SF2704 Clustering and persistence
Homework 2

Metrics and pseudometrics.

Let X be a finite set. A *pseudometric* on X is a function $d\colon X \times X \to \mathbf{R}$ such that:

- $d(a, b) = d(b, a)$ for any a and b in X;
- $d(a, b) \geq 0$;
- $d(a, b) + d(b, c) \geq d(a, c)$ for any a, b and c in X.

A *pseudo-metric* on X is called a *metric* if in addition the following condition is satisfied:

- $d(a, b) = 0$ if and only if $a = b$.

We say that pseudometrics (X, d) and (Y, d') are *equivalent* if there are functions $f\colon X \to Y$ and $g\colon Y \to X$ such that:

- $d(a, b) = d'(f(a), f(b))$ for any a and b in X,
- $d'(a, b) = d(g(a), g(b))$ for any a and b in Y.

Exercise 0

Show that being equivalent is en equivalence relation on the set of all pseudometrics.

Exercise 1

Let (X, d) and (Y, d') be metrics. Show that they are equivalent as pseudometrics if and only if they are equivalent as metrics, i.e., if there is a bijection $f\colon X \to Y$ such that $d(a, b) = d'(f(a), f(b))$ for all a and b in X.

Exercise 2

Let d be a pseudometric on X. We say that two elements a and b in X are related and write $a \sim_d b$ if $d(a, b) = 0$. Show that:

1. $\sim_d$ is an equivalence relation on X. We use the symbol X_d to denote the set of equivalence class of the relation $\sim_d$.

2. The function $d\colon X \times X \to \mathbf{R}$ factors as:

$$X \times X \underbrace{\xrightarrow{\text{quotient}} (X_d) \times (X_d) \xrightarrow{\overline{d}}}_{d} \mathbf{R}$$

3. The function $\overline{d}\colon X_d \times X_d \to \mathbf{R}$ is a metric on X_d.

4. The pseudometrics (X, d) and $(X_d, \overline{d})$ are equivalent.

5. Two pseudometrics (X, d) and (Y, d') are equivalent if and only if $(X_d, \overline{d})$ and $(Y'_d, \overline{d'})$ are equivalent, i.e., if there is a bijection $f\colon X_d \to Y_d$ such that $\overline{d}(a, b) = \overline{d'}(f(a), f(b))$ for any a and b in X/d.

Exercise 3

Show that two pseudometrics (X, d) and (Y, d') are equivalent if and only if there are maps $f\colon X \to Y$ and $g\colon Y \to X$ such that $d(x, g(y)) = d'(f(x), y)$ for any x in X and y in Y.

We will use the symbol $\mathcal{P}M$ to denote the set of equivalence classes of all pseudometrics. Recall that $\mathcal{M}_X$ denotes the set of equivalence classes of metrics on the set X.

Consider the following functions:

$$\Sigma\colon \coprod_{X \text{ a set}} \mathcal{M}_X \to \mathcal{P}M \qquad\qquad \Omega\colon \mathcal{P}M \to \coprod_{X \text{ a set}} \mathcal{M}_X$$

where:

1. Σ maps the equivalence class represented by a metric (X, d) to the class of pseudometrics represented by (X, d).

2. Ω maps the equivalence class represented by a pseudometric (X, d) to the class of metrics represented by $(X_d, \overline{d})$.

Exercise 4

Show that Σ and Ω are well defined and that the compositions $\Sigma\Omega$ and $\Omega\Sigma$ are the identity functions. Conclude that the sets $\coprod_{X \text{ a set}} \mathcal{M}_X$ and $\mathcal{P}M$ are in bijections.

We will use Exercise 3 to define a metric on $\mathcal{P}M$. Define:

$$\mu\colon \mathcal{P}M \times \mathcal{P}M \to \mathbf{R}$$

be the formula:

$$\mu((X,d),(Y,d')) =$$

$$\frac{1}{\sqrt{2}}\min\left\{\sqrt{\sum_{x\in X, y\in Y}(d(x,g(y)) - d'(f(x),y))^2} \mid \text{ over all } f\colon X\to Y, g\colon Y\to X\right\}$$

Exercise 5

Prove that $\mu\colon \mathcal{P}M \times \mathcal{P}M \to \mathbf{R}$ is a metric.

Exercise 6

Calculate a distance between the following metrics: