Variants of Gromov-Wasserstein

The Gromov-Wasserstein distance is highly useful for quantifying differences in cell morphology, and a number of variants of Gromov-Wasserstein distance have been proposed in the literature. This page explains two such variants, "Unbalanced Gromov-Wasserstein" and "Fused Gromov-Wasserstein". Here we focus on the mathematics; for applications see Tutorial 5.

Gromov-Wasserstein

Before we treat the variants, we will formally define Gromov-Wasserstein distance.

The Gromov-Wasserstein distance deals with cell models of the form (X, μ), where X is a metric space (a set of points throughout the body of the cell, together with a measure of distance between them, which we represent as a square pairwise distance matrix) and μ is a probability measure, a weight function which associates to each point in X a non-negative weight such that all weights sum to one. For neurons, one could set this value proportional to the measured radius of the dendrite at the observed point; for cells with image data, we could set this mass proportional to the pixel intensity of a certain image channel at the given point, measuring the concentration of a protein. In what follows, we will use the uniform probability measure for simplicity.

Given (X, μ) a cell with n points and (Y, ν) a cell with m points, an $n \times m$ matrix T with non-negative real entries is a **strict coupling** between X and Y if $\sum_j T_{ij} = X_i$ and $\sum_i T_{ij} = Y_j$ for all i, j.

For a strict coupling T, define

$\mathcal{G}(T) = \sum_{ijk\ell} \left(d^{X(x_i, x_j)} - d^{Y(y_k, y_\ell)} \right)^2 T_{ik} T_{j\ell}$

and

$d_{\mathrm{GW}}((X, \mu), (Y, \nu)) = \min_T \frac{1}{2} \sqrt{\mathcal{G}(T)}$

as T ranges over strict couplings.

We call $\mathcal{G}(T)$ the Gromov-Wasserstein **cost** of the coupling T; the Gromov-Wasserstein **distance** is

$d_{\mathrm{GW}}((X, \mu), (Y, \nu)) = \min_T \frac{1}{2} \sqrt{\mathcal{G}(T)}$

as T ranges over strict couplings between (X, μ) and (Y, ν). (The square root is necessary here in order for d_{GW} to satisfy the triangle inequality. The functions in the `CAJAL.run_gw` module report Gromov-Wasserstein **distances** rather than Gromov-Wasserstein **costs** which is inconsistent with some other libraries, such as the Python Optimal Transport library.)

Unbalanced Gromov-Wasserstein

The big-picture idea behind unbalanced Gromov-Wasserstein is that it is less sensitive to small changes in morphology than ordinary GW. If GW answers the question "How well can we align these two cell morphologies?" then UGW answers the question "How well can we align a large chunk of cell 1 with a large chunk of cell 2, where we want to maximize a weighted sum of the size of the pieces being matched and the goodness-of-fit of the match". In situations where it is safe to discard small pieces of a cell without this substantially changing the morphology, then unbalanced GW might be more robust than GW.

The definition of Gromov-Wasserstein distance involves searching through all possible "couplings" between two cells. The notion of "coupling" employed here is rather rigid and inflexible - cells are regarded as having unit mass, and the couplings are required to satisfy a "conservation of mass" law, that is, all mass in the first cell must be paired with corresponding mass in the second cell. If two neurons are modelled as point clouds with 100 points, then each point will be modelled as having

mass 0.01 units, and a valid coupling must satisfy the property that each point in one cell should have 0.01 units worth of mass associated to it from the other cell.

Suppose we have two neurons, which are absolutely identical except that an additional dendrite is present in one which is not present in the other. This would be biologically interesting, and it is plausible that considering such embeddings of one neuron into another would help us to capture important biological similarities. But Gromov-Wasserstein does not recognize such embeddings as valid cell couplings, because it violates "conservation of mass" - all the mass from the first neuron is paired with a fraction of the mass of the second neuron, and the extra dendrite of the other neuron is not paired with anything. The optimal GW transport plan would likely bear no trace of the structural equivalence between the first neuron and a fragment of the second.

The Unbalanced Gromov-Wasserstein distance allows for such embeddings - transport plans which are permitted to create or destroy mass, at the expense of paying a sharp additional penalty cost. The Unbalanced Gromov-Wasserstein paper by Séjourné, Vialard, and Peyré provides some useful examples of situations where the extra flexibility of unbalanced Gromov-Wasserstein makes it more tolerant of small differences between objects.

The mathematical framework behind UGW

The unbalanced Gromov-Wasserstein distance deals with cell models of the form (X, μ), where X is a metric space (a set of points throughout the body of the cell, together with a measure of distance between them, which we represent as a square pairwise distance matrix) and μ is a "measure", a weight function which associates to each point in X a non-negative weight. For neurons one could set this value to, for example, the measured radius of the dendrite at the observed point; we will use the uniform probability measure for simplicity.

Formally, given (X, μ) a cell with n points and (Y, ν) a cell with m points, we define the "unbalanced GW cost" of a transport plan $T \in (\mathbb{R}^+)^{n \times m}$ as $\mathcal{L}(T) = \mathcal{G}(T) + \rho_1 \circ [KL]^{\otimes 2}(x \texttt{ mid } y) + \rho_2 \circ [KL]^{\otimes 2}(x \texttt{ mid } y)$

and the unbalanced Gromov-Wasserstein distance

$\mathrm{U\ G\ W}((X, \mu), (Y, \nu)) = \min_{T \in (\mathbb{R}^{\leq 0})^{n \times m}} \mathcal{L}(T)$