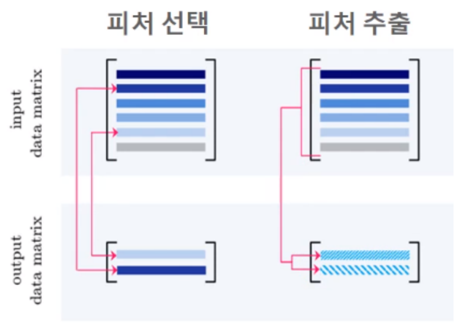
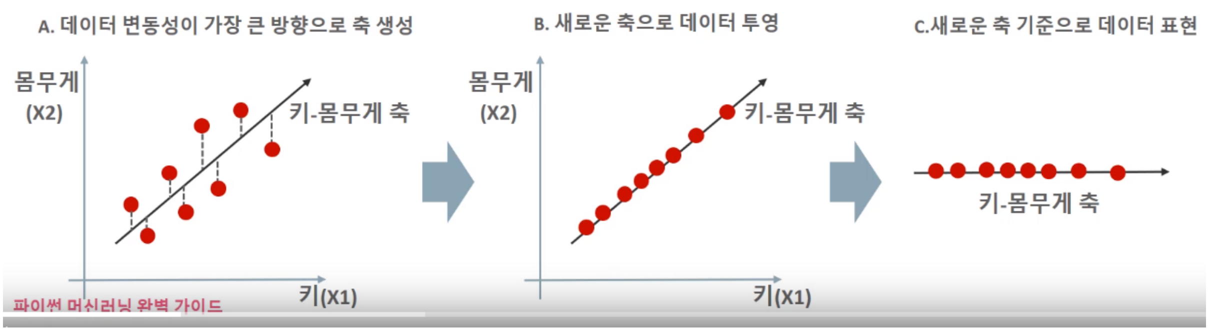
Dimensionality Reduction (2)

PCA (주성분 분석)

: 차원축소 중 변수추출 방법. 원 데이터의 분산을 최대한 보존하는 새로운 축을 찾아 고차원 공간의 표본들을 선형 연관성이 없는 저차원 공간으로 변환하는 기법

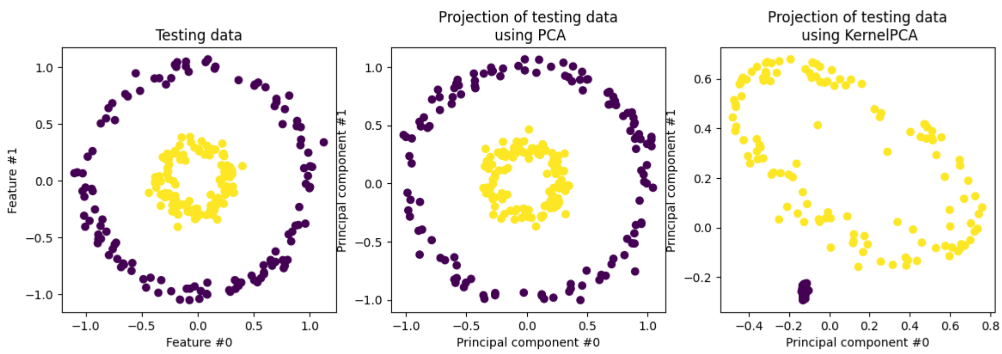




- 하나의 주성분 (PC1)을 새로운 축으로 선형 변환
- 공분산 행렬의 고유벡터 = 주성분 (Principal Component)

