Axiomatic Characterisation of Projection Rules: An Open Question

P. N. Karthik
Department of ECE,
Indian Institute of Science,
Bangalore - 560012
Email: periyapatna@iisc.ac.in

Rajesh Sundaresan
Department of ECE and
Robert Bosch Centre for Cyber Physical Systems,
Indian Institute of Science,
Bangalore - 560012
Email: rajeshs@iisc.ac.in

***Abstract*—This paper studies the close interplay between two commonly used approaches for describing projection rules: (a) function-minimisation approach, and (b) axiomatic approach. As one of the findings of this study, this paper reports an interesting connection that the topic of conservative vector fields has with the axiomatic approach for describing projection rules. This leads to a question on conservative vector fields that is of independent mathematical interest. While answers to this question are known only in a few instances, a general solution to this question is currently not available. Contrary to papers that present concrete results in connection with a selected problem, the main purpose of this paper is to bring to light the above mathematical question.**

I. Introduction

In this paper, we study the classical problem of recovering an unknown vector from a set of linear measurements. Formally, suppose that $x \in \mathbb{R}^n$ satisfies a set of k linear constraints of the form $a_i^T x = b_i$ for $i = 1, 2, \ldots, k$, where for each i, $a_i \in \mathbb{R}^n$, $a_i \neq 0$, and $b_i \in \mathbb{R}$. The goal is to recover x from the set

$$L = \{w \in \mathbb{R}^n : Aw = b\}, \tag{1}$$

where A denotes the $(k \times n)$ matrix whose i^{th} row is a_i, and b denotes the vector whose i^{th} component is b_i for $i = 1, 2, \ldots, k$. The typical approach towards solving this problem involves defining a cost function $F : \mathbb{R}^n \to \mathbb{R}$, and picking x as any vector satisfying $x \in \arg\min_{w \in L} F(w)$. For example, in the method of least squares, $F(w) = ||w||$, the 2-norm of w, in which case x is the vector in L with the least 2-norm. In the method of maximum entropy proposed by Jaynes [1], x is chosen as the vector in L having nonnegative components that sum up to 1, and having the largest Shannon entropy. Thus, given a function F, for any set L of the form (1), the minimisation (or maximisation) of F over L prescribes a rule for selecting an element from L (in the above examples, this element is the unknown vector x satisfying the constraints in L). We refer to such rules as projection rules.

It is then natural to ask if every projection rule may be obtained by the minimisation (or maximisation) of some function F. Answers to this question were obtained by Csiszár in [2], wherein a series of axioms is first presented, following which it is shown that for any projection rule satisfying one or more of such axioms, there exists a function F whose minimisation describes the projection rule. Therefore, we have at hand two approaches for describing projection rules: (a) function-minimisation approach, and (b) axiomatic approach as in [2]. Our interest is in understanding the relation between these two approaches. For this, throughout this paper, we study questions such as, "Given a function whose minimisation describes a projection rule, what axioms does this projection rule satisfy?", or, "Given a set of axioms for a projection rule, under what conditions is the projection rule also described by the minimisation of some function?", and seek answers to these questions on the same lines as [2].

Of the many axioms appearing in [2], those that are of interest to us are 'regularity', 'locality' and 'subspace transitivity', appearing in [2, Defs. 2-3, Def. 6(i)]. In relation to the questions posed in the previous paragraph, an important result related to the aforementioned axioms is [2, Theorem 3], which states that any regular, local and subspace transitive projection rule may be described by the minimisation of some Bregman's divergence [3]. On this note, we now bring to the reader's attention a recent work of Kumar and Sundaresan that studies projections with respect to a class of divergences known as relative α-entropy, that form a generalisation of KL divergence; here, $\alpha > 0$, and $\alpha \neq 1$ (see [4, Definition 1]). The relevance of their work to this paper arises from a key property of relative α-entropy identified in [4, Section VI], which is that when $\alpha < 1$, the projection rule described by the minimisation of relative α-entropy over any set of n-length probability vectors satisfies the axioms of regularity and subspace transitivity, but not locality.

