About the Upper Bound of the Chiral Index of Multivariate Distributions

Michel Petitjean

DSV/iBiTec-S/SB2SM (CNRS URA 2096), CEA Saclay, 91191 Gif-sur-Yvette Cedex, France.
http://petitjeanmichel.free.fr/itoweb.petitjean.skewness.html

Abstract. A family of distributions in R^d having a chiral index greater or equal to a constant arbitrarily close to $1/2$ is exhibited. It is deduced that the upper bound of the chiral index lies in the interval $[1/2;1]$, for any dimension d.

Keywords: chiral index. multivariate skewness. Wasserstein distance.
PACS: 02.50.Sk, 02.40.Dr

INTRODUCTION

Symmetry is viewed since centuries as a dichotomic concept: there is or there is not symmetry. Chirality is related to indirect symmetry, as stated by Lord Kelvin in 1904 [1]. Intuitively, some symmetric physical systems having degenerated energy levels and offering a continuous separation of its energy levels induced by, or inducing a symmetry breaking, may be such that symmetry could itself offer continuous variations. More generally, the need of symmetry and chirality measures has given rise to an important mathematical literature in various areas (see [2] for a review).

The chiral index χ of a distribution has been defined in [3]. It is a real number returning a value in the interval $[0;1]$, the value 0 characterizing an achiral distribution (improperly called symmetric, such as a Gaussian). The chiral index of the distribution of a random vector is just a measure of its degree of skewness.

The original definition of χ involves a random variable (X_c,X), defined over a probability space $(C,A,P_c) \otimes (R^d,B,P)$. C is a non empty set called the space of colors, A is a σ-algebra defined on C, and P_c is a probability measure on (C,A). B is the Borel σ-algebra of R^d, and P is a probability measure on (R^d,B).

In its more general form, the chiral index can indeed reach the maximal value 1, and examples of maximal chirality distributions have been exhibited [3, 4].

When the distribution in the space of colors is such that the X_c is almost surely constant, i.e. $P_c(X_c = c_0) = 1$, the chiral index depends only on the dsitribution P of the random vector X. In this situation, it is simpler to work only with the random variable X over the probability space (R^d,B,P). The chiral index is just a multivariate skewness measure, which is null if and only if the d-variate distribution P of X has indirect symmetry.

The upper bound of $\chi(P)$ was shown to be $1/2$ in the univariate case [3], and was shown to lie in the interval $[1-1/\pi; 1-1/2\pi]$ in the bivariate case [5]. No results are currently available in dimension 3 and higher. We show that the upper bound lies in the

CP1073, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering—28th International Workshop*
edited by M. de Souza Lauretto, C. A. de Bragança Pereira, and J. M. Stern

interval $[1/2;1]$ in dimension $d \geq 3$, by exhibiting a family of d-variate distributions having a chiral index arbitrarily close to $1/2$.

FRAMEWORK AND NOTATIONS

We consider two random vectors X and $\tilde{X}$ in R^d, such that $\tilde{X}$ is distributed as $Q \cdot R \cdot (X + t)$, where t is a translation, R a rotation, and Q an inversion operator, i.e. an orthogonal (d,d) matrix with $det(Q) = -1$. Let V be the variance matrix of the distribution. The trace $T = Tr(V)$ is its inertia, and it is assumed to be finite and non null. We denote by $\{W\}$ the set of joint distributions of the couple $(X,\tilde{X})$, and the quote indicates the matricial transposition operator. The chiral index is defined from the Wasserstein distance Δ [6], between the distribution P of X and the distribution $\tilde{P}$ of $\tilde{X}$, D being minimized for all rotations R and translations t applied to $\tilde{X}$:

$$\chi = d \cdot D^2/4T \tag{1}$$
$$D = Min_{\{R,t\}}\Delta \tag{2}$$
$$\Delta^2 = Inf_{\{W\}}E(X-\tilde{X})' \cdot (X-\tilde{X}) \tag{3}$$

The chiral index of $P(X)$ does not depend on the inversion Q, and it is insensitive to rotations, translations, inversions, and scaling of X. The Wasserstein distance is minimized when the expectation $E(X-\tilde{X})$ is null. Thus, we can assume without loss of generality, that $E(X) = 0$. Of course the chiral index is null if and only if the distribution is symmetric, in the sense of an indirect symmetry (i.e. a mirror symmetry).

For clarity, the distributions satisfying to the condition (4) are called here isoinertial distributions, where σ is any positive real constant and I is the identity matrix:

$$V = \sigma^2 \cdot I \tag{4}$$

When the condition (4) is satisfied, it stands for any rotation, translation, inversion, and scaling of X. It should be pointed out that the maximal value $\chi = 1$ is reachable only for isoinertial distributions when the general model of chirality is involved: see equations (3.9) and (3.10) in [3], and see [7] for a presentation of this general model. The model of chirality considered is just a particular case of the general model. It is why we conjecture that, in the case of d-variate distributions, the upper bound of the chiral index could be asymptotically reached inside a family of isoinertial distributions.

