Distance to a Subset

Recall that if X is a set, a *distance function* or *metric* on X is a function

$$d\colon \mathrm{X} \times \mathrm{X} \longrightarrow [0, \infty)$$

satisfying

(*i*) $d(x, y) = 0$ if and only if $x = y$.

(*ii*) $d(x, y) = d(y, x)$ for all x, $y \in \mathrm{X}$ (symmetry).

(*iii*) $d(x, z) \leqslant d(x, y) + d(y, x)$ for all x, y, $z \in \mathrm{X}$ (the triangle inequality).

Given a metric on X, for each $x \in \mathrm{X}$ and $\varepsilon > 0$ we define the ball of radius ε around x by :

$$\mathrm{B}_\varepsilon(x) = \left\{ y \in \mathrm{X} \;\middle|\; d(x, y) < \varepsilon \right\}.$$

The collection

$$\mathcal{B} = \left\{ \mathrm{B}_\varepsilon(x) \;\middle|\; x \in \mathrm{X},\, \varepsilon > 0 \right\}$$

forms a base for a topology on X, called the *metric topology*. The set X along with this topology is called a *metric space*.

Distance to a subset.

Let (X, d) be a metric space, and $\mathrm{A} \subseteq \mathrm{X}$ a subset. A commonly used construction is the "distance function to A", the function $d(\cdot, \mathrm{A})\colon \mathrm{X} \longrightarrow [0, \infty)$ defined by

$$d(x, \mathrm{A}) = \inf \left\{ d(x, z) \;\middle|\; z \in \mathrm{A} \right\} \qquad \text{for each } x \in \mathrm{X}.$$

Proposition (Elementary properties of the distance function) :

(*a*) $d(x, \mathrm{A}) = 0$ if and only if $x \in \overline{\mathrm{A}}$.

(*b*) For each x, $y \in \mathrm{X}$,

$$|d(x, \mathrm{A}) - d(y, \mathrm{A})| \leqslant d(x, y).$$

(*c*) $d(\cdot, \mathrm{A})\colon \mathrm{X} \longrightarrow [0, \infty)$ is a continuous function.

Part (*b*) is a "uniform continuity" statement for the distance function, and immediately implies (*c*). We will first prove (*a*), then record the deduction of (*c*) from (*b*), then stop to discuss a small complication in the argument for (*b*), and then prove (*b*).

Proof. (*a*) Given the definition of the distance function in terms of the infimum, we have

$$\begin{aligned} d(x, \mathrm{A}) = 0 \quad &\Longleftrightarrow \quad \text{for all } \varepsilon > 0 \text{ there exists a } z \in \mathrm{A} \text{ such that } d(x, z) < \varepsilon \\ &\Longleftrightarrow \quad \text{for all } \varepsilon > 0,\ \mathrm{B}_\varepsilon(x) \cap \mathrm{A} \neq \varnothing \\ &\Longleftrightarrow \quad \text{for all open sets V containing } x,\ \mathrm{V} \cap \mathrm{A} \neq \varnothing \\ &\Longleftrightarrow \quad x \in \overline{\mathrm{A}}. \end{aligned}$$

The last implication is by our characterization of points in the closure of A, and the second last because the B_ε form a base for the metric topology.

Part (*c*) follows directly from (*b*). Given an $x_0 \in \mathrm{X}$ and an $\varepsilon > 0$, for any point $y \in \mathrm{B}_\varepsilon(x_0)$ we have by (*b*) that

$$|d(x_0, \mathrm{A}) - d(y, \mathrm{A})| \leqslant d(x_0, y) < \varepsilon. \quad \square$$

Before proving (*b*) let us pause and discuss a minor complication. If for every x there were a point $z \in \mathrm{A}$ so that $d(x, \mathrm{A}) = d(x, z)$ (i.e., a point which "computes the distance to A"), here is how the proof would go : Given x, let $z \in \mathrm{A}$ be a point such that $d(x, z) = d(x, \mathrm{A})$. Then, using the triangle inequality,

$$d(y, z) \leqslant d(x, z) + d(x, y)$$

or

$$d(y, z) \leqslant d(x, \mathrm{A}) + d(x, y).$$

Since $d(y, \mathrm{A})$ is the infimum over all $d(y, z')$ with $z' \in \mathrm{A}$, we know that $d(y, \mathrm{A}) \leqslant d(y, z)$. Combining this with the previous inequality gives

$$d(y, \mathrm{A}) \leqslant d(x, \mathrm{A}) + d(x, y) \quad \text{or} \quad d(y, \mathrm{A}) - d(x, \mathrm{A}) \leqslant d(x, y).$$

Reversing the roles of x and y (and going through the argument again) we would get

$$d(x, \mathrm{A}) - d(y, \mathrm{A}) \leqslant d(x, y),$$

and combining these gives

$$|d(x, \mathrm{A}) - d(y, \mathrm{A})| \leqslant d(x, y).$$

Because the definition of $d(x, \mathrm{A})$ is in terms of the infimum, and not the minimum (i.e., since we don't know that the function $d(x, \cdot)$ achieves a minimum value on A), that argument does not work exactly as stated.

