A Game-theoretic Approach for Enumerating Dominant Sets

Samuel Rota Bulò, Andrea Torsello and Marcello Pelillo

Dipartimento di Informatica.
Università “Ca’ Foscari” di Venezia.
{srotabul,torsello,pelillo}@dsi.unive.it

Abstract. Recently, a novel graph-theoretic notion of a cluster has been proposed, i.e., the “dominant set”, which captures the two basic requirements of a cluster, namely internal coherency and external incoherency. In this paper, we tackle the problem of finding several dominant sets using the replicator dynamics. Specifically, we adopt a game-theoretic perspective to this iterative extraction: Game dynamics are used to locate individual dominant sets, and after each extraction the similarity matrix is transformed in such a way as to make the located cluster unstable under the dynamics, without affecting the remaining groups. This guarantees that once found, a cluster will not be extracted again.

1 Introduction

Recently, a novel graph-theoretic notion of a cluster has been proposed, i.e., the “dominant set”, which captures the two basic requirements of a cluster, namely internal coherency and external incoherency [1, 2]. In this framework clusters are in a one-to-one relation with the maxima of a quadratic function, where the non-zero entries of the maxima correspond to elements within a cohesive group. The approach is capable of automatically detect the number of groups present as this is determined by the number of maxima in the objective function. Furthermore, the approach works as a multi figure/ground discrimination algorithm since it only extracts cohesive groups, leaving spurious entries unclustered.

In this paper, following [3], we adopt a game-theoretic perspective to the iterative extraction of the dominant sets: Game dynamics are used to locate individual dominant sets, and after each extraction the similarity matrix is transformed in such a way as to make the located cluster unstable under the dynamics, without affecting the remaining groups. This guarantees that once found, a cluster will not be extracted again.

2 Dominant Sets Framework

In the pairwise clustering framework the data to be clustered are represented (possibly implicitly) as an undirected edge-weighted graph with no self-loops $G = (V, E, w)$, where $V = \{1, \ldots, n\}$ is the vertex set, $E \subseteq V \times V$ is the edge
set, and $w: E \to \mathbb{R}_+^*$ is the (positive) weight function. Vertices in $G$ correspond to data points, edges represent neighborhood relationships, and edge-weights reflect similarity/dissimilarity between pairs of linked vertices. As customary, we represent the graph $G$ with the corresponding weighted adjacency (or similarity/dissimilarity) matrix, which is the $n \times n$ nonnegative, symmetric matrix $A = (a_{ij})$ defined as:

$$a_{ij} = \begin{cases} w(i, j) , & \text{if } (i, j) \in E \\ 0 , & \text{otherwise} \end{cases}$$

The dominant set framework has been presented in [1]. Let $S \subseteq V$ be a non-empty subset of vertices and $i \in V$. The (average) weighted degree of $i$ w.r.t. $S$ is defined as:

$$awdeg_S(i) = \frac{1}{|S|} \sum_{j \in S} a_{ij}$$

(1)

where $|S|$ denotes the cardinality of $S$. Moreover, if $j \notin S$ we define $\phi_S(i, j) = a_{ij} - awdeg_S(i)$ which is a measure of the similarity between nodes $j$ and $i$, with respect to the average similarity between node $i$ and its neighbors in $S$.

Let $S \subseteq V$ be a non-empty subset of vertices and $i \in S$. The weight of $i$ w.r.t. $S$ is

$$w_S(i) = \begin{cases} 1 , & \text{if } |S| = 1 \\ \sum_{j \in S \setminus \{i\}} \phi_{S \setminus \{i\}}(j, i) w_{S \setminus \{i\}}(j) , & \text{otherwise} \end{cases}$$

(2)

while the total weight of $S$ is defined as:

$$W(S) = \sum_{i \in S} w_S(i) .$$

(3)

Intuitively, $w_S(i)$ gives us a measure of the overall similarity between vertex $i$ and the vertices of $S \setminus \{i\}$ with respect to the overall similarity among the vertices in $S \setminus \{i\}$, with positive values indicating high internal coherency.

