Isotonic regression and isotonic projection

A. B. Németh
Faculty of Mathematics and Computer Science
Babeš Bolyai University, Str. Kogálniceanu nr. 1-3
RO-400084 Cluj-Napoca, Romania
email: nemab@math.ubbcluj.ro

S. Z. Németh
School of Mathematics, The University of Birmingham
The Watson Building, Edgbaston
Birmingham B15 2TT, United Kingdom
email: nemeths@for.mat.bham.ac.uk

Abstract

The note describes the cones in the Euclidean space admitting isotonic metric projection with respect to the coordinate-wise ordering. As a consequence it is showed that the metric projection onto the regression cone (the cone defined by the general isotonic regression problem) admits a projection which is isotonic with respect to the coordinate-wise ordering.

1. Introduction

The isotonic regression problem and its solution is intimately related to the metric projection into a cone of the Euclidean vector space. In fact the isotonic regression problem is a special quadratic optimization problem. It is desirable to relate the metric projection onto a closed convex set to some order theoretic properties of the projection itself, which can facilitate the solution of some problems. When the underlying set is a convex cone, then the most natural is to consider the order relation defined by the cone itself. This approach gives rise to the notion of the isotonic projection cone, which by definition is a cone with the metric projection onto it isotonic with respect to the order relation endowed by the cone itself. As we shall see, the two notions of the isotonicity: the one related to the regression problem, the second to the metric projection are at the first sight rather different. The fact that the two notions are in fact intimately related (this relation constitute the subject of this note) is somewhat accidental: it derives from semantical reasons.
The relation of the two notions is observed and exploited in the paper [1]. There was exploited the fact that the totally ordered regression cone is an isotonic projection cone too.

The problem occurs as a particular case of the following more general question: How does a closed convex set in the Euclidean space which admits a metric projection isotonic with respect to some vectorial ordering on the space look like?

It turns out, that the problem is strongly related with some lattice-like operations defined on the space, and in particular with the Euclidean vector lattice theory. ([4]) When the ordering is the coordinate-wise one, the problem goes back in the literature to [8], [9], [10], [2] and [5]. However, we shall ignore these connections in order to simplify the exposition. Thus the present note besides proving some new results has the role to bring together some previous results and to present them in a simple unified form.

2. Preliminaries

Denote by $\mathbb{R}^m$ the $m$-dimensional Euclidean space endowed with the scalar product $\langle \cdot,\cdot \rangle : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, and the Euclidean norm $|| \cdot ||$ and topology this scalar product defines.

Throughout this note we shall use some standard terms and results from convex geometry (see e.g. [7]).

Let $K$ be a convex cone in $\mathbb{R}^m$, i.e., a nonempty set with (i) $K + K \subseteq K$ and (ii) $tK \subseteq K$, $\forall t \in \mathbb{R}_+ = [0, +\infty)$. The convex cone $K$ is called pointed, if $(-K) \cap K = \{0\}$. The cone $K$ is generating if $K - K = \mathbb{R}^m$. $K$ is generating if and only if int $K \neq \emptyset$. A closed, pointed generating convex cone is called proper.

For any $x, y \in \mathbb{R}^m$, by the equivalence $x \leq_K y \iff y - x \in K$, the convex cone $K$ induces an order relation $\leq_K$ in $\mathbb{R}^m$, that is, a binary relation, which is reflexive and transitive. This order relation is translation invariant in the sense that $x \leq_K y$ implies $x + z \leq_K y + z$ for all $z \in \mathbb{R}^m$, and scale invariant in the sense that $x \leq_K y$ implies $tx \leq_K ty$ for any $t \in \mathbb{R}_+$. If $\leq$ is a translation invariant and scale invariant order relation on $\mathbb{R}^m$, then $\leq = \leq_K$ with $K = \{x \in \mathbb{R}^m : 0 \leq x\}$. If $K$ is pointed, then $\leq_K$ is antisymmetric too, that is $x \leq_K y$ and $y \leq_K x$ imply that $x = y$.

