MEANS ON DISCRETE PRODUCT LATTICES

Gleb Beliakov[1] Humberto Bustince[2] Javier Fernandez[2] Daniel Paternain[2]

[1] School of Information Technology, Deakin University, Australia gleb@deakin.edu.au
[2] Dept. Automatics and Computation, UPNA, Pamplona, Spain {bustince, fcojavier.fernandez, daniel.paternain}@unavarra.es

Abstract

We investigate the problem of averaging values on lattices, and in particular on discrete product lattices. This problem arises in image processing when several color values given in RGB, HSL, or another coding scheme, need to be combined. We show how the arithmetic mean and the median can be constructed by minimizing appropriate penalties. We also discuss which of them coincide with the Cartesian product of the standard mean and median.

Keywords: Aggregation Functions, Mean, Median, Lattice.

1 INTRODUCTION

The need to aggregate several inputs into a single representative output arises in many practical applications. In image processing, it is often necessary to average the values of several neighboring pixels (to reduce the image size or apply a filter), or average pixel values in two different but related images (e.g., in stereovision). When the images are in color, typically coded as discrete RGB, CMY, or HSL values, then it is customary to average the values in the respective channels. Nevertheless, it would be of interest to find other ways to average color values.

In this paper we study averaging on product lattices (RGB or another color coding scheme is an example of a product lattice). We note previous works related to triangular norms on posets and lattices [5, 10] and on discrete chains [11]. In these works, associativity was required. In the setting we are going to develop, however, we will require an averaging behavior of the aggregation functions to be considered. We consider this can be a key requirement, specially from the point of view of image processing applications.

The structure of the paper is as follows. In Section 2 we provide preliminary definitions. In Section 3 we give the definitions of aggregation functions based on penalties, defined on product lattices. We discuss solutions to resulting optimization problems in Section 4. Conclusions are presented in Section 5.

2 PRELIMINARIES

2.1 Aggregation functions

The research effort concerning aggregation functions, their behavior and properties, has been disseminated throughout various fields including decision making, knowledge based systems, artificial intelligence and image processing. Recent books providing a comprehensive overview include [1, 3, 7, 13].

Definition 1 *A function $f : [a,b]^n \to [a,b]$ is called an aggregation function if it is monotone non-decreasing in each variable and satisfies $f(\mathbf{a}) = a$, $f(\mathbf{b}) = b$, with $\mathbf{a} = (a,a,\ldots,a), \mathbf{b} = (b,b,\ldots,b)$.*

Definition 2 *An aggregation function f is called averaging if it is bounded by the minimum and maximum of its arguments*

$$\min(\mathbf{x}) := \min(x_1,\ldots,x_n) \le f(x_1,\ldots,x_n)$$

$$\le \max(x_1,\ldots,x_n) =: \max(\mathbf{x}).$$

It is immediate (because of monotonicity) that averaging aggregation functions are idempotent (i.e., $\forall t \in [a,b] : f(t,t,\ldots,t) = t$) and vice versa. Then clearly the boundary conditions $f(\mathbf{a}) = a$, $f(\mathbf{b}) = b$ are satisfied.

Well known examples of averaging functions are the arithmetic mean and the median. It is known that the arithmetic means and the median are solutions to simple optimization problems, in which a measure of disagreement between the inputs is minimized (see [13]), [2, 4, 6, 14]. The main motivation is the following. Let $\mathbf{x}$ be the inputs and y the output. If all the inputs coincide $x = x_1 = \ldots, x_n$, then the output is $y = x$, and we have a unanimous vote. If some input $x_i \neq y$, then we impose a "penalty" for this disagreement. The larger the disagreement, and the more inputs disagree with the output, the larger (in general) is the penalty. We look for an aggregated value which minimizes the penalty.

Thus we need to define a suitable measure of disagreement, or dissimilarity.

Definition 3 *Let $P : [a, b]^{n+1} \to \Re$ be a penalty function with the properties*

i) $P(\mathbf{x}, y) \geq 0$ *for all* $\mathbf{x}, y$*;*

ii) $P(\mathbf{x}, y) = 0$ *if all* $x_i = y$*;*

iii) $P(\mathbf{x}, y)$ *is quasiconvex in* y *for any* $\mathbf{x}$*, i.e,* $P(\mathbf{x}, \lambda y_1 + (1-\lambda)y_2) \leq \max(P(\mathbf{x}, y_1), P(\mathbf{x}, y_2))$ *for all* $\lambda, y_1, y_2 \in [0, 1]$*.*

Observe that a quasiconvex function attains its minimum either at a single point or for a whole interval. The penalty based function is

$$f(\mathbf{x}) = \arg\min_y P(\mathbf{x}, y),$$

if y is the unique minimizer, and $y = \frac{a+b}{2}$ if the set of minimizers is the interval $[a, b]$.