Thus, while there appears in [2] a full axiomatic characterisation of all regular, local and subspace transitive projection rules, a similar characterisation for the class of all regular and subspace transitive projections that are non-local, which contains the family of relative α-entropies for all $\alpha < 1$, remains open. It is the characterisation of this class of projection rules that is the main subject of this paper. Towards this, our approach is to bring out the role played by the axiom of regularity, motivated by the in-depth exposition of the axioms of regularity and locality appearing in [2] that subsequently simplifies the presentation of the results therein. Therefore, the remainder of this paper deals with only the axiom of regularity.

As a first step in our study of the axiom of regularity, we discover an interesting connection with the topic of conservative vector fields (or vector-valued mappings), which we bring out in detail in Section III-A. Next, we identify a set of necessary conditions that must hold for any function whose minimisation describes a regular projection rule. With the help of a criterion from [2], we show that a regular projection rule may be described by the minimisation of a function which satisfies the necessary conditions along with the aforementioned criterion. Lastly, we look for conditions under which, given a regular projection, there exists a function whose minimisation describes the projection rule. We show that this leads us to the following question that is of independent mathematical interest: Given a continuous vector field $v : \mathbb{R}^n \to \mathbb{R}^n$, does there exist a nonzero, continuous scalar function $\lambda : \mathbb{R}^n \to \mathbb{R}$ such that the scaled vector field

$$u(y) = \lambda(y) \cdot v(y), \quad y \in \mathbb{R}^n, \tag{2}$$

is conservative, i.e., $u(y)$ may be expressed in the form $u(y) = \nabla G(y)$ for some continuously differentiable function G? While answers to the above question of conservative fields are known in a few instances, a general solution to the same question is not available currently.

Finally, under additional assumptions that $\lambda(\cdot)$ and $v(\cdot)$ in (2) are differentiable, we present a formulation of the above question on conservative vector fields in terms of a first order partial differential equation, a solution to which too is not available currently. We hope that this paper aids in eliciting ideas or approaches that provide answers to the above question.

In Section II, we set up the notations and some preliminary results from [2]. We then present, in Section III, our findings on the role of the axiom of regularity.

II. Notations and Preliminaries

Throughout, we fix $n \geq 3$ as in [2], and carry out all vector operations in $\mathbb{R}^n$. We treat all vectors as column vectors. We denote by $\mathscr{L}$ the collection of all linear (or affine) subsets of $\mathbb{R}^n$ of the form in (1). That is,

$$\mathscr{L} = \Big\{ L \subset \mathbb{R}^n : \ L \text{ is of the form (1) for some } k = 0, \ldots, n, \\ \text{for some } (k \times n) \text{ matrix } A, \text{ and some vector } b \in \mathbb{R}^k \Big\}, \tag{3}$$

where $k = 0$ corresponds to the case $L = \mathbb{R}^n$. Since any L of the form in (1) with matrix A having rank k is described by k linearly independent constraints, we say that such a set L has dimension equal to $(n-k)$.

Among all the sets in $\mathscr{L}$, we let $\mathscr{M}$ denote the collection of sets whose dimension is equal to $(n-1)$, i.e.,

$$\mathscr{M} = \Big\{ \{w \in \mathbb{R}^n : a^T w = b\} : a \in \mathbb{R}^n, \ a \neq 0, \ b \in \mathbb{R} \Big\}. \tag{4}$$

We now present a formal definition of a projection rule as given in [2].

Definition 1: (Projection rule) A *projection rule* Π is a family of rules indexed by $x \in \mathbb{R}^n$, and denoted by

$$\Pi = \{\Pi(\cdot|x) : x \in \mathbb{R}^n\},$$

where for each x, $\Pi(\cdot|x) : \mathscr{L} \to \mathbb{R}^n$ is a mapping that selects for each $L \in \mathscr{L}$ an element $\Pi(L|x) \in L$. Additionally, if $x \in L$, then $\Pi(L|x) = x$.

