Metric Preference Learning with Applications to Motion Imitation

Peter Kingston, Jason von Hinezmeyer, and Magnus Egerstedt

[1] School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. E-mail: kingston@gatech.edu
[2] Center for Puppetry Arts, Atlanta, GA 30309 E-mail: Jasonhines@puppet.org
[3] School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. E-mail: magnus@gatech.edu

Abstract. In order for engineered systems to produce behaviors that achieve aesthetic goals, one requires objective functions that accurately represent potentially subjective, human preferences as opposed to *a priori* given objectives. Starting from a collection of empirical, pairwise comparisons, we approach this issue by developing objective functions that are compatible with the expressed preferences. In addition, robust estimators for global optimizers to these functions are derived together with graph-theoretic simplification methods for the resulting systems of constraints and a limited-memory asymptotic observer that finds a globally-optimal alternative (e.g., motion). Two examples are presented involving the comparison of apples and oranges, and of human and synthetic motions.

Keywords: Preference Learning, Motion Imitation, Optimization

1 Introduction to Motion Imitation Through Puppetry

In entertainment and other artistic endeavors, subjective notions such as "style" and "aesthetics" play a key role. When control systems are incorporated in these settings, they are asked to generate behaviors that, rather than achieving a well-defined physical outcome, serve an aesthetic or communicative purpose. In these situations, the effectiveness of the control strategy is ultimately the degree to which it aligns with the subjective judgments of human observers.

One specific control application in which this issue has arisen is *robotic puppetry* (see e.g. [18, 27]), where marionettes with actuated strings – highly complex mechanical systems – are asked to perform expressive and aesthetically pleasing motions in the context of puppetry plays, as in Chapter **??????, by Murphey et al?????**. Given a particular human motion, how should the vastly-more-limited marionette move to not just mimic the human motion but also communicate the same emotional intent? Similar issues are encountered in controls applications like [8, 16], as well as in inverse reinforcement learning settings, e.g., [1, 32, 35].

In puppetry, as in other performing arts, not just any movement is artistically interesting or meaningful; What is the difference between someone just moving an object and someone animating it? What is the difference between someone folding their laundry and someone moving a shirt in a manner that gives it a personality and brings it to life? The movements of the puppet have the intention to give the puppet life and through those movements to communicate an idea. It is the movement of that puppet that is the heart of the communication. But just what is it about the movement that enables effective communication? Similar questions are pursed in Chapter **????? by Scholleing ??? and Chapter ???, by LaViers and Egerstedt**.

To answer this question in a quantitative manner, this Chapter draws inspiration from puppetry. And, it is important to note that puppets are defined by their limitations more than their abilities [7, 10, 19]. In fact, the limited movements of the puppet are a distillation of human movement – so-called motion caricature. Simply put, a puppet cannot do all of the movement of a living person, no matter how simple or complicated the puppet is in its design or construction. It is therefore necessary to distill the movement that the puppet can do to its very essence by first imitating the human motion, then simplifying it to make it executable on the puppet, and lastly, by exaggerating the motion to make it more expressive.

In order to produce such expressive movements, some guidelines have been developed through three main classes of puppet movements [33]. The first class of movements – the primary movements – deals with the overall puppet body and it captures basic movements such as "up", "down", "left", and "right". Any angling or spinning of the puppet body is also considered a primary movement. "Breath" is another major, primary movement that the puppet must perform to seem alive. In fact, most of the puppet's emotions and attitudes are expressed through the primary movement class, including the speed at which the puppet starts and stops or whether the movement is smooth and continuous or jerky and broken.

The secondary movement class is the next class of puppet movement. It considers the moving parts that are attached to the puppet, such as head, arms, and legs. The movements of these parts can multiply the emotions expressed by the puppet through, for example, a subtle tilt of the head or various positions of the arms. Finally, thetertiary movement class concerns puppets that have elements in their designs that are not directly controlled by the puppeteer. These can include, but are not limited to, costume elements such as capes or long sleeves, or hair which can be made of lightweight fibrous materials or feathers. In short, the tertiary class contains anything that can move on the puppet that can help it come alive.

In this paper, we try to take these observations about the Puppetry Arts and make then quantitative and mathematically precise. But, any such endeavor must ultimately allow for subjective, human observers to enter into the discussion. To accommodate this, this paper will indeed focus on how to imitate motions based on human preferences for which synthetic motions best capture

the original, human motion, with the ultimate objective of producing artistically meaningful motions. And we achieve this by making so-called preference learning applicable to motion sequences.

2 Preference Learning

The role of human preferences in these problems is unavoidable in that a system's output is aesthetically pleasing only if we think it is pleasing. In this work, we address is this rather unusual issue by developing techniques both for using empirical measurements to learn cost functions that are consistent with humans' aesthetic preferences, and for generalizing from these preference measurements to determine a globally best alternative. For the example of the marionette asked to mimic a human, this would mean finding the one marionette motion that best captures the subjective "essence" of a given human motion.

The idea of learning costs or rating functions from expressed preferences has received significant attention, and a number of related approaches and problems exist. We sketch a taxonomy of these approaches in the next few paragraphs, before highlighting their potential drawbacks, together with the novelties of the approach pursued here.

In *instance preference learning*, one is given a set of objects called *instances* or *alternatives* (usually points in a real vector space)[4] together with information about humans' preferences among them. The problem is to learn functions that generalize these expressed preferences to the entire space in some way. When the preference data take the form of values from an ordinal scale – e.g., "Good," "Fair," "Poor" – the problem is known as *ordinal regression* (e.g., [12, 23]). When they take the form of a collection of *pairwise comparisons* (i.e. answers to questions of the form "Which of these two options is better?"), we will refer to the problem as *preference learning* (e.g. [2, 11, 13, 15, 20, 22, 25]). Often, preference learning is done by constructing a real-valued ranking function over the instances (e.g., [20]), but in some cases, particularly when one wishes to allow intransitive preferences, one can seek merely to solve a binary classification problem that determines, for a given pair of instances, which is preferred (as in, e.g., [22]). Applications have included route selection by automotive GPS systems [20], food preference studies [6, 17], and the sorting of search results [15, 26], among many others.

It should be noted that pairwise comparison studies have the advantage over numerical or ordinal-ranking experiments of being less prone to *batch effects*, a psychological phenomenon in which people's rankings are only accurate among objects compared at around the same time [17]. Specific experimental protocols include two-alternative forced choice (2AFC), in which, when comparing two objects A and B, subjects must either respond "A is better than B" or "B is

[4] To the extent that a distinction is made between "instances" and "alternatives," it is that "instances" are the points that were shown to human judges, whereas "alternatives" may also include other points in the space besides those that were seen.

better than A;" three-alternative forced choice (3AFC), in which "the two are the same" is also an acceptable answer; and 4AFC, which also includes "the two are incomparable" as an answer. Our attention will be on 2- and 3- AFC experiments. In the related problem of *label preference learning* (e.g. [4, 24]), one attempts to learn a function that, when presented with an instance, returns a partial order over *labels* for that instance. That problem will fall somewhat outside the scope of this work.

For solving instance- and label- preference problems, *large-margin* approaches (e.g., [20, 23]) dominate in the literature. The *margin*, which these methods maximize, is the size of the largest error in the data that can be accommodated without contradicting the learned model. E.g., for linear classifiers, it is the distance from the decision hyperplane to the nearest instance, and for linear rating functions, it is the distance between level sets of the learned function. Ultimately, Support Vector Machine (SVM) algorithms (e.g., SMO [31]) are used to solve these problems. That said, competitive Gaussian-process [11] and least-squares [25] approaches also exist, which bring Bayesian and algebraic-topological interpretations, respectively.

The essential idea underlying the large-margin approaches is to develop algorithms that learn linear rating functions or classifiers, and to then generalize these algorithms to the nonlinear case through the use of *Mercer kernels* (the so-called *kernel trick* [5, 9]). It is an elegant and general approach, but what it lacks is the ability to easily guarantee that the resulting cost functions are convex, which is an obstacle to the efficient determination of globally-best alternatives. We will investigate techniques that sacrifice some of the generality of the kernel-linear approach in exchange for formulations that allow for the efficient determination of globally optimal alternatives. This paper unifies and builds on our work [28, 29], and, of earlier work on preference learning, is most closely related to the large-margin instance-preference-learning approaches of [17, 22, 23]. One of the methods we propose – a Chebyshev estimation scheme – is similar in that it also employs a constrained optimization approach and in a particular limiting case (but only then) can be reduced to an equivalent SVM classification problem. Our proposed approach differs from [17, 22, 23] in that it aims to find not only a rating function but also a globally best alternative, and to solve only computationally-efficient convex programs, which in turn motivates different (and sometimes more efficient) problem formulations and solutions.