A non-empty subset of vertices $S \subseteq V$ such that $W(T) > 0$ for any non-empty $T \subseteq S$, is said to be dominant if:

1. $w_S(i) > 0$, for all $i \in S$
2. $w_{S \cup \{i\}}(i) < 0$, for all $i \notin S$.

The two previous conditions correspond to the two main properties of a cluster: the first regards internal homogeneity, whereas the second regards external in-homogeneity. The above definition represents our formalization of the concept of a cluster in an edge-weighted graph.

Now, consider the following quadratic program, which is a generalization of the so-called Motzkin-Straus program [4]:

$$\begin{align*}
& \text{maximize } f(x) = x' Ax \\
& \text{subject to } x \in \Delta
\end{align*}$$

(4)
where
\[ \Delta = \{ \mathbf{x} \in \mathbb{R}^n : e' \mathbf{x} = 1, \mathbf{x} \geq 0 \} \]

is the standard simplex of \( \mathbb{R}^n \), and \( e \) is a vector of appropriate length consisting of unit entries. The support of a vector \( \mathbf{x} \in \Delta \) is defined as the set of indices corresponding to its positive components, that is \( \sigma(\mathbf{x}) = \{ i \in V : x_i > 0 \} \).

In [1], an intriguing connection between dominant sets and local solutions of Program (4) is established. Specifically, it is proven that if \( S \) is a dominant subset of vertices, then its (weighted) characteristic vector \( \mathbf{x}^S \), which is the vector of \( \Delta \) defined as
\[ x^S_i = \begin{cases} \frac{w_S(i)}{W(S)}, & \text{if } i \in S \\ 0, & \text{otherwise} \end{cases} \]

is a strict local solution of Program (4). Conversely, if \( \mathbf{x} \) is a strict local solution of Program (4) then its support \( S = \sigma(\mathbf{x}) \) is a dominant set, provided that \( w_{S \cup \{i\}}(i) \neq 0 \) for all \( i \notin S \).

3 A Game-theoretic Perspective

Let \( O = \{1, 2, \ldots, n\} \) be the set of pure strategies available to the players and \( A = (a_{ij}) \) the \( n \times n \) payoff or utility matrix [5] where \( a_{ij} \) is the payoff that a player gains when playing the strategy \( i \) against an opponent playing \( j \). In biological contexts, payoff are typically measured in terms of Darwinian fitness or reproductive success whereas in economics applications, they usually represent firms’ profits or consumers’ utilities.

A mixed strategy is a probability distribution \( \mathbf{x} = (x_1, x_2, \ldots, x_n)' \) over the available strategies in \( O \). Mixed strategies clearly lie in \( \Delta \). The expected payoff that a player obtains by playing the element \( i \) against an opponent playing a mixed strategy \( \mathbf{x} \) is \( u(e^i, \mathbf{x}) = (Ax)_i = \sum_j a_{ij}x_j \), where \( e^i \) is the vector with all components equal zero except for the \( i^{th} \)-component which is equal to 1. Hence, the expected payoff received by adopting a mixed strategy \( \mathbf{y} \) is \( u(\mathbf{y}, \mathbf{x}) = \mathbf{y}'Ax \).

Evolutionary game theory considers an idealized scenario wherein pairs of individuals are repeatedly drawn from a large population to play a two-player symmetric game. Each player is not supposed to behave rationally or have a complete knowledge of the details of the game, but he acts according to a pre-programmed pure strategy. This dynamic activates some selection process that results in the evolution of the fittest strategies.

