Graph matching is a generic and popular modeling tool for problems in computational sciences such as computer vision, computer graphics, medical imaging and machine learning.

In general, graph matching refers to several different optimization problems of the form:

$$\min_{X \in \mathcal{F}} E(X) \quad \text{s.t.} \quad X \in \mathcal{F}$$

where $\mathcal{F} \subseteq \mathbb{R}^{m \times n}$ is a collection of matchings between vertices of two graphs $G_a$ and $G_b$. Usually, $E(X) = X^T M X + a^T X$ is a quadratic function in $X$ and $\mathcal{F}$ is the set of permutation matrices. Often, $M$ quantifies the discrepancy between edge affinities exerted by the matching $X$. Edge affinities are represented by symmetric matrices $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{n \times n}$.

Multiple works have considered the relaxation:

$$E(X) = -tr(BX^TAX) \quad \text{s.t.} \quad X \in \mathcal{DS}$$

Where $\mathcal{DS}$ is the set of doubly-stochastic matrices, the convex-hull of the permutation matrix. This energy measures the inner product of the relabeled affinity matrix $AX^T$ with the matrix $B$. For example if the two graphs are isomorphic the minimum of this functional is obtained by the permutation $X$ that represents this isomorphism.

$E$ is in general indefinite, resulting in a non-convex relaxation. In case $E$ is concave the relaxation enjoys three key properties:

- Its solution is the same as the original problem.
- Its set of local optima are all permutations.
- For every descent direction the maximal step is optimal.

[Vogelstein et al., 2015, Lyzinski et al., 2016] suggest to locally optimize this relaxa\tion with the Frank-Wolfe algorithm and motivate it by proving that for a specific class of adjacency matrices, the optimal solution of the relaxation almost coincides with the GS optimal solution. [Vestner et al., 2017, Boyarski et al., 2017] make the useful observation that $E$ is concave for some important cases of affinities such as the heat kernels and Gausians.

**Conditionally concavely energies**

Conditionally concave energy $E(X)$ means that the restriction of the Hessian $M$ of the energy $E$ to the linear space

$$\text{lin}(DS) = \{X \in \mathbb{R}^{m \times n} \mid X 1 = 0, X^T 1 = 0\}$$

is negative definite. Where $\text{lin}(DS)$ is the linear part of the affine-hull of the doubly-stochastic matrices.

Our first result shows that there is a large class of affinity matrices resulting in conditionally concave $E$.

**Proposition 1.** Let $\Phi : \mathbb{R}^T \rightarrow \mathbb{R}$, $\Phi : \mathbb{R}^T \rightarrow \mathbb{R}$ be both conditionally positive (or negative) definite functions of order 1. For any pair of graphs with affinity matrices $A, B \in \mathbb{R}^{m \times m}$ so that

$$A_{ij} = \Phi(x_i - x_j), \quad B_{ij} = \Phi(y_i - y_j)$$

for some arbitrary $\{x_i \} \subseteq \mathbb{R}^T$, $\{y_i \} \subseteq \mathbb{R}^T$, the energy $E(X)$ is conditionally concave.

**Proof Idea:** we show that the projection onto $\text{lin}(DS)$ can be obtained as the Kno\cker product of a projection onto $\{|1|^T\}$.

**Application.** Matching graphs with Euclidean affinities:

$$A_{ij} = |x_i - x_j|, \quad B_{ij} = |y_i - y_j|$$

includes distances that can be isometrically embedded in Euclidean spaces such as distance differences, distances induced by deep learning embeddings, Mahalanobis distances and the spherical distances.

**Probabolically conditionally concavely energies**

An energy $E$ is called probably conditionally concave if it is rare to find a linear subspace $D$ of $\text{lin}(DS)$ so that the restriction of $E$ to it is convex, that is $M|D \succeq 0$.

The following theorem bounds the probability of finding uniformly at random a linear subspace $D$ such that the restriction of $M$ to $D \succeq 0$ is convex.

**Proposition 2.** Let $M \in \mathbb{R}^{m \times n}$ be a symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then, for all $t \in (0, \frac{\alpha}{\sqrt{n}})$:

$$P(|M| \cdot t \succeq 0 \mid \lambda_{\min} \leq \frac{\alpha}{\sqrt{n}}) \leq \frac{n}{n + 1} \left(1 - 2\lambda_{\min} t\right)^n$$

**Example.** Let $U$ be a unitary matrix and consider the matrix:

$$U \cdot \text{diag}(1 - 2\lambda_{\min} t, \ldots, 1 - 2\lambda_{\min} t) \cdot U^T$$

It is extremely unlikely to sample a convex direction in dimension $m = 300^2$, i.e., the probability will be $n \cdot 10^{-5}$.

**Application.** For example, matching problems in which the affinities $A, B$ describe geodesic distances on surfaces result in probably conditionally concave $E$.

We further use equation (6) to prove the following:

**Proposition 3.** There exists a rather general probability space of Hessians $M$ so that the probability of local minima of $E(X)$ to be outside $F$ is very small.

This probability space is of the form

$$\Omega = [A \lambda_{\max} U^T \mid U \in \Omega(m)], \quad \lambda_{\max} = \max \left\{ \lambda_1, \ldots, \lambda_{\max} \right\}$$

where $\lambda_1 \leq -b, b > 0$ are the negative eigenvalues; $0 \leq \mu_0 \leq a, a > 0$ are the positive eigenvalues, the ratio of positive to negative eigenvalues is a constant $p \in (0, 1/2)$ and $a \leq b$.

In case we want to optimize over the one-sided permutations, namely $\mathcal{F} = \{X \in [0, 1]^{n \times m} \mid 11 = 1\}$, some of the previous earned might not yield concave relaxations. In this case we devise a variation of the Frank-Wolfe algorithm using a concave search procedure. That is, in each iteration, instead of standard line search we subtract a convex energy from $E(X)$ that is constant on $\mathcal{F}$ until we find a descent step.

**Experiments**

In this paper we introduce the concepts of conditionally concave and probably conditionally concave energies $E(X)$ and show that these energies encapsulate many real world instances of the graph matching problem, including matching Euclidean graphs and graphs on surfaces.

**References**


