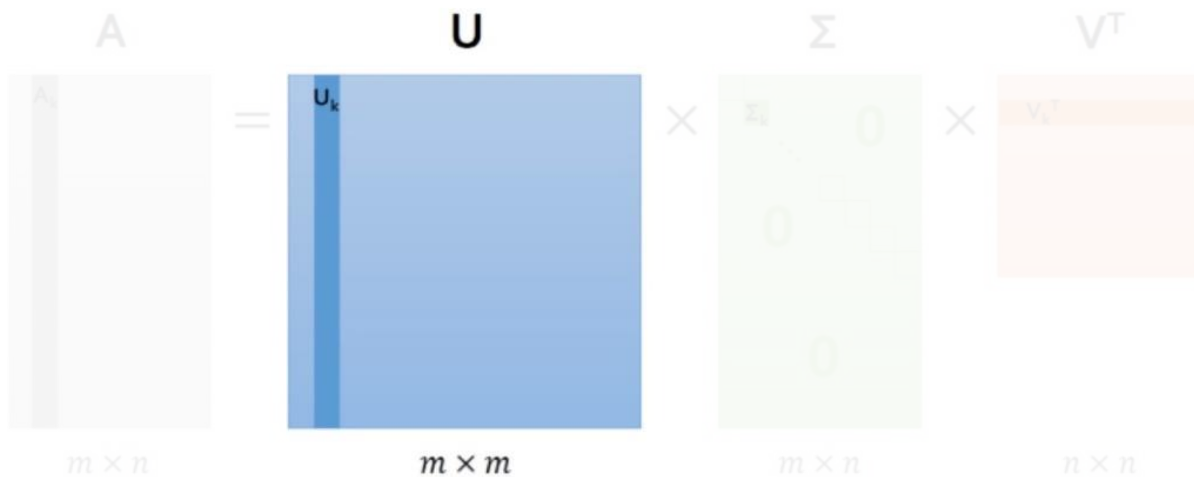
$A : m \times n$ rectangular matrix (직사각 행렬)

$U : m \times m$ orthogonal matrix (직교 행렬)

$\Sigma : m \times n$ diagonal matrix (대각 행렬)

$V^T : \mathrm{n} \times n$ orthogonal matrix (직교행렬)