A well-known formalization of the selection process is given by the replicator equations [5]:
\[ \dot{x}_i = x_i [u(e^i, \mathbf{x}) - u(\mathbf{x}, \mathbf{x})]. \]

If the payoff matrix is symmetric then \( \mathbf{x}'Ax \) is strictly increasing along any non-constant trajectory of any payoff-monotonic dynamics [5]. This result allows us to establish a one-to-one correspondence between the strict local solutions of program (4), namely characteristic vectors of dominant sets of a edge-weighted graph \( G \), and asymptotically stable points of the replicator dynamics with payoff matrix the weighted adjacency matrix of \( G \).
In order to obtain enumeration of dominant sets through a continuous formulation we move from undirected graphs to directed graphs, or, in other words, from symmetric payoff matrices to asymmetric payoff matrices. If we loosen the symmetry constraint, then all the results that bind local solutions to asymptotically stable points and dominant sets do not hold any longer, and $x'Ax$ is not a Lyapunov function for the dynamics.

The best replies against a mixed strategy $x$ is the set of mixed strategies $\beta(x) = \{ y \in \Delta : u(y, x) = \max_z u(z, x) \}$.

A mixed strategy $x$ is a Nash equilibrium if it is a best reply to itself, i.e. $\forall y \in \Delta, u(y, x) \leq u(x, x)$. This implies that for all $i \in \sigma(x)$, $u(e^i, x) = u(x, x)$, hence the payoff of every strategy in the support of $x$ is constant, while all strategies outside the support of $x$ earn a payoff that is less than or equal to $u(x, x)$. A strategy $x$ is said to be an evolutionary stable strategy (ESS) if it is a Nash equilibrium and for all $y \in \Delta$ such that $u(y, x) = u(x, x)$ we have that $u(x, y) > u(y, y)$. Intuitively, ESS are strategies such that any small deviation from them will lead to an inferior payoff.

Consider again Program (4), where $A$ is a symmetric matrix. It is well known that the set of ESS’s of a two-player game with payoff matrix $A$ is equal to the set of strict local maximizers of (4). But this also implies that ESS’s are in one-to-one correspondence with dominant sets of the edge-weighted graph having $A$ as weighted adjacency matrix. In [3], we show that this correspondence between ESS’s and dominant sets holds also for the general case of asymmetric affinities.

Through this change in perspective, we can move from a constrained maximization problem, to a game-theoretic setting. Instead of finding local solutions of a quadratic program, we look for ESS of a symmetric two-persons game. An advantage of this new approach is that we can generalize the Motzkin-Straus result also to non symmetric payoff matrices and, hence, edge-weighted directed graphs.

4 Continuous-based Enumeration

We have seen that the replicator dynamics is a simple and elegant tool for finding a dominant set in a edge-weighted graph. If however we reiterate it in order to extract other dominant sets, we have no guarantee that a new one will be found. The work we present overcomes this problem by rendering unstable the stationary points corresponding to dominant sets once they have been extracted using the replicator dynamics. We achieve this goal by repeated asymmetric extensions of the affinity matrix, in a way to preserve all other equilibria.

4.1 Binary Case: Enumeration of Maximal Cliques

In this subsection we will characterize our enumerative approach in presence of binary symmetric affinities, i.e. adjacency matrices of undirected graphs.
Given an undirected graph $G$ with adjacency matrix $A_G$ and $0 < \alpha < 1$, in [6], Bomze provides a one-to-one relation between local maximizers of (4) with $A = A_G + \alpha I$, where $I$ is the identity matrix, and the characteristic vectors of maximal cliques of $G$. This implies also that in the symmetric binary case, maximal cliques are indeed dominant sets.

Since the notion of dominant set still holds for asymmetric affinities, in [3] we inspected its graph-theoretic sense in the context of directed graphs, yielding an interesting result.

Let $G = (V, E)$ be a directed graph. A doubly-linked clique of $G$ is a set $S \subseteq V$ such that for all $u, v \in S$, $(u, v) \in E$ implies $(v, u) \in E$. The clique is saturated if there is no $t \in V \setminus S$ such that for all $s \in S$, $(t, s) \in E$.

