Voting with 2-dimensional preferences

Clay Thomas
claytont@princeton.edu

Yufei Zheng
yufei@cs.princeton.edu

January 14, 2019

Abstract

The study of voting rules often restricts attention to well-behaved classes of possible preferences of voters. We define a new class: 2-dimensional preferences, for which very good voting rules are possible. We argue that 2-dimensional preferences are in some ways more natural and expressive than more traditional classes such as single-peaked preferences. Furthermore, we give an almost-complete combinatorial classification of 2-dimensional preferences, and provide some additional results about the natural extension of d-dimensional preferences.

1 Introduction

We believe that d-dimensional preferences are a very natural setting, and that they give a beautiful geometric explanation for psychological concepts such as opinion and agreement.

Part I

Combinatorics

2 Definition of d-dimensional preferences

Let $\{x_1, \ldots, x_m\} = X \subseteq \mathbb{R}^d_{\geq 0}$ be any (ordered) set of m distinct points with nonnegative coordinates. We call X the set of *outcomes*. Given any $a \in \mathbb{R}^d_{\geq 0}$, define a order $>_a$ on X as follows: $x_i >_a x_j$ if and only if $\langle a, x_i \rangle > \langle a, x_j \rangle$. Define

$$P_d(X) = \{\succ \in R(X) | \exists a \in \mathbb{R}^d_{\geq 0} : x \succ y \iff x >_a y\}$$

Note that we consider only total orders, in particular, ties are not allowed.

Some first observations:

- If $d = 1$, then $|P(X)| = 1$, i.e. preferences are completely determined by the underlying set X.

- If $d = n$, then every linear preference on X can occur in $P(X)$.

    Proof. Let $X = \{e_i\}$ simply be the standard basis vectors. To induce an ordering $i_1, i_2, \ldots, i_n$, just create a preference vector a which gives weight $1/k$ to coordinate i_k. $\qquad\square$

The above hints that there should be some sort of continuum between $d = 1$ and $d = n$ of how complex the set $P(X)$ can be.

3 Lemmas

3.1 Lemmas based on the structure of X

These first few lemmas relate geometric properties of X to limitations on the structure of $>_a$ for a specific, fixed a.

Definition. *Let $x, y \in X \subseteq \mathbb{R}^d_{\geq 0}$. We say x dominates y, denoted $x \gg y$, if $x[k] > y[k]$ for each $k = 1, \ldots, d$.*

Proposition 1. *If $x \gg y$, then $x >_a y$ for any nonzero $a \in \mathbb{R}^d_{\geq 0}$.*

Proof. Simply observe $a_i x_i \geq a_i y_i$ for each $i = 1, \ldots, d$. Because $a \neq 0$, there is also some i where $a_i x_i > a_i y_i$. $\qquad\square$

We'll need to use the familiar geometric notion of a convex hull: ((ACTUALLY, I'M NOT SURE IF WE DO))

Definition. *For points $x_1, \ldots, x_n \in \mathbb{R}^d$, let $\mathrm{hull}(x_1, \ldots, x_n) = \{u_1 x_1 + \ldots + u_n x_n | 0 \leq u_i \leq 1, \sum_{i=1}^{n} u_i = 1\}$ denote the convex hull of $x_1, \ldots, x_n$.*

Lemma 2. *Let $z, x_1, \ldots, x_k \in \mathbb{R}^d_{\geq 0}$ and $a \in \mathbb{R}^d_{\geq 0} \setminus \{0\}$. If $z >_a x_i$ for $i = 1, \ldots, k$, then $z >_a w$ for any $w \in \mathrm{hull}(x_1, \ldots, x_k)$.*

Proof. We have $\langle a, z \rangle > \langle a, x_i \rangle$ for each $i = 1, \ldots, k$. If $w = u_1 x_1 + \ldots + u_n x_n$ and $\sum_i u_i = 1$, then $\langle a, w \rangle = u_1 \langle a, x_1 \rangle + \ldots + u_n \langle a, x_n \rangle < u_1 \langle a, z \rangle + \ldots + u_n \langle a, z \rangle = \langle a, z \rangle$ □

Proposition 3. *Let $z, x_1, \ldots, x_k \in \mathbb{R}^d_{\geq 0}$. Suppose that there exists $w \in \text{hull}(x_1, \ldots, x_k)$ such that $w \gg z$. Then no a satisfies $z >_a x_i$ for each $i = 1, \ldots, k$.*

Proof. For contradiction, suppose such an a exists. Then $z >_a w$ as well. However, because $w \gg z$, this is a contradiction. □

3.2 Lemmas relating different preference vectors

This next group of lemmas relates different different preference vectors. They will be some very useful for our results about voting schemes.