In [2] it was shown that any averaging aggregation function can be represented as a penalty based function. Further, the classical means, such as the arithmetic and the median are represented via the following penalty functions. The arithmetic mean is the solution to

$$minimize_y \sum_{i=1}^{n} (x_i - y)^2$$

whereas the median is a solution to

$$minimize_y \sum_{i=1}^{n} |x_i - y|.$$

In this work we will deal with penalty based functions defined on lattices, rather than the interval $[a, b]$.

2.2 Lattices

Definition 4 *Let L be a set. A lattice $\mathcal{L} = (\mathcal{L}, \leq, \wedge, \vee)$ is a poset with the partial order $\leq$ on L, and meet and join operations $\wedge, \vee$, if every pair of elements from L has both meet and join.*

Definition 5 *Let P be a poset. A chain in P is a totally ordered subset of P. The length of a chain is its cardinality.*

Definition 6 *If $\mathcal{L}_1 = (L_1, \leq_1, \wedge_1, \vee_1)$ and $\mathcal{L}_2 = (L_2, \leq_2, \wedge_2, \vee_2)$ are two lattices, their Cartesian product is the lattice $\mathcal{L}_1 \times \mathcal{L}_2 = (L_1 \times L_2, \leq, \wedge, \vee)$ with $\leq$ defined by*

$$(x_1, y_1) \leq (x_2, y_2) \Leftrightarrow x_1 \leq_1 x_2 \text{ and } y_1 \leq_2 y_2$$

and

$$(x_1, y_1) \wedge (x_2, y_2) = (x_1 \wedge_1 x_2, y_1 \wedge_2 y_2)$$

$$(x_1, y_1) \vee (x_2, y_2) = (x_1 \vee_1 x_2, y_1 \vee_2 y_2)$$

We will deal with Cartesian products of finite chains $\mathcal{C}$, which is precisely the type of product lattice representing colors in image processing, with the length of each chain typically being 256. We note that all finite chains of the same length are isomorph to each other, and hence we can represent them as non-negative integers $0, 1, \ldots, K$, and elements of product lattices as tuples $x = (x_1, x_2 \ldots, x_m)$, $x_i \in \mathcal{C}_i$.

Definition 7 *Let f_1, f_2 be two aggregation functions defined on sets X_1 and X_2 respectively. The Cartesian product of aggregation functions is $f = f_1 \times f_2 : X_1 \times X_2 \to Y_1 \times Y_2$ defined by*

$$f(x_1, x_2) = (f_1(x_1), f_2(x_2)).$$

Observe that this definition is given from a strict mathematical point of view, so no interaction or relation is assumed between f_1 and f_2.

3 MAIN DEFINITIONS

Following representations of the arithmetic mean and the median as penalty-based aggregation functions, we now define similar constructions on lattices.

Definition 8 *Let $\mathcal{L} = (L, \leq, \wedge, \vee)$ be a product lattice of chains. The distance between $x, y \in \mathcal{L}$ is defined as the length of a chain $\mathcal{C}$ with the least element $a = x \wedge y$ and the greatest element $b = x \vee y$ minus 1,*

$$d(x, y) = \ length(\mathcal{C}) - 1$$

We note that all the chains in Definition 8 have the same length, since we are dealing with product of chains. This definition is equivalent to the following

$$d(x, y) = \sum_{i=1}^{m} d_i(x_i, y_i) = \sum_{i=1}^{m} |x_i - y_i|,$$

where d_i is the distance on the i-th chain.

Definition 9 *Let $\mathcal{L}$ be a product lattice. Consider n elements $x_1, \ldots, x_n \in \mathcal{L}$, that need to be averaged. Let the penalty function be $P : Z^n_+ \to \Re_+$. The penalty based function on $\mathcal{L}$ is f given by*

$$f(x_1, \ldots, x_n) = \mu = \arg\min_y P(d(x_1, y), d(x_2, y), \ldots, d(x_n, y)),$$

where μ is rounded to the closest value in the lattice, if necessary.

Note that the minimum always exists. Also note that there can be several minimizers (see Example 1 in Section 4 for an example). From an applied point of view, these means that there are several elements which average the given ones. In this case one can take any minimizer . Finally, f is not necessarily monotone, i.e., an aggregation function.

