chap3. orthogonality

orthogonal vectors

왜 orthogonal vectors에 대해 배우는 것일까? 이상화 교수님께서는 independent basis와 easy calculation이라고 말씀하셨다.

- length of vector : $\parallel \vec{x} \parallel$
  $\parallel \vec{x} \parallel$ = $\vec{x}^T\vec{x}$

수직하는 두 벡터를 상상해보자. 우리는 직각삼각형의 모양이 만들어짐을 상상할 수 있다. 피타고라스의 정리에 의해

$$\parallel \vec{y} - \vec{x} \parallel^2 = \parallel \vec{x} \parallel^2 + \parallel \vec{y} \parallel^2$$

이고 이를 계산하면 $(\vec{y}-\vec{x})^T(\vec{y}-\vec{x}) = \vec{x}^T\vec{x} + \vec{y}^T\vec{y}$ 이고 전개하여 계산하면 $\vec{y}^T\vec{x} + \vec{x}^T\vec{y} = 0$이 나온다. 이 두 값은 같기 때문에 결론은 $\vec{x}^T\vec{y} = 0$이다. 즉, 수직인 두 벡터를 내적하면 0이다.

- $\vec{x}^T\vec{y}$
  1.$\vec{x}^T\vec{y} = 0$: 직각
  2.$\vec{x}^T\vec{y} < 0$: 둔각
  3.$\vec{x}^T\vec{y} > 0$: 예각
- if non-zero vector $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ are orthogonal then the vectors are linearly independent.
  (proof)
  orthogonal한 벡터들 $\vec{v}_i$의 linear combination
  $c_1\vec{v}_1 + c_2\vec{v}_2 + ... + c_n\vec{v}_n = 0$에서 모든 c_i가 0이면 linearly independent이다. 위 식에 $\vec{v}_i^T$를 곱해주면 $\vec{v}_i^T c_1\vec{v}_1 + \vec{v}_i^T c_2\vec{v}_2 + ... + \vec{v}_i^T c_n\vec{v}_n = 0$ 이고 orthogonal한 벡터들이기 때문에 다 0이 되고 $c_i\vec{v}_i^T = 0$ 만 남고 이는 $c_i \parallel \vec{x} \parallel^2 = 0$ 이다. 따라서 좌변이 0이 되기 위해서는 $c_i = 0$일여야만 하고 모든 c_i가 0 이므로 linearly independent이다,
- 결국 orthogonal vectors는 그 벡터들이 span하는 vector space basis vectors 가 된다.

orthogonal subspace

어떤 subspace의 모든 vectors가 다른 어떤 subspace의 모든 vectors와 orthogonal할 때 이를 orthogonal subspace라고 한다.

- row space ⊥ Null space (A_{m*n})
  $A\vec{x} = 0$ 의 모습을 생각해보면 이해완료
- col space ⊥ left null space
  $A^T\vec{x} = 0$의 모습을 생각해보면 이해완료

orthogonal complement subspace

$\vec{v} \in V, \vec{w} \in W \ in \ R^n$ 일 때, $v \perp w$ 이고 $dim(V) + dim(W) = n$인 경우를 orthogonal complement subspace라고 한다.

projection onto line

그림을 그리지 못해서 아쉽지만 정리를 해본다.
원점을 지나는 벡터 $\vec{a}, \vec{b}$를 상상해보자. $\vec{b}$에서 $\vec{a}$로 수선의 발을 내린다고 하면 $(projection\ \vec{b}\ onto\ \vec{a})$ 수선의 발을 p라고 하고 원점과 p의 거리($\vec{p}$의 length)를 k라고 하자. $(\vec{b} - k\vec{a}) \perp \vec{a}$이므로 $\vec{a}^T(\vec{b} - k\vec{a}) = 0$이고 이를 통해

$$k = \frac{\vec{a}^T\vec{b}}{\vec{a}^T\vec{a}}$$

$$\vec{p} = k\vec{a} = \frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}\vec{b}$$

위에서 $\frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}$ 을 projection marix P라고 한다.

- projection matrix : P
  $P^T = P, P^2 = P$

projection & least square

이전에 벡터를 벡터에 projection 하였다면 이제 space에 projection해보자. 미지수의 수보다 방적식의 수가 많은 경우 $A\vec{x} = \vec{b}$ 에서 $\vec{x}$를 어떻게 구할 수 있을까? exact한 해를 구할 수 없기 때문에 근접한 해를 구한다. $\vec{x}'$은 $A\vec{x}'$을 $\vec{b}$에 가장 가깝게 만들어 주는 값이다.