Kernel PCA

: 비선형 함수인 커널 함수를 이용하여 비선형 데이터를 고차원 공간으로 매핑하는 기법. 이때 데이터는 선형 분류기에 적합하게 변형됨

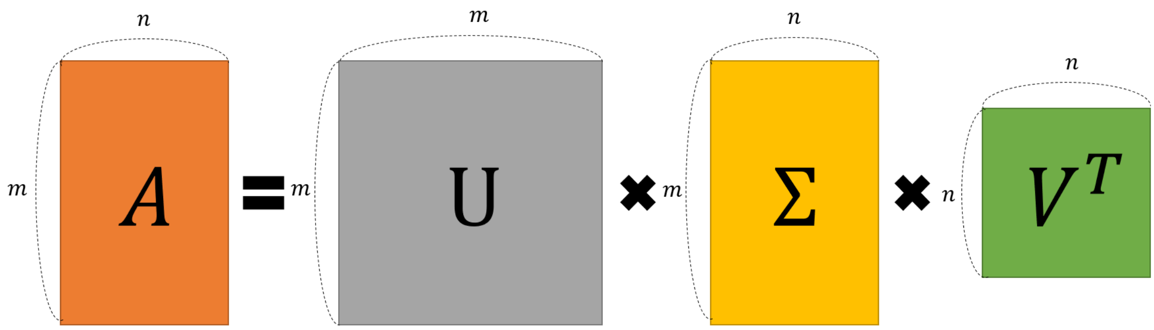


- 커널 종류: gaussian, polynomial, sigmoid, linear

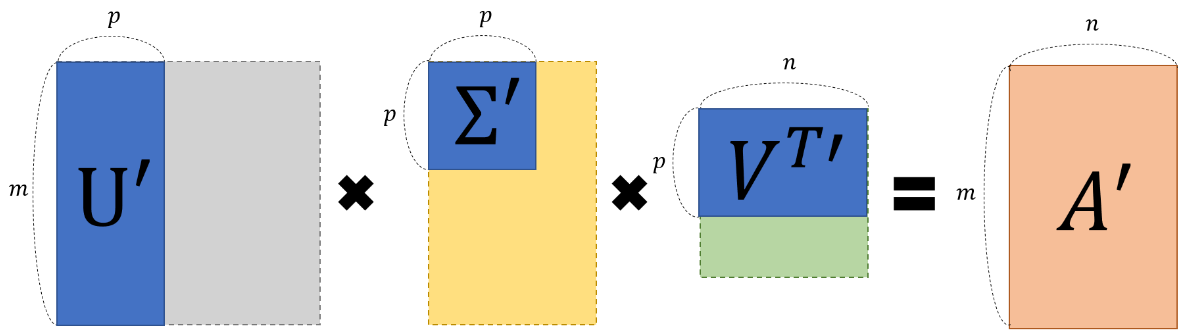
SVD (특이값 분해)

: 행렬 분해 방법 중 하나. 임의의 $m \times n$차원의 행렬 A에 대하여 다음과 같이 행렬을 분해할 수 있음

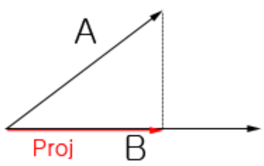
$$A = U\Sigma V^T$$



- SVD를 통해 얻어진 U, Sigma, V 행렬의 일부를 사용해 A'를 부분복원함. 중요한 정보들을 부분복원하면 결과적으로 사진 표현하고자 하는 내용을 살릴 수 있음



Projection (투영)



Proj = 벡터 A를 벡터 B에 투영한 벡터. 수선을 그어 생긴 직각삼각형을 기반으로 삼각함수를 활용해 공식 유도

$$Proj = (\frac{A \cdot B}{|B|^2})B$$

Determinant (행렬식)

: 행렬을 대표하는 값

Ex) 3x3 행렬 A의 행렬식

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

$$detA = a_{11}detA_{11} - a_{12}detA_{12} + a_{13}detA_{13}$$

$$= 1 \cdot det \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5 \cdot det \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0 \cdot det \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix}$$

$$= 1(0-2) - 5(0-0) + 0(-4-0) = -2$$

- $detA_{11}$=A에서 1행과 1열을 제외한 행렬의 행렬식
- 행렬식의 성질

> (1) 행렬 A의 임의의 행에 스칼라 곱을 한 뒤 다른 행에 더해 B를 만들었을 때 두 행렬의 행렬식은 같다.
>
> (2) 행렬 A의 임의의 행을 다른 행과 바꾸어 B를 만들었을 때 $detB = -detA$
>
> (3) 행렬 A의 임의의 행에 스칼라 곱을 해 B를 만들었을 때 $detB = kdetA$
>
> (4) **삼각행렬(triangular matrix)**의 행렬식은 주 대각원소들의 곱과 같다.
>
> (5) 행렬 A가 **가역(invertible)**임과 $detA \neq 0$임은 동치입니다.
>
> (6) $detA^T = detA$
>
> (7) $detAB = (detA)(detB)$

참고

https://ratsgo.github.io/machine learning/2017/04/24/PCA/

https://ml-explained.com/blog/kernel-pca-explained

https://angeloyeo.github.io/2019/08/01/SVD.html

https://sonagi87174.tistory.com/19

https://ratsgo.github.io/linear algebra/2017/05/21/determinants/