We consider a random vector X with finite inertia, V_X is the variance matrix of X, assumed to be of full rank. L_X is the diagonal matrix of eigenvalues of V_X, and U_X is a rotation matrix of eigenvectors of V_X, such that: $V_X \cdot U_X' = U_X' \cdot L_X$. The centered random vector $L_X^{-1/2} \cdot U_X \cdot (X - E(X))$ is isoinertial, because its variance matrix is the identity I.

In order to introduce some other useful notations, we build below an isoinertial finite discrete distribution of n points in R^d, from an arbitrary array Y of n lines and d columns

($n > d$), such that the line i of Y is a point to which the mass (i.e. the probability) m_i is attached ($m_i > 0$). We define the square diagonal matrix M of order n, such that the diagonal element is $M_{i,i} = m_i, i = 1..n$. It follows that $\mathbf{1}' \cdot M \cdot \mathbf{1} = 1$. We define also the square matrix B of order n, which operates as a mass-centering operator: $B = I - \mathbf{1} \cdot \mathbf{1}' \cdot M$, i.e. $(\mathbf{1}' \cdot M) \cdot Y$ is the transposed center of mass, the array $B \cdot Y$ is the mass-centered array, because $\mathbf{1}' \cdot M \cdot B = 0$ and thus $(\mathbf{1}' \cdot M) \cdot (B \cdot Y) = 0$. In addition, we assume that $(B \cdot Y)$ is of full rank.

The variance matrix of $(B \cdot Y)$ is: $V_Y = (Y' \cdot B') \cdot M \cdot (B \cdot Y)$. Let U be the rotation applied to the points of the centered array $B \cdot Y$, such that V_Y is diagonalized: $V_Y \cdot U' = U' \cdot L$, where L is the diagonal matrix of the eigenvalues of V_Y. It follows that the matrix Y_I in equation (5), which has n lines and d columns, defines a finite discrete isoinertial distribution, such that the line i of Y_I is a point to which the mass m_i is attached.

$$Y_I = B \cdot Y \cdot U' \cdot L^{-1/2} \tag{5}$$

The proof is as follows. The full rank matrix Y_I is centered: $(\mathbf{1}' \cdot M) \cdot Y_I = 0$ because $\mathbf{1}' \cdot M \cdot B = 0$, and the variance of Y_I is: $V = (L^{-1/2} \cdot U \cdot Y' \cdot B') \cdot M \cdot (B \cdot Y \cdot U' \cdot L^{-1/2})$, and since we have $U \cdot (Y' \cdot B' \cdot M \cdot B \cdot Y) \cdot U' = L$, thus $V = I$.

When $m_i = 1/n$ for $i = 1..n$ and $n = d+1$, we get the $d+1$ equally weighted vertices of an isoinertial simplex, which is known to be regular and achiral: see appendix 2 in [8].

ATTEMPT TO BUILD MAXIMAL CHIRALITY DISTRIBUTIONS

We consider now a family of finite discrete distributions Z parametrized by the positive quantity ε, Z being asymptotically isoinertial when ε is tending to zero. We set $\mu = 1/\varepsilon$, $n = d+1$, and, using the notations introduced in the previous section, we define $Z = B \cdot Y$, with:

$$Y = \left(\begin{array}{cccc} 0 & 0 & \ldots & 0 \\ \hline \mu & 0 & \ldots & 0 \\ 0 & \mu^2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \mu^d \end{array}\right) \qquad M = \frac{1}{c}\left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & \varepsilon^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \varepsilon^{2d} \end{array}\right)$$

The normalizing constant $c = (1 + \varepsilon^2 + \ldots + \varepsilon^{2d})$ is such that $\mathbf{1}' \cdot M \cdot \mathbf{1} = 1$. Z is centered, and its variance matrix is $V = Y' \cdot B' \cdot M \cdot B \cdot Z$. From the definition of B, we have $B' \cdot M \cdot B = M - M \cdot \mathbf{1} \cdot \mathbf{1}' \cdot M$, and then come V and $T = Tr(V)$:

$$V = \frac{1}{c} I - \frac{1}{c^2} \begin{pmatrix} \varepsilon^2 & \varepsilon^3 & \ldots & \varepsilon^{d+1} \\ \varepsilon^3 & \varepsilon^4 & \ldots & \varepsilon^{d+2} \\ \vdots & \vdots & & \vdots \\ \varepsilon^{d+1} & \varepsilon^{d+2} & \ldots & \varepsilon^{2d} \end{pmatrix} \tag{6}$$