The set

$$K = \text{cone}\{x_1, \ldots, x_m\} := \{t^1 x_1 + \ldots + t^m x_m : t^i \in \mathbb{R}_+, i = 1, \ldots, m\}$$

with $x_1, \ldots, x_m$ linearly independent vectors is called a simplicial cone. A simplicial cone is closed, pointed and generating.

The dual of the convex cone $K$ is the set

$$K^* := \{y \in \mathbb{R}^m : \langle x, y \rangle \geq 0, \forall x \in K\},$$

with $\langle \cdot , \cdot \rangle$ the standard scalar product in $\mathbb{R}^m$.

The cone $K$ is called self-dual, if $K = K^*$. If $K$ is self-dual, then it is a generating, pointed, closed convex cone.
In all that follows we shall suppose that $\mathbb{R}^m$ is endowed with a Cartesian reference system with the standard unit vectors $e_1, \ldots, e_m$. That is, $e_1, \ldots, e_m$ is an orthonormal system of vectors in the sense that $\langle e_i, e_j \rangle = \delta_{ij}$, where $\delta_{ij}$ is the Kronecker symbol. Then, $e_1, \ldots, e_m$ form a basis of the vector space $\mathbb{R}^m$. If $x \in \mathbb{R}^m$, then

$$x = x^1 e_1 + \ldots + x^m e_m$$

can be characterized by the ordered $m$-tuple of real numbers $x^1, \ldots, x^m$, called the coordinates of $x$ with respect the given reference system, and we shall write $x = (x^1, \ldots, x^m)$.

The set

$$\mathbb{R}_+^m = \{ x = (x^1, \ldots, x^m) \in \mathbb{R}^m : x^i \geq 0, i = 1, \ldots, m \}$$

is called the nonnegative orthant of the above introduced Cartesian reference system. A direct verification shows that $\mathbb{R}_+^m$ is a self-dual cone.

Besides the non-negative orthant, given a Cartesian reference system, the important class of regression cones should be mentioned. Let $(V = \{1, \ldots, m\}, E)$ be a directed graph of vertices $V$ and edges $E \subset V \times V$ and without loops (a so called simple directed graph). (If $(i, j) \in E$, then $i$ is called its tail, $j$ is called its head.) Then we shall call the set

$$K_E = \{ x \in \mathbb{R}^m : x^i \leq x^j, \forall (i, j) \in E \}$$

the regression cone defined by the relations $E$.

If $(V, E)$ is connected directed simple graph for which each vertex is the tail respective a head of at most one edge, then $K_E$ is called monotone cone. In this case $K_E$ can be written (after a possible permutation of the standard unit vectors) in the form

$$K_E = \{ x \in \mathbb{R}^m : x^1 \leq x^2 \leq \ldots \leq x^m \}.$$ 

A hypersubspace or a hyperplane through the origin, is a set of form

$$H(u, 0) = \{ x \in \mathbb{R}^m : \langle u, x \rangle = 0 \}, \ u \neq 0. \quad (1)$$

For simplicity the hypersubspaces will also be denoted by $H$. The nonzero vector $u$ in the above formula is called the normal of the hyperplane.

A hyperplane (through $a \in \mathbb{R}^m$) is a set of form

$$H(u, a) = \{ x \in \mathbb{R}^m : \langle u, x \rangle = \langle u, a \rangle, \ u \neq 0 \}. \quad (2)$$

A hyperplane $H(u, a)$ determines two closed halfspaces $H_-(a, u)$ and $H_+(u, a)$ of $\mathbb{R}^m$, defined by

$$H_-(u, a) = \{ x \in \mathbb{R}^m : \langle u, x \rangle \leq \langle u, a \rangle \}.$$
and
\[ H_+(u, a) = \{ x \in \mathbb{R}^m : \langle u, x \rangle \geq \langle u, a \rangle \}. \]

The cone \( K \subset \mathbb{R}^m \) is called \textit{polyhedral} if it can be represented in the form
\[ K = \cap_{k=1}^n H_-(a_k, 0). \]  \hspace{1cm} (3)

If \( \text{int} \ K \neq \emptyset \), and the representation (3) is irredudant, then \( K \cap H(a_k, 0) \) is an \( m-1 \)-dimensional convex cone \( (k = 1, \ldots, n) \) and is called a \textit{facet of K}.

