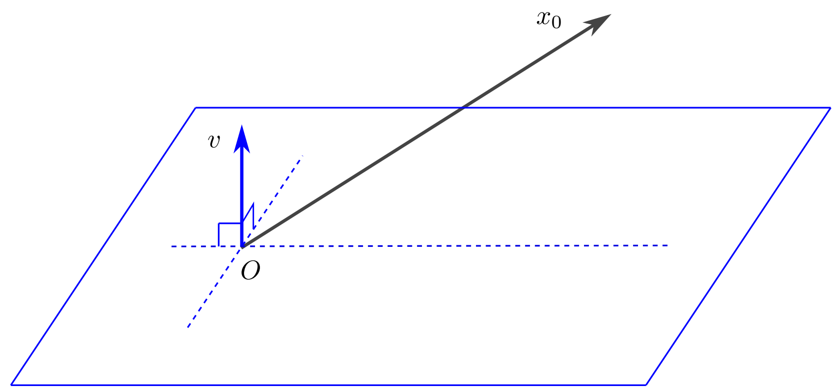
ASE2030 Linear Algebra and Statistics: Midterm exam (3 problems)

시험 시작 전, 다음의 '학생 명예선서(Honor Code)'를 답지 맨 위에 적고 서명하시오.
"나는 정직하게 시험에 응할 것을 서약합니다."
"By signing this pledge, I promise to adhere to exam requirements and maintain the highest level of ethical principles during the exam period."

1) *Operations around a plane (12 points).* 다음 그림과 같이 원점을 지나는 n차원 공간 상의 한 평면은, 그 평면과 수직하며 크기가 1인 normal vector v로 ($\|v\| = 1$) 표현할 수 있다. 참고로 해당 평면 위에 있지 않은 한 벡터 x_0를 그림에 함께 표시하였다.

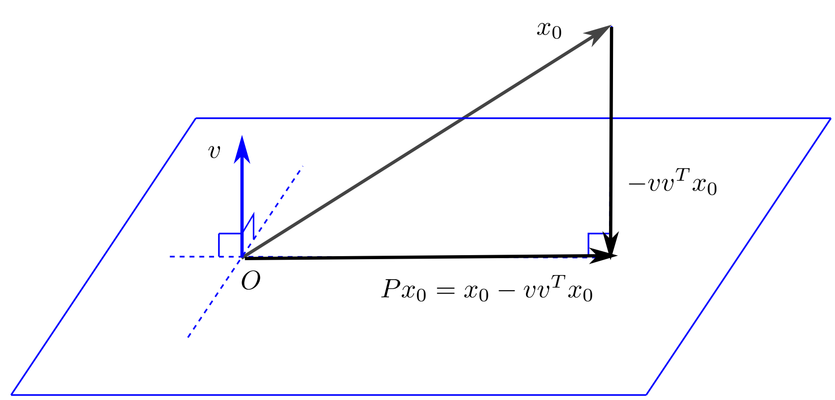


이와 같은 벡터 v에 대해 다음의 두 행렬을 정의하자.

$$P = I - vv^T$$
$$R = I - 2vv^T$$

a) 행렬 P에 의한 연산 Px_0는 벡터 x_0를 어디로 이동시키는지, 위의 그림에 덧붙여 명확하게 설명하시오 (3pts).

Solution. 벡터 vv^Tx_0는 x_0의 v방향 성분이므로, Px_0는 x_0에서 평면에 수직한 방향의 성분을 제거한 평면상의 벡터가 된다. 참고로, Px_0는 평면상의 모든 점 중 x_0와 가장 가까운 점이 되며, 이러한 벡터 Px_0를 평면 v에 대한 x_0의 projection 이라 부른다.



b) 행렬 P로 정의되는 linear dynamical system, $x_{t+1} = Px_t$에 대해 초기치 x_0가 위의 그림과 같이 주어졌을 때, 이후의 trajectory인 $x_1, x_2, x_3, \ldots$는 어디를 지나는가? 즉, x_k는 (k는 자연수) 무엇인가 (3pts)?

