

- Norm은 크기를 나타낸다.

Handwritten chalkboard showing vector subtraction and a 2D coordinate system. The equations are:

$$x - y = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

On the right, a 2D coordinate system shows two vectors: $\begin{bmatrix} 1 \\ 2 \end{bmatrix} = x$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix} = y$. A third vector, representing the difference $x - y$, is drawn from the tip of y to the tip of x .

Handwritten chalkboard showing the formula for the 2-norm:

$$\sqrt{(1-2)^2 + (2-1)^2}$$

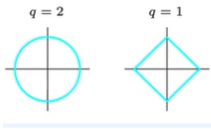
- 2-norm : 벡터와 벡터 사이의 거리를 나타낸다. (피타고라스의 정리)
- 1-norm : 거리를 나타내지는 않는다.

Handwritten chalkboard showing the general formula for the p-norm:

$$(|x_1|^p + |x_2|^p + |x_3|^p + \dots)^{1/p}$$

- L1-distance의 그래프를 보면 좌표 축을 경계로 각이 져있다. 좌표축에 영향을 받는다는 것을 나타낸다. 하지만 L2-distance는 좌표축에 상관 없이 일관된 원형을 보여주고 있다. 이것은 L2-distance보다 L1-distance가 입력 데이터의 feature에 더 큰 영향을 받는다는 것을 보여준다.
- 따라서, 입력데이터의 feature가 중요한 의미를 지닌다면 L1-distance가 더 적합할 것이고, L2는 general한 결과를 얻을 때 더 잘 어울릴 수 있다. (중요하지 않은 feature가 존재하면, 상수값을 계속 빼니까)

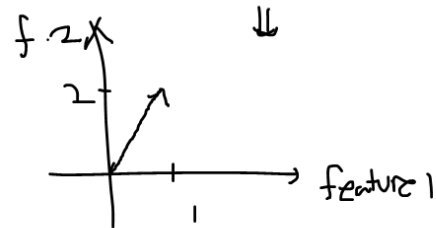
- L1-distance는 L2에 비해 outlier(이상치)의 영향을 덜 받는, Robust한 특성을 지니지만 0에서 미분 불가능



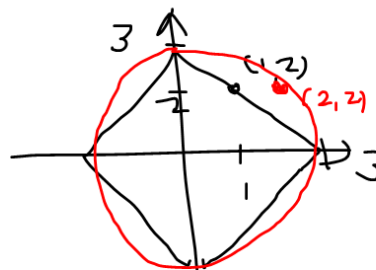
$$2) \quad x_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, x_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \boxed{\vec{v}_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}}$$

$$1) \quad x_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow x_1 - x_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \vec{v}_1$$

\swarrow feature 1 \searrow feature 2



$$\text{If } L1\text{-Norm} = 3 \\ (L2\text{-Norm} = 3)$$



If $\vec{v} = (2,2)$,
 가지는 x_1 과 $L1$ 관점에서는
 false, $L2$ 로는 True
 \Rightarrow outlier에 관대

+ L1은 feature값에 더 민감하다는 걸 알 수있다. \Rightarrow L2보다 좌표 축에 민감하다.