The simplicial cone and the regression cones are polyhedral.

3. Metric projection and isotone projection sets

Denote by \( P_D \) the projection mapping onto a nonempty closed convex set \( D \subset \mathbb{R}^m \), that is the mapping which associate to \( x \in \mathbb{R}^m \) the unique nearest point of \( x \) in \( D \) ([10]):
\[ P_Dx \in D, \text{ and } \| x - P_Dx \| = \inf\{ \| x - y \| : y \in D \}. \]

Given an order relation \( \leq \) in \( \mathbb{R}^m \), the closed convex set is said an \textit{isotone projection set} if from \( x \leq y, x, y \in \mathbb{R}^m \), it follows \( P_Dx \leq P_Dy \).

If \( \leq = \leq_K \) for some cone \( K \), then the isotone projection set \( D \) is called \( K \)-\textit{isotone}.

If the cone \( K \) is \( K \)-isotone then it is called an \textit{isotone projection cone}.

For \( K = \mathbb{R}_+^m \) we have \( P_Kx = x^+ \) where \( x^+ \) is the vector formed with the non-negative coordinates of \( x \) and 0-s in place of negative coordinates. Since \( x \leq_K y \) implies \( x^+ \leq_K y^+ \), it follows that \( \mathbb{R}_+^m \) is an isotone projection cone.

We have the following geometric characterization of a closed, generating isotone projection cones (Theorem 1 and Corollary 1 in [1]):

\textbf{Theorem 1} The closed generating cone \( K \subset \mathbb{R}^m \) is an isotone projection cone if and only if its dual \( K^* \) is a simplicial cone in the subspace it spans generated by vectors with mutually non-acute angles.

4. The positive orthant and its isotonic projection subcones

If \( \mathbb{R}_+^m \) is the positive orthant of a Cartesian system, then we have the following proposition (Corollaries 1 and 3 in [3]):

\textbf{Proposition 1} Let \( C \) be a closed convex set with nonempty interior of the coordinatewise ordered Euclidean space \( \mathbb{R}^m \). Then, the following assertions are equivalent:

(i) The projection \( P_C \) is \( \mathbb{R}_+^m \)-isotone;
where each hyperplane \( H(a_i, b_i) \) is tangent to \( C \) and the normals \( a_i \) are nonzero vectors \( a_i = (a_{i1}, \ldots, a_{in}) \) with the properties \( a_k^i a_l^i \leq 0 \) whenever \( k \neq l, \ i \in \mathbb{N} \).

**Example 1** Consider the space \( \mathbb{R}^3 \) endowed with a Cartesian reference system, and suppose
\[
K_1 = H_-((-2, 1, 0), 0) \cap H_-((1, -2, 0), 0) \cap H_-((0, 0, -1), 0),
\]
and
\[
K_2 = H_-((-2, 1, 0), 0) \cap H_-((1, -2, 0), 0) \cap H_-((0, 1, -1), 0).
\]
Then \( K_1 \) and \( K_2 \) are simplicial cones in \( \mathbb{R}_+^3 \), \( x = (1, 1, 2) \in \text{int} \ K_i, \ i = 1, 2 \). Since
\[
K_1 = \text{cone}\{(-2, 1, 0), (1, -2, 0), (0, 0, -1)\}^\perp
\]
and
\[
K_2 = \text{cone}\{(-2, 1, 0), (1, -2, 0), (0, 1, -1)\}^\perp,
\]
using the main result in [3] we see that \( K_1 \) is itself an isotone projection cone, while \( K_2 \) is not. Obviously, \( K_1 \) and \( K_2 \) are both \( \mathbb{R}_+^3 \)-isotone projection sets.

**Example 2** Let us consider the space \( \mathbb{R}^3 \) endowed with a Cartesian reference system. Let \( L = \mathbb{R}_+^3 \). Consider the vectors
\[
a_1 = (-2, 1, 0), \ a_2 = (1, -2, 0), \ a_3 = (-2, 0, 1), \ a_4 = (1, 0, -2), \ a_5 = (0, -2, 1), \ a_6 = (0, 1, -2).
\]
Then,
\[
K = \cap_{i=1}^6 H_-(a_i, 0) \subset \mathbb{R}_+^3
\]
is by Proposition [4] an \( L \)-isotone projection cone with six facets.

