딥러닝 과제2

1 $A = \begin{pmatrix} 1 & -2 & 3 & 5 \\ 2 & 2 & -1 & 0 \\ 3 & 0 & 1 & 2 \\ 1 & 0 & 2 & 0 \end{pmatrix}$의 2A, A^T, A^{-1}을 쓰시오. 또한 A가 선형독립인지 아닌지 판단하고 그 이유를 작성하시오. (Python, Matlab 사용 가능)

$$2A = \begin{bmatrix} 2 & -4 & 6 & 10 \\ 4 & 4 & -2 & 0 \\ 6 & 0 & 2 & 4 \\ 2 & 0 & 4 & 0 \end{bmatrix} \quad A^T = \begin{bmatrix} 1 & 2 & 3 & 1 \\ -2 & 2 & 0 & 0 \\ 3 & -1 & 1 & 2 \\ 5 & 0 & 2 & 0 \end{bmatrix} \quad A^{-1} = \begin{bmatrix} -0.235 & -6.235 & 0.588 & -0.059 \\ 0.294 & 0.794 & -0.735 & 0.324 \\ 0.118 & 0.118 & -0.294 & 0.529 \\ 0.294 & 0.294 & -0.235 & -0.176 \end{bmatrix}$$

역행렬이 존재하므로 선형독립

2 놈을 계산하시오.(%0.3f)

%0.3f

(1) $x = (3 \ \ -4 \ \ -1.2 \ \ 0 \ \ 2.3)^T$의 1차, 2차, 3차 놈과 최대 놈

$Norm|_{p=1} = 10.5$

$Norm|_{p=2} = 5.633$

$Norm|_{p=3} = 4.716$

$Norm|_{p=4} = 4.0$

(2) $\begin{pmatrix} 2 & 1 \\ 1 & 5 \\ 4 & 1 \end{pmatrix}$의 프로베니우스 놈 $Norm = 6.928$

P_4

3 윷놀이에서 P( )=0.4, 0.5, 0.6인 3가지 경우에 대해 답하시오.

(1) 각 경우의 확률분포를 구하시오.

(2) 각 경우의 엔트로피를 구하시오.

$P(X=도) = {}_4C_1 \cdot p^3 \cdot (1-p)^1 =$

$P(X=개) = {}_4C_2 \cdot p^2 \cdot (1-p)^2 =$

$P(X=걸) = {}_4C_3 \cdot p^1 \cdot (1-p)^3 =$

$P(X=윷) = {}_4C_0 \cdot p^0 \cdot (1-p)^4 =$

$P(X=모) = {}_4C_4 \cdot p^4 \cdot (1-p)^0 =$

P	0.4	0.5	0.6
도	0.154	0.25	0.346
개	0.346	0.375	0.346
걸	0.346	0.25	0.154
윷	0.130	0.063	0.026
모	0.026	0.063	0.130
H	1.995	2.033	1.995

(3) P( )=0.4와 P( )=0.5의 교차 엔트로피와 P( )=0.4와 P( )=0.6의 교차 엔트로피를 구하시오. 어느 것이 큰지 확인하고 그 이유를 설명하시오.

2.112   2.460

(4) P( )=0.4와 P( )=0.5의 KL 다이버전스와 P( )=0.4와 P( )=0.6의 KL 다이버전스를 구하시오. 어느 것이 큰지 확인하고 그 이유를 설명하시오.

0.117   0.465

확률 분포 간의 거리가 0.4 – 0.6이 더 멀기 때문에 KLD가 더 크다. 같은 원리로 $H(P,Q) = H(P) + KLD$이기 때문에 $H(P)$가 $P=0.4$일때 같아서 교차 엔트로피도 0.4–0.6이 더 크다.

4 다음 합성함수에 대해 답하시오.

$$f(x) = 2\left(\frac{1}{4}(1-2x)^2 - 1\right)^3 - 3\left(\frac{1}{4}(1-2x)^2 - 1\right)^2 - 3$$

(1) 식 (2.53)에 따라 $i(x)$와 $h(x)$를 쓰시오.