$$T = (d/c) - (c-1)/c^2 \quad (7)$$

Z is asymptotically isoinertial, because: $lim_{(\varepsilon \to 0)}\{c\} = 1$, and then: $lim_{(\varepsilon \to 0)}\{V\} = I$. We set $\tilde{Z} = Z \cdot Q' \cdot R'$, and from equations (1)-(3), we express the chiral index χ_Z of the centered finite discrete distribution Z as follows:

$$\chi_Z = \frac{d}{4T} Min_{\{R,W\}} \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} W_{i,j} d_{i,j}^2 \quad (8)$$

In equation (8), the joint density is a bistochastic matrix W, and $d_{i,j}$ is the distance between the point i (line i of Z) and the point j (line j of $\tilde{Z}$). The set of bistochastic matrices $\{W\}$ is closed and bounded (and convex): it is why there is at least one optimal bistochatic matrix. Equation (8) is rewritten in matricial form:

$$\chi_Z = \frac{d}{2T}[T - Max_{\{R,W\}} Tr(Z' \cdot W \cdot Z \cdot Q' \cdot R')] \quad (9)$$

Noting that $W \cdot \mathbf{1} = M \cdot \mathbf{1}$ and $\mathbf{1}' \cdot W = \mathbf{1}' \cdot M$, we have:

$$B' \cdot W \cdot B = W - M \cdot \mathbf{1} \cdot \mathbf{1}' \cdot M \quad (10)$$
$$Z' \cdot W \cdot Z = Y' \cdot W \cdot Y - Y' \cdot M \cdot \mathbf{1} \cdot \mathbf{1}' \cdot M \cdot Y \quad (11)$$

We set $A = Y' \cdot M \cdot \mathbf{1} \cdot \mathbf{1}' \cdot M \cdot Y$, and we note that the elements of A are all non negative. Thus the quantity: $Tr(A \cdot Q' \cdot R')$ takes values in the interval $[-\alpha^2; +\alpha^2]$, where the constant $\alpha^2 = (\mathbf{1}' \cdot A \cdot \mathbf{1})$ does not depend on R and W. Then, we get from (11) the inequality (12), and from (9) and (12) we get (13)

$$Tr(Z' \cdot W \cdot Z \cdot Q' \cdot R') \leq Tr(Y' \cdot W \cdot Y \cdot Q' \cdot R') + \alpha^2 \quad (12)$$
$$\chi_Z \geq \frac{d}{2T}[T - \alpha^2 - Max_{\{R,W\}} Tr(Y' \cdot W \cdot Y \cdot Q' \cdot R')] \quad (13)$$

Furthermore, we note that:

$$A = \frac{1}{c^2} \begin{pmatrix} \varepsilon \\ \vdots \\ \varepsilon^d \end{pmatrix} \cdot \begin{pmatrix} \varepsilon \\ \vdots \\ \varepsilon^d \end{pmatrix}' \quad (14)$$

$$lim_{(\varepsilon \to 0)}\{\alpha^2\} = 0 \quad (15)$$

We set $S = Tr(Y' \cdot W \cdot Y \cdot Q' \cdot R')$, and we denote by $q_{i,j}$ the element at line i and column j of the matrix $(R \cdot Q)$. Each element $q_{i,j}$ is upper bounded by 1 in absolute value. We express S with the elements of W:

$$S = \sum_{i=1}^{i=d} \sum_{j=1}^{j=d} (\mu^{i+j} W_{i+1,j+1}) q_{i,j} \quad (16)$$

Since each element $W_{i,j}$ is upper bounded by $\sum_{i=1}^{i=d} W_{i,j}$ and also by $\sum_{j=1}^{j=d} W_{i,j}$, we have:

$$W_{i+1,j+1} \leq Min\{\varepsilon^{2i}/c; \varepsilon^{2j}/c\} \quad (17)$$
$$(\mu^{i+j} W_{i+1,j+1}) q_{i,j} \leq \varepsilon^{|i-j|}/c \quad (18)$$

We use now the condition $\varepsilon \leq 1$, and we report it in (18) to get (19), and from (16) and (19) we get (20):

$$(\mu^{i+j} W_{i+1,j+1}) q_{i,j} \leq \varepsilon/c \qquad i \neq j \quad (19)$$
$$S \leq (d(d-1)\varepsilon/c) + \sum_{i=1}^{i=d} (\mu^{2i} W_{i+1,i+1}) q_{i,i} \quad (20)$$
$$Max_{\{R,W\}} S \leq (d(d-1)\varepsilon/c) + Max_{\{R,W\}} [\sum_{i=1}^{i=d} (\mu^{2i} W_{i+1,i+1}) q_{i,i}] \quad (21)$$

The term to be maximized in the right term of equation (21) is a linear combination of the quantities $q_{i,i}$, which are the diagonal elements of the orthogonal matrix $(R \cdot Q)$, this latter having a determinant equal to -1. The matrix Q being a fixed constant orthogonal matrix, it is equivalent to work with the unknown matrix (RQ) rather than with the unknown rotation R, because there is a one to one correspondence between the set of unknowns $\{R\}$ and the set of unknowns $\{(RQ)\}$).