The name 'projection' rule is used to convey the idea that

- $x \in \mathbb{R}^n$ is some prior guess,
- L is the set of possibilities consistent with some observations, that is, the ground truth $w \in \mathbb{R}^n$ when projected on a_i yields b_i for $i = 1, 2, \ldots, k$, where a_i and b_i constitute the i^{th} row of A and b respectively in the representation (1) for L,
- $\Pi(L|x)$ is the selection from L for the prior guess x, which is also consistent with the observations, and
- if the prior guess is itself in L, then the selection from L is simply the prior guess.

Thus, $\Pi(L|x)$ could be viewed as the 'projection' of the vector x onto the set L.

Often, projection rules stem from some minimisation principle. This is made precise in the following definition.

Definition 2: Given a projection rule Π and a function $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ written as $(x, y) \mapsto F(y|x)$, we say that F *generates* Π if, for each $x \in \mathbb{R}^n$, the following conditions hold:

1) The point x is the global minimum of the function $F(\cdot|x) : \mathbb{R}^n \to \mathbb{R}$.
2) The function $F(\cdot|x) : \mathbb{R}^n \to \mathbb{R}$ admits a unique minimum over each $L \in \mathscr{L}$.
3) For each $L \in \mathscr{L}$, if $x \in \mathbb{R}^n$ is the prior guess, then

$$\Pi(L|x) = \arg\min_{y \in L} F(y|x). \tag{5}$$

That is, for each $L \in \mathscr{L}$, the projection of x onto set L corresponds to the unique minimum of $F(\cdot|x)$ over L.

Without loss of generality, we may consider $F(x|x) = 0$ for each $x \in \mathbb{R}^n$, since the function $\tilde{F}$ defined as $\tilde{F}(y|x) = F(y|x) - F(x|x)$ also generates the same projection rule as that of F. Some example functions F considered in the literature are given below (see [2], [3] and the references therein):

1) The least squares projection rule is the one generated by the function $F(y|x) = \sum_{i=1}^{n} (x_i - y_i)^2$.
2) The I-divergence projection rule is the one generated by the function $F(y|x) = \sum_{i=1}^{n} y_i \log \frac{y_i}{x_i}$, where x, y are constrained to be n-length probability vectors. The principle of maximum entropy [1] may then be obtained as a special case of the I-divergence projection rule when $x = (1/n, \ldots, 1/n)^T$.
3) Kumar and Sundaresan [4] study projections with respect to the function $F(y|x) = \mathscr{I}_\alpha(y, x)$, where $\mathscr{I}_\alpha$ is a parametric extension of relative entropy, and is defined in [4, Section II, Eqn. (6)].

We now present the definition of the axiom of regularity for any projection rule Π, as appearing in [2].

Definition 3: A projection rule $\Pi = \{\Pi(\cdot|x) : x \in \mathbb{R}^n\}$ is said to be *regular* if for each $x \in \mathbb{R}^n$ the following axioms are satisfied:

1) Consistency: For any $L, L' \in \mathscr{L}$ such that $L' \subset L$, if $\Pi(L|x) \in L'$, then $\Pi(L'|x) = \Pi(L|x)$.
2) Distinctness: For any $L, \bar{L} \in \mathscr{M}$, the implication
$$L \neq \bar{L} \implies \Pi(L|x) \neq \Pi(\bar{L}|x) \tag{6}$$
holds, unless both L and $\bar{L}$ contain the vector x.
3) Continuity: The restriction of Π to the collection of all sets of any fixed dimension is continuous.

The axiom of consistency formalises the idea that if $\Pi(L|x)$, selected on the basis of the constraints in L, also satisfies the additional constraints in L', then the additional constraints provide no reason for changing the original selection. The axioms of distinctness and continuity are self-explanatory.