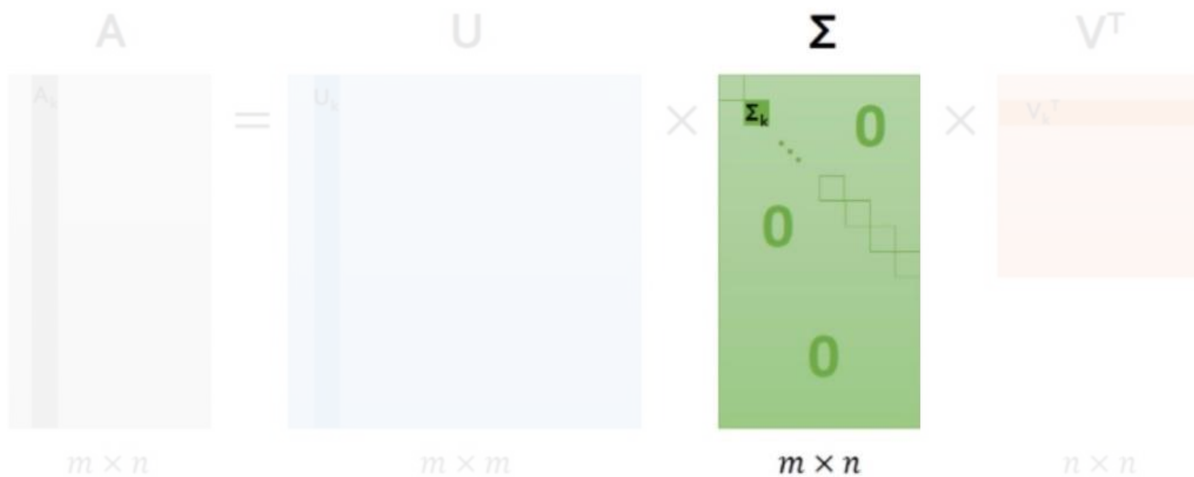
특이값 분해(SVD)의 개념



AA^T의 고유벡터로 구성된 직교행렬

$$(AA^T = U(\Sigma\Sigma^{\mathrm{T}})U^T)$$

특이값 분해(SVD)의 개념



AA^T, A^TA를 고유값 분해해서 나오는 고유값들의

Square root를 대각원소(특이값)로 하는

$m \times n$ 직사각 대각행렬

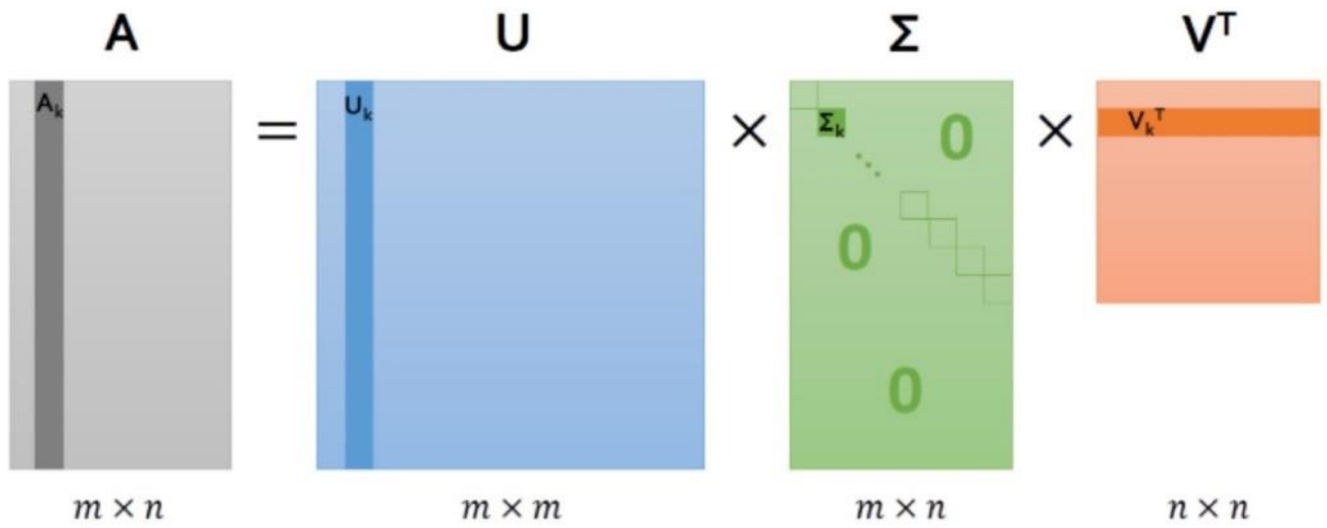
5 특이값 분해(SVD)

특이값 분해(SVD)의 개념

$A^T A$의 고유벡터로 구성된 직교행렬

$(A^T A = V(\Sigma\Sigma^T)V^T)$

특이값 분해(SVD)



