Earth Mover's Distance and The Applications

Hu Ding

Computer Science and Engineering, Michigan State University

The Motivations

- It is easy to compare two single objects: the pairwise "distance". For example:
  1. two points in Euclidean space;
  2. two colors in color space *"Color Science: Concepts and Methods, Quantitative Data and Formulae" by Wyszecki and Stiles*.

The Motivations

- It is easy to compare two single objects: the pairwise "distance". For example:
  1. two points in Euclidean space;
  2. two colors in color space *"Color Science: Concepts and Methods, Quantitative Data and Formulae" by Wyszecki and Stiles*.
- How about two distributions of objects?
  - Hausdorff distance (many-to-one matching);
  - one-to-one matching;
  - Kullback-Leibler (KL) divergence (from information theory), but the correspondence should be given in advance;
  - and many others.

The Motivations

- It is easy to compare two single objects: the pairwise "distance". For example:
  1. two points in Euclidean space;
  2. two colors in color space *"Color Science: Concepts and Methods, Quantitative Data and Formulae" by Wyszecki and Stiles*.
- How about two distributions of objects?
  - Hausdorff distance (many-to-one matching);
  - one-to-one matching;
  - Kullback-Leibler (KL) divergence (from information theory), but the correspondence should be given in advance;
  - and many others.
- Here we introduce a very general & robust measurement, **Earth Mover's Distance (EMD)** *"The Earth Mover's Distance as a Metric for Image Retrieval" by Rubner et al.*

Earth Mover's Distance

- Given two weighted point-sets $A = \{a_1, \cdots, a_n\}$ and $B = \{b_1, \cdots, b_m\}$ in some metric space (e.g., $\mathbb{R}^d$) with the weights $\{\alpha_i \mid 1 \leq i \leq n\}$ and $\{\beta_j \mid 1 \leq j \leq m\}$, which are called the "signatures":
  1. Color image: the histogram bins in color space.
  2. Grey image: the histogram bins in one dimensional space.
  3. Or represented by a set of weighted feature points in $\mathbb{R}^2$.

Earth Mover's Distance

- Given two weighted point-sets $A = \{a_1, \cdots, a_n\}$ and $B = \{b_1, \cdots, b_m\}$ in some metric space (e.g., $\mathbb{R}^d$) with the weights $\{\alpha_i \mid 1 \leq i \leq n\}$ and $\{\beta_j \mid 1 \leq j \leq m\}$, which are called the "signatures":
  1. Color image: the histogram bins in color space.
  2. Grey image: the histogram bins in one dimensional space.
  3. Or represented by a set of weighted feature points in $\mathbb{R}^2$.
- For any pair of a_i and b_j, define $||a_i - b_j||$ as the "ground distance" (i.e., the distance in the metric space).

Earth Mover's Distance (cont.)

- $EMD(A, B) = \min_F \sum_{i=1}^{n} \sum_{j=1}^{m} ||a_i - b_j|| f_{ij}$ where $F = \{f_{ij}\}$ is a feasible flow from A to B, s.t.,
  - $\forall f_{ij} \geq 0$;
  - $\sum_{i=1}^{n} f_{ij} \leq \beta_j$;
  - $\sum_{j=1}^{m} f_{ij} \leq \alpha_i$;
  - $\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} = \min\{\sum_{i=1}^{n} \alpha_i, \sum_{j=1}^{m} \beta_j\}$.

Earth Mover's Distance (cont.)

- $EMD(A, B) = \min_F \sum_{i=1}^{n} \sum_{j=1}^{m} ||a_i - b_j|| f_{ij}$ where $F = \{f_{ij}\}$ is a feasible flow from A to B, s.t.,
  - $\forall f_{ij} \geq 0$;
  - $\sum_{i=1}^{n} f_{ij} \leq \beta_j$;
  - $\sum_{j=1}^{m} f_{ij} \leq \alpha_i$;
  - $\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} = \min\{\sum_{i=1}^{n} \alpha_i, \sum_{j=1}^{m} \beta_j\}$.
- Intuitively, EMD measures the least amount of work for moving a mass of earth (A) to a collection of holes (B).

Earth Mover's Distance (cont.)

- Comparing with other measurements:
  - Many-to-one matching: all α_is are $+\infty$ and all β_js are 1.
  - One-to-one matching: all α_is and β_js are 1.
  - KL-divergence: need to know the correspondence in advance.

Earth Mover's Distance (cont.)

- Comparing with other measurements:
  - Many-to-one matching: all α_is are $+\infty$ and all β_js are 1.
  - One-to-one matching: all α_is and β_js are 1.
  - KL-divergence: need to know the correspondence in advance.
- EMD is more general and robust, but suffers from high computational complexity (we will see later).

Earth Mover's Distance (cont.)