We now note an important consequence of a projection rule Π being regular. This is based on the results of [2, Lemmas 1-2]: for every $x, y \in \mathbb{R}^n$ such that $y \neq x$, there exists a unique set $L = L(y|x) \in \mathscr{M}$ of dimension $(n-1)$ with normal $a = a(y|x) \in \mathbb{R}^n \setminus \{0\}$ such that $\Pi(L|x) = y$. From (4), $L(y|x)$ may be expressed as
$$L(y|x) = \{w \in \mathbb{R}^n : a(y|x)^T(w-y) = 0\} \tag{7}$$
since b in (4) is given by $a(y|x)^T y$ due to the fact that $y \in L(y|x)$. Since $L(y|x)$ is completely specified by its normal vector $a(y|x)$ as in (7), for each $x \in \mathbb{R}^n$, we then have the mapping $y \mapsto a(y|x)$, $y \neq x$.

We now present a lemma which shows that the above identified mapping $y \mapsto a(y|x)$, $y \neq x$, is continuous. We omit the proof of the lemma since it is a straightforward extension of the proof of [2, Lemma 2].

Lemma 1: Let Π be a regular projection rule. Fix $x \in \mathbb{R}^n$ and let $\mathscr{L}^0(x)$ denote the collection
$$\mathscr{L}^0(x) = \left\{L \subset \mathbb{R}^n : \Pi(L|x) = x\right\}. \tag{8}$$
Then, for each $x \in \mathbb{R}^n$, the restriction of $\Pi(\cdot|x)$ to $\mathscr{M} \setminus \mathscr{L}^0(x)$ is a homeomorphism onto the set $\mathbb{R}^n \setminus \{x\}$. Furthermore, the mapping $y \mapsto a(y|x)$, $y \neq x$, is continuous.

Henceforth, we shall only use the continuity of the mapping $y \mapsto a(y|x)$, $y \neq x$.

Remark 1: Note that while $a(y|x)$ is specified for each $y \neq x$, $a(x|x)$ has not been defined. Let us define $a(x|x)$ as the vector
$$a(x|x) := \lim_{\substack{y \to x \\ y \neq x}} a(y|x), \tag{9}$$
if the limit exists. If not, by suitable rescaling such as $a(y|x) \leftarrow a(y|x)||y-x||$, we can ensure that the definition in (9) holds. In this rescaled case, $a(x|x)$ will be 0. This defines, for each $x \in \mathbb{R}^n$, the mapping $y \mapsto a(y|x)$ for all y, thus leading to $a(\cdot|x)$ being a vector field, i.e., a mapping from $\mathbb{R}^n \to \mathbb{R}^n$. Further, Lemma 1 and (9) together imply that for each $x \in \mathbb{R}^n$, the vector field $a(\cdot|x)$ is continuous. We therefore have the family of continuous vector fields $\{a(\cdot|x) : x \in \mathbb{R}^n\}$, indexed by $x \in \mathbb{R}^n$.

We end this section with some definitions on vector fields.

Definition 4: A vector field $v : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *conservative* if there exists a differentiable function $V : \mathbb{R}^n \to \mathbb{R}$ such that $v(y) = \nabla V(y)$ for all $y \in \mathbb{R}^n$. The function V may be viewed as a potential function.

As an example, we note that the vector field given by $v(y) = y$ is conservative since $v(y) = \nabla V(y)$, with $V(y) = \frac{1}{2}||y||^2$.

Definition 5: Given a vector field $v : \mathbb{R}^n \to \mathbb{R}^n$ and a function $\lambda : \mathbb{R}^n \to \mathbb{R}$, the *$\lambda$-scaling* of v is defined as the vector field $u : \mathbb{R}^n \to \mathbb{R}^n$ given by
$$u(y) = \lambda(y)\, v(y), \quad y \in \mathbb{R}^n. \tag{10}$$

As outlined in Section I, our interest is in obtaining axiomatic characterisation of a larger class of projections rules than that considered by Csiszár in [2]. In this direction, noting that regularity is the most fundamental axiom considered in [2], we wish to identify some of the key implications of having regularity as an axiom. More precisely, we seek answers for the following questions.

1) What are the implications of the axiom of regularity alone?
2) Does every regular projection rule stem from the minimisation of some function?

The next section describes our attempts towards obtaining answers to the above questions.