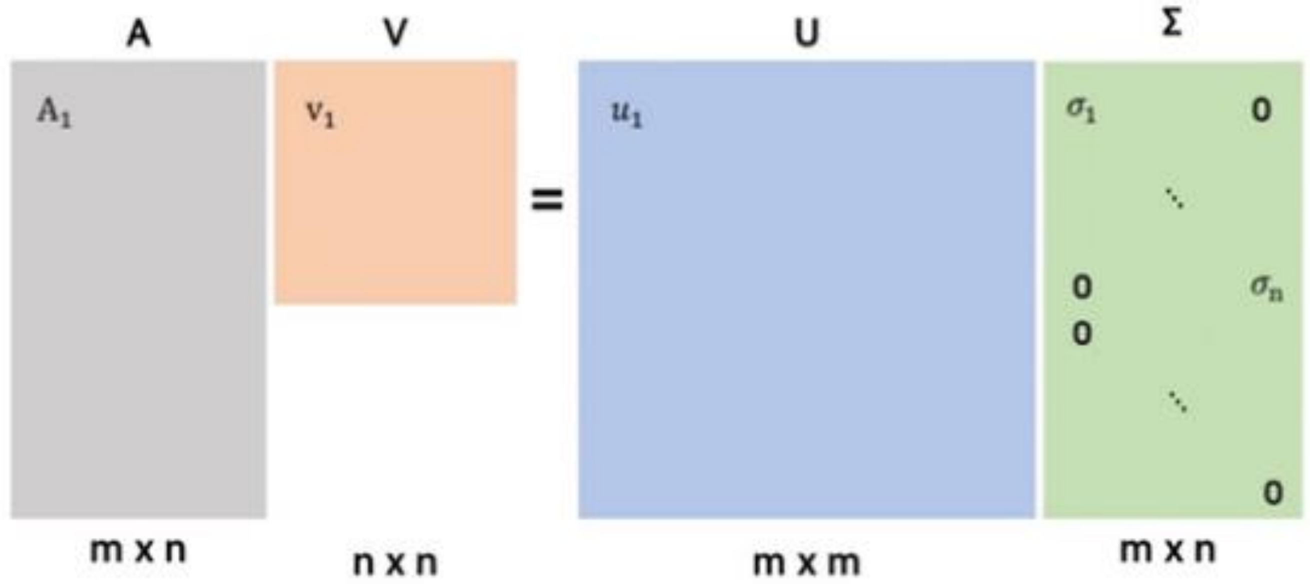
$$A = U\Sigma V^T$$

양변에 직교행렬 V를 곱해보자!

$$(V^T V = V^{-1} V = I)$$

$$AV = U\Sigma$$

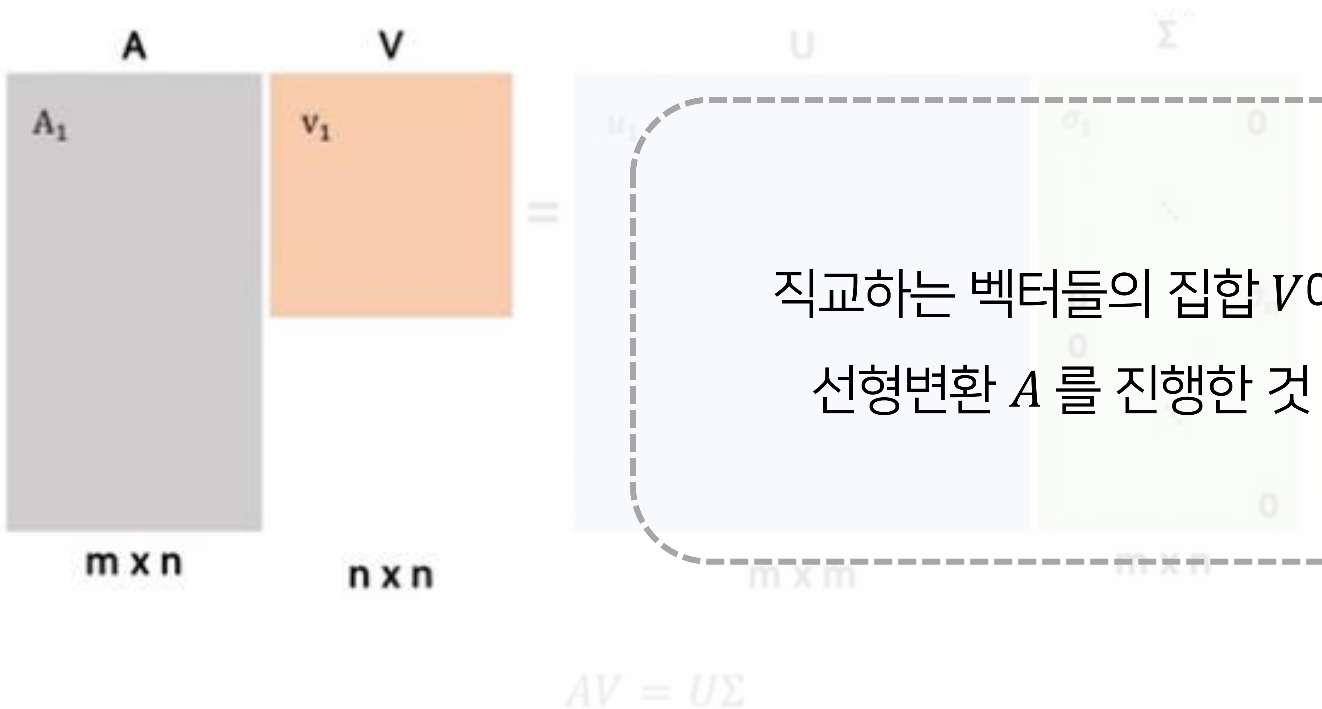
특이값 분해(SVD)



$$AV = U\Sigma$$

5 특이값 분해(SVD)