Theorem 1 *The function in Definition 9, is an averaging (and hence idempotent) function.*

PROOF. Clearly $\bigwedge x_i = a \leq \mu \leq b = \bigvee x_i$, because for any x_i, $d(x_i, a) > d(x_i, t)$ with $t < a$, and similarly at the other end. Also f is idempotent, since every averaging function is. □

A special case of penalty based functions was considered in [12], called dissimilarity functions, where the penalty P is given by

$$P(\mathbf{x}, y) = \sum_{i=1}^{n} K(x_i - y), \qquad (1)$$

where K is a convex function with the unique minimum $K(0) = 0$. In this case the penalty based function is monotone, i.e., an aggregation function. By adapting this definition to our case we have

Theorem 2 *The function in Definition 9, with P given by*

$$P(\mathbf{x}, y) = \sum_{i=1}^{n} K(d(x_i, y)),$$

(rounded if necessary) is an averaging aggregation function on a lattice.

PROOF. We only need to prove monotonicity. The proof follows the same lines as in [12], see also [2]. A convex function K, increasing on $[0, \infty)$ has the property

$$K(u) - K(v) \leq K(u') - K(v')$$

if $u-v \leq u'-v'$ and $u, v, u', v' \geq 0$, $u' > u$. Consider $\mathbf{x}$ and $\mathbf{x}'$, such that $x'_{ij} = x_{ij}$ for all i, j except one pair, $x'_{kl} = x_{kl} + 1$. We need to show that if $P(\mathbf{x}, y^*) \leq P(\mathbf{x}, y)$ for all $y \in \mathcal{L}$, then $P(\mathbf{x}', y^*) \leq P(\mathbf{x}', y)$ for all $y < y^*$, that is, the argument minimizing our function should increase.

Suppose $y < y^*$. Take

$$u = \sum_{j=1}^{m} = \sum_{j \neq l} |x_{kj} - y^*_j| + |x_{kl} + 1 - y^*_l| \ ,$$

$$v = \sum_{j=1}^{m} |x_{kj} - y^*_j| \ ,$$

$$u' = \sum_{j \neq l} |x_{kj} - y_j| + |x_{kl} + 1 - y_l|$$

and

$$v' = \sum_{j=1}^{m} |x_{kj} - y_j| \ .$$

Now,

$$u - v = |x_{kl} + 1 - y^*_l| - |x_{kl} - y^*_l| \ ,$$

which is equal to $+1$ if $y^*_l \leq x_{kl}$ or -1 otherwise. Also

$$u' - v' = |x_{kl} + 1 - y_l| - |x_{kl} - y_l|$$

which is also either $+1$ or -1, but because $y < y^*$, it takes value -1 only if $u-v = -1$. Hence $u-v \leq u'-v'$.

Now from $P(\mathbf{x}, y^*) \leq P(\mathbf{x}, y)$ we have

$$P(\mathbf{x}', y^*) = \sum_{i \neq k} K(d(x_i, y^*)) + K(u)$$

$$\leq \sum_{i \neq k} K(d(x_i, y)) + K(u') = P(\mathbf{x}', y)$$

□

Below we provide definitions for some specific instances of penalty based aggregation, based on the analogs with the classical means. In all cases we have penalties in the form (1), so Theorem 2 applies. We also round the results when necessary.

Definition 10 *Let $P(\mathbf{x}, y) = \sum_{i=1}^{n} d(x_i, y)^2$. Then the resulting aggregation function is the arithmetic mean.*

Definition 11 *Let $P(\mathbf{x}, y) = \sum_{i=1}^{n} w_i d(x_i, y)^2$, and $\mathbf{w}$ be a weighting vector, $w_i \geq 0$, $\sum w_i = 1$. Then the resulting aggregation function is a weighted arithmetic mean.*

Definition 12 *Let $P(\mathbf{x}, y) = \sum_{i=1}^{n} d(x_i, y)$. Then the resulting aggregation function is the median.*

4 SOLUTION TO PENALTY MINIMIZATION PROBLEMS

Consider now how to obtain solution to minimization problem in Definition 9. First, consider the arithmetic mean. We have the problem

$$minimize_y \sum_{i=1}^{n} \left(\sum_{j=1}^{m} |x_{ij} - y_j| \right)^2,$$

where x_{ij} denotes the j-th component of the i-th tuple $x_i \in \mathcal{L}$. We note that this problem is convex in y. We also note that the solution is different from the Cartesian product of the means, as the following example illustrates, and the differences are not just due to the rounding problem.

