[PCA 직관적 정리]

(중요한 열은..)
PCA로 선형독립인 열을 남기는
축소, 압축 작업!

→ DS에서, 열벡터간 (데이터의) 선형종속이 있어, 그대로 쓸 순 없고, 버리기 아까울 때

· PCA : 차원 축소 의미 (노이즈 제거, 데이터, 연산량 ↓)

· latent variable로 데이터를 압축 해준다.
(차원축소)

ex) '꽃의 크기'라는 잠재적 변수가 '꽃의 폭', '꽃의 높이'라는 데이터로 세분화되어 표현된 것일 뿐.

* latent variable : 바로 측정되지는 않지만, 측정된 데이터들의 기저에 숨어서 측정 데이터를 결정짓는 것
PCA는 이 잠재변수들을 찾아 공간을 구성한다.
(더 낮은 차원으로)

* PCA의 가정 : $u_i = W^T x_i$
(u_i = 잠재변수 = 측정데이터 벡터 (x_i)의 각 원소를 $\underbrace{\text{조합해}}_{\text{가중치 벡터} = W}$ 만든 것)

측정 데이터와 잠재변수가
선형관계에 있다는 가정.

ex) $\underbrace{u_i}_{i\text{번째 꽃 크기}} = w_1 \underbrace{x_{i1}}_{\text{꽃 길이 관측 data}} + w_2 \underbrace{x_{i2}}_{\text{꽃 폭 관측 data}}$

· latent variable 찾기 ⟺ 이를 구성하는 가중치 벡터 찾기.
(w_1, w_2)

어떻게? 역시나 선형예제 근사!
= k차원 근사 (로우랭크 근사문제)
= 투영 문제 ⟺ 차원축소 문제
\+
"근사성능 높이기 위해 직선이 원점을 지나야 한다는 제한조건 없앰"

PCA : N개의 m차원 데이터벡터 $x_1, \cdots, x_n$에서 특정벡터 x_0를 뺀, $x_1 - x_0, \cdots, x_n - x_0$를 정규직교 ~~[illegible]~~ 기저벡터 $w_1, \cdots, w_k$로 이루어진 k차원 공간에 투영하여

$x_1, \cdots, x_n$ 데이터와 가장 비슷한 $x_1^{\parallel w}, \cdots, x_n^{\parallel w}$를 만들기 위한

정규직교 ~~[illegible]~~ 기저벡터 $w_1, \cdots, w_k$와 상수벡터 x_0를 찾는다.

(ex) · x_0는 $\frac{1}{N}(x_1 + \cdots + x_n)$ 평균벡터

· $w_1, \cdots, w_k$는 내림차순 k개 특이값에 대응하는 우특이벡터 $v_1, \cdots, v_k$이다.

[PCA 수학적 정리]

1) $\underset{k\times1}{\hat{\vec{x_i}}} = \underset{k\times m}{W}\,\underset{m\times1}{\vec{x_i}} \iff \underbrace{\begin{bmatrix} \hat{x}_{i1} \\ \vdots \\ \hat{x}_{ik} \end{bmatrix}}_{\hat{\vec{x_i}}}^{k\times1} = \underbrace{\begin{pmatrix} w_{11} & w_{1m} \\ w_{k1} & w_{km} \end{pmatrix}}_{W}^{k\times m} \underbrace{\begin{bmatrix} x_{i1} \\ \vdots \\ x_{im} \end{bmatrix}}_{\vec{x_i}}^{m\times1}$

$\hat{X} = XW^T \quad \underset{n\times k}{\begin{bmatrix} -\hat{x}_1- \\ \vdots \\ -\hat{x}_n- \end{bmatrix}} = \underset{n\times m}{\begin{bmatrix} -x_1- \\ \vdots \\ -x_n- \end{bmatrix}} \underset{m\times k = (k\times m)^T}{W^T}$

2) SO, $\hat{x}_i$와 x_i
$\hat{X}$와 X를 가장 유사하게 해주는
W (변환행렬)을 찾는다.

3) 비교는 x_i와 $U\hat{x}$로
그 전에! W의 역기능하는 U도 찾아야함.
(m→k차원) (k→m차원)
원상복구, 차원 회귀

* x_i와 $\hat{x}_i$는 서로 차원이 달라 직접적 비교불가.
ex) (3,2,1,4) (2,1)
∴ $\hat{x}_i$를 다시 원래 차원으로 돌려서 비교해보자!

$\hat{\hat{x}}_i = U\hat{x}_i \longleftrightarrow x_i$
(m차원) (k차원) m차원
비교!

* $WU = I$,
$\hat{\hat{x}} = U\hat{x}$,
$W\hat{\hat{x}} = \hat{x} = WU\hat{x}$
$\therefore WU = I$

4) so, $\underset{\hat{x}}{\arg\min} \| x - U\hat{x} \|^2$

$\| x - U\hat{x} \|^2 = (x - U\hat{x})^T (x - U\hat{x}) = (x^T - \hat{x}^T U^T)(x - U\hat{x})$
$= x^T x - 2x^T U\hat{x} + \hat{x}^T\hat{x}$

$\hat{x}$로 미분.

$-2U^T x + 2\hat{x} = 0$

$\iff \hat{x} = U^T x, \quad \hat{x} = Wx$

$\iff U^T = W, \quad U = W^T, \quad W^T W = I$

5) W찾기는 rank-k 근사로! $\arg\min \| X - XW^T W \|^2$