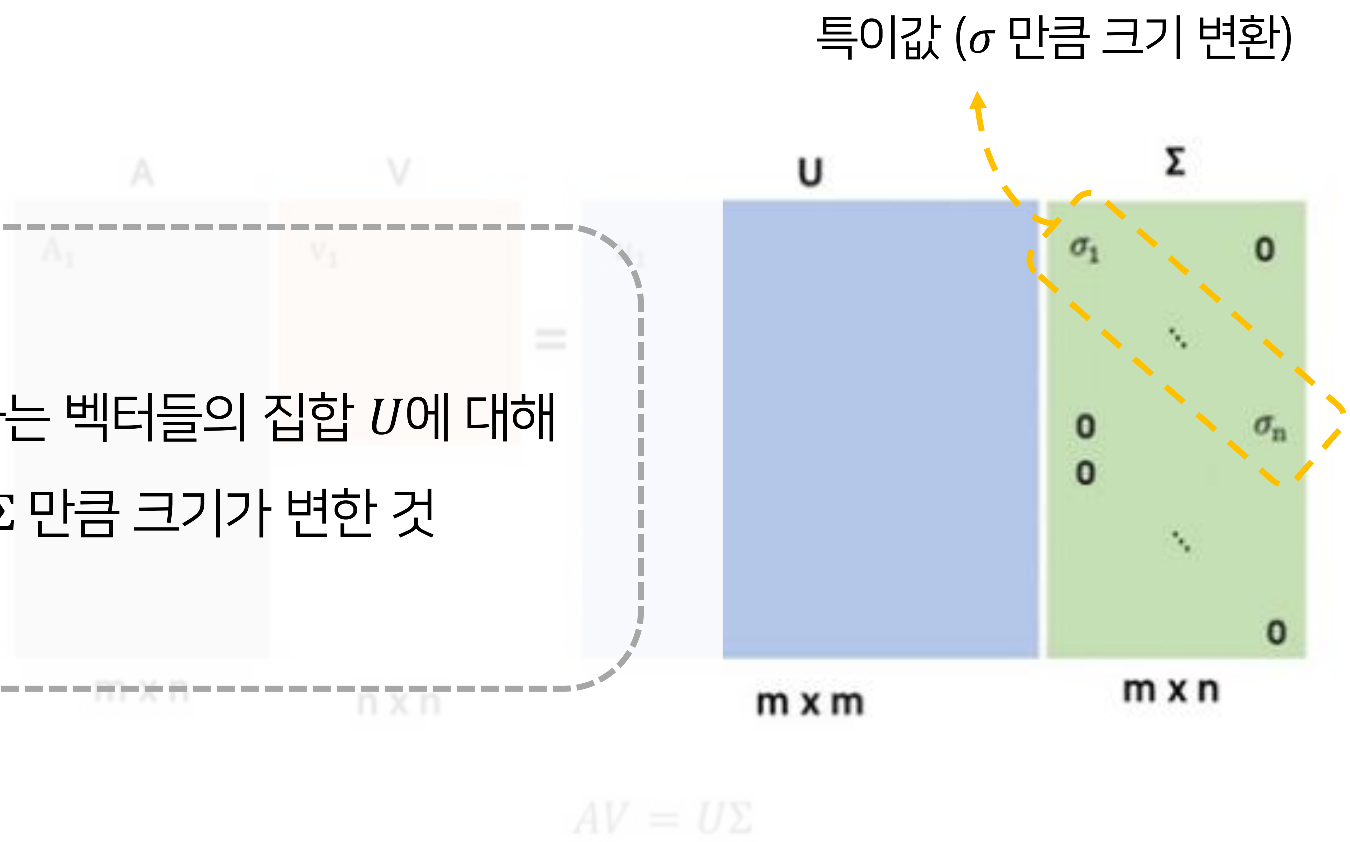
특이값 분해(SVD)



직교하는 벡터들의 집합 V에

선형변환 A 를 진행한 것

특이값 분해(SVD)