Indeed, \( \langle a_1, x \rangle \leq 0 \) and \( \langle a_2, x \rangle \leq 0 \) imply that \( x^1 \geq 0 \) and \( x^2 \geq 0 \). We can show similarly that for \( x \in K \) it holds \( x^3 \geq 0 \). Thus \( K \subset L \). For \( y = (1, 1, 1) \) we have \( \langle a_i, y \rangle < 0 \). Hence \( y \in \text{int} \ K \). The sets \( H(a_i, 0) \cap K, \ i = 1, \ldots, 6 \) are different facets of \( K \).

We shall show next that the cone in Example [2] is in some sense extremal among the \( L = \mathbb{R}_+^3 \) isotone subcones in \( L \). More precisely we have

**Proposition 2** If \( K \) is a generating cone in \( \mathbb{R}^m \), then it is \( L = \mathbb{R}_+^m \)-isotone, if and only if is of the form
\[
K = \cap_{k<l}(H_-(a_{kl1}, 0) \cap H_-(a_{kl2})), \ k, l \in \{1, \ldots, m\}
\]
where \( a_{kl} \) are nonzero vectors with \( a_k^l a_{kl} \leq 0 \) and \( a_j^l = 0 \) for \( j \notin \{k, l\} \), \( i = 1, 2 \). Hence \( K \) possesses at most \( m(m-1) \) facets. There exists a cone \( K \) of the above form with exactly \( m(m-1) \) facets.
Proof.

We have by Proposition 1 that

\[ K = \cap_{i \in \mathcal{J}} H_{-}(a_i, 0), \]

where \( \mathcal{J} \subset \mathbb{N} \) is a set of indexes and where each hyperplane \( H(a_i, 0) \) is tangent to \( K \) and the normals \( a_i \) are nonzero vectors \( a_i = (a_i^1, ..., a_i^m) \) with the properties \( a_i^k a_i^l \leq 0 \) whenever \( k \neq l, \ i \in \mathbb{N} \).

First of all we introduce the notation

\[ \mathcal{A}_{kl} = \{ i : a_i^j = 0, \ j \notin \{k, l\}, \ k, l \in \{1, ..., m\}, \ k < l. \]

(In Example 2 \( \mathcal{A}_{12} = \{1, 2\}, \ \mathcal{A}_{13} = \{3, 4\}, \ \mathcal{A}_{23} = \{5, 6\} \).

We claim that

\[ \mathcal{A}_{kl} \neq \emptyset, \ k < l, \ \text{and} \ \cup_{k<l} \mathcal{A}_{kl} = \mathcal{J}. \]  

This follows from the structure of the normals \( a_i \). Indeed if \( a_i \) possesses two non-zero components, say \( a_i^k \) and \( a_i^l, \ k < l \), then \( i \in \mathcal{A}_{kl} \). If it has only one non-zero component, say \( a_i^k \) with \( k < m \), then \( i \in \mathcal{A}_{km} \), or only one nonzero component \( a_i^m \) then \( i \in \mathcal{A}_{km} \) for \( k < m \).

Let us see that

\[ \cap_{i \in \mathcal{A}_{kl}} H_{-}(a_i, 0) = H_{-}(a_{i_1}, 0) \cap H_{-}(a_{i_2}, 0), \]

where \( H_{-}(a_{i_1}, 0) \) are among those in (6) and the case \( i_1 = i_2 \) is possible.