Definition. *For any vectors $a_1, \ldots, a_k \in \mathbb{R}^d_{\geq 0}$, define $\text{cone}(a_1, \ldots, a_k) = \{u_1 a_1 + \ldots + u_k a_k | u_i \geq 0 \forall j\}$. That is, the convex cone of vectors defined by $a_1, \ldots, a_k$.*

Proposition 4. *Suppose that for preference weights $a_1, \ldots, a_k$, we have $x >_{a_1} y, \ldots, x >_{a_k} y$. Then for any nonzero $b \in \text{cone}(a_1, \ldots, a_k)$, $x >_b y$ as well.*

Proof. Let $b = u_1 a_1 + \ldots + u_k a_k$. We get $\langle b, x \rangle = u_1 \langle a_1, x \rangle + \ldots + u_k \langle a_k, x \rangle > u_1 \langle a_1, y \rangle + \ldots + u_k \langle a_k, y \rangle = \langle b, y \rangle$, as desired. □

((INSERT LEMMAS ABOUT SLOPES))

4 Bounding the number of 2-dimensional preferences

5 Impossibility Results

The following preference set may be important in high dimension. We call it d-CYCLE

$$
\begin{array}{ccccccccccccc}
1 & > & 2 & > & 3 & > & \ldots & > & d-2 & > & d-1 & > & d \\
2 & > & 3 & > & 4 & > & \ldots & > & d-1 & > & d & > & 1 \\
3 & > & 4 & > & 5 & > & \ldots & > & d & > & 1 & > & 2 \\
& & & & & & \vdots & & & & & & \\
d-1 & > & d & > & 1 & > & \ldots & > & d-4 & > & d-3 & > & d-2 \\
d & > & 1 & > & 2 & > & \ldots & > & d-3 & > & d-2 & > & d-1
\end{array}
$$

Theorem 5. ?? *d-CYCLE is not $(d-1)$-dimensional for any d.*

Proof. Suppose for contradiction there existed $a_1, \ldots, a_d \in \mathbb{R}^{d-1}_{\geq 0}$ such that a_i yielded the ith line of d-CYCLE (in particular, the line whose favorite point is i). The vectors $\{a_i\}$ cannot be linearly independent. Thus, there exists a linear combination of vectors $u_1 a_1 + \ldots + u_n a_n = 0$. Let $S = \{a_i | u_i > 0\}$ and let $T = \{a_i | u_i < 0\}$. Note that, because every coordinate of each a_i is nonnegative (and no a_i is zero) neither S nor T are empty. We have

$$\sum_{a_i \in S} u_i a_i = \sum_{a_i \in T} -u_i a_i$$

That is, there exists a nonzero $b \in \text{cone}(S) \cap \text{cone}(T)$.

Now, I claim that b satisfies the following:

$$1 >_b 2 >_b 3 >_b \ldots >_b d - 1 >_b d >_b 1$$

Proof: For each pair $i > i + 1$, observe that the inequality is satisfied for all the vectors a_i except for one of them (namely, a_{i+1} which ranks $i + 1$ highest). Thus, for either S or T, every vector a_i in the set has the opinion $i >_{a_i} i + 1$. Thus, every vector in the cone of that set has the opinion $i > i + 1$ as well, in particular, $i >_b i + 1$. Of course, the above argument works for the pair $d > 1$ as well.

Thus, transitivity gives us $1 >_b 1$, a contradiction. $\square$

This is one of our few results which work for any dimension. We'll see that 3-Cycle being impossible for 2 dimensional preferences is a necessary condition for many of the nice properties of 2 dimensional preferences (especially theorem 17). Intuitively, it is possible that "cyclic disagreements" are a core issue for different applications with ordinal preferences (we consider voting here, but other cases include matching or resource allocation). Put another way, the imposibility of the $d + 1$ cycle in d dimensions gives a very nicely parametrized example of preferences getting more complicated as dimension increases.