The resulting algorithms operate on batch data, and simultaneously use all of the information available. We give a graph-theoretic method, based on the work of [3] on transitive reductions, to efficiently simplify the resulting optimization problems prior to their solution. Additionally, for situations in which a great deal more data are available in real-time, we also present a limited-memory asymptotic observer, which trades some efficiency in the use of data for constant memory requirements and very cheap measurement updates, and which nevertheless guarantees probabilistic convergence to a global optimum.

The paper is organized as follows: After giving a problem formulation, we describe an algebraic-graph-theoretic simplification method for the resulting sys-

tems of constraints, before introducing the optimization problems in more detail in both *direct* and *instance vector expansion* forms, and an example is given to demonstrate the applicability of the methods to LQ-type optimal control problems. Next we show how in a particular limiting case (but only then) a natural generalization is equivalent to a certain SVM classification problem, before describing a novel limited-memory asymptotic observer. The approaches are demonstrated with two examples; the first of these compares apples relative to an orange, and the second involves the comparison of human and synthetic motions.

3 Problem Formulation

At the core of preference learning is a collection of empirical, pairwise comparisons. The assumption is that these comparisons reflect an underlying rating function. Hence, given a sequence of pairwise comparisons between points in an inner product space, we wish to find (1) a real-valued rating function that is consistent with those preferences, and (2) a global optimum to this function – the best point in the metric space. By solving these two problems we would have recovered what the underlying source for the comparisons is.

Formally, let $(X, \langle\cdot,\cdot\rangle)$ be the inner product space, and $S = \{(x_i^1, x_i^2)\}_{i=1}^N = \{s_1, \cdots, s_N\} \subset X \times X$ the sequence of comparisons; a pair (x_i^1, x_i^2) appears in the sequence S if and only if x_i^1 is preferred to x_i^2. The first item we seek, given some assumptions about its parametric form, is a function $f : X \to \mathbb{R}$ such that

$$f(x^1) \leq f(x^2) \Leftrightarrow (x^1, x^2) \in S \ . \tag{1}$$

That is, we adopt the convention that lower scores are better; hence we will refer to f as a *cost function*.

The second item we seek is a global minimizer to f,

$$\bar{x} \triangleq \text{argmin}_x f(x) \tag{2}$$

which represents the best possible point in the inner product space.

Crucially, we would like to be able to determine f and $\bar{x}$ entirely by convex optimization – both so that the resulting problems are computationally efficient, and to ensure that any minima we find are in fact global optima. Although the SVM methodology employed in e.g. [17, 23] finds f as the solution to a convex program, its use of *Mercer kernels* (the so-called *kernel trick* – see e.g. [5, 9]) introduces nonlinearities that prevent the determination of $\bar{x}$ by convex programming. Yet without the kernel trick and using the SVM approach, one arrives at linear cost functions for which minima do not exist. What we will present in this paper is instead a set of convex programs that provide a useful compromise between these extremes, and which only reduce to an SVM classification problem in a particular limiting case. These formulations will allow us to entertain the idea of a unique "best" point in X, and at the same time determine what it is by convex programming.

4 The Preference Graph

The *preference graph* $\mathcal{G} = (V, S)$ corresponding to the comparison sequence S is the directed graph whose vertex set $V = \{x_1^1, x_1^2, \cdots, x_N^1, x_N^2\} \subset X$ is the collection of all unique points that have been compared, and whose edge set is S. We will index the vertices as $V = \{x_1, \cdots, x_M\}$, where $M \leq 2N$ is the cardinality of V.

If (1) is to hold with strict inequality, then we note immediately that the graph $\mathcal{G}$ must be *acyclic*, and thus represent a *partial order*. When nonstrict inequalities are allowed, however, then we may permit cycles, and moreover $\mathcal{G}$ can be replaced by a smaller, equivalent acyclic graph. This has the practical significance of allowing redundant constraints to be eliminated on purely graph-theoretic grounds, thereby speeding up later optimization steps. This is constructed, following [3], in the following way:

A *cell* is defined to be an equivalence class of vertices; two vertices $v_1, v_2 \in V$ belong to the same cell (denoted $v_1 \sim v_2$) if and only if there exist directed paths in $\mathcal{G}$ from v_1 to v_2 *and* from v_2 to v_1. The quotient graph $\mathcal{G}/\sim$ is the directed acyclic graph whose vertices are these equivalence classes, and in which the directed edge (C_1, C_2) exists between two cells C_1 and C_2 whenever there exist vertices $v_1 \in C_1$ and $v_2 \in C_2$ such that there is a directed path in $\mathcal{G}$ from v_1 to v_2.

Since any two vertices in the same cell must by (1) have the same cost, one may optimize using only the constraints represented by the edges of this quotient graph, and discard the rest. Hence without loss of generality we will assume that $\mathcal{G}$ is acyclic; when it is not it should be understood that we will actually work with $\mathcal{G}/\sim$.

Additional constraints can be eliminated via the *transitive reduction*. Formally, using Aho's definition [3], $\mathcal{G}^t$ is the transitive reduction of a graph $\mathcal{G}$ if,

1. there is a directed path from vertex u to vertex v in $\mathcal{G}^t$ if and only if there is a directed path from u to v in $\mathcal{G}$, and
2. there is no graph with fewer arcs than $\mathcal{G}^t$ satisfying condition 1.

In the case of a directed acyclic graph, the reduction $\mathcal{G}^t$ (which is unique) is a subgraph of $\mathcal{G}$. It was shown in [3] that computation of the transitive reduction is of the same complexity as transitive closure, and hence matrix multiplication; thus, the transitive reduction can be found in $O(n^{\log_2 7})$ steps using Strassen's algorithm [34], or, in principle, $O(n^{2.376})$ steps using the Coopersmith-Winograd algorithm [14]. (See, e.g., [21, 30]). Moreover, if $\mathcal{G}$ contains cycles, then the algorithm given in [3] can compute $(\mathcal{G}/\sim)^t$ with the same complexity.

In short, by working with the transitive reduction of the quotient graph, we are able to eliminate redundant constraints on purely graph-theoretic grounds, before even knowing the form of the cost function f. The reduction is illustrated by Figure 1.

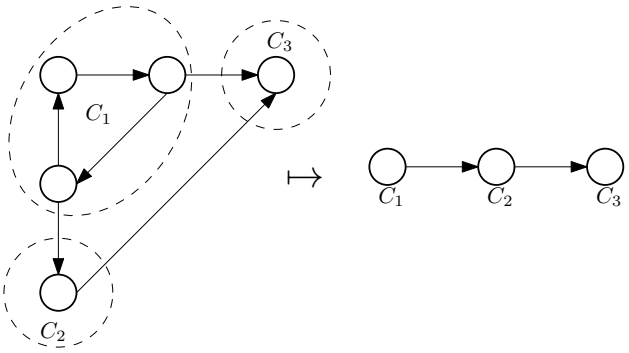

Fig. 1. The original preference graph $\mathcal{G}$ (left), and the corresponding transitively-reduced quotient graph, $(\mathcal{G}/\sim)^t$ (right). The vertices of $(\mathcal{G}/\sim)^t$, labeled $C_1, \cdots, C_3$, are sets of vertices in $\mathcal{G}$ called *cells* (dashed circles and ellipses, left).