Denote with \( \mathbb{R}_{kl} \) the bidimensional subspace in \( \mathbb{R}^m \) endowed by the \( k \)-th and \( l \)-th axis. Then we have the representation

\[ \cap_{i \in \mathcal{A}_{kl}} H_{-}(a_i, 0) = \mathbb{R}_{kl}^+ \times (\cap_{i \in \mathcal{A}_{kl}} H_{-}(a_i, 0)) \cap \mathbb{R}_{kl}. \]

Now, \( \cap_{i \in \mathcal{A}_{kl}} H_{-}(a_i, 0)) \cap \mathbb{R}_{kl} \) must be a two dimensional cone in \( \mathbb{R}_{kl} \) (since \( K \) is generating), hence it must have one or two extremal rays. That is the intersection can be expressed by one or two terms, that is, we can suppose that \( 1 \leq \text{card} \mathcal{A}_{kl} \leq 2 \) and (5) is proved.

With these remarks we can assert that the formula (6) becomes

\[ K = \cap_{k<l} (\cap_{i \in \mathcal{A}_{kl}} H_{-}(a_i, 0)) = \cap_{k<l} (H_{-}(a_{kl1}, 0) \cap H_{-}(a_{kl2}, 0)), \]

where \( a_{kl1}, a_{kl1} \leq 0 \) and \( a_{klj}^j = 0 \) for \( j \notin \{k, l\}, \ i = 1, 2 \).

From formula (6) it follows that in the representation (5) of \( K \) there are at most \( m(m-1) \) facets \( H(a_i, 0) \cap K \) of \( K \).

Using the construction in Example 2 we can construct a \( K \) with exactly \( m(m-1) \) facets. To this end, let for \( k < l \) \( a_{kl1} \) be the vector with \( a_{kl1}^k = -2, \ a_{kl1}^l = 1 \) and \( a_{kl1}^j = 0 \) for \( j \notin \{k, l\} \), and \( a_{kl2} \) be the vector with \( a_{kl2}^k = 1, \ a_{kl2}^l = -2 \) and \( a_{kl2}^j = 0 \) for \( j \notin \{k, l\} \). We have that the vectors \( a_{kl1} \) are pairwise nonparallel. Putting these vectors in the representation (5) we get a proper subcone of \( \mathbb{R}^m_+ \) which is \( \mathbb{R}^m_+ \)-isotone and possesses exactly \( m(m-1) \) facets. Indeed, we must see that in this case the representation (5) is irredundant. But this follows from the fact that \( K \subset \mathbb{R}^m_+ \) is a polyhedral cone with
x = (1,1,...,1) an interior point. Hence some of $F_{kli} = H(a_{kli},0) \cap K$ must be facets of $K$. Now, from the special feature of $a_{kli}$ it follows that the sets $F_{kli}$ are structurally equivalent and if one of them is a facet, then all of them are so.

The proof implies also that $K$ must be a polyhedral cone and if its representation (6) is irredundant, than the set $\mathcal{J}$ must be finite.

\[ \square \]

Remark 1 The representation (9) can be redundant, even if the original one in (6) is irredundant. Indeed, $\mathbb{R}_m^m$ must be of form (6) and its irredundant representation contains $m$ terms, while its equivalent form (9) formally contains much more terms. In this case (9) can contain \( m(m-1)/2 \) terms. But even a minimal “dual” decomposition of $\mathbb{R}_m^m$ is of cardinality \( \lceil m/2 \rceil \) and hence it contains \( 2\lceil m/2 \rceil \) halfspaces.

5. Every isotonic regression cone is an $\mathbb{R}_m^m$-isotonic projection set

Projecting $y \in \mathbb{R}^m$ into $K$ given by (9) we have to solve the following quadratic minimization problem:

\[
P_K y = \text{argmin} \left\{ \sum_{i=1}^{m} (x^i - y^i)^2 : a_{kli}^k x^k + a_{lji}^l x^l \leq 0, \ a_{kli}^k x^k + a_{lji}^l x^l \leq 0, \ k < l \right\}, \quad (10)
\]

where the cases $a_{kli}^k = 0$, or $a_{lji}^l = 0$ are not excluded.

By using Proposition 1 we see that, from

\[
u \preceq_{\mathbb{R}_m^m} v,
\]

it follows that

\[
P_K u \preceq_{\mathbb{R}_m^m} P_K v.
\]

A particular case of this projection problem is the so called isotonic regression problem. This problem is as follows:

For a given $y \in \mathbb{R}^m$ get

\[
\text{iso}(y) := P_K y = \text{argmin} \left\{ \sum_{i=1}^{m} (x^i - y^i)^2 : x^i \leq x^j, \ \forall (i,j) \in E \right\},
\]

where $(V = \{1,...,m\}, E)$ is a directed simple graph.