$$min \ ||A\vec{x}' - \vec{b}||$$

여기서 $\vec{x}'$를 구하는 과정이 least square이다. 이전에 배운대로 projection을 하면 된다. $A\vec{x}$는 A의 column vector linear combination이므로 $\vec{b}$를 A의 column vector가 span하는 vector space에 projection하면 된다. 그 수선의 발에 해당하는 vector가 $A\vec{x}'$이고 $(\vec{b} - A\vec{x}') \perp A$이므로 이전에 했던대로 계산하면

$$\vec{x}' = (A^TA)^{-1}A^T\vec{b}$$

$$\vec{p} = A\vec{x}' = A(A^TA)^{-1}A^T\vec{b}$$

$$projection \ matrix = A(A^TA)^{-1}A^T$$

Gram-schmidt orthonormalization

그람슈미츠 정규직교화과정은 임의의 벡터들을 통해 정규직교벡터들로 표현하는 것이다. $\vec{a_1}, \vec{a_2}, ..., \vec{a_n}$ vector들이 span하는 vector space를 정규직교벡터들로 표현할 수 있게 된다. 그림이 없어서 이해는 어렵지만 수식을 정리해봤다.

$\vec{a_1}, \vec{a_2}, ..., \vec{a_n}$ 들이 있다.

(1) $\vec{a}_1$ 를 크기 1로 normalize :

$$\frac{\vec{a}_1}{||\vec{a}_1||} = \vec{q}_1$$

(2) project $\vec{a}_2$ onto $\vec{q}_1$, 수선의 발에 해당하는 벡터는

$$\frac{\vec{q_1}^T\vec{a}_2}{\vec{q_1}^T\vec{q_1}}\vec{q}_1 = (\vec{q}_1^T\vec{a}_2)\vec{q}_1$$

(3) 위에서 구한 벡터를 이용하여

$$\frac{\vec{a}_2 - (\vec{q}_1^T\vec{a}_2)\vec{q}_1}{||\vec{a}_2 - (\vec{q}_1^T\vec{a}_2)\vec{q}_1||} = \vec{q}_2$$

여기서 $\vec{q}_1 \perp \vec{q}_2$ 이다.

(4) 위에서 구한 vector로 $\vec{a}_2 = (\vec{q}_1^T\vec{a}_2)\vec{q}_1 + (\vec{q}_2^T\vec{a}_2)\vec{q_2}$

$$\frac{\vec{a}_3 - ((\vec{q}_1^T\vec{a}_3)\vec{q}_1 + (\vec{q}_2^T\vec{a}_3)\vec{q}_2)}{\left\|\vec{a}_3 - ((\vec{q}_1^T\vec{a}_3)\vec{q}_1 + (\vec{q}_2^T\vec{a}_3)\vec{q}_2)\right\|} = \vec{q}_3$$

여기서 $\vec{q}_3 \perp \vec{q}_2\ and\ \vec{q}_2$ 이다.

이러한 과정을 반복하면서 우리는 orthogonal한 vector들을 통해 원래 vector들을 표현할 수 있고 그 vector space를 span하는 orthonomal basis를 구할 수 있다.

- 위에서 배운 걸 이용하여 for orthonomal vectors $\vec{q}_1, \vec{q}_2, ..., \vec{q}_n \in R^n$ (sqaure) , $\vec{x} = \sum_{i=1}^{n} c_i\vec{q}_i$ 여기서 c_i 를 구하고 $\vec{x}$ 를 표현하고 싶다. 양변에 $\vec{q}_i^T$ 를 곱하면

$$\vec{q}_i^T\vec{x} = c_i\vec{q}_i^T\vec{q}_i = c_i$$

이고 따라서 $\vec{x} = \sum c_i\vec{q}_i = \sum(\vec{q}_i^T\vec{x})\vec{q}_i$ 으로 표현할 수 있다. 이번에는 square가 아니라 rectangular한 경우를 생각해보자. A_{m*n} 이라고 하면 m < n 인 경우만 가능하다 (left inverse). m > n 경우는 basis가 안되기 때문이다.

A = QR 분해

위에서 벡터를 orthonomal vectors로 표현한 것을 이용한 분해 방법이다. $\vec{a}_i$를 행렬 A의 col vector라고 하면

$$A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & .... & \vec{a}_n \end{bmatrix} = \begin{bmatrix} (\vec{q_1}^T\vec{a}_1)\vec{q}_1 & (\vec{q_1}^T\vec{a}_2)\vec{q}_1 + (\vec{q_2}^T\vec{a}_2)\vec{q}_2 & ... & (\vec{q_1}^T\vec{a}_n)\vec{q}_1 + ... + (\vec{q_1}^T\vec{a}_n)\vec{q}_n \end{bmatrix}$$

$$= \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & ... & \vec{q}_n \end{bmatrix} \begin{bmatrix} (\vec{q}_1^T\vec{a}_1) & (\vec{q}_1^T\vec{a}_2) & ... & (\vec{q}_1^T\vec{a}_n) \\ 0 & (\vec{q}_2^T\vec{a}_2) & ... & (\vec{q}_2^T\vec{a}_n) \\ ... & ... & ... & ... \\ 0 & 0 & ... & (\vec{q}_n^T\vec{a}_n) \end{bmatrix} = QR$$