5 Metric Costs

Colloquially, when comparing various alternatives, we often speak of options as being "closer to what we would like," or of being "far from perfect." Motivated by this everyday use of geometric language, in [28] we considered *metric costs*, which have the form,

$$f(x) = ||x - \bar{x}||^2 \ . \tag{3}$$

In short, it is assumed that there exists some single best point $\bar{x}$ in X, and one alternative is preferred over another if and only if it is closer to that point. Moreover, costs of this type are nicely compatible with the infinite-dimensional alternatives encountered in control applications, as demonstrated in the following, motivating example:

Example 1 (LQ Tracking).
We consider a situation in which a human user would like a linear system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{4}$$

to perform a tracking task, with an LQ-type cost functional,

$$J(y,u) = \frac{1}{2}\int_0^T \left[(y-\bar{y})^T Q(y-\bar{y}) + (u-\bar{u})^T R(u-\bar{u})\right] dt \tag{5}$$

where $Q = Q^T \succ 0$, $R = R^T \succ 0$, and $T > 0$. Subjectivity enters because the pair $\bar{x} \triangleq (\bar{y}, \bar{u})$ is known only implicitly to the user, and it is the goal of preference learning to estimate what it is. It is well-known that (5) can be written as

$$\frac{1}{2}||(y,u) - (\bar{y},\bar{u})||^2_{\mathrm{LQR}} \tag{6}$$

where $||\cdot||_{\text{LQR}}$ is the norm induced by the inner product

$$\langle (y_1, u_1), (y_2, u_2) \rangle_{\text{LQR}} = \int_0^T \left[y_1^T Q y_2 + u_1^T R u_2 \right] dt \ . \tag{7}$$

If both y and u are measured, this is enough to apply directly the preference learning methods to be introduced in the subsequent sections, without knowledge of the dynamics (4). However, if these dynamics are known, then together with (7), they induce an inner product on the control inputs, which enables us to work in that smaller vector space. This is done in the following way:

Defining the linear operator $\mathcal{L} : L^2([0,T], \mathbb{R}^m) \to L^2([0,T], \mathbb{R}^p)$ by,

$$(\mathcal{L}u)(t) \triangleq C \int_0^t e^{A(t-\tau)Bu(\tau)} d\tau \tag{8}$$

we define a second inner product,

$$\langle u_1, u_2 \rangle_{\text{LQR}'} \triangleq \langle (\mathcal{L}u_1, u_1), (\mathcal{L}u_2, u_2) \rangle_{\text{LQR}} \ . \tag{9}$$

With these definitions, the cost can be written,

$$J(u) = \frac{1}{2} ||u - \bar{u}||_{\text{LQR}'} \tag{10}$$

and the problem, by taking advantage of knowledge of the system dynamics, has been reduced from that of finding a pair $(\bar{y}, \bar{u})$, to that of simply finding $\bar{u}$. Moreover, note that (9) can be written in a more standard double-integral form as,

$$\langle u_1, u_2 \rangle_{\text{LQR}'} = \int_0^T \int_0^T u_1^T(\tau_1) M(\tau_1, \tau_2) u_2(\tau_2) d\tau_1 d\tau_2 \ , \tag{11}$$

where

$$M(\tau_1, \tau_2) = \delta(\tau_1 - \tau_2) R + \int_{T - |\tau_2 - \tau_1|}^T B^T e^{A^T(t-\tau_1)} c^T Q c e^{A(t-\tau_2)} B dt \tag{12}$$

and δ denotes the Dirac delta distribution. In this way, system dynamics are incorporated directly into the preference learning framework.

Under an assumption of metric costs, what does an individual response (x^1, x^2) tell us about the location of $\bar{x}$? The following are equivalent:

1. $(x_i^1, x_i^2) \in S$
2. $f(x^1) \leq f(x^2)$
3. $\left\langle x_i^2 - x_i^1, \bar{x} \right\rangle - \frac{1}{2} \left\langle x_i^2 - x_i^1, x_i^2 + x_i^1 \right\rangle < 0.$

In words, each comparison constrains $\bar{x}$ to lie within a particular halfspace of X. Defining,

$$d_i \triangleq x_i^2 - x_i^1 \tag{13}$$

$$\mu_i \triangleq \frac{1}{2} \left(x_i^1 + x_i^2 \right) \tag{14}$$

$$b_i \triangleq \langle d_i, \mu_i \rangle \ , \tag{15}$$

the totality of what we know, then, about where $\bar{x}$ might lie is summarized by the inclusion over all the comparison halfspaces,

$$\bar{x} \in P \triangleq \bigcap_{i=1}^{N} \{x \mid \langle d_i, x \rangle - b_i < 0\} \ . \tag{16}$$

The set P, if it is bounded, is a polytope in X. In [28], we stated this system of inequalities and gave an asymptotic observer that converges to $\bar{x}$ under certain assumptions. Here, we ask another question: Out of all the points in this polytope, which is "best?" Two cases will be relevant: Either a particular linear program is bounded, in which case such a platonic ideal point $\bar{x}$ exists; or it is unbounded, in which case only an ideal *direction* or *ray* exists. The next subsections introduce this linear program and address these two cases in turn.

5.1 Bounded Case

When P is bounded, a natural choice for $\bar{x}$ is the *incenter* or *Chebyshev center* of the polytope,

$$\bar{x} = \text{argmin}_x \max_i \frac{1}{||d_i||} \left(\langle d_i, x \rangle - b_i \right) \ . \tag{17}$$

This is the point that is maximally far away from the closest constraint plane, as illustrated by Figure 2. In other words, when P is nonempty, $\bar{x}$ is the point that can be perturbed as much as possible without contradicting any of the preferences expressed in S; and when P is empty, it is the "compromise" point whose worst constraint violation is minimal. I.e., with the definition (17), if the constraints are feasible (i.e., if $P \neq \emptyset$), then $\bar{x} \in P$. This can be viewed as minimizing the ∞-norm of the vector of constraints. The minimization problem (17) is feasible even when P is empty, in which case its solution is the point whose worst constraint violation is as small as possible. Equivalently, the problem (17) can be rewritten in epigraph form as the linear program,

$$\begin{aligned} (\bar{z}, \bar{x}) \ = \ & \text{argmin}_{(z,x)} z \\ \text{s.t.} \quad & ||d_i|| z \geq \langle d_i, x \rangle - b_i \end{aligned} \tag{18}$$

which is always feasible (but possibly unbounded), and satisfies $\bar{z} > 0 \iff P = \emptyset$.

The formulation (17) also has the advantage that it can be solved with a complexity that grows not with the dimensionality of the space X containing the points, which may be quite large, but instead with the number of instances seen, which is typically much smaller; Theorem 1 states the fact that enables this simplification:

Theorem 1. *If (17) has a global minimizer, then it has a global minimizer in* $\text{aff}\left\{x_1^1, x_1^2, \cdots, x_N^1, x_N^2\right\}$.

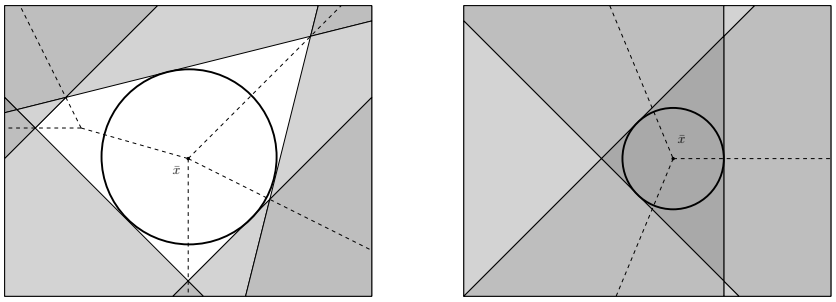

Fig. 2. Two examples for $X = \mathbb{R}^2$. Shades of gray indicate the number of violated constraints (points in darker regions violate more constraints), and discontinuities in the derivative of the piecewise-linear function $x \mapsto \max_i \frac{1}{||d_i||} (\langle d_i, x \rangle - b_i)$ are indicated by dashed lines. In the first example (left), $P \neq \emptyset$ (white region), and $\bar{x}$ is its incenter, the point maximally far away from the closest of the constraint surfaces (thin, solid lines) - i.e., it is the center of the largest inscribed sphere (thick, solid curve). In the second example (right), $P = \emptyset$, and the resulting optimum, $\bar{x}$, is the point whose worst constraint violation is minimal.

Proof. Let x be a global minimum to (17), and $\bar{x}$ be the projection of x onto $\text{aff}\left\{x_1^1, x_1^2, \cdots, x_N^1, x_N^2\right\}$; i.e., $\bar{x} = x + \delta$ with $\delta \perp \text{span}\{d_1, \cdots, d_N\}$. Then for all $i \in \{1, \cdots, N\}$, since $\langle d_i, \delta \rangle = 0$ and by linearity of the inner product, $\frac{1}{||d_i||} \langle d_i, \bar{x} \rangle - b_i = \frac{1}{||d_i||} \langle d_i, x \rangle - b_i$, and hence the value of the objective function in (17) is the same at either x or $\bar{x}$.