In fact in this case $K = K_E$ where $K_E$ is the regression cone defined in Section 2, hence iso($y$) is the metric projection of $y$ into the regression cone $K_E$. 

7
To compare with this with the general projection problem \([10]\), we observe that the restrictions on \(x\) are of form

\[a_{ij}^i x^i + a_{ij}^j x^j \leq 0\]

with \(a_{ij}^i = 1\) and \(a_{ij}^j = -1\), \((i, j) \in E\). Thus we have established the

**Proposition 3** Every regression cone \(K_E\) is an \(\mathbb{R}_+^\infty\)-isotone projection set.

We have further that

**Proposition 4** The regression cone \(K_E\) is an isotone projection cone if and only if in the oriented graph \((V, E)\) does not exist different edges with same tail or different edges with same head, that is, edges of form \((i, j)\) and \((i, k)\) with \(j \neq k\), or edges of form \((i, j)\) and \((k, j)\) with \(i \neq k\).

**Proof.** Assume e. g. that \((1, 2), (1, 3) \in E\). Then the corresponding normals are

\[a_{1,2} = (1, -1, 0, \ldots, 0)\]

and

\[a_{1,3} = (1, 0, -1, 0, \ldots, 0).\]

Then \(a_{1,2}\) and \(a_{1,3}\) are normals in the irreducible representation of \(K_E\), and \(\langle a_{1,2}, a_{1,3} \rangle > 0\). Thus according to Corollary 1 in \([1]\) \(K_E\) cannot be an isotone projection cone. Conversely, if there are no vertices with the above type multiplicity property, then the normals in the irreducible representation of \(K_E\) (which in fact generates \(-K_E^*\)) form pair-wise non-acute angles, hence by the same result \(K_E\) is an isotone projection cone.

\[\Box\]

**Corollary 1** If \(K_E\) is an isotone projection cone, then \((V, E)\) splits in disjoint union of connected simple graphs with vertices being the tails or heads of at most one edge. The single (up to a permutation of the canonical basis) isotone projection regression cone \(K_E\), with \((V, E)\) a directed connected simple graph, is the monotone cone.

**References**

[1] A. Guyader, N. Jégou, A. B. Németh, and S. Z. Németh. A geometrical approach to iterative isotone regression, 2014.

[2] G. Isac. On the order monotonicity of the metric projection operator. In Approximation theory, wavelets and applications (Maratea, 1994), volume 454 of NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., pages 365–379. Kluwer Acad. Publ., Dordrecht, 1995.
[3] G. Isac and A. B. Németh. Monotonicity of metric projections onto positive cones of ordered Euclidean spaces. *Arch. Math.*, 46(6):568–576, 1986.

[4] A. B. Németh and S. Z. Németh. Lattice-like operations and isotone projection sets. *Linear Algebra and its Applications*, 439(10):2815–2828, 2013.

[5] H. Nishimura and E. A. Ok. Solvability of variational inequalities on Hilbert lattices. *Preprint*, pages 1–28, 2012.

[6] M. Queyranne and F. Tardella. Bimonotone linear inequalities and sublattices of $\mathbb{R}^n$. *Linear Algebra Appl.*, 413:100–120, 2006.

[7] R. T. Rockafellar. *Convex Analysis*. Princeton: Princeton Univ. Press, 1970.

[8] D. M. Topkis. The structure of sublattices of the product of $n$ lattices. *Pacific J. Math.*, 65:525–532, 1976.

[9] A. F. Veinott. Representation of general and polyhedral sublattices and sublattices of product spaces. *Linear Algebra Appl.*, 114/115:172–178, 1981.

[10] E. Zarantonello. Projections on convex sets in Hilbert space and spectral theory, I: Projections on convex sets, II: Spectral theory. *Contrib. Nonlin. Functional Analysis, Proc. Sympos. Univ. Wisconsin, Madison*, pages 237–424, 1971.