- Comparing with other measurements:
  - Many-to-one matching: all α_is are $+\infty$ and all β_js are 1.
  - One-to-one matching: all α_is and β_js are 1.
  - KL-divergence: need to know the correspondence in advance.
- EMD is more general and robust, but suffers from high computational complexity (we will see later).
- In addition, the continuous version of EMD is the famous "Monge-Kantorovich problem" in transportation theory. **Leonid Kantorovich** won the Nobel prize in economics due to his contribution in this area.

The Algorithms for EMD

- $EMD(A,B) = \min_F \sum_{i=1}^{n} \sum_{j=1}^{m} ||a_i - b_j|| f_{ij}$ where $F = \{f_{ij}\}$ is a feasible flow from A to B, s.t.,
  - $\forall f_{ij} \geq 0$;
  - $\sum_{i=1}^{n} f_{ij} \leq \beta_j$;
  - $\sum_{j=1}^{m} f_{ij} \leq \alpha_i$;
  - $\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} = \min\{\sum_{i=1}^{n} \alpha_i, \sum_{j=1}^{m} \beta_j\}$.
- It is easy to observe that EMD is special min cost max flow problem. So directly, we have

The Algorithms for EMD

- $EMD(A, B) = \min_F \sum_{i=1}^{n} \sum_{j=1}^{m} ||a_i - b_j|| f_{ij}$ where $F = \{f_{ij}\}$ is a feasible flow from A to B, s.t.,
  - $\forall f_{ij} \geq 0$;
  - $\sum_{i=1}^{n} f_{ij} \leq \beta_j$;
  - $\sum_{j=1}^{m} f_{ij} \leq \alpha_i$;
  - $\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} = \min\{\sum_{i=1}^{n} \alpha_i, \sum_{j=1}^{m} \beta_j\}$.
- It is easy to observe that EMD is special min cost max flow problem. So directly, we have
  - linear programming by network simplex algorithm;
  - or several other combinatorial algorithms (see Prof. David Karger's online course `https://www.youtube.com/channel/UCtv9PiQVUDzsT4yl7524DCg`).

The Algorithms for EMD

- $EMD(A,B) = \min_F \sum_{i=1}^{n} \sum_{j=1}^{m} ||a_i - b_j|| f_{ij}$ where $F = \{f_{ij}\}$ is a feasible flow from A to B, s.t.,
  - $\forall f_{ij} \geq 0$;
  - $\sum_{i=1}^{n} f_{ij} \leq \beta_j$;
  - $\sum_{j=1}^{m} f_{ij} \leq \alpha_i$;
  - $\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij} = \min\{\sum_{i=1}^{n} \alpha_i, \sum_{j=1}^{m} \beta_j\}$.
- It is easy to observe that EMD is special min cost max flow problem. So directly, we have
  - linear programming by network simplex algorithm;
  - or several other combinatorial algorithms (see Prof. David Karger's online course `https://www.youtube.com/channel/UCtv9PiQVUDzsT4yl7524DCg`).
- But the drawback: high running time (usually at least cubic $O(VE)$), not utilizing the geometric properties.

The Geometric Algorithms for EMD

- A heuristic algorithm called *thresholded EMD* by pruning the long distances $||a_i - b_j||$ so as to reduce the size of E. *"Fast and Robust Earth Mover's Distances" by Pele and Werman*.

The Geometric Algorithms for EMD

- A heuristic algorithm called *thresholded EMD* by pruning the long distances $||a_i - b_j||$ so as to reduce the size of E. *"Fast and Robust Earth Mover's Distances" by Pele and Werman*.
- Several approximate EMD algorithms with theoretical quality guarantee, such as *" A near-linear time ϵ-approximation algorithm for geometric bipartite matching" by Sharathkumar and Agarwal*.
  - Random shift + Quadtree to compress the pairwise distances.
  - The running time is nearly linear on $\max\{\sum_{i=1}^{n} \alpha_i, \sum_{j=1}^{m} \beta_j\}$.

The Applications

- Pattern matching. *"The Earth Mover's Distance under Transformation Sets" by Cohen and Guibas; "Supervised Earth Mover's Distance Learning and Its Computer Vision Applications" by Wang and Guibas.*

The Applications

- Pattern matching. *"The Earth Mover's Distance under Transformation Sets" by Cohen and Guibas; "Supervised Earth Mover's Distance Learning and Its Computer Vision Applications" by Wang and Guibas.*
- Image retrieval (embedding + LSH). *"Fast Image Retrieval via Embeddings" by Indyk and Thaper; "Fast Contour Matching Using Approximate Earth Mover's Distance" by Grauman and Darrell.*

The Applications (cont.)

- Deep learning. *"Wasserstein GAN" by Arjovsky et al.*, *"A Geometric View of Optimal Transportation and Generative Model" by Lei et al.*

The Applications (cont.)

- Deep learning. *"Wasserstein GAN" by Arjovsky et al.*, *"A Geometric View of Optimal Transportation and Generative Model" by Lei et al.*
- Word Mover's Distance for analyzing documents. *"From Word Embeddings To Document Distances" by Kusner et al*.

Thank You!

Any Question?