In Section 5.1, which follows, we describe an efficient *Instance Vector Expansion* form suggested by Theorem 1, and obtain problems that have optima when P is bounded. For the case when the set P is unbounded, we will generalize the solution concept and find optimal *directions* rather than points, in Section 5.2.

Instance Vector Expansion Since $\bar{x} \in \text{aff}\left\{x_1^1, x_1^2, \cdots, x_N^1, x_N^2\right\}$, i.e., the affine span of the constituent points, the optimization problem (17) can be solved as a finite-dimensional problem even when X is not finite-dimensional, by expanding $\bar{x}$ in terms of a finite-dimensional basis, as described by the following theorem:

Theorem 2. *The point*

$$\bar{x} = \sum_{k=1}^{N} \bar{c}_k d_k + x^* \tag{19}$$

solves the optimization problem (17), where

$$x^* = \mathit{argmin}_x \left\{ ||x||^2 \mid x \in \text{aff}\left\{x_1^1, x_1^2, \cdots, x_N^1, x_N^2\right\} \right\} \ , \tag{20}$$

and $\bar{c}$ is found by solving

$$\begin{aligned} (\bar{z}, \bar{c}) \ &= \ \mathit{argmin}_{(z,c)} z \\ \text{s.t.} \quad & G^{dd} c - Dz \leq \beta \ , \end{aligned} \tag{21}$$

with $D = (||d_1||, \cdots, ||d_N||)$, $\beta \in \mathbb{R}^N$ *defined by*

$$\beta_i \triangleq \langle d_i, \mu_i \rangle \tag{22}$$

and $G^{dd} \in \mathbb{R}^{N \times N}$ *being the Gramian,*

$$G^{dd} \triangleq \begin{bmatrix} \langle d_1, d_1 \rangle & \cdots & \langle d_1, d_N \rangle \\ \vdots & \ddots & \vdots \\ \langle d_N, d_1 \rangle & \cdots & \langle d_N, d_N \rangle \end{bmatrix} . \tag{23}$$

Proof. Defining x^* by (20), one can write any x in the affine span of the data in the form (19). Substituting the expansion (19) into (18) and noting that by Hilbert's Projection Theorem $x^* \perp d_i$ for all $i \in \{1, \cdots, N\}$, one obtains (21).

Remark 1. We also note at this point that (19) can be written,

$$x = \sum_{k=1}^{M} \left(\text{indeg}_c(x_k) - \text{outdeg}_c(x_k)\right) x_k + x^* \tag{24}$$

$$\triangleq \sum_{k=1}^{M} \xi_k x_k + x^* \tag{25}$$

by treating c as a vector of edge weights to the preference graph, and denoting the weighted in- and out-degrees of a given node x_k by $\text{indeg}_c(x_k)$ and $\text{outdeg}_c(x_k)$ respectively. Precisely,

$$\text{indeg}_c(x_k) \triangleq \sum_{i | x_i^2 = x_k} c_i \tag{26}$$

$$\text{outdeg}_c(x_k) \triangleq \sum_{i | x_i^1 = x_k} c_i \ . \tag{27}$$

Remark 2. Moreover, β can be written,

$$\beta_i = e_i^T G^{\mu d} e_i \ , \tag{28}$$

where $G^{\mu d} \in \mathbb{R}^{N \times N}$ is the cross-Gramian

$$G^{\mu d} \triangleq \begin{bmatrix} \langle d_1, \mu_1 \rangle & \cdots & \langle d_1, \mu_N \rangle \\ \vdots & \ddots & \vdots \\ \langle d_N, \mu_1 \rangle & \cdots & \langle d_N, \mu_N \rangle \end{bmatrix} \tag{29}$$

and e_i denotes the i-th element of the natural basis.

Remark 3. Note that the problem (21) depends only on inner products of the various d_i and u_i vectors, and hence the problem can be solved even when X is infinite-dimensional. Precisely, $\frac{N(N+1)}{2} + N^2 \sim O(N^2)$ inner products must be computed to build the matrices G^{dd} and $G^{\mu d}$, where N is the number of

comparisons. Alternatively, the relevant matrices can also be produced directly from inner products of elements of S, as

$$G^{dd} = K^{22} - K^{21} - K^{12} + K^{11} \tag{30}$$

$$G^{\mu d} = \frac{1}{2}(K^{22} + K^{21} - K^{12} - K^{11}) \tag{31}$$

where each matrix $K^{lm} \in \mathbb{R}^{N\times N}$ is defined by

$$K^{lm}_{ij} = \left\langle x^l_i, x^m_j \right\rangle \tag{32}$$

and can be built by indexing into the single Gramian (or kernel) matrix $K \in \mathbb{R}^{M\times M}$ defined,

$$K_{ij} = \langle x_i, x_j \rangle \ . \tag{33}$$

Moreover, $D = (\sqrt{G^{dd}_{11}}, \sqrt{G^{dd}_{22}}, \sqrt{G^{dd}_{33}}, \cdots, \sqrt{G^{dd}_{NN}})$.

Finally, $\bar{x}$ can be reconstructed using (19) and

$$x^* = \sum_{i=1}^{M} \alpha_i x_i \tag{34}$$

$$\alpha = \frac{1}{\mathbf{1}^T K^\dagger \mathbf{1}} K^\dagger \mathbf{1} \tag{35}$$

where $K^\dagger$ denotes the Moore-Penrose pseudoinverse of K, and $\mathbf{1} = (1, 1, \cdots, 1) \in \mathbb{R}^M$.

In particular, the costs of the presented instances can be reconstructed as,

$$f(x_k) = (e_k - \xi - \alpha)^T K (e_k - \xi - \alpha) \tag{36}$$

where ξ is related to c by (24), (26), and (27).

With a method thus in hand to efficiently compute Chebychev centers when P is bounded, we now turn our attention to the case when it is not, so that we have a meaningful solution in all cases.

5.2 Unbounded Case: The minimax-rate problem

When P is nonempty but unbounded, we ask a slightly different question: What is the "point at infinity," or *direction*, that is best? More precisely, what we seek in this case is a unit vector

$$\bar{v} = \text{argmin}_{v\in X | ||v||=1} \lim_{t\to\infty} \frac{1}{t} \left[\max_i \frac{1}{||d_i||} \left(\langle d_i, tv \rangle - b_i \right) \right] \tag{37}$$

$$= \text{argmin}_{v\in X | ||v||=1} \max_i \frac{1}{||d_i||} \langle d_i, v \rangle \tag{38}$$

or equivalently,

$$(\bar{p}, \bar{v}) = \operatorname{argmin}_{v \in X, p \in \mathbb{R}} p \quad \text{s.t.} \quad \begin{cases} ||d_i|| p \geq \langle d_i, v \rangle & \forall i \in \{1, \cdots, N\} \\ ||v||^2 \leq 1 \end{cases} . \tag{39}$$

As before, an instance vector expansion is possible:

Theorem 3. *Letting $v = \sum_{k=1}^{N} c_k d_k$, the optimization problem (39) is equivalent to*

$$(\bar{p}, \bar{c}) = \operatorname{argmin}_{(p,c)} p \quad \text{s.t.} \quad \begin{cases} G^{dd} c - Dp \leq 0 \\ c^T G^{dd} c \leq 1 \end{cases} \tag{40}$$

with the matrices G^{dd} and D as defined in the previous subsection.

Proof. The proof takes the form of that of Theorem 2.

The problem (40) is a finite-dimensional second-order cone program (SOCP), which can be solved efficiently.

The cost function for the unbounded case arises from a similar limit process to (38), as

$$f(x) = \lim_{t \to \infty} \left(\frac{1}{t} ||x - vt||^2 - t \right) \tag{41}$$

$$= \lim_{t \to \infty} \left[\frac{1}{t} \left(||x||^2 - 2 \langle x, vt \rangle + ||vt||^2 \right) - t \right] \tag{42}$$

$$= -2 \langle x, v \rangle \tag{43}$$

which can be evaluated at the instances as,

$$f(x_k) = -2 e_k^T K \xi \ . \tag{44}$$

QP Form and Relation to SVMs In the unbounded case, so long as $\operatorname{int} P$ is nonempty, the minimization problem (38) can be rewritten as an equivalent quadratic program (QP), which will make the relationship to the usual SVM approach very clear. In fact, (38) is equivalent in this case to a particular SVM classification problem (which differs from but is related to that studied in e.g. [17] and [23]).