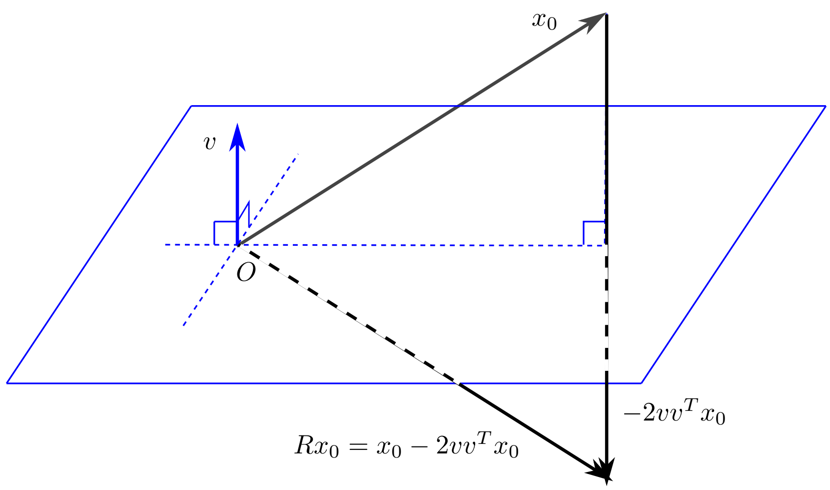
Solution. $x_1 = Px_0$에서 이미 v방향 성분이 제거되었으므로, 모든 $k = 2, 3, \ldots$에 대해 $x_k = x_1 = Px_0$가 된다. 또한,

$$\begin{aligned} P^2 &= (I - vv^T)(I - vv^T) = I - 2vv^T + vv^Tvv^T \\ &= I - 2vv^T + v(v^Tv)v^T = I - vv^T = P \end{aligned}$$

이므로, 자연수 k에 대해 $P^k = P$가 되어 $x_k = P^kx_0 = Px_0$임을 확인할 수도 있다.

c) 행렬 R에 의한 연산 Rx_0는 벡터 x_0를 어디로 이동시키는지, 위의 그림에 덧붙여 명확하게 설명하시오 (3pts).

Solution. Rx_0는 x_0에서 $-vv^Tx_0$를 두 번 빼 준 벡터이므로, R에 의해서 x_0는 평면에 대해 대칭인 벡터로 이동된다. 벡터 Rx_0는, 평면 v에 대한 x_0의 reflection이라 할 수 있다.



d) 행렬 R로 정의되는 linear dynamical system, $x_{t+1} = Rx_t$에 대해, 초기치 x_0가 위의 그림과 같이 주어졌을 때, 이후의 trajectory인 $x_1, x_2, x_3, \ldots$는 어디를 지나는가? 즉, x_k는 (k는 자연수) 무엇인가 (3pts)?

Solution. x_1은 x_0를 평면 반대쪽 대칭점으로 옮기고, x_2는 x_1을 다시 반대편 대칭점으로 옮기므로 $x_2 = x_0$가 된다. 즉, 짝수 k에 대해 $x_k = x_0$가 되고, 홀수 k에 대해 $x_k = x_1$이 된다. b)에서와 비슷한 방법으로, $R^2 = I$임을 보이고, k가 짝수일때 $R^k = I$, k가 홀수일 때 $R^k = R$임을 확인할 수도 있다.

$$\begin{aligned} R^2 &= (I - 2vv^T)(I - 2vv^T) = I - 4vv^T + 4vv^Tvv^T \\ &= I - 4vv^T + 4v(v^Tv)v^T = I \end{aligned}$$

2) *PQ factorization (10 points).* Linearly independent한 row들로 구성된 행렬 $A \in \mathbb{R}^{m\times n}$를 생각하자. 행렬 A의 row들을 $a_1^T, \dots, a_m^T$로 표현한다.

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}$$

이와 같은 행렬 A는 항상 다음과 같이 $A = PQ$로 분해될 수 있는데, 여기서 P는 정사각형 lower triangular 행렬이며, Q는 orthonormal한 row들로 구성되어 있다.

$$A = PQ = \begin{bmatrix} p_1^T \\ \vdots \\ p_m^T \end{bmatrix} \begin{bmatrix} q_1^T \\ \vdots \\ q_m^T \end{bmatrix}$$

a) 행렬 Q의 생김새와 특징 등에 대해 설명하시오 (2pts).

Solution. Q는 A와 크기가 같으며, A가 정사각형 또는 wide하므로, Q도 그렇다. Q의 row들이 orthonormal하므로 Q^T가 orthogonal하며 $QQ^T = I$가 된다.

b) 행렬 P와 Q의 원소들을 어떻게 계산 수 있는지, (수업을 제대로 들은) 여러분의 동료들이 그 수치들을 계산해낼 수 있을 정도로 구체적으로 설명하시오 (4 pts).