직교하는 벡터들의 집합 U에 대해

Σ 만큼 크기가 변한 것

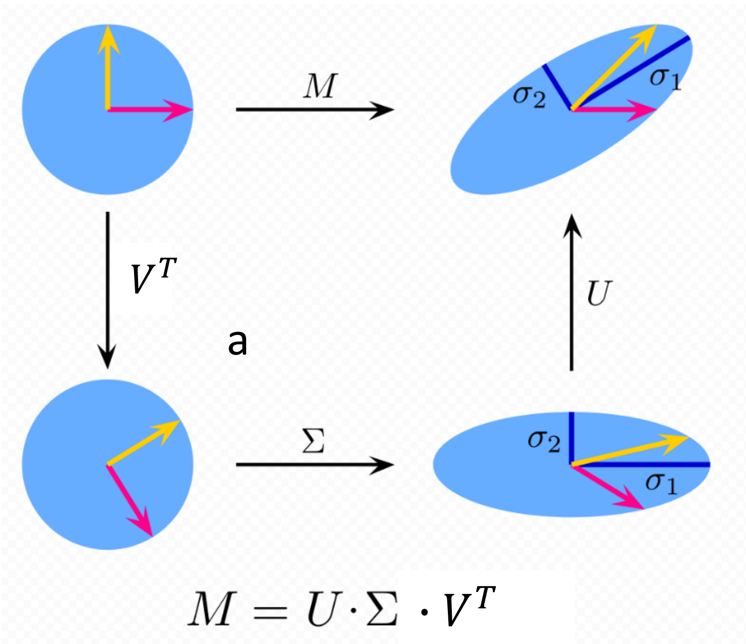
특이값 분해(SVD)

V 에 있는 열벡터를 행렬 A 를 통해 선형변환 했을 때,

크기는 σ 만큼 변하지만

여전히 직교하는 벡터(U의 열벡터)를 찾을 수 있다

특이값 분해의 기하학적 의미



$$x \rightarrow V^{T}x \rightarrow \Sigma V^{T}x \rightarrow U\Sigma V^{T}x\ (Mx)$$

① V^T : 정의역의 표준 기저에서

다른 기저로 기저 변환 (회전변환)

② Σ : 새로운 기저에서 특이값(σ)만큼 크기 변환

③ U: 회전을 통해 기저 변환 (회전변환)

특이값 분해의 이점

SVD로 분해한 행렬 $A = U\Sigma V^T$

$$= \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_m \\ | & | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \end{pmatrix} \begin{pmatrix} - & v_1^T & - \\ - & v_2^T & - \\ & \vdots & \\ - & v_n^T & - \end{pmatrix}$$

$$= \sigma_1 u_1 v_1^T + \cdots + \sigma_m u_m v_m^T$$

특이값 분해의 이점

SVD로 분해한 행렬 $A = U\Sigma V^T$

$$= \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_m \\ | & | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \end{pmatrix} \begin{pmatrix} - & v_1^T & - \\ - & v_2^T & - \\ & \vdots & \\ - & v_n^T & - \end{pmatrix}$$

행렬 A 와 같은 크기

$$= \sigma_1 u_1 v_1^T + \cdots + \sigma_m u_m v_m^T$$

A 와 같은 크기를 갖는 행렬들의 합으로 생각할 수 있음

이때 각 원소의 크기는 σ 값에 의해 결정

5 특이값 분해(SVD)