Defining,

$$w = \frac{1}{p} v \tag{45}$$

and restricting our attention to negative values for p (since when $\operatorname{int} P$ is nonempty, $p^* < 0$), we note that

$$\operatorname{argmin}\ p = \operatorname{argmax}\ p^2 = \operatorname{argmin}\ \frac{1}{p^2} = \operatorname{argmin}\ ||w||^2 \ . \tag{46}$$

Additionally, the constraints in (39) can be replaced by,

$$\left\langle \frac{d_i}{||d_i||}, w \right\rangle \geq 1 \tag{47}$$

which results in the standard unbiased SVM problem,

$$\begin{aligned} \bar{w} \;&=\; \operatorname{argmin}_w ||w||^2 \\ \text{s.t.} \quad &\left\langle \frac{d_i}{||d_i||}, w \right\rangle \geq 1 \quad \forall i \in \{1, \cdots, N\} \; . \end{aligned} \tag{48}$$

This is equivalent to (39) in the unbounded case except when $\operatorname{int} P = \emptyset$; then, since $\bar{p} = 0$, $\bar{w}$ from (45) is undefined, but the solution to the SOCP problem (40) nevertheless exists.

The *minimax-rate* problem (48) differs from the SVM problem considered in e.g. [17] and [23] by the factor of $\frac{1}{||d_i||}$ included in each constraint. The difference is that whereas the standard SVM approach attempts to classify differences using a maximum-margin separating hyperplane, the minimax-rate approach finds the direction that maximizes the rate of constraint satisfaction; this is illustrated in Figure 3.

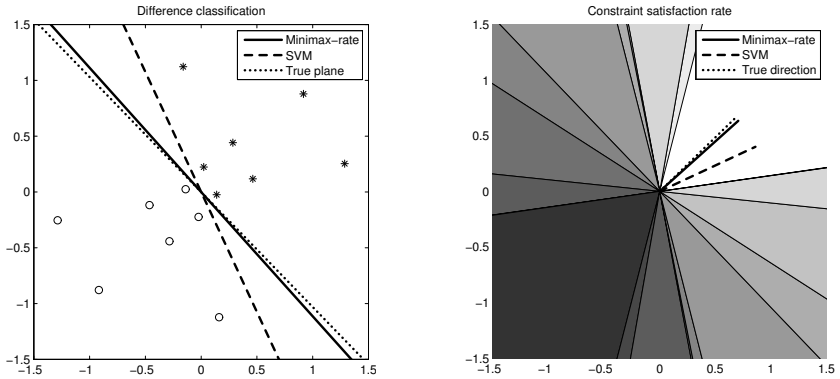

Fig. 3. A number of uniformly-randomly selected points in $[-1, 1] \times [-1, 1] \subset \mathbb{R}^2$ are compared according to a point at infinity (i.e., a linear cost function) (dotted), and both the traditional SVM (dashed) and the minimax-rate (solid) approaches are used to produce estimates of this direction from the comparisons. From the difference-classification point of view (top), one wishes to separate the vectors $\{d_i\}_{i=1}^N$ (displayed as "o"s) from the vectors $\{-d_i\}_{i=1}^N$ (displayed as "*"s). From the minimax-rate point of view (bottom), one wishes to find the direction that maximizes the rate of constraint satisfaction (the numbers of violated constraints are represented by shades of gray; the white region is feasible). The traditional SVM solution separates the positive from the negative differences with a larger margin (top), but the minimax-rate solution stays as far from the edge of the constraint cone as possible (bottom).

6 An asymptotic observer for metric cost models

The preference learning methods described so far have been batch processes. Preference data are collected, and then an optimization problem is solved to produce optimal alternatives. Although the convex formulations chosen allow fairly large problems to be solved efficiently in this way, the memory requirements do grow linearly in the number of observations. In this section, instead of a batch algorithm, we consider a stream algorithm that continually adjusts its estimate of the optimal alternative. This algorithm has constant memory and time-per-update requirements, so it is more suitable for use in embedded applications, where, e.g., a small microcontroller can continually adjust system operation to improve subjective performance.

Formally, suppose we have access to a very long (infinite) sequence of comparisons $S = \{(x_k^1, x_k^2)\}_{k=1}^{\infty} = \{s_1, s_2, \cdots\} \subset X \times X$, perhaps as the result of passive monitoring over an extended period of time, and we would like to know the features $\bar{x}$ of the ideal alternative. If alternatives are presented at random to the human, can we construct an asymptotic observer for $\bar{x}$ which can avoid storing all of the very (infinitely) many constraints implied by this sequence? It turns out that the answer is yes, and exactly such an observer is given by,

$$\tilde{x}_{k+1} = \begin{cases} P^k \tilde{x}_k + \frac{\alpha_k b_k}{d_k^T d_k} d_k \text{ if } d_k^T \tilde{x}_k - b_k > 0 \\ \tilde{x}_k \qquad \text{otherwise} \end{cases} \tag{49}$$

$$P_k = I - \alpha^k \frac{d_k d_k^T}{d_k^T d_k} \tag{50}$$

for any sequence of observer gains $\alpha_k \in (0, 2)$ (and d_k, b_k defined by (13-15)), regardless of $\tilde{x}_0$. That is, $\tilde{x}_k$ converges to $\bar{x}$ in probability as $k \to \infty$, given a few assumptions; we will prove this shortly in Theorem 4. Moreover, note that, although (49-50) are broken down into separate expressions for clarity of presentation, they are in fact all functions of $\tilde{x}^k$, so this observer can be implemented with only $\dim\{X\}$ real memory elements.

Geometrically, the observer (49-50) operates through a series of projections (or under/over-projections, if $\alpha_k \neq 1$), as illustrated in Figure 4, with each projection bringing the estimate $\tilde{x}_k$ of the ideal closer to the true ideal, $\bar{x}$. That the resulting sequence does indeed converge to $\bar{x}$ is guaranteed by Theorem 4, which follows.

Before continuing, we now state a useful lemma, whose geometric interpretation is that comparisons between distances relative to reference points can be interchanged with signed point-plane distance tests.

Lemma 1. *Let x^1, x^2, $\bar{x}$ be any vectors in an inner product space $(X, \langle \cdot, \cdot \rangle)$, and let $\bowtie$ be a binary relation from the set, $\{=, <, >, \leq, \geq\}$. Then,*

$$\bar{x} \in \{x \mid \langle d, x \rangle - b \bowtie 0\} \iff ||x^1 - \bar{x}|| \bowtie ||x^2 - \bar{x}||$$

where $d = x^2 - x^1$, and $b = \frac{1}{2} \langle d, x^1 + x^2 \rangle$.

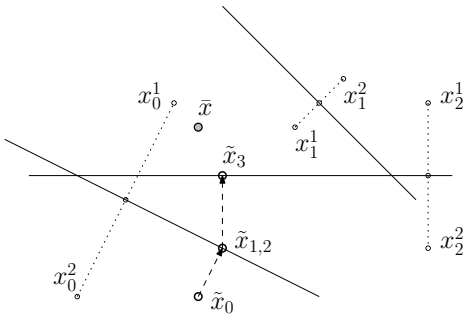

Fig. 4. A series of the observer's estimates, with $\alpha_k = 1\ \forall k$. The initial estimate is $\tilde{x}_0$, and the true ideal is given by $\bar{x}$. In step 0, the observer projects $\tilde{x}_0$ onto the plane (solid line) corresponding to the measured output $s_0 = (x_0^1, x_0^2)$ to produce $\tilde{x}_1$. In step 1, the observer makes no changes to its estimate, because $\tilde{x}_1$ is on the correct side of the plane corresponding to s_1; hence $\tilde{x}_2 = \tilde{x}_1$. In step 2, the observer projects $\tilde{x}_2$ onto the plane corresponding to s_2 to create the estimate $\tilde{x}_3$, which is yet closer to $\bar{x}$.

Proof. The proof of this is based on the Polarization Identity and is straightforward.

