Pareto Backbone Results

Jordan Rozum

October 25, 2022

1 Terminology

Consider weighted digraphs $(G_0, G_1, \ldots, G_m)$ corresponding to layers of a multigraph $\mathcal{G}$ with inter-layer edge set $\mathcal{E} = \mathcal{E}(\mathcal{G})$. A node belonging to layer l is denoted by a superscript, e.g., u^l is node u of G_l. A path is a sequence of edges in the multigraph $P_i = (u_i^{l_i}, v_i^{l_i})_i$, where either $v_i^{l_i} = u_{i+1}^{l_{i+1}}$ or $(v_i^{l_i}, u_{i+1}^{l_{i+1}}) \in \mathcal{E}$. The weight of an edge is determined by the weight function w that sums the (strictly positive) weights of the edges along an input path.

The **multidistance** of a path is the tuple

$$M(P) = \left(\sum_{i:l_i=0} w(P_i), \sum_{i:l_i=1} w(P_i), \ldots, \sum_{i:l_i=m} w(P_i) \right).$$

We abbreviate this $M(P) = (M_0, M_1, \ldots, M_m)$. We define the partial order $\prec$ on multidistances. We say that $M(P) \prec M(P')$ if $M_l \leq M'_l$ for all l and if there is at least one value of l for which the inequality is strict; in this case, we say that P' *dominates* P. If neither $M(P) \prec M(P')$ nor $M(P') \prec M(P)$, then the two multidistances are incomparable, and we write $M(P) \sim M(P')$. We extend these comparisons to P for notational brevity, e.g., $P \prec P'$ is equivalent to $M(P) \prec M(P')$.

A path P is **pareto minimal** (or shortest) if does not dominate any paths with the same endpoints. A pareto minimal path's multidistance is *pareto efficient*. Note that there can be multiple pareto shortest paths with distinct multidistances; these multidistances will be incomparable, and the set of such multidistances is called the **pareto distance set** between the endpoints.

A multidistance can be converted into a traditional distance by computing a weighted sum of its its entries. The weights correspond to a weighting of the layers. For weight vector W, the distance of a path is

$$dist_W(P) = \sum_{l=0}^{m} M_l W_l,$$

where $M = M(P)$. We assume that $W_l > 0$ for all l.

Note that if a path P between two nodes is a shortest path (i.e., has minimal distance) for some weight W, then $M(P)$ is an element of the pareto distance set. Otherwise there is a path P' between the same endpoints with $M'_l < M_l$ for some l and no j such that $M_j < M'_j$, and thus $dist_W(P') < dist_W(P)$ (recall that $W_l > 0$ for all l by assumption).

However, the reverse is not true. For example, consider the four pareto minimal path multidistances:

$$
\begin{aligned}
M^0 &= (2, 1, 1) \\
M^1 &= (1, 2, 1) \\
M^2 &= (1, 1, 2) \\
M^3 &= (3, 0, 0)
\end{aligned}
$$

It is straightforward to show that while all four multidistances are incomparable, the first, M^0, cannot yield a minimal distance. To see this, consider an arbitrary weight vector $W = (a, b, c)$ with $a, b, c > 0$. For M^0 to be minimal, it must satisfy

$$
\begin{aligned}
2a + b + c &\leq a + 2b + c \\
2a + b + c &\leq a + b + 2c \\
2a + b + c &\leq 3a + 0b + 0c,
\end{aligned}
$$

but the first two inequalities imply $a \leq b, c$, while the last implies $b + c \leq a$. These are impossible to satisfy simultaneously. It remains true, however, that M^0 can be shorter than any *individual* path under a carefully chosen weighting. No such weighting can make the corresponding path shorter than all others.

The multidistance M^3 in this example, however, illustrates an important special case: its path is confined to a single layer. Because of this, the pareto shortest path criterion reduces to the requirement that no other multidistance has a smaller value in the entry corresponding to that layer. Thus, a weighting can always be chosen to make such paths have minimal distance. To do so, simply assign a very small weight to that layer, and a very large weight to all others. We formalize this as a lemma:

Lemma 1.1. *Let a path P have a multidistance $M(P)$ that is zero in all but one entry. Then if P is an element of the pareto distance set between two nodes, there is a layer weighting W for which $dist_W(P)$ is the shortest distance between the endpoints of P.*