특이값 분해의 이점

SVD로 분해한 행렬 $A = U\Sigma V^T$

행렬 A 와 같은 크기

$$= \sigma_1 u_1 v_1^T + \cdots + \sigma_m u_m v_m^T$$

A 행렬을 여러 Layer로 쪼개서 생각할 수 있다

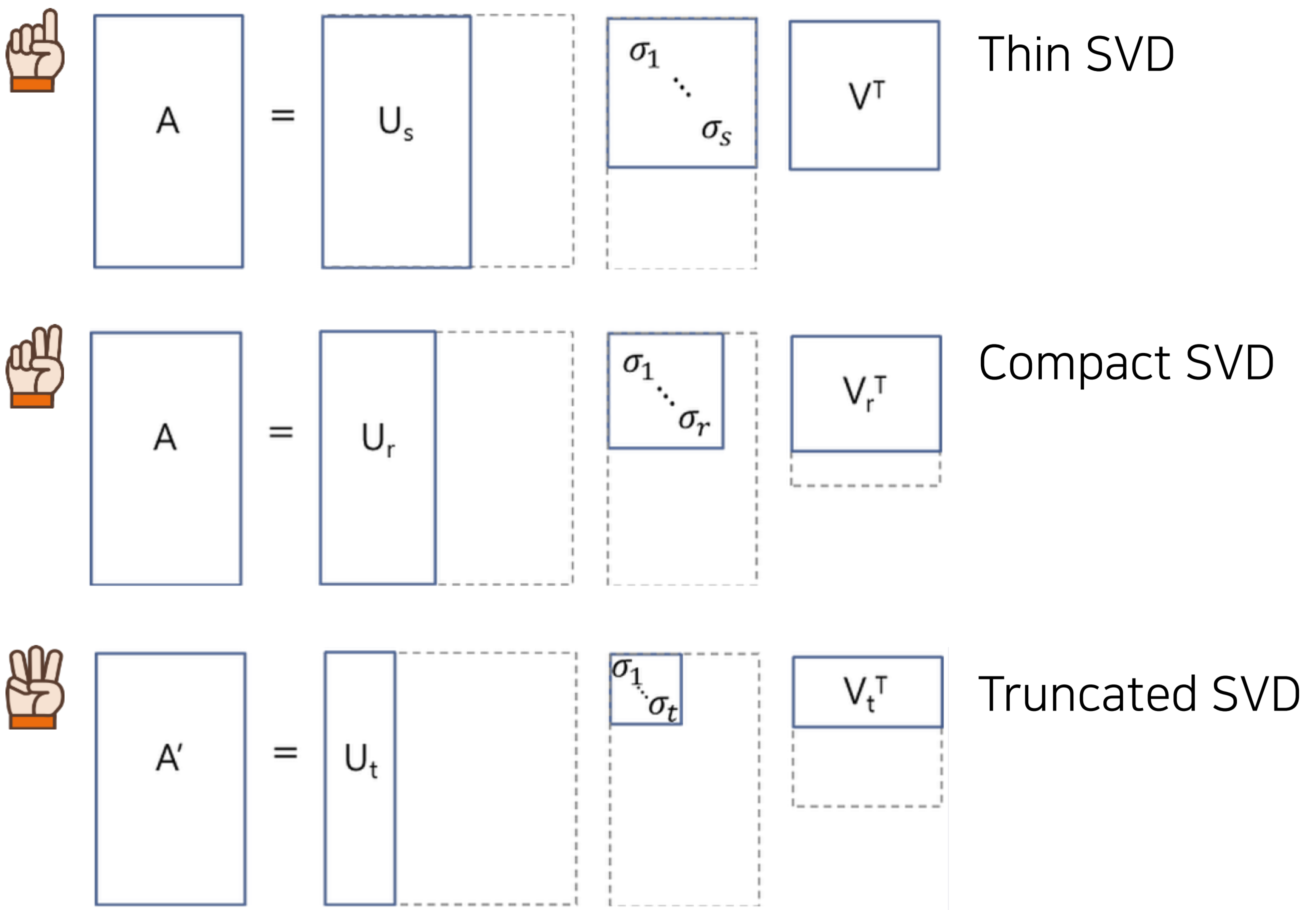
A 와 같은 크기를 갖는 행렬들의 합으로 생각할 수 있음

이때 각 원소의 크기는 σ 값에 의해 결정

5 특이값 분해(SVD)