**Theorem 1.** Let $G = (V, E)$ be a directed graph with adjacency matrix $A$, $S \subseteq V$ is a saturated doubly-linked clique of $G$ if and only if $x^S$ is an ESS for a two-player game with payoff matrix $B = A + \alpha I$, where $0.5 < \alpha < 1$.

We have already seen that if we consider an undirected graph $G$ and the payoff matrix $A_G + \alpha I$ with $0 < \alpha < 1$, then the ESS's of the related two-player game are in one-to-one correspondence with maximal cliques of $G$. However if we strengthen the constraint on $\alpha$ to lay between 0.5 and 1, then we can see that the concept of saturated doubly-linked clique is a direct generalization to the asymmetric case of the concept of maximal clique, i.e. ESS's are in one-to-one correspondence with saturated doubly-linked cliques.

In order to render a given ESS $x$ unstable, it is sufficient to drop the Nash condition for $x$. A simple way to do it without affecting other equilibria, is to add a new strategy $z$ that is a best reply to $x$, but to no other ESS. This way, $x$ will be no longer asymptotically stable.

Let $G = (V, E)$ be an undirected graph and $G' = (V, E')$ be its directed version where for all $(u, v) \in E$, $(u, v), (v, u) \in E'$. Given a set $\Sigma$ of maximal cliques of $G$, we extend $G'$ obtaining the $\Sigma$-extension $G_\Sigma$ of $G$. The extension is as follows. For each clique $S \in \Sigma$, we create a new vertex $v$, called $\Sigma$-vertex, and put edges from $v$ to each vertex in $S$ and from each vertex outside $S$ to $v$. After this operation, each $\Sigma$-vertex $v$ dominates a particular clique $S$ of $\Sigma$. Further, each vertex not in $S$ dominates the $\Sigma$-vertex $v$ so that it cannot form a new asymptotically stable strategy.

The next theorem guarantees, given a set of maximal cliques $\Sigma$ of an undirected graph $G$, that there exists a one-to-one correspondence between the set of maximal cliques of $G$ not in $\Sigma$ and the set of ESS's of a two-player symmetric game associated to the extended graph $G_\Sigma$.

**Theorem 2.** Let $G = (V, E)$ be an undirected graph, $\Sigma$ be a set of maximal cliques of $G$ and $A$ be the adjacency matrix of the $\Sigma$-extension $G_\Sigma$ of $G$. Let $\Phi$ be a two person symmetric game with payoff matrix $A + \alpha I$ with $0.5 < \alpha < 1$. Then $x$ is an ESS equilibrium of $\Phi$ if and only if it is the characteristic vector of a maximal clique of $G$ not in $\Sigma$.

**Proof.** ($\Rightarrow$) By Theorem 1, if $x$ is an ESS of $\Phi$ then it is the characteristic vector of a saturated doubly-linked clique $S$ of $G_\Sigma$. By construction of $G_\Sigma$, the only
possible doubly-linked cliques are subsets of $V$, therefore $S$ is a clique of $G$. It
is also maximal and not in $\Sigma$ because otherwise it would not be saturated.

$(\Leftarrow)$ Consider $S \notin \Sigma$ a maximal clique of $G$. Then by construction of $G^\Sigma$,
it is a saturated doubly-linked clique of $G^\Sigma$ and hence, by Theorem 1, $x^S$ is an
ESS equilibrium of $\mathcal{G}$. \hfill $\square$

The continuous-based enumerative algorithm uses this result in the following
way. We iteratively find an asymptotically stable point through the replicator
dynamics. If we have an ESS, then we have found a new maximal clique. After
that, we extend the graph by adding the newly extracted clique to $\Sigma$, hence
rendering its associated strategy unstable, and reiterate the procedure until we
have enumerated the selected number of maximal cliques.

4.2 General Case: Enumeration of Dominant Sets

In order to obtain enumeration of dominant sets we iteratively render unstable
all previously extracted ESS’s by adding new strategies that are best replies to
the previous ESS’s, but to no other. This way the previous equilibria will no
longer be asymptotically stable.

