DSE 220: Machine learning

Worksheet 2 — Solutions

1. We have $x = (-1, 1, -1, 1)$ and $x' = (1, 1, 1, 1)$.

   (a) $\|x - x'\|_2 = \sqrt{8}$

   (b) $\|x - x'\|_1 = 4$

   (c) $\|x - x'\|_\infty = 2$

2. We have $x = (1, 2, 3, 4)$ in $\mathbb{R}^4$.

   (a) $\|x\|_1 = 10$

   (b) $\|x\|_2 = \sqrt{30}$

   (c) $\|x\|_\infty = 4$

3. Shapes of metrics.

   (a) ℓ_2: ball

   (b) ℓ_1: diamond

   (c) ℓ_∞: box

4. The points in $\mathbb{R}^2$ with $\|x\|_1 = \|x\|_2 = 1$ are $\{(1, 0), (-1, 0), (0, 1), (0, -1)\}$.

5. Metric or not?

   (a) $\mathcal{X} = \mathbb{R}$ and $d(x, y) = x - y$: Not a metric. Violates positivity and symmetry.

   (b) Hamming distance: Metric.

   (c) Squared Euclidean distance: Not a metric. Violates triangle inequality.

6. Let d_1 and d_2 be any two metrics on a space $\mathcal{X}$, and let d be their sum: $d(x, y) = d_1(x, y) + d_2(x, y)$. Then d is a metric. All four properties can be verified directly.

   (P1) $d(x, y) \geq 0$ because it is the sum of two nonnegative values.

   (P2) Pick any x, y.

   $$
   \begin{aligned}
   d(x, y) = 0 &\iff d_1(x, y) + d_2(x, y) = 0 \\
   &\iff d_1(x, y) = 0 \text{ and } d_2(x, y) = 0 \text{ (since both nonnegative)} \\
   &\iff x = y
   \end{aligned}
   $$

   (P3) $d(x, y) = d_1(x, y) + d_2(x, y) = d_1(y, x) + d_2(y, x) = d(y, x)$.

   (P4) For any x, y, z,

   $$
   \begin{aligned}
   d(x, z) &= d_1(x, z) + d_2(x, z) \\
   &\leq (d_1(x, y) + d_1(y, z)) + (d_2(x, y) + d_2(y, z)) \\
   &= (d_1(x, y) + d_2(x, y)) + (d_1(y, z) + d_2(y, z)) \\
   &= d(x, y) + d(y, z)
   \end{aligned}
   $$