III. The Role of Regularity

In this section, we identify an important connection between the axiom of regularity and that of conservative vector fields. Then, we list a set of necessary conditions for any function to generate a regular projection rule. With an additional criterion from [2], we show that any function that satisfies the above necessary conditions and the added criterion generates a regular projection rule. We end the section with a discussion on the fundamental question of whether, given a regular projection rule, there exists a function that generates the projection rule.

A. Regularity and Conservative Vector Fields

Suppose that $F = \{F(y|x) : x, y \in \mathbb{R}^n\}$ is continuously differentiable, and generates a projection rule Π. Then, it immediately follows that F satisfies the conditions 1-3 given in Def. 2. Suppose, in addition, that Π is regular. This then defines the family of vector fields $\{a(\cdot|x) : x \in \mathbb{R}^n\}$, where in addition, the vector field $a(\cdot|x)$ is continuous for each $x \in \mathbb{R}^n$. The purpose of this section is to demonstrate that for each $x \in \mathbb{R}^n$, there exists a scaling of the vector field $a(\cdot|x)$ which makes it conservative, with the potential function given by $F(\cdot|x)$.

Let us fix $x \in \mathbb{R}^n$. Then, for each $y \neq x$, since (5) holds, it follows by the theory of Lagrange multipliers that the relation
$$\nabla F(y|x) = \lambda(y|x)\, a(y|x), \quad y \neq x, \tag{11}$$

holds for some $\lambda(y|x) \in \mathbb{R}$ which may possibly be zero. This specifies the Lagrange multipliers $\lambda(y|x)$ for all $y \neq x$, thus leaving us to specify $\lambda(x|x)$. For this, we note that $\nabla F(x|x) = 0$ since F is continuously differentiable, and x is the global minimum of $F(\cdot|x)$. Further, from Remark 1, we have that $a(x|x)$ is well defined. If $a(x|x) \neq 0$, then we define $\lambda(x|x)$ as

$$\lambda(x|x) = \lim_{\substack{y \to x \\ y \neq x}} \frac{||\nabla F(y|x)||}{||a(y|x)||} = \frac{||\nabla F(x|x)||}{||a(x|x)||} = 0. \tag{12}$$

Else, if $a(x|x) = 0$, we set $\lambda(x|x)$ as

$$\lambda(x|x) = \begin{cases} \lim\limits_{\substack{y \to x \\ y \neq x}} \lambda(y|x), & \text{if this limit exists,} \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

Then, (12) and (13) imply that the mapping $\lambda(\cdot|x) : \mathbb{R}^n \to \mathbb{R}$ is well defined for each $x \in \mathbb{R}^n$. Further, this mapping is continuous everywhere, except possibly at x. Finally, (12) and (13) imply that (11) holds when $y = x$, thus proving that for each $x \in \mathbb{R}^n$, the $\lambda(\cdot|x)$-scaling of $a(\cdot|x)$ is a conservative vector field, with potential function $F(\cdot|x)$.

B. *Necessary Conditions for a Function to Generate a Regular Projection Rule*

We now build on the developments in the previous sections, and list the conditions that must necessarily be satisfied by any function F that generates a regular projection rule. These conditions are given in the following proposition.

Proposition 1: Let $\Pi : \mathscr{L} \to \mathbb{R}^n$ be a regular projection rule defining the family of vector fields $\{a(\cdot|x) : x \in \mathbb{R}^n\}$. Suppose that $F = \{F(y|x) : x, y \in \mathbb{R}^n\}$ is continuously differentiable and generates Π. Then, for each $x \in \mathbb{R}^n$, the following conditions hold.

1) $F(\cdot|x)$ satisfies the conditions in items 1 and 2 of Definition 2.
2) For each $y \neq x$, there exists a unique set $L(y|x) \in \mathscr{M}$ with normal $a(y|x)$ such that y is the unique minimiser of $F(\cdot|x)$ over $L(y|x)$, i.e.,

$$y = \arg \min_{w \in L(y|x)} F(w|x). \tag{14}$$

3) The vector field $a(\cdot|x)$ is continuous.
4) There exists a function $\lambda(\cdot|x) : \mathbb{R}^n \to \mathbb{R}$ such that the $\lambda(\cdot|x)$-scaling of the vector field $a(\cdot|x)$ is conservative, and $\lambda(\cdot|x)$ is continuous everywhere, except possibly at x.