Let $\Sigma$ be a tuple of ESS’s of a game with payoff matrix $A$. So for example
if $x$ and $y$ are ESS’s of a doubly symmetric game then $\Sigma = (x, y)$ and with $\Sigma_i$
we select the $i$-th ESS.

We denote the barycenter of the simplex-face spanned by the vertices in $C$
with $b^C \in \Delta$.

$$b^C_i = \begin{cases} \frac{1}{|C|} & i \in C \\ 0 & \text{otherwise} \end{cases}$$

The $\Sigma$-extension $A^\Sigma = (a^\Sigma_{ij})$ of the payoff matrix $A$ is defined as follows.

$$a^\Sigma_{ij} = \begin{cases} a_{ij} & \text{if } i, j \in [1, n] \\ \alpha & \text{if } j > n \text{ and } i \notin \sigma(\Sigma_{j-n}) \\ \beta & \text{if } i, j > n \text{ and } i = j \\ \frac{1}{|\Sigma_{i-n}|} \sum_{k \in \Sigma_{i-n}} a_{kj} & \text{if } i > n \text{ and } j \in \sigma(\Sigma_{i-n}) \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha > \beta$ and $\beta = \max_{i,j} a_{ij}$.

**Lemma 1.** Let $\Phi^\Sigma$ be a two-player game with $\Sigma$-extended payoff matrix $A^\Sigma$ and
let $x$ be an ESS of $\Phi^\Sigma$. If $i$ is a $\Sigma$-strategy such that $i \in \Sigma_{i-n} \cap \sigma(x)$
then there exists $j \in \Sigma_{i-n} \cap \sigma(x)$.

**Proof.** Suppose that $\Sigma_{i-n} \cap \sigma(x) = \emptyset$. Then $(A^\Sigma x)_i = \beta x_i$. But for all $z \in \sigma(x) \cap \{1, \ldots, n\}$ we have that

$$(A^\Sigma x)_z > \alpha x_i > \beta x_i = (A^\Sigma x)_i$$

because by construction of the $\Sigma$-extension, $\alpha > \beta$. This contradicts the hypoth-
esis that $x$ is an ESS. \hfill $\square$
Lemma 2. Let $\Phi^\Sigma$ be a two-player game with $\Sigma$-extended payoff matrix $A^\Sigma$ and let $x$ be an ESS of $\Phi^\Sigma$. Then $\sigma(x)$ does not contain $\Sigma$-strategies.

Proof. Suppose $i$ a $\Sigma$-strategy such that $i \in \sigma(x)$. By lemma 1 there exists $j \in \Sigma_i \cap \sigma(x)$. Take $z = x + \varepsilon (e^i - e^j)$, where $0 < \varepsilon \leq x_j$.

Clearly, we have $z' A^\Sigma x = x' A^\Sigma x$.

The second condition for asymptotic stability is:

$$x' A^\Sigma z > z' A^\Sigma z = x' A^\Sigma z + \varepsilon(e^i - e^j)' A^\Sigma z$$

$$(e^i - e^j)' A^\Sigma z = (e^i - e^j)' A^\Sigma x + \varepsilon(e^i - e^j)' A^\Sigma (e^i - e^j) < 0$$

$$= a_{ii}^\Sigma - a_{ij}^\Sigma - a_{ji}^\Sigma + a_{jj}^\Sigma = \beta - (b_{\sigma^{(\Sigma_i \cap \Sigma_j)}}(A^\Sigma)_z < 0.$$

However, $\beta = \max_{i,j} a_{ij} \geq (b_{\sigma^{(\Sigma_i \cap \Sigma_j)}}(A^\Sigma)_z$, hence, we have a contradiction and we conclude the proof. \hfill \square

From here on we will use the following notation. Let $\Phi$ be a two-player doubly symmetric game with payoff matrix $A$ and $\Phi^\Sigma$ be a two-player game with payoff matrix $A^\Sigma$. If $x$ is a mixed strategy of $\Phi$ then $\bar{x}$ is a mixed strategy of $\Phi^\Sigma$ obtained from $x$ by setting the components relative to $\Sigma$-strategies to 0.