Theorem 4. *Let $\bar{x} \in X$ be the ideal alternative, and $S = \{(x_k^1, x_k^2)\}_{k=1}^{\infty} = \{s_1, s_2, \cdots\}$ a sequence of pairs of i.i.d. random vectors drawn according to a probability density function p on $\{(x^1, x^2) \in X \mid ||x^1 - \bar{x}|| < ||x^2 - \bar{x}||\}$ which is nonzero in an open ball $B(\bar{x}, r) = B_r$ around $\bar{x}$. Then, the asymptotic observer given by (49-50) converges to $\bar{x}$ in probability.*

Proof. **1.** *If $\langle d_k, \tilde{x}_k \rangle - b_k > 0$, then $||\tilde{x}_{k+1} - \bar{x}|| < ||\tilde{x}_k - \bar{x}||$.* The distances $||\tilde{x}_k - \bar{x}||$ and $||\tilde{x}_{k+1} - \bar{x}||$ are related through the Polarization Identity by (where $\Delta_k = \tilde{x}_{k+1} - \tilde{x}_k$),

$$||\tilde{x}_{k+1} - \bar{x}||^2 = ||\tilde{x}_k + \Delta_k - \bar{x}||^2 = \\ ||\tilde{x}_k - \bar{x}||^2 + ||\Delta_k||^2 + 2\langle \tilde{x}_k - \bar{x}, \Delta_k \rangle$$

so, it order to show that $||\tilde{x}_{k+1} - \bar{x}|| < ||\tilde{x}_k - \bar{x}||$, it is sufficient to demonstrate

$$||\Delta_k||^2 + 2\langle \tilde{x}_k - \bar{x}, \Delta_k \rangle < 0. \tag{51}$$

From (49, 50),

$$\begin{aligned} \Delta_k &= \left(I - \alpha^k \frac{d_k d_k^T}{d_k^T d_k} \right) \tilde{x}_k + \frac{\alpha_k b_k}{d_k^T d_k} d_k - \tilde{x}_k \\ &= \frac{\alpha}{\langle d_k, d_k \rangle} \left(b_k - \langle d_k, \tilde{x}_k \rangle \right) d_k \end{aligned} \tag{52}$$

so, substituting Δ_k into (51) (and dropping the superscript indices k),

$$\frac{\alpha^2}{\langle d, d \rangle} \left(b - \langle d, \tilde{x} \rangle \right)^2 + 2 \frac{\alpha}{\langle d, d \rangle} \left(b - \langle d, \tilde{x} \rangle \right) \langle d, \tilde{x} - \bar{x} \rangle < 0$$

or equivalently, so long as $\alpha > 0$ (as we require),

$$-\left(\langle d, \tilde{x}\rangle - b\right)\left[\alpha\left(b - \langle d, \tilde{x}\rangle\right) + 2\langle d, \tilde{x} - \bar{x}\rangle\right] < 0. \tag{53}$$

Since by assumption $\langle d, \tilde{x}\rangle - b > 0$, this is satisfied iff the second factor is positive; that is,

$$\begin{aligned}\alpha\left(b - \langle d, \tilde{x}\rangle\right) + 2\langle d, \tilde{x} - \bar{x}\rangle = \\ \alpha b + (2-\alpha)\langle d, \tilde{x}\rangle - 2\langle d, \bar{x}\rangle > 0.\end{aligned} \tag{54}$$

Since $\langle d, \tilde{x}\rangle > b$, and by Lemma 1, $\langle d, \bar{x}\rangle \leq b$, this is satisfied so long as $\alpha \in (0, 2)$, as we require.

2. *The sequence* $e_k = ||\tilde{x}_k - \bar{x}_k||,\ k = 0, 1, 2, ...$ *is nonincreasing.* In the second case of (49), $\tilde{x}_{k+1} = \tilde{x}_k$; this is nonincreasing. In the first case, $\langle d_k, \tilde{x}_k\rangle - b_k > 0$, so $e_{k+1} < e_k$ by point 1 above.

3. g.l.b.$(e_k) = 0$ *with unit probability.* By positivity of $||\cdot||$, zero is a lower bound. To show that this is the greatest such bound, consider some $\epsilon > 0$ and suppose that, at iteration m, $||\tilde{x}_m - \bar{x}|| = \epsilon$. Now, let $z = \min(r, \epsilon/2)$, and consider the open balls $B_1 = B(c_1, z/4), B_2 = B(c_2, z/4)$, where the center points c_1, c_2 are defined,

$$c_j = \bar{x} + \frac{\tilde{x} - \bar{x}}{||\tilde{x} - \bar{x}||}\frac{(2j-1)}{4}z$$

for $j \in \{1, 2\}$ (see Figure 5); additionally, let $x^1 \in B_1, x^2 \in B_2$. Then by Lemma 1, we can confirm that $\bar{x}$ and $\tilde{x}$ are on opposite sides of the plane corresponding to (x^1, x^2) (and hence, that a projection will occur) by verifying that,

$$||x^2 - \tilde{x}|| < ||x^1 - \tilde{x}|| \tag{55}$$

$$||x^2 - \bar{x}|| > ||x^1 - \bar{x}||. \tag{56}$$

Considering the first of these, we note by the triangle inequality,

$$||x^2 - \tilde{x}|| \leq ||x^2 - c_2|| + ||c_2 - \tilde{x}|| < \frac{1}{4}z + ||c_2 - \tilde{x}||$$

whereas, by the inverse triangle inequality,

$$\begin{aligned}||x^1 - \tilde{x}|| &\geq |\ ||x^1 - c_1|| + ||c_1 - \tilde{x}||\ | \\ &\geq ||c_1 - \tilde{x}|| = \frac{1}{2}z + ||x^2 - c_2||\end{aligned}$$

so this is indeed the case. Considering the second inequality (56), we have likewise,

$$||x^1 - \bar{x}|| \leq ||x^1 - c_1|| + ||c_1 - \bar{x}|| < \frac{1}{4}z + \frac{1}{4}z = \frac{1}{2}z$$

and

$$||x^2 - \bar{x}|| \geq |\ ||x^2 - c_2|| - ||c_2 - \bar{x}||\ | \geq \frac{3}{4}z$$

so this inequality holds as well. Therefore, *any* x^1, x^2 from B_1, B_2 are associated with a plane that separates $\tilde{x}$ from $\bar{x}$ and hence triggers a projection. Since B_1 and B_2 have nonzero measure, and are subsets of B_r in which $p(\cdot)$ is nonzero, then the probabilities for this iteration $P_1 = \Pr(\text{“a point is selected in } B_1\text{”})$ and $P_2 = \Pr(\text{“a point is selected in } B_2\text{”})$ are both nonzero, and therefore, since the s^k are independent, $P_{\text{both}} = \Pr(\text{“one point is selected in } B_1 \text{ and the other is selected in } B_2\text{”}) = P_1 P_2$ is nonzero, and the probability that this occurs for *at least* one iteration $k > m$ is given by $1 - \prod_{k=m}^{\infty} \left(1 - P_{\text{both}}^k\right) = 1$ or in other words, with probability one, there exists a $q > m$ such that $\langle d_q, \tilde{x}_q \rangle - b_q > 0$. Then, by point 1, $||\tilde{x}_q - \bar{x}|| < ||\tilde{x}_m - \bar{x}|| = \epsilon$, and so ϵ, with unit probability, cannot be a lower bound. Since e_k is a nonincreasing sequence in $\mathbb{R}$ and g.l.b.$(e_k) = 0$, e_k converges to 0 and thus $\tilde{x}$ converges to $\bar{x}$ in probability.

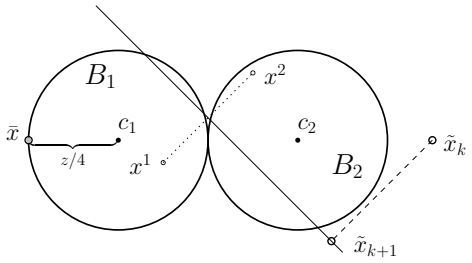

Fig. 5. If $x^1 \in B_1$ and $x^2 \in B_2$, then $||\tilde{x}_{k+1} - \bar{x}|| < ||\tilde{x}_k - \bar{x}||$.