We partition the set of unknowns $\{(RQ), W\}$ into two disjoint subsets, one containing the unknowns such that all $q_{i,i}$ are non negative: $\{(RQ)^+, W\}$, the other, $\{(RQ)^-, W\}$, being its complement to $\{(RQ), W\}$.

From (17), we get:

$$Max_{\{(RQ)^+,W\}} [\sum_{i=1}^{i=d} (\mu^{2i} W_{i+1,i+1}) q_{i,i}] \leq Tr(RQ)/c \quad (22)$$

The eigenvalues of (RQ) are either -1, $+1$, or couples of conjugated complex numbers. It follows that $Tr(RQ) \leq (d-2)$, and the inequality (22) becomes:

$$Max_{\{(RQ)^+,W\}} [\sum_{i=1}^{i=d} (\mu^{2i} W_{i+1,i+1}) q_{i,i}] \leq (d-2)/c \quad (23)$$

Now, maximizing on $\{(RQ)^-, W\}$ leads to set $W_{i+1,i+1} = 0$ when a negative $q_{i,i}$ value is encountered. Since at least one negative $q_{i,i}$ value is ensured, and using again (17), we get the inequality:

$$Max_{\{(RQ)^-,W\}} [\sum_{i=1}^{i=d} (\mu^{2i} W_{i+1,i+1}) q_{i,i}] \leq (d-1)/c \quad (24)$$

Using (23) and (24) gives (25), which is reported in (21) to get the upper bound of S:

$$Max_{\{(RQ),W\}}[\sum_{i=1}^{i=d}(\mu^{2i}W_{i+1,i+1})q_{i,i}] \leq (d-1)/c \tag{25}$$

$$Max_{\{R,W\}}S \leq (d(d-1)\varepsilon/c)+(d-1)/c \tag{26}$$

We report now (26) in (13) to get a lower bound of the chiral index:

$$\chi_Z \geq \frac{d}{2T}[T-\alpha^2-[(d(d-1)\varepsilon/c)+(d-1)/c]] \tag{27}$$

When ε is arbitrarily small, using (7) and (15), it comes from (27) that χ_Z is greater or equal to a constant arbitrarily close to $1/2$. Thus we have exhibited a family of distributions for which the upper bound of the chiral index cannot be smaller than $1/2$.

CONCLUSION

The upper bound of the chiral index, taken over the set of d-variate distributions with finite and non null inertia, has been shown to lie in the interval $[1/2;1]$. Despite that this interval has been shortened for $d=2$, calculating this upper bound is an open problem for $d \geq 2$.

ACKNOWLEDGMENTS

I thank very much Professor Julio Stern and the MaxEnt 2008 Organizing Committee for having invited me to present this work at the MaxEnt 2008 Conference.

REFERENCES

1. Lord Kelvin, *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*, C. J. Clay and Sons, Cambridge University Press Warehouse, London, 1904, Appendix H., sect. 22, footnote p. 619.
2. M. Petitjean, "Chirality and Symmetry Measures: A Transdisciplinary Review," *Entropy* **5**[3], 271–312 (2003).
3. M. Petitjean, "Chiral Mixtures," *J. Math. Phys.* **43**[8], 4147–4157 (2002).
4. M. Petitjean, "On the Root Mean Square Quantitative Chirality and Quantitative Symmetry Measures," *J. Math. Phys.* **40**[9], 4587–4595 (1999).
5. D. Coppersmith, M. Petitjean, "About the Optimal Density Associated to the Chiral Index of a Sample from a Bivariate Distribution," *Compt. Rend. Acad. Sci. Paris, Série I (mathématiques)* **340**[8], 599–604 (2005).
6. S. T. Rachev, *Probability Metrics and the Stability of Stochastic Models*, Wiley, New-York, 1991.
7. M. Petitjean, "From Shape Similarity to Shape Complementarity: toward a Docking Theory," *J. Math. Chem.* **35**[3], 147–158 (2004).
8. M. Petitjean, "About Second Kind Continuous Chirality Measures. 1. Planar Sets," *J. Math. Chem.* **22**[2-4], 185–201 (1997).