Theorem 3. Let $\Phi$ be a two-player doubly symmetric game with payoff matrix $A$ and let $\Sigma$ be a tuple of ESS’s of $\Phi$. Furthermore let $\Phi^\Sigma$ be a two-player game with payoff matrix $A^\Sigma$. Then $x$ is an ESS of $\Phi$ not in $\Sigma$ if and only if $\bar{x}$ is an ESS of $\Phi^\Sigma$.

Proof. ($\Rightarrow$) Let $x$ be an ESS of $\Phi$ not in $\Sigma$. We will prove that $\bar{x}$ is an ESS also for $\Phi^\Sigma$.

In first step we prove that $\bar{x}$ is a Nash equilibrium by checking that for all pure strategies $i$, $(A^\Sigma \bar{x})_i \leq \bar{x}' A^\Sigma \bar{x}$. It is trivial to verify that $\forall i = 1 \ldots n$,

$$(A^\Sigma \bar{x})_i = (A \bar{x})_i \leq x'A \bar{x} = \bar{x}' A^\Sigma \bar{x}'$$

Consider now the case when $i$ is a $\Sigma$ strategy, i.e. $n \leq i \leq n + |\Sigma|$ and let $z = \Sigma_{i-n}$

$$(A^\Sigma \bar{x})_i = b^{(\Sigma_i \cap \Sigma_j)} A^\Sigma \bar{x} \leq \bar{x}' A^\Sigma \bar{x}$$

Here we have the equality if and only if $z$ is a best reply to $x$. This implies that $z = x$ since $x$ and $z$ are ESS. However, this case is not possible because by hypothesis $x \notin \Sigma$.

This implies that any $y$ best reply to $\bar{x}$ must have $\sigma(y) \subseteq \{1, \ldots, n\}$. Hence, it suffices to check the second condition for evolutionary stability.

$$y' A^\Sigma \bar{x} = \bar{x}' A^\Sigma \bar{x} \Rightarrow \bar{x}' A^\Sigma y = x'A y > y'A y = y' A^\Sigma y.$$
This proves that $\bar{x}$ is an ESS of $\Phi^\Sigma$.

(⇐) Let $x$ be an ESS of $\Phi^\Sigma$. By lemma 2 we have that $\sigma(x)$ does not contain $\Sigma$-strategies. Let $y$ be the mixed strategy obtained removing the $\Sigma$-strategies. It is trivial to verify that $y$ is an ESS of $\Phi$. We show that $y$ cannot be in $\Sigma$, otherwise there would exist a $\Sigma$-strategy $i$ such that

$$(Ay)_i = (A^\Sigma x)_i = b^{\sigma(x)'} A^\sigma x = x' A^\Sigma x$$

and by the second condition for evolutionary stability

$$(x' A^\Sigma)_i = 0 > a^\Sigma_{ii} = \beta$$

arriving at a contradiction. Concluding we have an ESS $y$ of $\Phi^\Sigma$ only if $y = \bar{x}$ where $x$ is an ESS of $\Phi$ not in $\Sigma$.

We use this result to enumerate the dominant sets in the following way: We iteratively find new dominant sets by looking for an asymptotically stable point using the replicator dynamics. After that, we extend the graph by adding the newly extracted set to $\Sigma$, hence rendering its associated strategy unstable, and reiterate the procedure until we have enumerated all the groups and hence are unable to find new dominant sets.

5 Conclusions

In this paper we presented a game-theoretic approach for enumerating dominant sets of an edge-weighted graph. Game dynamics are used to locate individual dominant sets and after each extraction, the similarity matrix is transformed in such a way as to make the located cluster unstable under the replicator dynamics, without affecting the remaining groups.

References