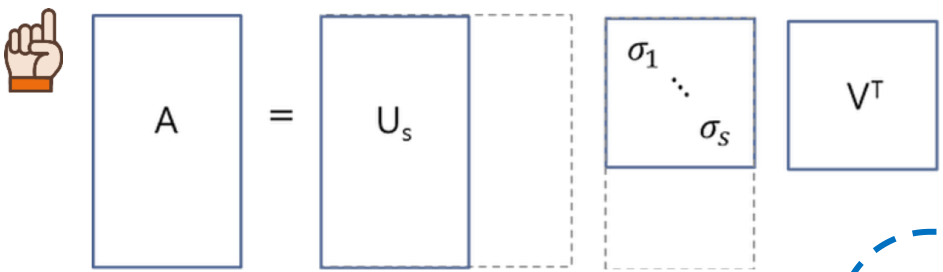
특이값 분해의 다양한 형태

기존 Full SVD는 연산량이 많아 간략화



특이값 분해의 다양한 형태

기존 Full SVD는 연산량이 많아 간략화



Thin SVD

Compact SVD

Truncated SVD

Σ 행렬 아랫부분(비대각 파트)과

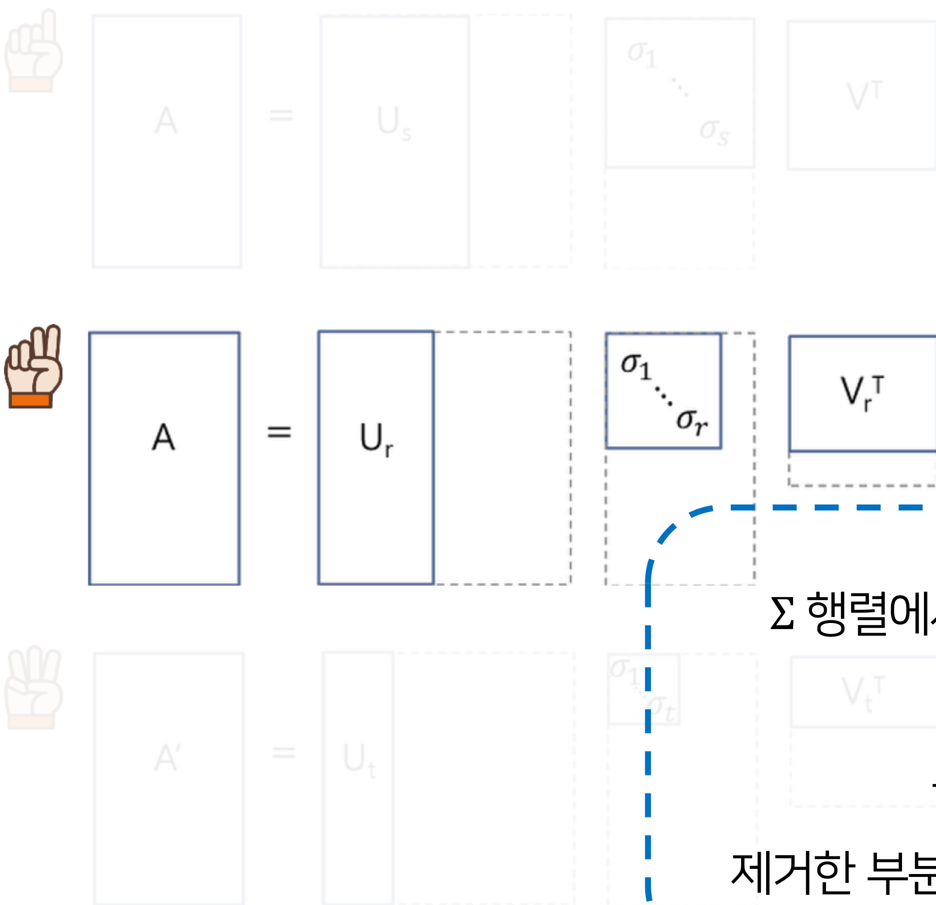
U 에서 이에 대응되는 부분 제거

제거한 부분은 연산을 진행해도 0이므로

A 복원 가능

특이값 분해의 다양한 형태

기존 Full SVD는 연산량이 많아 간략화



Thin SVD

Compact SVD

Σ 행렬에서 비대각 파트, 특이값이 0인 부분 제거

대응하는 U 와 V^T 요소 제거

→ 특이값이 양수인 부분만 남김

제거한 부분은 연산을 진행해도 0이므로 A **복원 가능**

5 특이값 분해(SVD)

특이값 분해의 다양한 형태

기존 Full SVD는 연산량이 많아 간략화

Thin SVD

Compact SVD

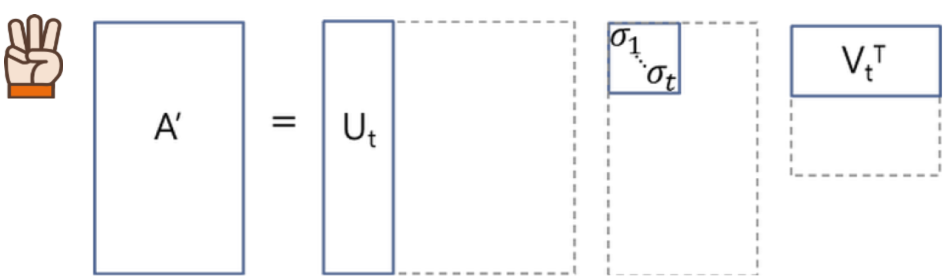
특이값 가운데 상위 t개 선정

A **복원은 불가능**하지만,

데이터를 압축해 행렬 A 근사 가능

→ 데이터의 핵심만 사용, 연산속도 ↑

데이터 압축, 불필요한 노이즈 제거 등에 활용



Truncated SVD

Truncated SVD 이미지 압축

원본 이미지, 행렬 A (600 * 367)