We note here an important point about $\lambda(\cdot|x)$. The theory of Lagrange multipliers does not guarantee that $\lambda(y|x) \neq 0$ when $y \neq x$. Thus, if $\lambda(\cdot|x) \equiv 0$, then the $\lambda(\cdot|x)$-scaling of $a(\cdot|x)$ is the zero vector field, which is trivially conservative. It is therefore interesting to look for conditions under which the $\lambda(\cdot|x)$-scaling of $a(\cdot|x)$ is conservative for $\lambda(\cdot|x)$ nonzero except possibly at x.

C. *Existence of a Nonzero Scaling*

A careful examination of the proof of [2, Theorem 1], Csiszár reveals that if Π is a *regular and local* projection rule, then there exists, for each $x \in \mathbb{R}^n$, a function $\lambda(\cdot|x)$ with the property

$$y \neq x \implies \lambda(y|x) \neq 0, \tag{15}$$

i.e., the axioms of regularity and locality together imply that the Lagrange multipliers $\lambda(y|x)$ are nonzero whenever $y \neq x$. This prompts us to ask the following question: Does regularity alone guarantee that $\lambda(\cdot|x)$ satisfies the condition in (15), also ensuring that the $\lambda(\cdot|x)$-scaling of $a(\cdot|x)$ is conservative? While we postpone the discussion of this question to Section III-E, we show in the following section that if $\lambda(\cdot|x)$ satisfies (15), then any function F satisfying the conditions of Proposition 1, along with (11), generates a regular projection rule.

D. *Sufficient Conditions for a Function to Generate a Regular Projection Rule*

The main result of this section is the following proposition.

Proposition 2: Let $F = \{F(y|x) : x, y \in \mathbb{R}^n\}$ be continuously differentiable. For each x, let $L(y|x)$, $y \neq x$, be the set in (7) with normal vector $a(y|x)$, and let $a(x|x)$ be as in (9). Suppose that F satisfies conditions 1-3 of Proposition 1. Further, suppose that for each $x \in \mathbb{R}^n$, there exists $\lambda(\cdot|x) : \mathbb{R}^n \to \mathbb{R}$ satisfying (15) such that the $\lambda(\cdot|x)$-scaling of $a(\cdot|x)$ is conservative, with potential function $F(\cdot|x)$. Then, the projection rule Π defined by (5) is regular.

Proof: We begin by showing that Π, as defined by (5), satisfies consistency. For the rest of this proof, we fix an arbitrary $x \in \mathbb{R}^n$. Pick $L, L' \in \mathscr{L}$ such that $L' \subset L$. Suppose that $\Pi(L|x) \in L'$. Then, it follows from (5) that $F(\Pi(L|x)|x) \geq F(\Pi(L'|x)|x)$. Since $\Pi(L'|x) \in L$ trivially, we also have $F(\Pi(L|x)|x) \leq F(\Pi(L'|x)|x)$. Combining these, we get $F(\Pi(L|x)|x) = F(\Pi(L'|x)|x)$. However, from the condition in item 2 of Definition 2, since $F(\cdot|x)$ attains a unique minimum over every $L \in \mathscr{L}$, it follows that $\Pi(L'|x) = \Pi(L|x)$, thus proving that Π satisfies consistency.

Next, we show that Π satisfies distinctness. Towards this, pick $L, \bar{L} \in \mathscr{M}$ such that $L \neq \bar{L}$, and suppose that $x \notin L \cap \bar{L}$ (for otherwise, by consistency, $\Pi(L) = \Pi(\bar{L}) = x$). Let $\Pi(L|x) = y$ and $\Pi(\bar{L}|x) = \bar{y}$. Then, we may write $L = L(y|x)$ and $\bar{L} = L(\bar{y}|x)$, where the sets $L(y|x)$ and $L(\bar{y}|x)$ are as in (7), with normal vectors $a(y|x)$ and $a(\bar{y}|x)$ respectively.