An example of the estimate trajectory in feature space generated by such an observer is given in Figure 6. For this example, $X = \mathbb{R}^2$, and features were drawn from a uniform distribution in the square $[-20, 20] \times [-20, 20]$. The estimate evolves from its initial condition, $\tilde{x}_0 = (-15, 15)$ to near the ideal $\bar{x} = (17, 0)$.

7 Applications

7.1 Apples and Oranges

To demonstrate the application of the metric preference learning formulation, photos of nine apples were shown to an audience of thirteen people in a number of pairwise experiments. (The fruit is shown in Figure 7.)

Each apple was described by a 15-dimensional feature vector, containing (1-3) the average color in HSB (hue, saturation, brightness) color space, (4-6) the average color in RGB color space, (7) the color variance, (8-10) width, height, and the ratio of the two, (11-12) stem length, and angle relative to apple, (13-14) dimple angle and depth, and (15) roundness. This represents a collection of many conceivable characteristics of an apple that may make it more or less

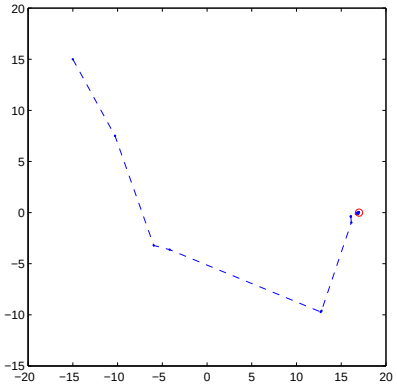

Fig. 6. Example estimate trajectory for observer (49-50) for $\alpha^k = \alpha = 1$, with $X = \mathbb{R}^2$. The estimate begins at $\tilde{x}_0 = (-15, 15)$, and approaches the ideal $\bar{x} = (17, 0)$.

Fig. 7. Depicted are the 9 apples used to generate comparisons with the single orange.

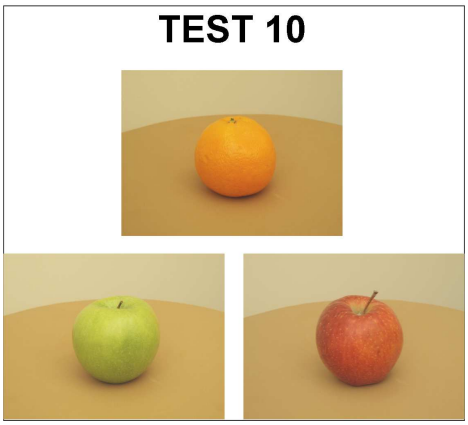

Fig. 8. An example of a pairwise comparison between two apples, relative to the orange.

orangelike, and the idea is to learn a cost function without making assumptions a priori about which of these features are most salient.

The partial order over the apples was thus generated by having a number of people make a number of randomly selected, pairwise comparisons (as the one depicted in Figure 8). Represented as a preference graph, the results of these experiments are given as Figure 9.

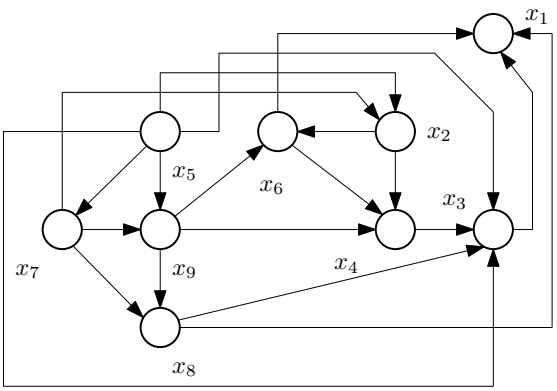

Fig. 9. The preference graph corresponding to the apple experiments.

For these data, the minimization problem (21) is unbounded and hence we find an optimal direction via (19). Solving (40), we obtain the optimum,

$\bar{v}_1 = -0.0252$	(Hue)	$\bar{v}_9 = -0.2380$	(Height)
$\bar{v}_2 = -0.0844$	(Saturation)	$\bar{v}_{10} = -0.0472$	(Width/Height)
$\bar{v}_3 = 0.1374$	(Brightness)	$\bar{v}_{11} = -0.0409$	(Stem Length)
$\bar{v}_4 = 0.3572$	(Red)	$\bar{v}_{12} = -0.5017$	(Stem Angle)
$\bar{v}_5 = 0.1137$	(Green)	$\bar{v}_{13} = 0.6683$	(Dimple Angle)
$\bar{v}_6 = 0.1856$	(Blue)	$\bar{v}_{14} = 0.0996$	(Dimple Depth)
$\bar{v}_7 = 0.0442$	(Variance)	$\bar{v}_{15} = -0.0472$	(Roundness)
$\bar{v}_8 = -0.1593$	(Width)		

which has the interpretation that dimple angle and redness are important orangelike qualities, and that large stem angles are perceived as un-orangelike.

7.2 Amoebas and Humans

To understand the comparison of higher-dimensional objects and in particular motions, another experiment was performed in which an audience of 25 people was asked to perform pairwise comparisons of different motions of a computer-animated amoeba, relative to the motion-captured movement of a human who danced the bhangra. An example of one such question is illustrated in Figure 11. In this manner, a preference graph was generated as before, with 12 vertices (the amoeba motions) and 20 edges; this is shown in Figure 10.

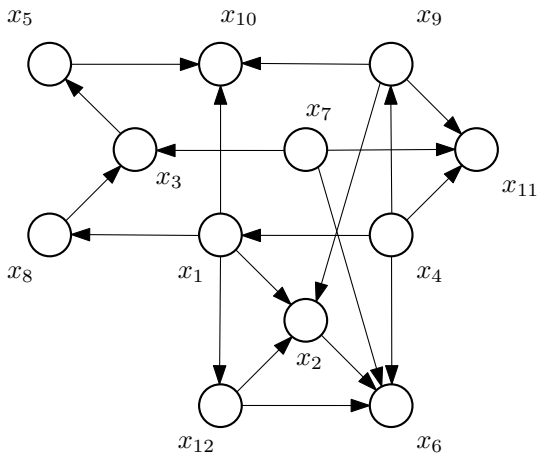

Fig. 10. The DAG corresponding to the amoeba experiments.

Inner products between the various amoeba motions were computed by rasterizing the motions to binary videos, blurring each frame of the result, and computing the standard Euclidean inner product of these (extremely large) [Frame

`Width`]×[`Frame Height`]×[`Number of Frames`]-dimensional vectors. The resulting inner product is a relaxation of set overlap area, and its corresponding metric returns small distances between videos in which mass is located at nearly, but not necessarily exactly, the same locations at the same times. We note that the sheer size of the rasterized representation highlights the advantage of the instance vector expansion described in section 5.1, without which the optimization problem simply could not be realistically solved.

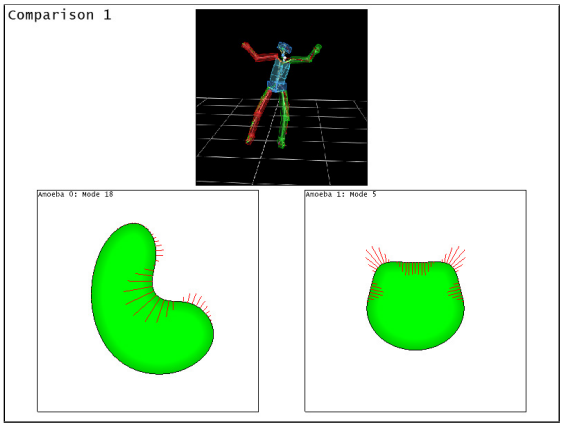

Fig. 11. Each question took the form, "Which of the two 'amoebas' (bottom) looks more like the [motion capture data from a] human dancer (top)?"