5 특이값 분해(SVD) 응용

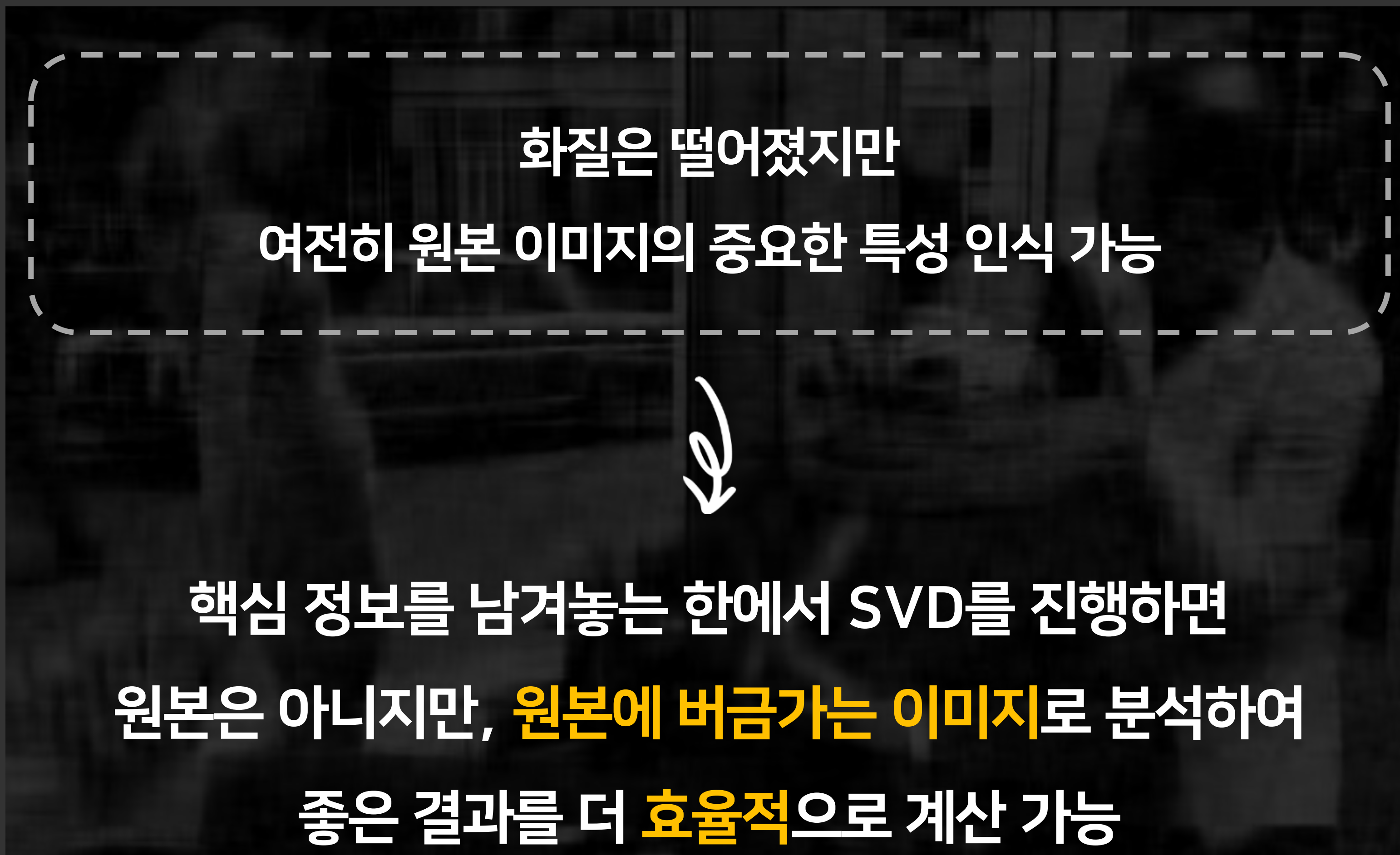
Truncated SVD 이미지 압축

Truncated SVD, 특이값 367개 중 100개만 사용

Truncated SVD 이미지 압축

Truncated SVD, 특이값 367개 중 20개만 사용

Truncated SVD 이미지 압축



특이값 367개 중 20개만 사용

다음 주 예고

1. 주성분 분석

2. 특이값 분해 응용

3. 커널과 커널트릭

감사합니다