$g(x) = 2x^3 - 3x^2 - 3$

$h(x) = \frac{1}{4}x^2 - 1$

$i(x) = 1 - 2x$

$$\left.\begin{aligned} f'(x) &= g'\big(h(x)\big)h'(x) \\ f'(x) &= g'\Big(h\big(i(x)\big)\Big)h'\big(i(x)\big)i'(x) \end{aligned}\right\} \quad (2.53)$$

$f(x)=g(h(i(x)))$ $\quad \frac{\partial g}{\partial h}=6h^2-6h$

$\frac{\partial h}{\partial i}=\frac{1}{2}i^{-}$

$\frac{\partial i}{\partial x}=-2$

(2) 연쇄법칙을 이용하여 $f'(x)$를 구하시오.

(3) $f'(0)$과 $f'(2.1)$을 계산하시오.

-7.875 | 16.773

%.03f

$f'(x)=\frac{\partial g}{\partial h}\cdot\frac{\partial h}{\partial i}\cdot\frac{\partial i}{\partial x}=(6h^2-6h)\cdot\frac{1}{2}i^{-}\cdot(-2)$

$=(6(\frac{1}{4}i^{-2}-1)^2-6(\frac{1}{4}i^{-2}-1))\cdot\frac{1}{2}i^{-}\cdot(-2)$

$=(6(\frac{1}{4}(1-2x)^2-1)^2-6(\frac{1}{4}(1-2x^2)-1))\cdot\frac{1}{2}(1-2x)\cdot(-2)$

5 다음 함수에 대해 답하시오.

$$f(\mathbf{x}) = f(x_1, x_2) = 2x_1^2 + 3x_1x_2 + 2x_2^2 - 4x_1 + 2x_2 - 24$$

$\frac{\partial f}{\partial x_1}=4x_1+3x_2-4$

$\frac{\partial f}{\partial x_2}=3x_1+4x_2+2$

(1) 최소점과 최솟값을 분석적[a]으로 구하시오. $\vec{x}=(\frac{22}{7}, -\frac{20}{7})^T$,

(2) 난수를 생성하여 초깃값 $\mathbf{x}_0 = (1.0, 0.9)^T$를 얻었다고 가정하고, 식 (2.58)을 연속적으로 적용하여(즉, 수치적 방법으로) 얻는 점 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$을 구하시오. 이때 학습률 $\rho = 0.1$을 사용하시오. (1)에서 구한 최소점을 향해 이동하는지 확인하시오.

$$\boldsymbol{\theta} = \boldsymbol{\theta} - \rho\mathbf{g} \quad (2.58)$$

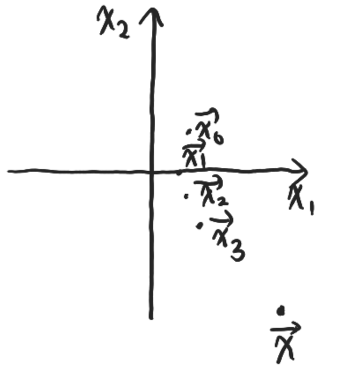
a) 머신러닝에서 사용하는 분석적 해/수치적 해의 의미:

- 분석적 해(Analytical solution): 함수의 극대값/극소값을 도함수를 통해 1번에 구하는 것 (예, $f'(x) = 0$ 을 만족하는 x가 분석적 해가 됨
- 수치적 해(numerical solution): 점진적으로 반복을 통해(iterative) 함수의 극대값/극소값의 근사값을 찾는 방법 (예: gradient descent)

$\vec{x}_1=\begin{bmatrix}1.0\\0.9\end{bmatrix}-0.1\begin{bmatrix}2.7\\8.6\end{bmatrix}\approx\begin{bmatrix}0.7\\0\end{bmatrix}$

$\vec{x}_2=\begin{bmatrix}0.7\\0\end{bmatrix}-0.1\begin{bmatrix}-1.2\\4.1\end{bmatrix}\approx\begin{bmatrix}0.8\\-0.4\end{bmatrix}$

$\vec{x}_3=\begin{bmatrix}0.8\\-0.4\end{bmatrix}-0.1\begin{bmatrix}-2\\2.8\end{bmatrix}\approx\begin{bmatrix}1\\-0.7\end{bmatrix}$



구한 벡터들을 좌표평면 상에 표현해보면 점점 최솟값인 $\vec{x}$에 다가가고 있음을 확인할 수 있다.