The minimization problem (21) with the resulting data turns out to be unbounded and hence we again find an optimal direction via (19). We obtain the coefficient expansion for the optimal direction,

$$\bar{v} = \sum_{k=1}^{M} \xi_k x_k \tag{57}$$

where

$$\begin{aligned} \xi = 10^3(\ & 1.4918, -3.6556, -0.1390, 0.3113, \\ & -1.1243, -0.1771, \ \ 2.6335, 0.5878, \\ & \ 1.8362, -1.7319, -0.2999, 0.2672 \ ) \ . \end{aligned}$$

What this means is that, in order to look as much like it is dancing the bhangra as possible, an amoeba should as its first priority aspire to be as much like amoeba 7 ($\xi_7 = 2.6335$) and as dissimilar from amoeba 2 ($\xi_2 = -3.6556$) as possible, and that it should to a lesser extent model itself after amoebas 1 and 9 ($\xi_1 = 1.4918, \xi_9 = 1.8362$) while avoiding the aesthetically unappealing moves of amoebas 5 and 10 ($\xi_5 = -1.1243, \xi_{10} = -1.7319$). Although this does not explain why, psychologically, e.g. amoeba 7 – which pulses with two upward-pointing, armlike protrusions – is preferred to amoeba 2 – in which a bulge on

one side moves in and out – it does produce both a consistent cost structure, and an estimate for an amoeba motion that will be preferred to all others in the larger space of motions.

8 Concluding Remarks

In this chapter, we investigated the problem of motion preference learning under the assumption of an underlying *metric cost model*; here, the alternatives being compared are points in a metric space, and human judges are assumed to prefer one point to another if and only if it is closer to some fixed but unknown best alternative that they may not have been shown. This assumption appears to be a good one for the examples considered and the features chosen, in that the feasible set P in this case is nonempty.

Based on the metric cost assumption, a Chebyshev estimator was given for the best point for the case when P is bounded, and a natural generalization, the minimax-rate estimator, was developed for when P is unbounded. In the first case, the solution was found, with an efficiency rivaling standard quadratic SVMs, as the solution to a linear program; and in the second case the problem was shown to in fact reduce to a particular SVM classification problem.

In order that the estimators for the bounded and unbounded cases be applicable to situations in which the compared alternatives inhabit high- or infinite-dimensional metric spaces, the optimization problems were additionally given in an instance vector expansion form, which results in optimization problems whose size is proportional not to the dimensionality of the metric space, but only to the number of comparisons available. This is particularly relevant in a controls context, when the alternatives being compared are the signals input to or produced by continuous-time systems, which inhabit infinite-dimensional function spaces.

Finally, for the case when a large amount of data are available, or stream- rather than batch- processing is desired, a limited-memory asymptotic observer was given that avoids the need to store all the (infinitely) many constraints.

In all cases, optimal cost functions *and* points/directions were found efficiently by convex programming. The result is an efficient minimax estimator for the best possible alternative.

Acknowledgments

This work was supported by the U.S. National Science Foundation through Creative IT Grant #0757317. The human-study was conducted within the Georgia Institute of Technology, Institute Review Board Protocol H08162 - "Perceived Similarity Study." We would like to thank Akhil Bahl for his assistance in producing the motion capture data used in the synthetic amoeba experiment.

References

1. P. Abbeel and A.Y. Ng. Apprenticeship learning via inverse reinforcement learning. *Proceedings of the International Conference on Machine Learning*, 2004.
2. S. Agarwal. Ranking on graph data. *ICML*, ACM Press, pp. 25–36, 2006.
3. A.V. Aho, M.R. Garey, and J.D. Ullman. The transitive reduction of a directed graph. *SIAM Journal on Computing*, Vol. 1, pp. 131–137, 1972.
4. F. Aiolli and A. Sperduti. Learning preferences for multiclass problems. *Advances in Neural Information Processing Systems 17*, MIT Press, pp.17–24, 2004.
5. M. Aizerman, E. Braverman, and L. Rozonoer, Theoretical foundations of the potential function method in pattern recognition learning, *Automation and Remote Control*, Vol. 25, pp. 821–837, 1964.
6. A. Bahamonde, J. Diez, J. Quevedo, O. Luances, and J. del Coz. How to learn consumer preferences from the analysis of sensory data by means of support vector machines. *Trends in Food Science and Technology*, Vol. 18, pp. 20–28, 2007.
7. B. Baird. *The Art of the Puppet.* Mcmillan Company, New York, 1965.
8. J. Black, E. Iyasere, and J. Wagner. Creation of a driver preference objective metric to evaluate ground vehicle steering systems. *American Control Conference*, 2011.
9. B. Boser, I. Guyon, and V. Vapnik. An training algorithm for optimal margin classi ers. *Fifth Annual Workshop on Computational Learning Theory*, pp.144–152, 1992.
10. P. Callahan. *Puppets, Stop the Flap!: The Arpeggio Method of Mouth Puppet Manipulation.* Arpeggio, 1994.
11. W. Chu and Z. Ghahramani. Preference learning with gaussian processes. *Proceedings of the 22nd international conference on machine learning*, New York, NY, pp. 137–144, 2005.
12. W. Chu and S.S. Keerthi. Support vector ordinal regression. *Neural Computation*, 2007.
13. W.W. Cohen, R.E. Schapire, and Y. Singer. Learning to order things. *Journal of Artificial Intelligence Research*, Vol. 10, pp. 243–270, 1999.
14. D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *STOC '87: Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pp.1–6, New York, NY, USA, 1987.
15. C. Cortes, M. Mohri, and A. Rastogi. Magnitude-preserving ranking algorithms. *Proceedings of the Twenty-fourth International Conference on Machine Learning*, 2007.
16. C. Canudas de Wit. Fun-to-drive by feedback. *European Journal of Control: Fundamental issues in Control*, 2005.
17. J. Diez, J. Jose del Coz, O. Luaces, and A. Bahamonde. Clustering people according to their preference criteria. *Expert Syst. Appl.*, Vol. 34, pp. 1274–1284, 2008.
18. M. Egerstedt, T. Murphey, and J. Ludwig. Motion Programs for Puppet Choreography and Control. *Springer-Verlag*, pp. 190–202, April 2007.
19. L. Engler and C. Fijan. *Making Puppets Come Alive.* Taplinger Publishing Company, New York, 1973.
20. C.N. Fiechter and S. Rogers. Learning subjective functions with large margins. *Proceedings of the Seventeenth International Conference on Machine Learning*, pp. 287–294, 2000.
21. M.J. Fischer and A.R. Meyer. Boolean matrix multiplication and transitive closure. *Twelfth Annual Symposium on Switching and Automata Theory*, pp. 129–131, 1971.
22. R. Herbrich, T. Graepel, P. Bollmann-Sdorra, and K. Obermayer. Supervised learning of preference relations. *Proceedings FGML-98, German National Workshop on Machine Learning*, pp. 43–47, 1998.

23. R. Herbrich, T. Graepel, and K. Obermayer. Large margin rank boundaries for ordinal regression. *Advances in large margin classifiers*, MIT Press, pp. 115–132, 2000.
24. E. Hullermeier, J. Frankranz, W. Cheng, and K. Brinker. Label ranking by learning pair-wise preferences. *Artificial Intelligence*, Vol. 172, pp. 1897–1917, 2008.
25. X. Jiang, L.-H. Lim, Y. Yao, and Y. Ye. Statistical ranking and combinatorial hodge theory. *Mathematical Programming*, Vol. 127, pp. 203–244, 2011.
26. T. Joachims. Optimizing search engines using click-through data. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 133–142, 2002.
27. E. Johnson and T. Murphey. Dynamic modeling and motion planning for marionettes: Rigid bodies articulated by massless strings. *IEEE International Conference on Robotics and Automation*, pp. 330–335, 2007.
28. P. Kingston and M. Egerstedt. Comparing apples and oranges through partial orders: An empirical approach, *American Control Conference*, 2009.
29. P. Kingston and M. Egerstedt. Motion preference learning. *American Control Conference*, 2011.
30. I. Munro. *Efficient determination of the strongly connected components and the transitive closure of a graph.* 1971.
31. J.C. Platt. Sequential minimal optimization: A fast algorithm for training support vector machines. *Advances in kernel methods*, MIT Press, 1998.
32. Q. Qiao and P.A. Beling. Inverse reinforcement learning with gaussian process. *American Control Conference*, 2011.
33. A. Roser. *Gustaf Und Sein Ensemble: Beschreibungen eines Puppenspielers.* Bleicher Verlag, Gerlingen, Germany, 1979.
34. V. Strassen. Gaussian elimination is not optimal. *Numer. Math*, Vol. 13, pp. 354–356, 1969.
35. U. Syed, M. Bowling, and R.E. Schapire. Apprenticeship learning using linear programming. *Proceedings of the International Conference on Machine Learning*, 2008.