Proof. Without loss of generality, we take $M(P) = (M_0, 0, \ldots, 0)$. Consider each path P' between the endpoints of P. Because P is a pareto shortest path, $P' \not< P$. This implies that either $M(P) = M(P')$ or $\sum_{l>0} M'_l > 0$. In the first case, no weighting can yield $dist_W(P') < dist_W(P)$. Otherwise we consider the path P' with $\sum_{l>0} M'_l > 0$ minimal, and denote this quantity σ. The weighting $W = (1/M_0, 2/\sigma, \ldots, 2/\sigma)$ then gives $dist_W(P) = 1$ and $dist_W(P') \geq 2$ for $M(P) \neq M(P')$, proving the lemma. $\square$

The **pareto backbone** is obtained by eliminating each edge that whose multidistance is not among the the pareto distance set between its parent and child nodes. We denote this $\mathcal{B}(\mathcal{G})$. This generalizes the concept of a graph backbone of a graph G, which eliminates all edges that are not part of a shortest path and is denoted $B(G)$. In fact, the pareto backbone of a multigraph with layers $(G_0, G_1, \ldots, G_m)$ is the multigraph whose layers are the graph backbones of each layer, as we now formally state and prove.

Theorem 1.2. *The pareto backbone $\mathcal{B}(\mathcal{G})$ of a multigraph $\mathcal{G}$ with layers $(G_0, G_1, \ldots, G_m)$ is the multigraph $\mathcal{G}'$ with layers $(B(G_0), B(G_1), \ldots, B(G_m))$.*

Proof. The proof proceeds in two parts.

($\mathcal{G}' \subseteq \mathcal{B}(\mathcal{G})$) Consider an edge $(u,v) \in B(G_l)$. By definition, the edge (u,v) is a shortest path between u and v in G_l and is of length $w((u,v))$. The multidistance in $\mathcal{G}$ along the path P consisting of only this edge is zero in all entries except entry l, in which it is $w((u,v))$. If a path P' connecting u and v in $\mathcal{G}$ satisfies $P' \prec P$, then its l^{th} entry must be less than $w((u,v))$ and all other entries must be zero. This implies that P' is a path confined to G_l that has a lower weight than $w((u,v))$, contradicting the assumption $(u,v) \in B(G_l)$.

($\mathcal{B}(\mathcal{G}) \subseteq \mathcal{G}'$) Consider (u,v) in $\mathcal{B}(\mathcal{G})$. By definition, $M((u,v))$ is a pareto shortest path between u and v. Because this path is of length one, it is confined to a single layer, and thus is a shortest path within that layer. Because the edge is a shortest path in its layer, it is part of that layer's backbone. ☐

A corrolary of this result relates the pareto backbone and pareto distance set between two nodes.

Corollary 1.2.1. *The pareto distance set between nodes u and v in $\mathcal{G}$ is equal to the pareto distance set between u and v in $\mathcal{B}(\mathcal{G})$.*

Proof. Denote the pareto distance set computed in $\mathcal{G}$ by D and the pareto distance set computed in $\mathcal{B}(\mathcal{G})$ by D'.

($D' \subseteq D$) Because $\mathcal{B}(\mathcal{G})$ is a subgraph of $\mathcal{G}$, it follows that every path in $\mathcal{B}(\mathcal{G})$ is also in $\mathcal{G}$, and therefore D' is a subset of D.

($D \subseteq D'$) Consider a path P connecting u to v in $\mathcal{G}$ whose multidistance is in D. Because $D' \subseteq D$, $M(P)$ is incomparable to all multidistances in D'. Consider an arbitrary edge of P, (P_i, P_{i+1}) in a layer G_l. Because P is pareto minimal, there is no path in G_l from P_i to P_{i+1} that is shorter than $w((P_i, P_{i+1}))$, and therefore (P_i, P_{i+1}) is in the backbone of G_l. By the previous theorem, this implies that (P_i, P_{i+1}) is an edge of $\mathcal{B}(\mathcal{G})$. Applying this argument ot all edges of P reveals that P is a path in $\mathcal{B}(\mathcal{G})$. Because P is a path in $\mathcal{B}(\mathcal{G})$ and $M(P)$ is incoparable to all multidistances in D', $M(P)$ is an element of D'. As this argument applies for arbitrary paths with multidistance in D', it follows that D is a subset of D'. ☐

This theorem and its corollary have significant computational implications. They imply that a pareto backbone and its associated pareto distance sets can be computed from the backbones of individual layers.