Suppose now that $y = \bar{y} \neq x$. Then, by (11), we have

$$\nabla F(y|x) = \lambda(y|x) \cdot a(y|x), \tag{16}$$
$$\nabla F(\bar{y}|x) = \lambda(\bar{y}|x) \cdot a(\bar{y}|x), \tag{17}$$

where $\lambda(y|x), \lambda(\bar{y}|x) \neq 0$ from (15). Noting that the left hand sides of (16) and (17) are identical, it follows that

$$a(\bar{y}|x) = \frac{\lambda(y|x)}{\lambda(\bar{y}|x)} a(y|x). \tag{18}$$

This implies, from the representation in (7), that the sets $L(y|x)$ and $L(\bar{y}|x)$ are identical, resulting in a contradiction.

Therefore, it follows that $\Pi(L|x) \neq \Pi(\bar{L}|x)$, proving that Π satisfies distinctness.

Finally, we note from item 3 of Proposition 1 that the mapping $y \mapsto a(y|x)$ is continuous. Using this and the fact that any $L \in \mathscr{L}$ of the form (1) having dimension $(n-k)$ can be constructed from an intersection of k sets, each of the form (7), it follows that Π satisfies the axiom of continuity. This completes the proof of the proposition. ■

E. Existence of a Function That Generates a Regular Projection Rule

The developments in the present section have thus far dealt with conditions under which a given function F generates a regular projection rule Π. We now ask the following question: Given a regular projection rule Π, does there exist a function F satisfying the conditions of Proposition 2 such that F generates Π? The conditions in the statement of Proposition 2 then immediately lead us to a second question of whether regularity alone implies, for each $x \in \mathbb{R}^n$, the existence of a nonzero function $\lambda(\cdot|x)$ that is continuous everywhere except possibly at x, such that the $\lambda(\cdot|x)$-scaling of $a(\cdot|x)$ is conservative. It is clear that the answer to the second question above identifies, for each $x \in \mathbb{R}^n$, a potential function $F(\cdot|x)$, which serves as an answer to the first question. However, answers to the second question are currently not known, except in a few instances. Since the second question above is of independent mathematical interest, we formally state two versions of the same in the following section, thereby concluding our paper.

IV. An Open Question On Conservative Vector Fields

Fix an arbitrary $x \in \mathbb{R}^n$. The discussion in Section III-A implies that the function $\lambda(\cdot|x)$ is continuous, except possibly at x. While our search is for any such $\lambda(\cdot|x)$, it is clear that any $\lambda(\cdot|x)$ that is continuous everywhere, or additionally differentiable, is also a solution of interest. On this note, we state below two versions of the question of conservative fields pointed out in Section III-E, one for each of the cases when $\lambda(\cdot|x)$ is continous and when $\lambda(\cdot|x)$ is differentiable. The first version below is for the case when $\lambda(\cdot|x)$ is continuous. We use $v(\cdot)$ to identify the vector field $a(\cdot|x)$, and simply write $\lambda(\cdot)$ to denote $\lambda(\cdot|x)$.

Question 1:
Given a continuous vector field $v : \mathbb{R}^n \to \mathbb{R}^n$, does there exist a continuous function $\lambda : \mathbb{R}^n \to \mathbb{R}$ such that the vector field

$$u(y) = \lambda(y) \cdot v(y), \quad y \in \mathbb{R}^n, \tag{19}$$

is conservative, i.e., there exists a continuously differentiable function $G : \mathbb{R}^n \to \mathbb{R}$ such that $u(y) = \nabla G(y)$ for all $y \in \mathbb{R}^n$?