5.1 $d = 2$

Proposition 6. *Let $P = P_2(X)$ be any set of 2 dimensional preferences, where $|X| \geq 3$. It is not possible for every $x \in X$ to be both the favorite and least favorite outcome of some preverence in P.*

Proof. If any point dominates another, then the dominated point cannot be favorite. Thus we may assume that no point in X is dominated by any other point. Pick two points $x \neq y$ such that x has the largest first coordinate (i.e. x_1) among all of X, and y has the highest second coordinate (i.e. y_2). Consider the following figure:

```
 y-------------
 |’’--.._    2|
 |1      ’’-._|
 |-----------x
```
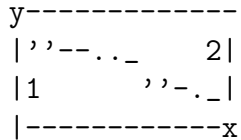
The remaining point $z \neq x, y$ must lie in region 1 or 2, as all other regions either violate the assumption that $x_1 > z_1$ and $y_2 > z_2$ or cause one point to dominate another. If z lies in region 1, it can never be the favorite, by proposition 3. If z lies in region 2, it can never be the least favorite, again by proposition 3. ((TODO: MAYBE SPELL OUT HOW HULLS GET INVOLVED ETC)) $\square$

This proposition, combined with Theorem **??**, gives a complete combinatorial classification of two dimensional preferences on exactly three outcomes: ((TODO: BRING SANDWICH UP HERE OR SOMETHING IDK ITS HARD TO SEPARATE THE COMBINATORIAL AND VOTING DISCUSSION))

Theorem 7. *A set of preferences P on exactly three outcomes is 2 dimensional if and only if it does not contain 3-Cycle or Sandwich (up to relabeling).*

Proof. We have seen that 3-CYCLE and SANDWICH are not 2 dimensional. Thus, it suffices to show that for any set of preferences which does not contain a 3-CYCLE, the only way for the preference set to fail to be 2 dimensional is if it contains SANDWICH.

The following lists all preference on three outcomes, organized by column into the two copies of 3-CYCLE they form.

$$1 > 2 > 3 \qquad\qquad\qquad 1 > 3 > 2$$
$$2 > 3 > 1 \qquad\qquad\qquad 3 > 2 > 1$$
$$3 > 1 > 2 \qquad\qquad\qquad 2 > 1 > 3$$

At least one element from each of these cycles must not be present in P. By relabeling, we can without loss of generality assume that the preference with $1 > 2 > 3$ is not in P. Now, we simply need to look at three cases based on which element of the other cycle is not in P.

- Suppose $1 > 3 > 2$ is not in P. Up to relabeling, all the remaining preferences form a copy of BADCOMPROMISE, where 1 is the point which is never a voter's favorite.

- Suppose $3 > 2 > 1$ is not in P. In this case, the remaining preferences form a copy of SANDWICH, where 3 is the outcome which is always either favorite or least favorite.

- Suppose $2 > 1 > 3$ is not in P. Up to relabeling, all the remaining preferences form a copy of GOODCOMPROMISE, where 2 is the point which is never a voter's least favorite.

□

((WE BELIEVE THAT THE MATERIAL ABOUT SLOPES MAY PROVIDE A MORE GENERAL COMBINATORIAL CLASSIFICATION OF 2D PREFS. SHOW SOME NECESSARY CONDITIONS WE CAN PROVE. CONJECTURE THAT A RELATED THING IS AN IFF))

Part II
Voting

6 Definitions

6.1 Voting

Consider a set of preferences P (i.e. linear orders) on a set of outcomes M. We follow the convention that m is the number of outcomes and n is the number of voters. We have the following definitions:

- A *social welfare function on P* is a function $F : P^n \to P$

- A *social choice function on P* is a function $f : P^n \to M$

- A welfare function F is *unanimous* if, for any $\succ \in P$, we have $F(\succ, \ldots, \succ) = \succ$

- A welfare function F is a *dictatorship* if there exists an i such that $F(\succ_1, \ldots, \succ_n) = \succ_i$

- A choice function f is a *dictatorship* if there exists an i such that $f(\succ_1, \ldots, \succ_n) = a_i$, where a_i is the favorite outcome of $\succ_i$

- A welfare function F satisfies *in dependence of irrelevant alternatives* if, whenever $a \succ_i b \iff a \succ_i' b$ and $\succ = F(\succ_1, \ldots, \succ_n), \succ' = F(\succ_1', \ldots, \succ_n')$, we get $a \succ b \iff a \succ' b$

- A choice function f is *incentive compatible* if, for any $\succ_1, \ldots, \succ_n, i, \succ_i'$, we have $f(\succ_1, \ldots, \succ_n) \succ_i f(\succ_1, \ldots, \succ_i' \ldots, \succ_n)$