Example 1 *Let* $\mathcal{L}$ *be the product of two chains* $0, \ldots, 10$. *Take the mean of* $(10, 10), (8, 0), (3, 2)$. *The Cartesian product of means gives* $(7, 4)$, *with the objective value* $9^2 + 5^2 + 6^2 = 142$. *The solutions to the minimization problem are* $(9, 2)$ *with the objective* $9^2 + 3^2 + 6^2 = 126$ *and* $(8, 3)$ *with the same objective.*

While we could not obtain a closed form solution, we note that starting from any $y \in \mathcal{L}$, and in particular the Cartesian product of means or the medians, and performing coordinate descent (because of the convexity of the objective), one can reach the minimum algorithmically.

For the median consider the problem

$$minimize_y \sum_{i=1}^{n}\sum_{j=1}^{m} |x_{ij} - y_j| = \sum_{j=1}^{m}\sum_{i=1}^{n} |x_{ij} - y_j|.$$

Each term in the inner sum depend on y_j only, thus the problem is equivalent to m separate problems

$$\min_{y_j} \sum_{i=1}^{n} |x_{ij} - y_j|.$$

The solution is the median function. Hence the minimum is achieved at $y = (Med(x_{\cdot,1}), \ldots, Med(x_{\cdot,m}))$, i.e., the result is the Cartesian product of the medians.

5 CONCLUSIONS

Based on the representation of the classical mean and the median as solution to minimization problems, we have defined the mean and the median on discrete product lattices in the same way. We have shown that the median becomes the Cartesian product of the medians defined on discrete chains, and that the mean is not the Cartesian product of the respective means.

The product lattices that we have considered in this work appear in the image processing setting. The main application that we intend to develop of this work is aggregation of colors in this particular setting of image processing. We hope that our theoretical background applied to image processing, specially when considering median instead of means, will provide some advantages. There are two reasons. First, the median on a product lattice can be conveniently computed as the Cartesian product of the medians, whereas for the mean a more complex algorithm is required. Secondly, the median is much less sensitive to noisy values. In image processing non-Gaussian noise (e.g., salt and pepper noise) is quite common, and the use of the median allows one to filter out such noise efficiently.

An interesting line of research is to investigate the relation to aggregation of linguistic labels, in particular linguistic OWA [8, 9].

Acknowledgments

This research was partially supported by Grant TIN2007-65981.

References

[1] G. Beliakov, A. Pradera, and T. Calvo. *Aggregation Functions: A Guide for Practitioners.* Springer, Heidelberg, Berlin, New York, 2007.

[2] T. Calvo and G. Beliakov. Aggregation functions based on penalties. *Fuzzy Sets and Systems*, in press doi:10.1016/j.fss.2009.05.012, 2009.

[3] T. Calvo, G. Mayor, and R. Mesiar, editors. *Aggregation Operators. New Trends and Applications.* Physica-Verlag, Heidelberg, New York, 2002.

[4] T. Calvo, R. Mesiar, and R. Yager. Quantitative weights and aggregation. *IEEE Trans. on Fuzzy Systems*, 12:62–69, 2004.

[5] B. De Baets and R. Mesiar. Triangular norms on product lattices. *Fuzzy Sets and Systems*, 104:61–75, 1999.

[6] C. Gini. *Le Medie.* Unione Tipografico-Editorial Torinese, Milan (Russian translation, Srednie Velichiny, Statistica, Moscow, 1970), 1958.

[7] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap. *Aggregation Functions.* Cambridge University press, Cambridge, 2009.

[8] F. Herrera and E. Herrera-Viedma. Aggregation operators for linguistic weighted information. *IEEE Trans. on Syst., Man, and Cybernetics*, 27:646–656, 1997.

[9] F. Herrera and E. Herrera-Viedma. Linguistic decision analysis: Steps for solving decision problems under linguistic information. *Fuzzy Sets and Systems*, 115:67–82, 2000.

[10] S. Jenei and B. De Baets. On the direct decomposability of t-norms on product lattices. *Fuzzy Sets and Systems*, 139:699–707, 2003.

[11] G. Mayor and J. Monreal. Additive generators of discrete conjunctive aggregation operations. *IEEE Trans. on Fuzzy Systems*, 15:1046–1052, 2007.

[12] R. Mesiar. Fuzzy set approach to the utility, preference relations, and aggregation operators. *Europ. J. Oper. Res.*, 176:414–422, 2007.

[13] Y. Torra, V. Narukawa. *Modeling Decisions. Information Fusion and Aggregation Operators.* Springer, Berlin, Heidelberg, 2007.

[14] R. Yager and A. Rybalov. Understanding the median as a fusion operator. *Int. J. General Syst.*, 26:239–263, 1997.