We note here that answers to the above question are known in some specific instances such as when the given projection rule satisfies, in addition to regularity, additional axioms such as locality and subspace transitivity. For the least squares projection rule, it can be shown that $a(y) = y$, in which case $\lambda(y) = 1$ for all $y \in \mathbb{R}^n$, which yields $F(y) = \frac{1}{2}||y||^2$. Similarly, answers to the above question are known for the I-divergence projection rule. However, for the regular, non-local and subspace transitive projection rule generated by relative α-entropy, where $\alpha < 1$, no solution is known.

We now state a version of the above question assuming that $\lambda(\cdot|x)$ is also differentiable for each $x \in \mathbb{R}^n$. For this, we require that the vector field $a(\cdot|x)$ is differentiable for each $x \in \mathbb{R}^n$. This requirement is not very restrictive, and is satisfied for the least squares projection rule (for which $a(y|x) = y - x$) and the I-divergence projection rule (for which the i^{th} component of $a(y|x)$ is given by $a(y|x)_i = 1 + \log \frac{y_i}{x_i}$, $i = 1, 2, \ldots, n$).

We pick coordinates $1 \leq i < j < k \leq n$ arbitrarily. Then, focusing attention only to the coordinates i, j, k in (19), we get

$$\begin{bmatrix} u_i(y) \\ u_j(y) \\ u_k(y) \end{bmatrix} = \lambda(y) \begin{bmatrix} v_i(y) \\ v_j(y) \\ v_k(y) \end{bmatrix}, \quad y = (y_i, y_j, y_k) \in \mathbb{R}^3. \tag{20}$$

From the property that $\mathbb{R}^3$ is a simply connected set, it follows that the left hand side of (20) is a conservative vector field if and only if its curl (with respect to the coordinates y_i, y_j, y_k) is the zero vector field. Then, evaluating the curl of the right hand side of (20) and setting it to zero yields a system of 3 first order partial differential equations in the coordinates $y = (y_i, y_j, y_k)$, as below.

$$\begin{aligned} v_i(y)\frac{\partial \lambda(y)}{\partial y_j} - v_j(y)\frac{\partial \lambda(y)}{\partial y_i} &= \lambda(y)\left(\frac{\partial v_j(y)}{\partial y_i} - \frac{\partial v_i(y)}{\partial y_j}\right), \\ v_j(y)\frac{\partial \lambda(y)}{\partial y_k} - v_k(y)\frac{\partial \lambda(y)}{\partial y_j} &= \lambda(y)\left(\frac{\partial v_k(y)}{\partial y_j} - \frac{\partial v_j(y)}{\partial y_k}\right), \\ v_k(y)\frac{\partial \lambda(y)}{\partial y_i} - v_i(y)\frac{\partial \lambda(y)}{\partial y_k} &= \lambda(y)\left(\frac{\partial v_i(y)}{\partial y_k} - \frac{\partial v_k(y)}{\partial y_i}\right). \end{aligned} \tag{21}$$

Suppose that $\lambda_{ijk}(y_i, y_j, y_k)$ is any solution to (21). Since (19) is conservative if and only if (21) holds for all choices of coordinates $y_i, , y_j, y_k \in \mathbb{R}$, we arrive at the following question.

Question 2:
Does there exist a consistent solution $\lambda(\cdot) : \mathbb{R}^n \to \mathbb{R}$, with the property $\lambda(y) = \lambda_{ijk}(y_i, y_j, y_k)$ for all distinct i, j, k, such that the vector field (19) conservative?

While solutions to the above question are known in some cases such as those of least squares and I-divergence projection rules, the existence of solutions in more general cases remains open.

References

[1] E. T. Jaynes, "On the rationale of maximum-entropy methods," *Proceedings of the IEEE*, vol. 70, no. 9, pp. 939–952, 1982.

[2] I. Csiszar, "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *The annals of statistics*, vol. 19, no. 4, pp. 2032–2066, 1991.

[3] L. M. Bregman, "The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming," *USSR computational mathematics and mathematical physics*, vol. 7, no. 3, pp. 200–217, 1967.

[4] M. A. Kumar and R. Sundaresan, "Minimization problems based on relative α-entropy I: Forward projection," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 5063–5080, September 2015.