But, for any $\varepsilon > 0$ there is a point $z_0 \in \mathrm{A}$ such that $d(x, z_0) \leqslant d(x, \mathrm{A}) + \varepsilon$ (this follows directly from the definition of the distance in terms of the infimum). Using this we can copy the argument above, and arrive at an inequality not involving z_0 which is "only off by ε". Since the inequality does not involve z_0, it holds for all $\varepsilon > 0$, and so must hold in the limit, which will return us to one of the inequalities above.

Proof of (*b*) : By the definition of the distance function in terms of the infimum, for each $\varepsilon > 0$ there is a $z_0 \in \mathrm{A}$ (depending on ε) such that $d(x, z_0) \leqslant d(x, \mathrm{A}) + \varepsilon$. By the triangle inequality we therefore have

$$d(y, z_0) \leqslant d(x, z_0) + d(x, y) \leqslant d(x, \mathrm{A}) + \varepsilon + d(x, y).$$

Since $d(y, \mathrm{A}) = \inf\{d(y, z) \mid z \in \mathrm{A}\}$, we have $d(y, \mathrm{A}) \leqslant d(y, z_0)$ and so

$$d(y, \mathrm{A}) \leqslant d(x, \mathrm{A}) + d(x, y) + \varepsilon$$

and therefore

$$d(y, \mathrm{A}) - d(x, \mathrm{A}) \leqslant d(x, y) + \varepsilon.$$

This last statement holds for all $\varepsilon > 0$, and so must hold in the limit as $\varepsilon \to 0^+$, giving us

$$d(y, \mathrm{A}) - d(x, \mathrm{A}) \leqslant d(x, y) \quad \text{or} \quad d(y, \mathrm{A}) - d(x, \mathrm{A}) \leqslant d(x, y).$$

Symmetrically, we have $d(x, \mathrm{A}) - d(y, \mathrm{A}) \leqslant d(x, y)$, and together these last two inequalities give

$$|d(x, \mathrm{A}) - d(y, \mathrm{A})| \leqslant d(x, y). \quad \square$$

Distance between two sets

Similarly, given subsets $\mathrm{A}, \mathrm{B} \subseteq \mathrm{X}$, we define

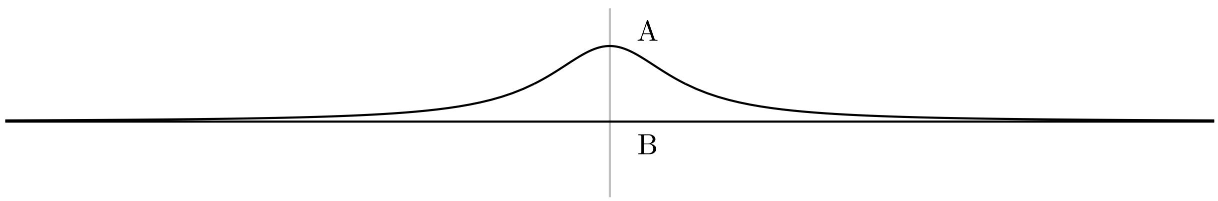
$$d(\mathrm{A}, \mathrm{B}) = \inf \left\{ d(z, w) \;\middle|\; z \in \mathrm{A},\ w \in \mathrm{B} \right\}.$$

This is a single number, and not a function. One can show that

$$\inf \left\{ d(z, \mathrm{B}) \;\middle|\; z \in \mathrm{A} \right\} = d(\mathrm{A}, \mathrm{B}) = \inf \left\{ d(w, \mathrm{A}) \;\middle|\; w \in \mathrm{B} \right\}.$$

One caution : Unlike property (*a*) for distance to a set, it is not true that $d(\mathrm{A}, \mathrm{B}) = 0$ if and only if $\overline{\mathrm{A}} \cap \overline{\mathrm{B}} \neq \varnothing$. One direction still holds : if $\overline{\mathrm{A}} \cap \overline{\mathrm{B}} \neq \varnothing$ then $d(\mathrm{A}, \mathrm{B}) = 0$.

A simple counterexample in the other direction is to take $\mathrm{X} = \mathbf{R}^2$, A to be the graph of $y = \frac{1}{1+x^2}$, i.e., the set $\left\{(x, \frac{1}{1+x^2}) \mid x \in \mathbf{R}\right\}$, and B to be the x-axis. Then both A and B are already closed in $\mathbb{R}^2$, $\mathrm{A} \cap \mathrm{B} = \varnothing$, but $d(\mathrm{A}, \mathrm{B}) = 0$.



If X is compact (in the metric topology, of course!) then it is true that $d(\mathrm{A}, \mathrm{B}) = 0$ if and only if $\overline{\mathrm{A}} \cap \overline{\mathrm{B}} \neq \varnothing$.