Solution. A^T의 column들이 linearly independent하므로, A^T의 QR factorization을 통해 P와 Q를 찾을 수 있다. A^T의 QR factors를 $\tilde{Q}$와 $\tilde{R}$이라 하면 ($\tilde{Q}$는 orthogonal, $\tilde{R}$은 square upper triangular),

$$A^T = \tilde{Q}\tilde{R}$$

양 변을 transpose하여 $A = \tilde{R}^T\tilde{Q}^T$를 얻을 수 있고, $\tilde{R}$과 $\tilde{Q}$로부터 문제에서 원했던 $P = \tilde{R}^T$, $Q = \tilde{Q}^T$를 찾을 수 있다.

c) A의 모든 row들을 합한 벡터와 P의 모든 row들을 합한 벡터는, 차원은 서로 다르지만 길이는 같음을 보이시오. 즉, 다음을 보이시오 (4 pts).

$$\|a_1 + a_2 + \cdots + a_m\| = \|p_1 + p_2 + \cdots + p_m\|$$

Solution. A^T와 P^T의 Gram matrix를 각각 G, H라 하면, $G = H$임을 알 수 있다.

$$G = AA^T = PQQ^TP^T = P(QQ^T)P^T = PP^T = H$$

그런데 $G_{ij} = a_i^Ta_j$이며, $H_{ij} = p_i^Tp_j$이므로,위의 관계에서 모든 i, j에 대해 $a_i^Ta_j = p_i^Tp_j$임을 알 수 있다. 그러므로 다음 결론을 얻는다.

$$\begin{aligned} \|a_1 + a_2 + \cdots + a_m\|^2 &= \sum_{i=1}^m\sum_{j=1}^m a_i^Ta_j = \sum_{i=1}^m\sum_{j=1}^m p_i^Tp_j \\ &= \|p_1 + p_2 + \cdots + p_m\|^2 \end{aligned}$$

3) *Orthogonality of sets (8 points).* 같은 차원의 벡터들로 구성된 두 집합 $\mathcal{X}$와 $\mathcal{Y}$에 대해, 각각의 집합에서 임의의 원소들을 꺼냈을 때 그 둘이 항상 수직이면, 두 집합이 수직이라고 말하며 다음과 같이 표현한다.

$$\mathcal{X} \perp \mathcal{Y}$$

즉, 임의의 $x \in \mathcal{X}$와 임의의 $y \in \mathcal{Y}$에 대해 $x \perp y$이면 ($x^T y = 0$이면) $\mathcal{X} \perp \mathcal{Y}$이다.

이제, linearly dependent한 column들로 구성된 행렬 $A \in \mathbb{R}^{m\times n}$와 그에 따라 정의되는 다음의 두 집합, range와 nullspace를 생각하자.

$$\begin{aligned} \text{range}(A) &= \{Ax \mid x \in \mathbb{R}^n\} \\ \text{null}(A) &= \{x \mid Ax = 0\} \end{aligned}$$

a) A의 nullspace는 무한히 많은 원소를 가지고 있음을 설명하시오 (3 pts).

Solution. A의 linearly dependent한 column들을 $a_1, \ldots, a_n$이라 하면,

$$\beta_1 a_1 + \beta_2 a_2 + \cdots \beta_n a_n = 0$$

을 만족하는 0이 아닌 $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$이 존재하고, 위의 관계는 $A\beta = 0$임을 의미하므로 $\beta \in \text{null}(A)$이다. 이 β에 임의의 scalar c를 곱한 $\beta' = c\beta$ 역시 $A\beta' = Ac\beta = cA\beta = 0$을 만족하므로 $\beta' \in \text{null}(A)$이다. 즉, null$(A)$는 무수히 많은 원소를 가지고 있다.

b) A^T의 range와 A의 nullspace가 수직임을 보이시오. 즉, 다음을 보이시오 (5 pts).

$$\text{range}(A^T) \perp \text{null}(A)$$

Solution. 임의의 $x \in \text{range}(A^T)$와 $y \in \text{null}(A)$를 생각하자. 그러면 $x = A^T z$를 만족하는 $z \in \mathbb{R}^m$이 존재하며, $Ay = 0$이다. x와 y의 inner product를 살펴보면,

$$x^T y = (A^T z)^T y = z^T A y = z^T (Ay) = z^T 0 = 0$$

x와 y는 항상 수직임을 알 수 있으므로, range$(A^T) \perp$ null(A)이다.