- A collection of preferences $\succ_1, \ldots, \succ_n$, has a Condorcet winner $x \in M$ if, for any other $y \in M$, we have $x \succ_i y$ for more than half of the indices $i = 1, \ldots, n$

Let $R(M)$ denote the set of all linear orders on M. Recall that when $P = R(M)$, there are many known impossibility results, the two most famous of which are:

Theorem 8 (Arrow's Impossibility Theorem). *If $|M| \geq 3$, then every unanimous social welfare function on M which satisfies independence of irrelevant alternatives is a dictatorship.*

Theorem 9 (Gibbard Satterthwaite). *If $|M| \geq 3$, then every surjective, incentive compatible social choice function on M is a dictatorship.*

6.2 Single-peaked preferences

Our main point of contrast will be the well-understood class of *single peaked* preferences.

Let $S \subseteq [0, 1]$ be a finite set of m points in the unit interval. We call S the set of *outcomes*. Let $R(S)$ denote the set of linear orders on X. A preference $\succ \in R(S)$ is called *single peaked* if there exists an outcome $p \in S$ (called the "peak" of $\succ$) such that $x < y < p \implies x \prec y$ and $p < y < x \implies x \prec y$. In other words, the preference has a favorite outcome, and its opinion strictly decreases as you move farther away from the favorite. Note that no assumption is made about outcomes on different sides of the peak. Define

$$P_{sp}(S) = \{ \succ \in R(S) | \succ \text{ is single peaked } \}$$

7 Single-peaked verses 2-dimensional preferences

When there are exactly three candidates, 2-dimensional preferences have strictly more expressive power than single-peaked preferences.

Proposition 10. *Every single-peaked preference set on $m = 3$ outcomes is a 2-dimensional preference set. In particular, up to relabeling it is a subset of $P_2(X)$, for $X = \{(3,0),(2,2),(0,3)\}$.*

Proof. Without loss of generality, relabel the outcomes of the single peaked preferences as $S = \{1, 2, 3\}$. The resulting preferences are given by GOOD-COMPROMISE. ((INSERT PICTURE OF GOOD COMPROMISE AS A 2D SET)) □

Furthermore, BADCOMPROMISE gives an example of a preference set which is 2-dimensional, but not single-peaked:

Proposition 11. *In a single peaked set of preferences, it is not possible for every outcome to be the lowest ranked outcome of some preference.*

Proof. The median outcome (via the standard order on $[0, 1]$) cannot be the lowest ranked. □

When $m > 3$, neither single-peaked nor 2-dimensional preferences are a subset of the other. One was to see this to to simply count the number of single-peaked preferences, and see that it grows much faster with m than 2-dimensional preferences do (recall that we showed that the maximal size of a 2-dimensional preference set is $O(m^2)$).

Proposition 12. *The number of single-peaked preferences for any set S of m outcomes is at least $2^{\Omega(m)}$*

Proof. Choose the median outcome of S to be the peak. Given a subset $T \subseteq [m-1]$ of size $|T| = \lfloor m/2 \rfloor$, we can associate a unique single-peaked preference as follows: treat the preference as an array of outcomes, ranked highest to lowest. Put the median of S at index 0. Let the outcomes to the left of the median occupy the indices corresponding to set T, and let those to the right occupy the other indices. There are $\binom{m-1}{\lfloor m/2 \rfloor} \geq 2^{\Omega(m)}$ such subsets T. □

A cleaner, more constructive way to see this is the following proposition

Proposition 13. *The preference set FLIPFLOP is single-peaked, but not 2-dimensional.*

Proof. To obtain a single-peaked representation, order the outcomes from left to right as follows: $3, 1, 2, 4$. All preferences will then be possible with peak 1 or 2.

Now we show that FLIPFLOP is not 2-dimensional. Note that neither 1 nor 2 can dominate the other, and the same holds for 3 and 4. Let a, b satisfy $1 =_a 2$ and $3 =_b 4$. Suppose a has angle smaller than b, the other case being symmetric.

$1 > 2 > 3$
$2 > 1 > 3$
$2 > 3 > 1$
$3 > 2 > 1$

GOODCOMPROMISE

$1 > 2 > 3$
$1 > 3 > 2$
$3 > 1 > 2$
$3 > 2 > 1$

BADCOMPROMISE

$1 > 2 > 3 > 4$
$1 > 2 > 4 > 3$
$2 > 1 > 3 > 4$
$2 > 1 > 4 > 3$

FLIPFLOP

$1 > 2 > 3 > 4$
$4 > 3 > 2 > 1$

REVERSE

$1 > 2 > 3$
$2 > 3 > 1$
$3 > 1 > 2$

3-CYCLE

$1 > 2 > 3$
$1 > 3 > 2$
$3 > 2 > 1$
$2 > 3 > 1$

SANDWICH

Consider the pairs $(1, 2)$ and $(3, 4)$ as they appear in the list in the upper bound section. Because all combinations of $1 > 2$ vs. $2 > 1$ and $3 > 4$ vs. $4 > 3$ occur, neither pair can be before or after the other in the list. This is a contradiction.

((REVISE IMPOSSIBILITY HALF)) □

When considering restricted domains of preference, it is perhaps more important to consider which preference sets *do not* fall into that domain. Indeed, if too many preferences are possible, then the imposibility results which hold for arbitrary preferences will apply.

As we've directly shown, 3-CYCLE is not 2 dimensional. This is already a good sign: 3-CYCLE is a classic example of the so-called Condorcet paradox. Despite each preference being a linear order, the majority of voters favor 1 over 2, 2 over 3, and 3 over 1. That is, "collective preference" is cyclic. Indeed, 3-CYCLE is a preference set with no Condorcet winner, and is often used as a basic counterexample for the existence of good voting schemes. The fact that cycle is impossible already seems like good news for voting schemes on two dimensional preferences.

Here's another case:

Proposition 14. SANDWICH *is neither single-peaked nor* 2 *dimensional*

Proof. SANDWICH is not single-peaked by proposition 11. It's not 2-dimensional by proposition □

Another similarity between SANDWICH. ((DISCUSS SETS WHICH ARE NEITHER SINGLE-PEAKED NOR 2-D))

((ARGUE THAT 2-D IS MORE NATURAL THAN SINGLE-PEAKED, EVEN THOUGH THERE ARE (ASYMPTOTICALLY) MORE SINGLE-PEAKED))

8 Voting for 2-dimensional preferences

Theorem 15. *The median-angle voting rule for* 2-*dimensional preferences is incentive compatible.*

Proof. Let u denote the median-angled preference vector under when all agents report their true preference. Suppose the manipulative voter i's true preference a has angle less than u (the other case is symmetric). The only way in which i can change the chosen preference vector is to increase its angle, say to some v. Note that $u \in \mathrm{cone}(v, a)$.

Suppose y is the outcome chosen by the non-manipulated median angle preference u. Now, in order for i to profit by his manipulation, there must be some x such that $x >_a y$, $y >_u x$, and $x >_v y$. However, this is a contradiction, via the lemma about cones preserving agreed upon opinions.

((PROOF NEEDS REVISION)) □

Theorem 16. *The median-angle social welfare rule satisfies independence of irrelevant alternatives*

Proof. □

9 2-dimensional entails a Condorcet winner

Theorem 17. *Among an odd number of* 2-*dimensional preferences, there is always a Condorcet winner, which is selected by the median angle voting scheme*

8

Proof. Let u be the median-angled preference vector, and let x be the favorite outcome of u. Given some outcome $y \neq x$, if $y \ll x$, then y is never preferred to x. So suppose neither x nor y dominate each other, and let w be the vector such that $x =_w y$. On one side of w, preferences prefer x to y, and on the other side, they prefer y to x. Because x is preferred by the median-angled voter, more than half of the preference vectors lie on the side of w which prefers x.

((PROOF NEEDS REVISION)) □

We note the following important consequence of the previous theorem: if preferences are known to lie in $P_2(X)$ for some X, but no details about X are known, then a voting protocol can simply ask voters for their list of preferences and look for the Condorcet winner (note that there is at most one Condorcet winner). Moreover, it hints that when agents act approximately via a weighted sum of two attributes, it is more likely that there is a Condorcet winner. ((NOTE: CONTEMPLATE THE FOLLOWING HOLE IN THE ABOVE LOGIC: WHAT IF AN AGENT STRATEGICALLY REPORTS A PREFERENCE THAT IS NOT TWO DIMENSIONAL FOR THE UNDERLYING SET X. FOR ONE, THE AGENT MY FORCE THE PROCEDURE TO FAIL BECAUSE IT WON'T HAVE A CONDORCET WINNER. BUT AD HOC WE DON'T KNOW HE CAN'T MANIPULATE THE CONDORCET WINNER))